ADVANCES IN INTELLIGENT AND SOFT COMPUTING 143

Min Zhu (Ed.)

Business, Economics, Financial Sciences, and Management

Springer

Advances in Intelligent and Soft Computing 143

Editor-in-Chief: J. Kacprzyk

Advances in Intelligent and Soft Computing

Editor-in-Chief

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

Further volumes of this series can be found on our homepage: springer.com

Vol. 126. Khine Soe Thaung (Ed.)
Advanced Information Technology in Education, 2012
ISBN 978-3-642-25907-4

Vol. 127. Tianbiao Zhang (Ed.)
Instrumentation, Measurement, Circuits and Systems, 2012
ISBN 978-3-642-27333-9

Vol. 128. David Jin and Sally Lin (Eds.)
Advances in Multimedia, Software Engineering and Computing Vol.1, 2011
ISBN 978-3-642-25988-3

Vol. 129. David Jin and Sally Lin (Eds.)
Advances in Multimedia, Software Engineering and Computing Vol.2, 2011
ISBN 978-3-642-25985-2

Vol. 130. Kusum Deep, Atulya Nagar, Millie Pant, and Jagdish Chand Bansal (Eds.)
Proceedings of the International Conference on Soft Computing for Problem Solving (SocProS 2011) December 20–22, 2011, 2012
ISBN 978-81-322-0486-2

Vol. 131. Kusum Deep, Atulya Nagar, Millie Pant, and Jagdish Chand Bansal (Eds.)
Proceedings of the International Conference on Soft Computing for Problem Solving (SocProS 2011) December 20–22, 2011, 2012
ISBN 978-81-322-0490-9

Vol. 132. Suresh Chandra Satapathy, P.S. Avadhani, and Ajith Abraham (Eds.)
Proceedings of the International Conference on Information Systems Design and Intelligent Applications 2012 (India 2012) held in Visakhapatnam, India, January 2012, 2012
ISBN 978-3-642-27442-8

Vol. 133. Sabo Sambath and Egui Zhu (Eds.)
Frontiers in Computer Education, 2012
ISBN 978-3-642-27551-7

Vol. 134. Egui Zhu and Sabo Sambath (Eds.)
Information Technology and Agricultural Engineering, 2012
ISBN 978-3-642-27536-4

Vol. 135. Honghua Tan (Ed.)
Knowledge Discovery and Data Mining, 2012
ISBN 978-3-642-27707-8

Vol. 136. Honghua Tan (Ed.)
Technology for Education and Learning, 2012
ISBN 978-3-642-27710-8

Vol. 137. Jia Luo (Ed.)
Affective Computing and Intelligent Interaction, 2012
ISBN 978-3-642-27865-5

Vol. 138. Gary Lee (Ed.)
Advances in Intelligent Systems, 2012
ISBN 978-3-642-27868-6

Vol. 139. Anne Xie and Xiong Huang (Eds.)
Advances in Electrical Engineering and Automation, 2012
ISBN 978-3-642-27950-8

Vol. 140. Anne Xie and Xiong Huang (Eds.)
Advances in Computer Science and Education, 2012
ISBN 978-3-642-27950-8

Vol. 141. Dehuai Zeng (Ed.)
Advances in Computer Science and Engineering, 2012
ISBN 978-3-642-27947-8

Vol. 142. Gary Lee (Ed.)
Advances in Computational Environment Science, 2012
ISBN 978-3-642-27956-0

Vol. 143. Min Zhu (Ed.)
Business, Economics, Financial Sciences, and Management, 2012
ISBN 978-3-642-27965-2

Min Zhu (Ed.)

Business, Economics, Financial Sciences, and Management

Editor
Min Zhu
Nanchang University
Nanchang Jaingxi
China, People's Republic

ISSN 1867-5662 e-ISSN 1867-5670
ISBN 978-3-642-27965-2 e-ISBN 978-3-642-27966-9
DOI 10.1007/978-3-642-27966-9
Springer Heidelberg New York Dordrecht London

Library of Congress Control Number: 2011946098

Printed on acid-free paper

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

Preface

The 2011 International Conference on Business, Economics, and Financial Sciences, Management (BEFM 2011) will be held on December 30–31, 2011 in Jeju Island, Korea.

Management in all business and organizational activities is the act of getting people together to accomplish desired goals and objectives using available resources efficiently and effectively. Management comprises planning, organizing, staffing, leading or directing, and controlling an organization (a group of one or more people or entities) or effort for the purpose of accomplishing a goal. Resorting encompasses the deployment and manipulation of human resources, financial resources, technological resources and natural resources.

Since organizations can be viewed as systems, management can also be defined as human action, including design, to facilitate the production of useful outcomes from a system. This view opens the opportunity to 'manage' oneself, a pre-requisite to attempting to manage others.

BEFM 2011 will be the most comprehensive Conference focused on the various aspects of advances in Business, Economics, and Financial Sciences, Management. Our Conference provides a chance for academic and industry professionals to discuss recent progress in the area of Business, Economics, and Financial Sciences, Management. The goal of this Conference is to bring together the researchers from academia and industry as well as practitioners to share ideas, problems and solutions relating to the multifaceted aspects of Business, Economics, and Financial Sciences, Management.

Topics cover Banking and Finance, Business and Industrial Marketing, Business Development, Business Ethics and Anti-corruption, Business Process Management, Computational Economics and Finance, Comparative Economic Systems, Consumer Marketing, Corporate Finance and Governance, Crisis Management, Development Planning and Policy, Economic Development, Economic Methodology, Economic Policy (Fiscal Policy/Monetary Policy etc), Economic Systems, Financial Economics, Financial Systems Development Growth, Aggregate Productivity, Household Behavior and Family Economics, Information Management, Innovation and Technology Management, International Economics, International Finance, Islamic Finance and Banking, Knowledge Management, Labor Economics, Law and Economics, Market Structure and Pricing, Measurement of Economic and Social Performance, Prices, Business Fluctuations, and Cycles, Production and Organizations, Public Choice, Public Economics, Public Finance, Regional Economic Development, Regulatory Economics, Strategic Management, Technological Change; Research and Development, Tourism and Hospitality Management, Transparency and Accountability, Welfare Economics.

We encompass lots of papers in the related fields. To ensure the high quality of the proceedings, each paper has been strictly peer-reviewed, and low quality ones have been rejected.

I express my deep gratitude to all contributors of this book who worked very hard to make this project successful. If our efforts came to the present result, it was not without ideas, encouragement and support from Springer.

Min Zhu
Publication Chair

Committee

Keynote Speaker Chairs

Andrew Hunter	Kuala Lumpur ACM Chapter, Malaysia
Gerald Schaefer	Loughborough University

Scholarship Committee Chairs

Qihai Zhou	Southwestern University of Finance and Economics, China
Min Zhu	Nanchang University, China

Program Committee Chairs

Dehuai Yang	Huazhong Normal University, China
Wensong Hu	Nanchang University, China

Publication Chair

Min Zhu	Nanchang University, China

International Committees

David Lee	American Applied Sciences Research Institute, USA
Wei Deng	American Applied Sciences Research Institute, USA
Ping He	Dalian Maritime University, China
Wei Liu	Wuhan University, China
Weitao Zheng	Wuhan University of Technology, China
Min Zhu	Nanchang University, China
Andrew Hunter	Kuala Lumpur ACM Chapter, Malaysia
Gerald Schaefer	Loughborough University
Dehuai Yang	Huazhong Normal University, China
Wensong Hu	Nanchang University, China
Yuanzhi Wang	Anqing Teachers' University, China
Junwu Zhu	University of Western Sydney, Australia
Minli Dai	Suzhou University, China
Zhenghong Wu	East China Normal University, China

Contents

Health Economics and Management Framework Based on Chongqing Regional Health Information Platform

Ping Cheng

The Key R&D Center for Finance of Chongqing,
Chongqing University of Technology, Chongqing, China
Chgpg2006@163.com

Abstract. The hospital health economy management mainly take is appraised the object the economic efficiency as the core goal, but regarding the hygienic economy management most core health resources disposition management, medical and health organization economy contents and so on management target and method has not involved nearly completely. Changes the work based on the national new medicine for the background medical department Chongqing four experiment site project, analyzes the new medicine to change the economical management question which under the background various medical and health organization faces, proposed one based on the Chongqing region health information platform hygienic economy management frame. The expectation through this article research, can promote our country's hygienic economy supervisory work, makes a contribution for our country's medicine health organizational reform.

Keywords: medical institutions, regional health information platform, economic management, framework.

1 Introduction

Over the past 40 years, the vast majority of countries in the world rapid increases in the cost of health care, health economics and management in the management of health care is becoming more and more prominent. Distribution of social resources to the health sector, as well as the resources within the health industry on how to configure, determines the overall accessibility of health services and social fairness, as one of the core elements of health service management.

Our country's medical and health organization overwhelming majority all is the public welfare institution, theoretically the medical and health organization the major part resources should certainly originate from the financial payment, but as a result of the historical reason, in our country reform and open policy flood tide, the medical service organization lost “the public welfare” gradually the nature, the hospital economic efficiency has once become the hospital superintendent and the public health policy-maker with emphasis consideration major event, has neglected the medical service organization essence was provides the public welfare for the people medical safeguard to serve this core question, thus caused the hospital construction and the national health service develops moves towards the wrong road, “saw a doctor

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 1–5.
springerlink.com

difficultly, to see a doctor expensively" becomes affects our country socialism enterprise to develop, the influence society harmoniously One of important reasons. CPC Central Committee and State Council in March 2009 release of "new medical reform" policy made it clear that the deepening health system reform must adhere to the guiding ideology of the public nature of public health, The "universal access to basic health services as the fundamental starting point and goal, from the reform program design, health system was established to service system should follow the principle of public welfare, health care system as the basic public goods to all people". At the same time published the programmed for the medical and health system focused on implementation in the near future (2009-2011), further clearly proposes that the key to resolve the masses more "difficulty and high cost of" problem. To solve the "high cost" we'll have to cut overall health care costs of the community to address the "difficulty" is solve the accessibility problems of medical services, these two issues are the core elements of health economics and management.

Our country Hospital informationization construction starts from on century 80's to start, at the end of the 90's enter the informationization construction the fast development time, in the recent ten years enter the comprehensive construction the high tide. The majority are from the hospital information system, the progressive development financial management began to medications essay management, clinical information management, individual hospitals, such as the "Digital hospital" to build a comprehensive hospital business management system. However, the process of building the hospital decision-makers in the hospital ignored the medical and health services in the "public interest" requirement, which the hospital management of major health economic evaluation of the object to be the core economic objectives, specific performance for all types of medical income, net of computable costs assigned to the department by a certain percentage or staff, so-called "performance appraisal", but the core of economic management for the health of medical and health resources management, health care institutions and methods of economic management indicators and almost entirely not involved.

Medical and health institutions of economic management of health care not only to monitor the economic performance within the organization, but also to a shift of resources within the region on how to configure the health care industry, health industry, how to configure the internal resources to healthcare facilities, medical staff and medical services within the scope of people's problems. Therefore, this article changes the work based on the national new medicine for the background medical department Chongqing four experiment site project (Ministry of Finance, the medical department 2010 [wealth society 2010) 270th document), includes specifically based on the electronic healthy file, the electronic medical record, the outpatient service overall plan management basic unit health information system constructs the experiment site project, the public hospital long-distance medical consultation system experiment site project, county level hospital ability constructs the experiment site project, the village clinic constructs the experiment site project and so on is a background, through analyzes the new medicine to change the economical management question which under the background various medical and health organization faces, proposed one based on the Chongqing region health information platform hygienic economy management frame. Hope that through this study, for the country's health system reform, to make a contribution.

2 Health Economics and Management Framework Based on Regional Health Information Platform

Take the Chongqing region health information platform as the foundation, analyzes health economical data which various medical and health organization produces, studies each kind of hygienic economic indicator and the hospital health economy management level relevance, seeks under the new situation the medical establishment health economy management suitable target and the inspection means, unifies various medical establishment medical service safeguard service to complete the situation, for the establishment city, the area county two level of government health resources disposition decision analysis system, is the realization “everybody enjoys the basic medical health care” the goal builds the solid foundation.

2.1 Health Demand Appraisal Target and Method

Hospital coverage within the actual medical needs of the population, and basic health care needs and population, age distribution, the spectrum of disease epidemiological indicators, such as personal or household income, medical care prices and other factors related to the study on the establishment of mathematical model for evaluating medical needs, the indicators for correlation with the actual occurrence of medical business, the filter that can reflect the residents ' medical needs indicators and methods.

2.2 Overall Medical Service Cost Appraisal Target and Method

Medical cost includes direct medical cost and indirect medical cost. The direct medical cost refers to the patient to receive a medical examination - cures the medical expense which the entire process produces, the indirect medical cost includes the opportunity cost which the patient receives a medical examination occurs (choice different medical establishment difference which possibly creates to seeing a doctor result), the time cost (except receives a medical examination beside time which itself spends, but also includes time and so on transportation, line, appointment, thus as well as treats many times causes disease to progress influence disease extension to turn over to with change and so on prognosis costs), the family member and the relatives and friends revolve the transportation which the patient receives a medical examination occurs, to delay work, nursing, the nutrition, miss a mealtime and so on the overhead costs and so on. This topic will choose 2~3 kind of different types diseases to take representative, will study disease overall medical service cost assessment method, the screening being suitable appraisal target, exploration establishment disease diagnoses the overall medical service cost the appraisal system.

2.3 Medical Effect Appraisal Target and Method

Health of the residents of quantitative currency value is often difficult to use the method for the assessment, and use of cost-effectiveness analysis method to study achieved the expected individual health goals, and explore the cost of health care commensurate with different medical institutions, establish a set goals for different grades and different levels of health-care institutions medical effectiveness evaluation indicators and methods.

2.4 New Medical Measure Appraisal Target and Method

Using the cost-utility analysis methods to study a single comprehensive medical measures of medical items or specific cost of each additional quality-adjusted life years, exploring material impact indicators and methods for evaluation of new medical projects, further explore the establishment of a new type of medical projects economic efficiency forecast and analysis of indicators and methodologies.

2.5 Region Health Economy Macroscopic Economy Monitoring Method

In Chongqing's regional health information platform based on the select similar business of diagnosis and treatment in medical institutions at all levels horizontal comparison of economic indicators, focusing on monitoring of similar indicators in more than four projects, diagnosis and treatment of patients, prescriptions, medicines, and other activities of economic indicators. To single prescription costs monitoring for cases, can statistics a medical institutions (for doctor), and a regional (for doctor or medical institutions), and city (for doctor, and medical institutions or County Health Council) in the single prescription over four min bit number of amount, over four min bit number amount of prescription is included monitoring range; and as filter focus monitoring of check project, both can select price over all check project in the over four min bit number of project, also can select open single frequency (number) over four min bit number of project for regulatory ; Such as drugs, you can select in the median frequency on more than a quarter of prescription drugs, you can also choose to, you can also select unit price, you can select indicators such as cost, monitoring of the same type of data than the previous four prescriptions, doctors, and so on. Through this class designed specifically for high-cost medical behavior of monitoring, sustainable, reduce health care costs, guaranteed limited medical health funds as possible can be used properly, thus filtering out regional health effective indicators of economic macro-management, establish a set of macroeconomic management.

2.6 Economic Evaluation of Medical and Health Institutions

Uses the crosswise contrast, the longitudinal contrast, the longitudinal contrast crosswise analysis and so on the many kinds of method appraisal similar medical establishment's hygienic economic effect. Crosswise contrast analysis method main appraisal single doctor makes out bill or list with this medical establishment, the home court county, the whole city average level comparison (inspects doctor's medical efficiency), single medical establishment and home court county, whole city scope average level comparison (inspection medical and health organization's medical efficiency), some area county and whole city scope average level comparison (inspection area county health bureau's supervising and managing level).The longitudinal contrast analysis mainly is this medical establishment when the time data and the history the data carries on the contrast same time, thus contrast analysis medical establishment health economy management level historical change.Then the longitudinal contrast crosswise analysis unifies the longitudinal contrast analysis and the crosswise contrast analysis, the contrast single patient, the single kind of disease, single doctor, the single medical establishment, the sole area county and the history in

the data average level compares same time, and carries on the comparison with the longitudinal contrast analysis result, may inspect doctor, the medical establishment health economic efficiency improvement degree.

3 Conclusion

Based on the regional health information platform of Chongqing health economic management framework can help build up a set of Chongqing medical institutions based on the health of economic management method suitable, sure a batch of effective to evaluate the effect of health resources use hospital microeconomics evaluation tool, single out a set of effective cost-benefit analysis, cost-effectiveness analysis, medical and health supplies-demand analysis, medical behavior many aspects such as the marginal cost of health resources evaluation of the use effect of the economic index, the competent industrial department of health of Chongqing city efficient allocation of resources provide health range decision support. In addition, in this paper on the basis of research achievements, can develop the relevant software system, according to the theme of the selected establish a series of theme database, according to the statistics of the different dimensions and particle size generate statistics analysis report at all levels, the internal relationship between the index, with more convenient and effective services in health economic management.

References

1. Dangqin: Discussion on cost accounting in hospital management under market economy. Accountants 2, 106–107 (2011)
2. Xiao, J.: How to strengthen hospitals under the conditions of the market economy management. Economist 2, 175 (2011)
3. Zhang, P., Song, P.: On financial management in hospitals under the new situation and method. The Chinese Market 1, 89–90 (2011)
4. Wei, R.: On the reform of medical hospital financial management strengthening. Financial Management 9, 56–58 (2009)
5. Guo, Z.: Problems Hospital financial management and countermeasure. Economic Research Guide 23 (2009)
6. Cai, J.: Problems Hospital financial management and countermeasure. Henan Journal of Preventive Medicine 3 (2007)
7. Xia, Y.: Cost accounting for hospital compliance issues. Health Economics Research 2 (2004)
8. Tang, Z.: Current situation, difficulties and countermeasures of hospital cost accounting. Health Economics Research 2 (2005)
9. Zhang, Y.: My views on the reinforcement of hospital cost accounting. Health Economics Research 12 (2004)

The Strategic Thinking of the Economic Transition in the Economic Zone on the West Side of the Taiwan Straits China

Pengyi Zhu*

School of Economics, Fujian Normal University, P.R. China, 350108
zhupengyi@hotmail.com

Abstract. The Economic Zone on the West Side of the Taiwan Straits, taking Fujian province as the main body, spans multiple administrative regions in China. In accordance with the law of market economy and intra-economic links, the zone will break through the boundaries of administrative regions and progressively establish inter-provincial economic regions, the economic transitions are an internal requirement of economic development in the zone. Based on the regional condition, the ways to accelerate the economic transition in the zone are analyzed from two aspects in this paper, including optimizing the benefit of the zone as a whole, changing the growth pattern of the traditional economy, The two ways will promote the transitions and sustainable development of the economy in the zone.

Keywords: Economic transition, strategic thinking, the Economic Zone on the West Side of the Taiwan Straits.

1 Introduction

Many scholars have interpreted the economy transition from different angles. It is considered that the economic transition means one kind of economic operation turn to another, including the transformation of the economic growth pattern, the driving force and pathway of economic development as well as the industrial distribution and space organization[1], or it has the idea that the economy transition has been focusing on economic life, associated with institutional qualitative change of ensemble, including change of eco-resources, historical tradition, religion, ideology and national policy[2]. In China, the most urgent task in the transitional process is structural adjustments[3],it could also say the economy transition is taking economic restructuring as paramount[4]. Its main purposes are changed from the labor-intensive industries to the capital-intensive industries, further to technique-intensive industries, resulting in the increase of regional competitiveness.

The Economic Zone on the West Side of the Taiwan Straits, taking Fujian province as the main part, covers around the southern part of Zhejiang province, the northern part of Guangdong province and part areas of Jiangxi province, and relies on the

* Author: Peng-yi Zhu (1968-), male, Associate Professor, School of Economics, Fujian Normal University. Major in Economics of Management Research.

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 7–12.
springerlink.com

circumlittoral center city and urban agglomerations in southeast China to form the economic zone opening to the world. The north of the zone borders on the Yangtze River Delta, the south connects the Pearl River Delta, the east is the Taiwan Strists separated with Taiwanese region, the west is linking the vast inland region through Jiangxi province, it is the important constituent for the coastal economic belts in China.

"Some opinions about supporting Fujian province to speed the construction for the Economic Zone on the West Side of the Taiwan Straits" made by the China State Council [5] (called "Opinion" for short) requires the zone to be constructed in a "harmonious region" and "sustainable development region".

At present, Fujian enters a moderately developed economies with per capital GDP of 4800 U.S.dollar, it is key period for going to developed economies. The economic transition has become an internal need. Based on the regional condition, the ways to accelerate the economic transition in the zone are analyzed from two aspects in this paper.

2 Optimizing Benefit of the Zone as a Whole

The Economic Zone on the West Side of the Taiwan Straits is consisting of multi-regions including several provinces and cities, whole optimization can only be effected through the collaborative process.

Synergetics considered that open systems can become the well-organized structure by means of synergistic effect of the sub-systems ,it alter the component of dispersion even mutual conflict into the well-organized join force with whole function[6]. While the whole function has some new properties, which are never possessed in the sub-system [7].

Under the background of the economic globalization, the economic advantage is hard to develop in separate district, and will be well develop in multi-districts formed high-efficient and high-ordering integration in the cooperative process, exploiting their regional whole advantage. The zone is facing this situation.

Lin Yi-fu considered[8], the developing country formulates corresponding policy based on the endowment factor to develop industry of comparative advantage, it will be able to carry on the industrial structural transition, and surpass the developed country.

There are economic differences in the location, industry and foundation between sub-regions of the Economic Zone on the West Side of the Taiwan Straits, it should be optimizing the benefit of the zone as a whole ,cooperating development on the basis of the regional divided the work, it is useful in complementation of advantages and realizing the win-win result. At present, the synergic relationship had been developing in the zone, Shantou city of Guangdong province, Wenzhou city of Zhejiang province as well as Hunan, Jiangxi, Sichuan, Guizhou provinces participated in the economic activities of the Economic Zone in different ways. Fujian also adopted a series measures to provide the good environment for cooperative development. But consolidation of cooperative relation requires operating mechanism of the government alliance and enterprises linkage.

The government alliance formed into the union of the administrative units under the principle of "voluntary participation, advantages complementation, mutual benefit and result in win-win", it is the most important to find out joint benefits in the alliance. The "Opinion" points out that the Economic Zone on the West Side of the Taiwan Straits

should become “the new open comprehensive passageway for the serving development of peripheral region ”,it will make the biggest joint benefits for the every member in the zone. thus, administrative organs in the zone should initiatively let a portion of the powers to the alliance, which owns moderate decision-making power, and formulates out the regional policy of industry, investment, finance and tax, benefiting cooperative development, forming “self-organization” of regional economic cooperative system, insuring the fair usage of “new open comprehensive passageway”, sharing the result of cooperative development, it still needs the certain binding force as “laws” in the process of the coorperative development, implementing the valid monitoring.

The enterprises linkages develop cooperative activities with the fluxion of productive factors and optimizing recombination of industry by means of network, resulting in optimized benefits of the industries as a whole. Such as the large enterprise groups and the trade associations of trans-regions break through limits of the administrative units, and unite relevant industries with divided the work and the cooperation in the different region, carry out optimization of the whole industrial chain for the biggest commercial interests. It needs to give the enterprises with the fair market environment, pledge the enterprises to the dominant position in the market.

3 Changing the Growth Pattern of the Traditional Economy

Nowadays, the model of economic development should be changed from the driving growth into the innovating growth, from great investment of the capital, energy, raw material and labor force into the technical progress.

The United States, Japan, Finland are widely recognized as the innovative countries in the world, their technology contribution to economic development generally reached above 70%. In Fujian province, the innovated force was rather weaker, the combination of technology and economic appeared in the loose status, the application, conversion and popularization of technological achievements were moving ahead slowly, the innovative ability needs to strengthen urgently.

The enterprise is the main body of innovation, it is necessary to strengthen the consciousness of innovation, arouse the new vitality of innovation, built up support ability of innovated technology by the various measures, enhance the technological element of product, strengthen the level of high-tech products, cultivate a batch of brand products with proprietary intellectual property rights and high affixture value, entering into the international high-end market. For example, the Fu-Yao company group was a small factory located village, it produced the water meter glass originally, they carried out the technical innovative strategy, annual investments of science and technology, the replacement and reformation of equipment and the personnel training had reached 7% of sale sum, some techniques have reached to the advanced level in the world nowadays, they have been developing into a transnational large enterprises with total assets of 8 billion RMB.

The enterprises should be the division of labor and cooperation in the innovative activity, the large enterprises should establish organization for research and development, the medium and small-sized enterprises should participate in division of work and cooperate with the large enterprises to accept the diffusion and transferring of technology from the large enterprises, sharing the innovative result.

The technique innovations express the characteristic of knowledge overflow and transfer and technological diffusion. .Porter considered [9] that any traditional industries with the knowledge and technique can become a technology-intensive industries with an international competitiveness. It demands to foster the development of knowledge service industries vigorously, speed up the construction of platform for technological innovation, playing a role of development ,cooperation, achievements conversion, intermediary services of science and technology, and sharing technological resources.“Straits Item Achievement Trade Fair, China”(brief named “6.18”) highlights the function of platform of technological innovation. In 2009, signing joint contract and the negotiated item had 5008 ,total investment reached 88.6 billion RMB, it promoted effectively the commercialization of research achievements,”6.18” has become the important carrier with technological innovation and economic structural adjustment on the West Side of the Taiwan Straits.

It is necessary to establish alliance constituting of the university, research institute and enterprises to research common, key and foresight technology for upgrading production level and adjusting industry structure. The industries would be changed from simple “processing workshop” to “innovation base”, from the scale speed growth to the quality benefit growth. The independent innovation ability for industries will be enhanced.

In order to technical innovation, the incubate enterprise groups of advantage industries such as photoelectricity, electronic information, bio-medicine, energy-save, environment protection and resource recycle were established in “Business Incubator for Returned Personnel of Xiamen”, more than 200 the medium and small -sized enterprises of science and technology intensity appeared “making wings from worms”, the incubated enterprises had 400 or more, they manifested combining innovation of science and technology with economy.

It is imperative to create social atmosphere for technical innovation, promote the supporting capacity of independent innovation in the region, guarantee the intelligent rights of scientific and technical personnel and the interests of brand product in the enterprises, reward the contributor who can be independent innovation, cultivate the innovated and enterprising talents, import the advanced technology and talented person, encourage a batch of leading enterprises of the province formed a coalition with the same trade of the overseas excellent enterprises, promoting the independent innovation ability. For example the Hai-Tian Light Textile Group of Quanzhou city cooperated with E.I.Du Pont Company of American produced the new fabric used corn as raw material, thus it has become the first in China and the second in the world to produce PTT short fiber manufacturer allowed by the E.I.Du Pont Company and the one of the main suppliers of Adidas, Nike, Li-ning, Anta which are well-known business enterprises at home and abroad. It received order form from domestic and international markets continuously.

The introducing talents and innovations were succeeded in Fuzhou Bonded-Logistic Park. It had attracted the international well-known shipping company and logistics enterprises, gathering a batch of logistics enterprises with purchase, transfer, dispatch and distribution business faced at global markets, importing the advanced service concept and management technique of modern logistics, promoting the competitive power of logistics industry. Thereby, it drives more enterprises of multinational manufacture to assemble on the Park.

4 Conclusion

At present, the environment of economic development has changed greatly ,the restriction of the resources and labor force has been becoming more and more stronger, the capital and technical function is getting greater and greater. At the dawn of 21st century, the global economy begins to go downwards and trade protectionism becomes popular again, the exchange rate of the Renminbi (RMB) has appreciated, the economic transition should be enforced in China. The Economic Zone on the West Side of the Taiwan Straits, taking Fujian province as the main body, spans multiple administrative regions in China. There are economic differences in the location, industry and foundation between sub-regions of the zone, it should be optimizing the benefit of the zone as a whole ,cooperating development on the basis of the regional divided the work, and achieving the win-win result by means of complementation of advantages. The consolidation of cooperative relation requires operating mechanism of the government alliance and enterprises linkage. The government alliance formed into the union of the administrative units under the principle of "voluntary participation, advantages complementation, mutual benefit and result in win-win", it is the most important to find out joint benefits in the alliance. The enterprises linkages develop cooperative activities with the fluxion of productive factors and optimizing recombination of industry by means of network, resulting in optimized benefits of the industries as a whole. Nowadays, the model of economic development should be changed from the driving growth into the innovating growth, from great investment of the capital, energy, raw material and labor force into the technical progress. The innovated force was rather weaker, the application, conversion and popularization of technological achievements were moving ahead slowly in the zone, the innovative ability needs to strengthen urgently. it is necessary to built up support ability of innovated technology by the various measures, enhance the technological element of product, strengthen the level of high-tech products, cultivate a batch of brand products with proprietary intellectual property rights and high affixture value, and enter into the international high-end market.

Acknowledgement. I am thankful and acknowledge the support under Fujian Macro-Science Foundation Programme (No. 2010R0046).

References

1. Huang, L.-H.: Strategic ideas of the transformation of the development pattern in Zhejiang province. Zhejiang Economy 12 (2008)
2. Zheng, Y.: Reflecting on and interpreting the theory and mechanism of economic transformation. Journal of Jishou University (Social Sciences Edition) 3 (2011)
3. Yuan, Z.-G., Shao, T.: Transition of Chinese economy and rebalance of global economy. Academic Monthly 1 (2010)
4. Hu, A.-G.: The correlation of social transition with economy transition in China. Qiushi Journal 4 (2011)

5. China State Council. Some opinions on supporting Fujian to speed the construction of the Economic Zone on the West Side of the Taiwan Straits. The Chinese Government Internet (May 14, 2009), http://www.gov.cn
6. Wang, W.-G.: Study on Coordinated Theory and Methods. Financial and Economic Publishing House, Beijing (2000)
7. Yu, C.-Z.: Entropy, Information with Cross Science. Yunnan Unviersity Publishing House, Kunming (1994)
8. Lin, Y.-F.: The develpoment and transformation: ideological trend, strategy and self-innovation ability. Journal of Beijing Jiaotong University: Social Sciences Edition 4 (2008)
9. Porter, M.E. (trans.), Qiu, R.-M., Li, M.-X.: Competitive Advantage of Country. CITIC Publishing House, Beijing (2007)

System Structure of Venture Capital Industry in Japan

Peng Liang[1] and Yu Huang[2]

[1] Beijing Technology and Business University, P.R. China, 100048
[2] College of Geography and Remote Sensing,
Beijing Normal University, Beijing, P.R. China, 100875

Abstract. This paper is the first to introduce an experience of the system structure of venture capital industry in Japan. In addition, this paper provides some context for understanding the Japanese venture capital industry. It briefly discusses Japan's economic difficulties and the role of venture capital in promoting economic growth. It then set forth a short history of Japanese venture capital. The heart of the paper is devoted to a discussion of key features of Japanese venture capital.

Keywords: Initial public offering, Venture capital, Entrepreneur, investment.

1 Introduction

Venture capital has a number of positive features when compared to other forms of innovation financing and especially debt financing. However the most distinguishing aspect of venture financing is the rendition of a number of value added services provided by the venture capitalist to its portfolio companies. In recent study by Kuemmerle that there are considerable differences in the evolution of venture capital system in three technologically developed countries, the U.S., Germany and Japan. It shows that the U.S. alone has a well-developed venture capital system. So in a sense, the practical and theoretical attractions have a number of countries to try to copy its mode, and develop their venture capital systems. The Japanese VC industry which had learned a lot from the U.S. has paid the closest attention to its development. But according to the study of Clark, there are three major obstacles for Japan to the copy of a US-style venture capital system, the first of these is the power and adaptability of large manufacturing corporations, the second great difficulty facing venture capitalist is the lifetime employment' system, the third obstacle is the aversion Japanese show to the contractual business relation on which US-style venture capital depends. So Japanese venture capital has learned a lot from the US-style, but some features of Japanese economic and social structure constitute significant obstacles to venture activity of an US type.

According to research, there are two reasons why the idea of a VC system has proved attractive to Japan. The first is that, in Japan as elsewhere, technical progress has given small business new scope, and lessened their disadvantages relative to big ones. The second is that Japan's particular circumstances are changing rapidly, in a way that suggests that a VC system may eventually be both successful and economically

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 13–20.
springerlink.com

productive. Although Japanese VC has gone through ups and downs, and grown into a trillion yen industry and a trillion yen amount of VC's total investment balance accounts only 0.2% of GDP in Japan. It shows that there is still an enough room for further development.

2 Brief History of VC in Japan

When the first VC firm was established, VC was not known in the Japanese business world. Ever in Japanese history, venture capital even expected to contribute to the revitalization of Japanese economy. The economic recovery that represents a new era must be led by new industries. New industries require some amount of new investment like that of VC. Let's go back to the development history of Japanese venture industry.

The history of Japanese venture capital industry traces back to the early1970s when rules for initial public offerings on the Over-The-Counter (OTC) market (later renamed JASDAQ OTC market) were liberalized. In 1963, Small Business Investment Companies (SBICs) were set up in Tokyo, Nagoya and Osaka. Capital was contributed into these SBICs by both local government institutions and local financial institutions and companies. Twelve years after the SBICs were founded, in 1975, MITI established the Venture Enterprise Center (VEC), the organization which conducted the surveys of the performance of medium-sized enterprises. It has become the main source of information about VC in Japan. Dating back ten years prior to this, there was completely no available information concerning VC in Japan. During this period, the first VC firm named Kyoto Enterprise Development (KED) was set up in Kyoto in 1972. Between 1972 and 1974, eight private VC firms(VCFs) were formed by major banks and security firms. In other words, the major Japanese financial institution formed VC subsidiaries. In 1973, the largest private VC firm that existing now named JAFCO was set up by Nomura Securities Firm. So from 1970 to 1973, in these three years, Japan experienced its first VC boom. In 1983 the Japanese government deregulated the over-the-counter (OTC) market act. This act made it easies for VCFs to make investments for new ventures managing toward IPO. So Japan stimulated a second VC boom. From 1980 to 1985, 68 VCFs were founded. Especially, 14 VCFs were set up in 1983, 26 VC firms were set up in 1984, 20 VCFs were founded in 1985[9]. The major players in this boom were the stockbrokers and regional banks. The six pre-existing firms expanded their operations. Their goal was to use venture investing to create relationship with the Small and Medium Enterprise (SMEs). Responding to this development new stock exchanges were established:Mothers (Market of the High-Growth and Emerging Stocks) in 1999 and NASDAQ Japan in 2000. They aimed at making it easier and faster for emerging companies to conduct initial public offerings. JASDAQ OTC also relaxed its more rigid initial listing requirements. As the new economy boom deflated and the economy started to slow down worldwide towards the end of 2000, the Japanese venture capital industry did not remain unaffected, either. In many respects, Japan is thought to have lost its former competitiveness and there is

perceived to be a paucity of original and effective solutions to the country's economic problems. As such, people refer to the 1990s as Japan's "lost decade." However, a number of changes have been occurring behind the scenes. In 1994, with the growth of the Internet and the upswing in the Silicon Valley economy, Japanese interest in the role of VC in facilitating new business formation and the support of start-ups was renewed. Japanese venture capital industry has welcomed its third boom. But with the financial crisis in Asia, Japanese economy development falls into the trough, and continues to downturn. This time, however, the boom occurred in an environment of continuing stagnation of the economy. But government has introduced a series of liberalizing reforms aimed at creating an environment more supportive of entrepreneurship and venture creation. These reforms entailed a series of amendments to the Commercial Code (where Japanese corporate law is found), as well as amendments to stock exchange and OTC rules which had created some benefits to the capital flow of newly and the high-tech enterprises.

3 System Structure of Japanese VC

3.1 Annual Investments in Japan and Western Countries

In 2008 venture capitals invested ¥137 billion, about 30% decrease from the previous year. 23 new funds were organized, which raised ¥84.8 billion and the average number of partnership members dropped down from 14.2 in 2007 to 4.5 in 2008. As a result, investment share of general partners was pushed up to 60.2% (36.6% in previous year). Between April 2009 and March 2010, 87.5 billion yen in venture capital (VC) was financed to 991 startup companies, a decrease of 36.8% and 25.9% respectively compared to the previous year. VC investment appears to be equally stagnant abroad, as the annual investment amount in 2009 declined at a similar rate in Western countries.

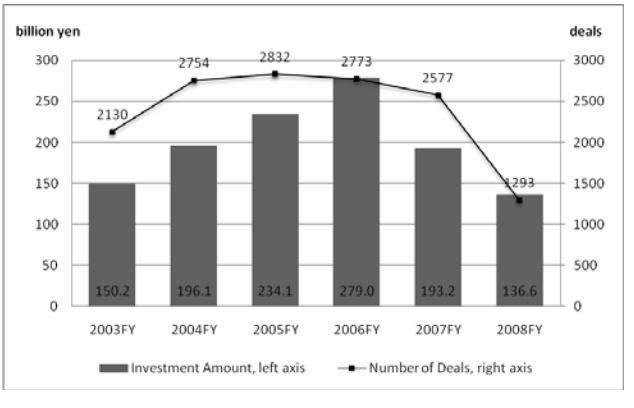


Fig. 1. Venture Capital Investment 2003 to 2008

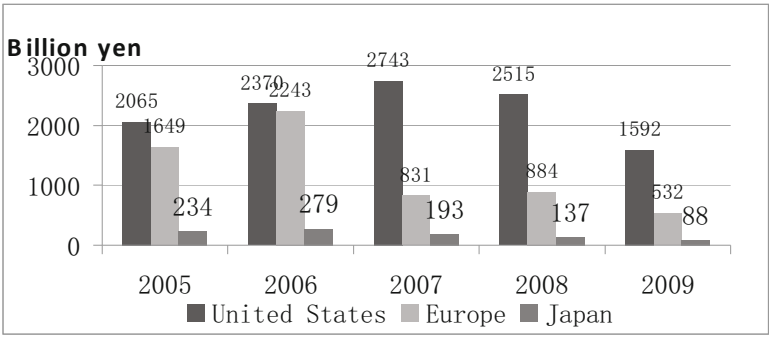


Fig. 2. Investment of US, Europe and Japan

3.2 Corporate-Affiliated Firms

Through independent partnership are used in Japan, the majority of the VC organizations are closely affiliated with corporations rather than independent. Many of the partnership are also, in fact, operated as corporate subsidiaries. These affiliated firms generally invest their own money rather than that of other investors. While the VC firms in the U.S, are usually limited partnerships in which one VC firm serves as general partner and manages a pool of invested money. The pool includes money invested by the general partner. These firms are called private independent firms.

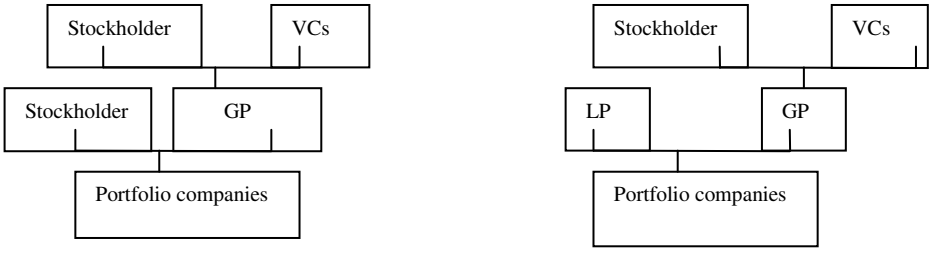


Fig. 3. VC Organization in Japan

3.3 Investment Stage

Traditionally Japanese VCs invest in the later stage of a firm's growth. They rarely invest in new business. They prefer to buy the shares of relatively well-established enterprises. According to 1998 survey by the VEC, the percentage of companies receiving investment that was from set up 5 years old was 26%, from 5 years to 10 years was 16.8%, and from 10 years to 20 years was 24.2%, while 20 years old or older was 30.4%.according to the figures provided by VEC, the percentage of the firms

funded by VCs were more than 10 years old was over 50%.while in the year 2009, the percentage from set up 5 years old was 15%, from 5 years to 10 years was 31%, and from 10 years to 20 years was 32%, while 20 years old or older was 22%.for the stage, there have no new trend.

3.4 Sectors of New Investment Recipient Companies

As the METI survey suggested, venture capital firms show keen interest in biotech ventures. The number of Japanese firms investing in this field is indeed increasing rapidly. While both initial and follow-on investment have plummeted, new investment in "Industrial/Energy and other " industries grew in both number and amount, and that in the number of "Clean technology" companies also improved from 4.6% in the previous year to 7.2%. From fig. 4, investment amount in "Clean technology" startups climbed from 4.1% to 5.8%.

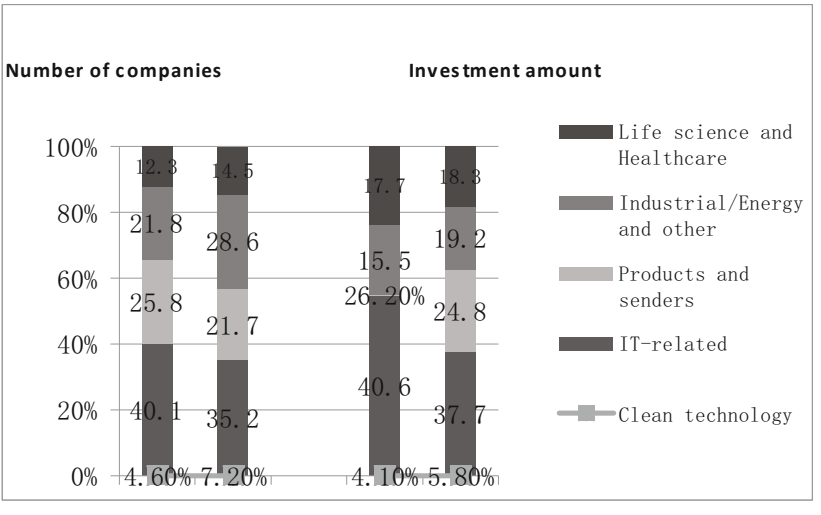


Fig. 4. Initial Investment by Industry

3.5 Investors by Business Category

In Japan there have some regulations that prohibited the pension funds to be invested in VC, and also families and person are unwilling to bear the risks to invest, so personal investment ration is relatively lower. VCFs rely heavily on debt financing from banks, and on money from cautious VCFs that are themselves subsidiaries of banks and securities companies. As the venture funs invested newly and additionally, looking at the breakdown of investors from the fig. 5

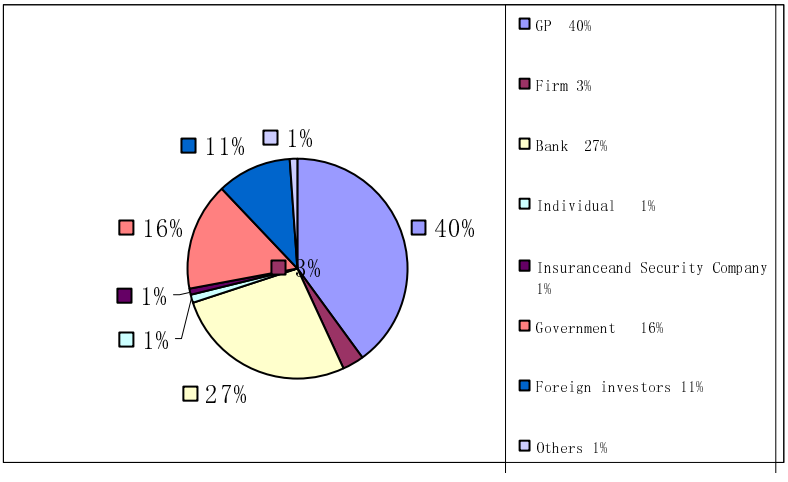


Fig. 5. Composition of Investors by Business (amount)
Source: Survey of Venture Capital Investment 2009

3.6 The Exit Way

There are three principal methods of exiting an investment: the first is through an initial public offering(IPO) of shares to outside investors, the second is by selling the portfolio company directly to another company, the third is by selling the company back to the entrepreneur. IPOs are by far the most profitable and prestigious option for the VCs. For Japanese VCFs, the main exit strategy for VC remains the IPO, and with the recent downturn in the startup stock markets, VC has shifted away from IPO and into MBOs.

Decline of the VC investment has followed the fall of IPO activity. A number of corporate and accounting scandals eroded public confidence in the securities market and consequently tightening the listing requirements reduced IPOs.

IPO activity has fallen sharply since 2007 after the second highest 188 companies have gone public in 2006. A number of corporate and accounting scandals eroded public confidence in the securities market and consequently tightening the listing requirements reduced IPOs. In 2009 Japan saw 19 IPOs, a 60% fall in deal numbers from previous year, which also saw big decrease. This level was the smallest number since 1978.

Table 1. Exit Situation of Investment Recipients

	2003	2004	2005	2006	2007	2008
IPO	427	495	328	345	249	66
Write off, Liquidation	186	159	131	179	166	202
Trade sales, Buyback	286	383	359	229	270	323
Other	30	4	43	15	24	79
Total	929	1041	861	768	709	670

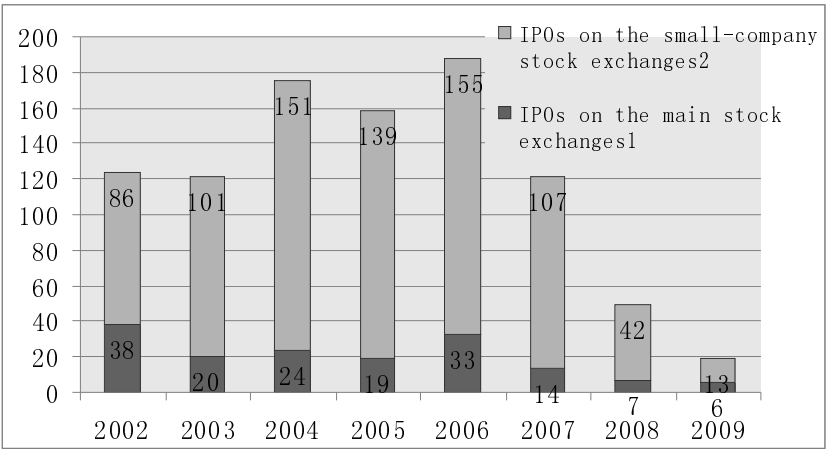


Fig. 6. Number of IPOs 2002 to 2009

1) TSE1, 2 (Tokyo), OSE1, 2 (Osaka), NSE1, 2 (Nagoya), FSE (Fukuoka), SSE (Sappro)
2) Jasdaq, Mothers (TSE), Hercules (OSE), Centrex (NSE), Q-Board (FSE), Ambitious (SSE), NEO (Jasdaq)

4 Conclusion

In the last few years, Japanese government has recognized the economic benefits of fostering the creation and development of venture firms, and has applied themselves assiduously to the task of creating a supportive environment. What's more, the government is ambitious to plan to create a large number of venture firms and to focus on strategic industries. However, Generally speaking, except the national life industry, the government is not suitable to participate in the industry management, not to mention the venture capital industry. It causes the Japanese venture capital industry operates too conservative, therefore the development is restricted. So, an adventurous, risk-taking approach should be adopted to venture development. Like the Japanese government, Chinese government has also played a central role in defining the institutional antecedents of Chinese venture capital industry through its control over related institutional systems, including banking, legal, innovation and other systems. The central government has been too directly involved in China's venture capital system. Generally speaking, except the national life industry, the government is not suitable to participate in the industry management, not to mention the venture capital industry. So the government should have less direct participation and intervention and give more support in terms of creating appropriate institutions for the venture capital industry.

References

1. Kenney, M., Han, K., Tanaka, S.: The Globalization of Venture Capital: The Case of TaiWan and Japan. In: Financial System, Corporate Investment in Innovation, and Venture Capital, pp. 53–83. Edward Elgar, Publishing Inc. (October 2004)
2. Walter, K.: Comparing Catalysts of Change: Evolution and Institutional Differences in the Venture Capital Industries in the US Japan and Germany. In: Comparative Studies of Technological Evolution, pp. 227–261 (August 2001)
3. Clark, R.: Venture Capital in Britain, America, and Japan, pp. 1–23. Croom Helm, London (1998)
4. Yasuyuki, H.: Venture capital in Japan. Nippon Keizai Shimbumshya, pp. 1–23 (October 1998)
5. Yasuyuki, H.: Investing Activities of Japanese Venture Capitalists and Venture Capital Firms. Unpublished Working Paper, pp. 1–13 (November 2007)
6. Kutsuna, K., Hasegawa, H., Yamamoto, K.: Handbook of Venture Capital, pp. 22–32. Central Economy Press (June 2006)
7. Shinbunsha, N.K. (ed.): Survey of Japanese Venture Capital Investment, pp. 1–45 (December 2009)

Identification about Key Element of Urban Core Competitiveness Based on Structural Equation Model*

Shui-ying Jin

Institute of Strategic Management,
Zhejiang Normal University, Jinhua 321004, China
jsy@zjnu.cn

Abstract. Based on the analysis about the relationships among urban competitiveness elements, the paper proposed an identification model for key element of urban core competitiveness. Then it selected prefecture level or above prefecture level cities' sample data and used structural equation model to identify the key element of urban core competitiveness. The result shows that urban intellectual capital constitutes the key element of urban core competitiveness. So, the accumulation and operation of intellectual capital should be paid great attention to promote urban core competitiveness.

Keywords: urban core competitiveness, identification model, structural equation model, intellectual capital.

1 Introduction

The center question of research on urban core competitiveness is "Where should we set about to cultivate urban core competitiveness?" To answer the question, we need to identify the key element of urban core competitiveness. After the core competitiveness theory was proposed by C. K. Prahalad and G.Hamel in 1990[1], most scholars believe that the core competitiveness lies in the organization's resources. Mahoney & Pandian (1992) considered core competitiveness as a function of resourses[2]. Henderson & Cockburn (1994) thought that the core competitiveness is a combination of organizational resources which can provide the greatest strategic value and differentiate themselves from competitors[3]. If the organization's resources include tangible assets and intangible assets, core competitiveness is a function or combination of resources (Wei Zhang, 2002) [4]. Therefore, we also believe that the urban core competitiveness comes from the resources of the city. According to the mainstream theory of intellectual capital, urban resources include physical capital (PHCA) resource and intellectual capital (IC) resource. Then what kind of resource constitutes the key element of the urban core competitiveness?

* The paper is sponsored by the humanities and social science research youth fund project of the education ministry (10YJC630105) and the soft science research project of Zhejiang province (2011C35056).

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 21–28.
springerlink.com

2 Identification Criteria about Key Elements of Urban Core Competitiveness

The elements of urban competitiveness can be depicted from inputs perspective and outputs perspective[5]. From inputs perspective, urban competitiveness depends on its resources (PHCA resource and IC resource). In the paper, IC resource is defined soft force of urban competitiveness. PHCA resource (including practical capital resources and financial capital resources) is defined hard force of urban competitiveness. According to the international mainstream view of four-factor structure, urban IC is divided into human capital (HUCA), innovation capital (INCA), process capital (PRCA), relational capital (RECA). Urban HUCA is citizen's comprehensive knowledge, education level, health status, skills and other capabilities needed to accomplish tasks. Urban INCA refers to the inputs and outputs in research, invention, entrepreneurship of the city to promote the creation of knowledge wealth. Urban PRCA is the program activities and related basic conditions embedded in the urban transportation, communications, and environmental construction system to create, share, and transmit knowledge. Urban RECA refers to the relationships with related interest groups to achieve its objectives[6]. From outputs perspective, the ultimate expression of urban competitiveness is its more sustainable development ability than other cities in economy, society and environment, which is called explicit urban competitiveness (EUC) in the article.

As for the identification about key element of urban core competitiveness, Wen-Jian Liu (2003) pointed out that the contributions of the component forces to urban competitiveness are different. Those elements with the highest contribution to the urban competitiveness constitute urban core competitiveness[7]. Because the EUC is the ultimate expression of urban competitiveness, the identification criteria about key element of urban core competitiveness can be described as: The element with the highest contribution to the EUC constitutes the key element of urban core competitiveness. In the following, we will use structural equation model (SEM) to analyze empirically the contributions of IC resource and PHCA resource to the EUC. The resource with the highest contribution will be recognized as the key element of the urban core competitiveness.

3 Research Method

3.1 Research Hypotheses and Research Model

With the coming of knowledge-based economy era and the gradual deepening of world economic integration, intellectual capital (IC) exceeds physical capital (PHCA) greatly both in quantity and quality. In the early 1990s, the U.S. management expert Peter Drucker illustrated in his book "post-capitalist society" that economic growth depends more directly on the investment and operation of "IC" in the knowledge economy era. IC has become the most important resource of an organization[8]. Therefore, in order to have a sustainable competitive advantage, the government of a city must attach great importance to cultivation and management of urban IC. A growing number of empirical studies have shown that the city competition in the 21st century will be fully reflected in IC, which is characterized in the talent and knowledge[9]. So, we propose the following two hypotheses:

H1. Urban IC has a significantly positive impact on explicit urban competitiveness (EUC).

H2. The impact of urban IC on EUC is significantly greater than the impact of PHCA on EUC.

Although IC is increasingly replacing PHCA as the main resource of urban development, but that does not mean that the development of society can do without PHCA. Firstly, PHCA is the physical infrastructure and condition to realize urban sustainable development. Secondly, a significant portion of IC exits in physical capital. Finally, IC must match with PHCA in order to play a role. So, we propose the third hypothesis:

H3. Urban PHCA has a positive impact on EUC.

It should be noted that there are complex correlations among various elements of intellectual capital (IC)- human capital (HUCA), innovation capital (INCA), process capital (PRCA), relational capital (RECA). Bontis (1998) showed that there is interaction among HUCA, RECA and structural capital (PRCA and INCA) through empirical studies[10]. For such a question with large number of latent variables and complex correlations among different factors, a second-order factor analysis model can be designed to make the relationship among the latent variables more clearly[11]. Thus, the paper constructs structural equation theoretical model (shown in Figure 1) to identify key element of urban core competitiveness.

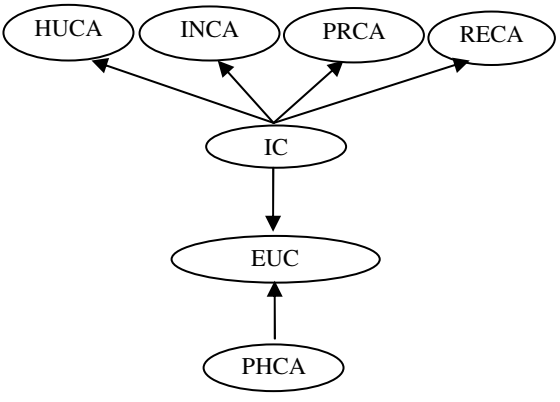


Fig. 1. Structural equation theoretical model to identify key element of urban core competitiveness

3.2 The Observation Indicators of Latent Variables

After considering the characteristics of Chinese cities and data availability, a specific observation indicator system (shown in Table 1) is built based on the research of Nick Bontis (2004) [6]and Chen Yu-Fen (2006) [12] and other scholars. As for the design and references of the indicator system, the author has written an article published in "Chinese Science and Technology Forum," 2009,(10) [5] .

Table 1. Observation indicator system of latent variables

Latent variables	First level indicator	Second level indicator	Second level indicator adjusted
HUCA	Education level	Number of colleges and universities per million people	Education level factor (Y1)
		Number of full-time teachers in colleges and universities per myriad people	
		Number of students in colleges and universities per myriad people	
	Health guarantee	Number of doctors per myriad people Number of hospital beds per myriad people	Health guarantee factor (Y2)
	Cultural life	Number of theaters per million people Public library collection	Cultural life factor (Y3)
INCA	Investments in innovation	Expenditure on education per capita	Investments in innovation factor (Y4)
		Expenditure on science per capita	
	Innovators	Research, technical services and geological prospecting industry employees per myriad people Information transmission, computer services and software industry professionals per myriad people	Innovators factor (Y5)
	Innovation output	Patent applications examined per myriad people Patents granted per myriad people	Innovation output factor (Y6)
PRCA	Postal and tele-communications	Mobile phone users per hundred people	Postal and telecommunications factor (Y7)
		Internet users per hundred people Post and telecommunication services per capita	
	Transportation	Total freight per capita Total passenger transport per capita	Transportation factor (Y8)
	Investment in environment construction	Total investment in environmental pollution control per capita Investment in urban environmental infrastructure construction per capita	Investment in environment construction factor (Y9)
RECA	International trade	Contracted foreign investment per capita	International trade factor (Y10)
		The actual foreign investment per capita Total import and export commodities per capita	
	Domestic trade	Sales of goods in retail trade industry per capita Total retail sales of social consumer goods	Domestic trade factor (Y11)
	Tourism contact	Revenue of domestic tourism per capita Foreign exchange tourism earnings per capita	Tourism contact factor (Y12)

Table 1. (*continued*)

PHCA	Practical capital	Forest cover (%)	Forest cover (X1)
		Annual ore per capita	Annual ore per capita (X2)
		Water per capita	Water per capita (X3)
		Land area per capita	Land area per capita (X4)
		Total amount of LPG per myriad people	Total amount of LPG per myriad people (X5)
	Financial capital	Deposits of financial institutions per capita	Deposits of financial institutions per capita (X6)
		Loans of financial institutions per capita	Loans of financial institutions per capita (X7)
EUC	Economic development status	Gross regional product per capita	Gross regional product per capita (Y13)
		Growth rate of gross regional product (%)	Growth rate of gross regional product (Y14)
	The level of social life	Engel coefficient of urban households (%)	Engel coefficient of urban households (Y15)
		Housing area per capita	Housing area per capita (Y16)
		Coverage of basic endowment insurance (%)	Coverage of basic endowment insurance (Y17)
	The intensity of environmental protection	Removal rate of industrial sulfur dioxide (%)	Removal rate of industrial sulfur dioxide (Y18)
		Urban sewage treatment rate (%)	Urban sewage treatment rate (Y19)
		Garbage treatment rate (%)	Garbage treatment rate (Y20)

3.3 Sample Selection and Data Sources

Because the administrative divisions of prefecture level or above prefecture level cities are relatively stable, and most indicator data of these cities are available from the relevant state ministries and commissions' statistics, 287 prefecture level or above prefecture level cities are selected as sample cities in this study. After removing the incomplete data and obvious outlier samples, the final sample consists of a total of 216 cities. For the small amount of data missing, EM estimation method is adopted to estimate the missing data by using SPSS 15.

Because of the limitation of time and energy, it is difficult to make scientific sample survey for all sample cities. All indicator data are from "Chinese City Statistical Yearbook 2010" and the public statistical data released by state ministries at the same.

3.4 Data Analysis and Processing

By descriptive statistics of sample data, we found that the value range of different indicator is significantly different. So we make natural logarithm transformation process for all indicator variables. After the process, except the total import and export commodities per capita is slightly larger than 10, the other indicators are all in the range of 0~10. This logarithmic transformation result is consistent with the idealized data standard proposed by Gauch [13].

In order to get a better model fitting results in the empirical analysis with structural equation model, we should avoid correlations between different observation indicators of the same latent variable[12]. Therefore, Bartlett and KMO tests are conducted on the observation indicators (after logarithm transformation) of the same latent variable. And factor analysis is made to extract the common factors as the new observation indicators of latent variable, which can reduce the correlations between observing indicators. The observation indicators failed in Bartlett and KMO tests are retained. As a result, the final observation indicators of each latent variable are shown as the adjusted secondary indicators in Table 1. Where, observation indicators of the endogenous latent variable are expressed with the letter Y, and observation indicators of the exogenous latent variable are expressed with the letter X.

4 Empirical Results

The article tests the reliability and validity of selected variables through confirmatory factor analysis using the Amos 7.0 software. Test results show that the variables in this study have good reliability and validity. The test results of model hypotheses are shown in Table 2. Figure 2 demonstrates the estimate result of each parameter, which value has been standardized. The empirical results indicate that hypotheses H1 and H2 are significantly true, but hypothesis H3 gets partial support.

Table 2. Hypotheses testing table

	Research hypotheses	Route γ	Testing result
H1	Urban IC has a significantly positive impact on EUC	0.866***	Confirmed
H2	The impact of urban IC on EUC is significantly greater than the impact of PHCA on EUC	——	Confirmed
H3	Urban PHCA has a positive impact on EUC	0.235	Partially confirmed

Note: Route coefficients of latent variables are standardized parameters. *** denotes $p<0.001$.

The empirical results indicate that IC is the most decisive element with the highest contribution to the EUC and constitutes the key element of urban core competitiveness. So, in the knowledge-based economy, a city should attach great importance to the accumulation of IC to foster and enhance its core competitiveness. The empirical results also show that the standardized loading coefficients of IC on HUCA, INCA, PRCA and RECA are 0.908, 0.723, 0.852 and 0.812 respectively, with 0.001 significance level. Thus, during the improvement and cultivation of urban core competitiveness, any element of IC can not be ignored. HUCA is the kernel to nurture and enhance the core competitiveness. The quality of HUCA will determine the operation effectiveness of other capital. INCA is the power to cultivate the urban core competitiveness. It is a natural choice to cultivate innovation in order to build a dynamic city, and enhance the urban core competitiveness. PRCA is the foundation to nurture and enhance urban core competitiveness. Construction and improvement of PRCA plays a vital role in the

development of a city. RECA is the guarantee to enhance the urban core competitiveness. The close relationship between the city and relevant interest groups can make its own advantages known in the wider area, which can cluster more advanced resources to the city, resulting in stronger urban core competitiveness.

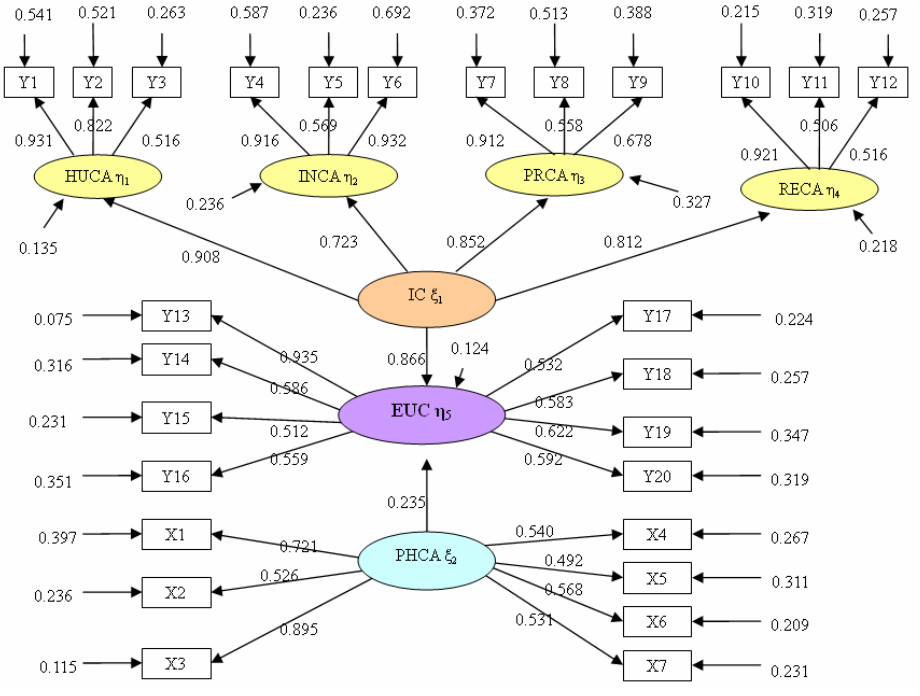


Fig. 2. The estimate result of each parameter in SEM

This study expands quantitative research methods to identify the key element of the urban core competitiveness. But there are many detailed questions have not been discussed due to the limitation of time and energy. For example, the article concludes that IC is key element of urban core competitiveness from the overall analysis. Due to the effect of historical, geographical and other conditions, there will be some difference in emphasis for different city's core competitiveness. So, what's the characteristic for the core competitiveness of a specific city? This issue needs to be studied in-depth later. With the development of actual condition, the theoretical and research framework in the article needs to be improved and refined continuously in the future study.

References

1. Prahalad, C.K., Hamel, G.: The Core Competence of the Corporation. Harvard Business Review 68(3), 79–91 (1990)
2. Mahoney, J.T., Pandian, J.R.: The Resource-based View within the Conversation of Strategic Management. Strategic Management Journal (13), 363–380 (1992)

3. Henderson, R., Cockburn, I.: Measuring Competence? Exploring Firm Effects in Pharmaceutical Research. Strategic Management Journal 15, 63–84 (1994)
4. Zhang, W.: Analysis of Core Competitiveness. Economic Management New Management (12), 10–17 (2002)
5. Jin, S.-Y., Li, W.-B., Wu, Y.-Y.: Intellectual Capital Theory: a New Perspective for Urban Competitiveness Study. Chinese Science and Technology Forum (10), 77–82 (2009)
6. Bontis, N.: National intellectual capital index: a united nations initiative for the Arab region. Journal of Intellectual Capital 5(1), 13–39 (2004)
7. Liu, W.-J., Zhao, L.-Y.: Culture Support of Urban Core Competitiveness. Journal of Hangzhou Municipal Party School (5), 63–67 (2003)
8. Drucker, P.: Post-capitalist Society. Shanghai Translation Publishing House, Shanghai (1998)
9. Wang, Z.-Z.: The New Path of Global knowledge cities and the urbanization process in China. Research on Urban Development 14(3), 13–19 (2007), 15(1), 13–15 (2008)
10. Bontis, N.: Intellectual Capital: An Exploratory Study that Develops Measures and Models. Management Decision 36(2), 63–76 (1998)
11. Yi, D.-H.: Methods and Application of SEM. Renmin University of China Press, Beijing (2008)
12. Chen, Y.-F.: The Measuring Indicator system of Regional IC. Statistical Research (2), 24–29 (2006)
13. Gauch, H.G.: Multivariate analysis in community ecology. Science Press, Beijing (1989)

Work-Integrated Learning at the Engineering School at Vongchavalitkul University

Sanguan Vongchavalitkul and Busaya Vongchavalitkul

Vongchavalitkul University, Nakhon-Ratchasima, Thailand
sanbus@hotmail.com

Abstract. An engineering school is under increasing pressure to re-evaluate the place of practice in its programs and there are increasing demands for workplace-based experiences to be built into undergraduate degrees. The paper reports on an extended responsive case study conducted in the Engineering School at Vongchavalit University, Thailand during 2008-2010. The findings point towards a work-integrated learning (WIL) model that can be used to develop courses based around a set of key challenges that learners need to face, criteria for the university's construction of worthwhile activities, and the principles of procedure that are required in order to implement them. It is suggested that university curriculum planners should concentrate on a process model of teaching and learning rather than on the more traditional models. This program since it commenced operation has received positive feedback from both placement providers and students. The work-integrated learning program offers a "win-win" situation for students, the university and placement providers.

Keywords: Work Integrated Learning (WIL) program, Engineering School.

1 Introduction

Work integrated learning (WIL) or cooperative education is a form of learning whereby periods of study are alternated with periods related to work in business, industry or government agencies. These way students are given the opportunity to effectively integrate the theory of the classroom with the practice and the responsibility of the workplace [1]. Work-integrated learning provides an opportunity for students to gain experience in the workplace where they apply the problem-solving skills and discipline-based theory learned in their formal education to an authentic context as a colleague and employee, with all the responsibilities and expectations such a role entails [2].

The terms Work-integrated learning and work-based learning are used to describe the range of educational programs that integrate formal learning and workplace experience. Examples of such programs are work-based projects, practice firms and co-operative education programs [3]. For two years Vongcavalitkul University has included a work-integrated learning or co-operative year after completion of the third year of the Bachelor of Engineering degree, with students working in paid discipline-related employment for 16 weeks before returning to the university to complete the final year of their degree.

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 29–36.
springerlink.com

Heron, J. [4], Dressler and Keeling [5], Fraser and Bosanquet [6] provided overviews of the outcomes attributed to co-operative education. Dressler and Keeling [5] found categorized and distinguished the academic benefits, personal benefits, career benefits and work skill development benefits as follows:

- Academic benefits (e.g. increased discipline thinking, increased motivation to learn, improved performance in the classroom),
- Personal benefits (e.g. increased communication skills, increased initiative, increased team work and co-operation),
- Career benefits (e.g. improved career identity and clarification, increased employment opportunities and increased salaries), and
- Work skills development benefits (e.g. development of positive work values and ethics, increased competence and increased technical knowledge and skills).

It is an old adage that people learn by doing. All genuine education comes through experience [2]. Experience in real life provides the four conditions for effective learning: a knowledge base, a motivational context, learning activity and interaction [7].

For WIL, the educational goals should be developed in the design of a work-based curriculum [2] as follows:

- Promote the development and negotiation of a program of activities.
- Address the diverse range of knowledge and skills possessed by students at the commencement of work-based learning.
- Locate the outcomes of work-based learning in a framework of levels and standards of achievement.

2 Work-Integrated Learning (WIL) in the Engineering School at Vongchavalitkul University

Work-integrated learning (WIL) is a successful program for engineering students at Vongchavalitkul University. This program is provided for undergraduate students to spend 16 weeks of industrial placement over summer during the second semester of their third year.

The program involves a course in the second semester of the third year in the Bachelor of Engineering conducted by Vongchavalitkul University. The occupational areas for which students are being prepared include engineering projects. The project involves students being situated in a number of quite different work environments, so it needs to be able to provide the flexibility of individually tailored experiences within an open course framework under the supervision of practitioners in the field. Academic staff members augment this supervision, with each student's learning structured according to a constantly renegotiable learning plan. Since its inception in 2008, this has evolved into a process-based, action-learning curriculum for each student.

Students do WIL by communicating with co-workers in the workplace and peers in the workshops. The key learning experiences are in the design of the tasks set and in the associated consultative process that occurs between the university faculty member, designated organizational supervisor and the student [8]. For the Vongchavalitkul

Engineering School, students begin the course by attending two-hour workshops prior to the practicum itself. These focus on the particular skills essential for achieving graduate employment: job searching, resume writing and job interview techniques. This provides students with an overview of the range of occupations available to them and is a specific requirement not usually regarded as a university criterion for achievement [8]. During the practicum, students attend the placement organization for around 16 weeks over the semester.

The WIL program in the Engineering School at Vongchavalitkul University consists of a number of activities to be carried out by students in the third year of their four-year engineering degree.

Table 1 shows the schedule of the placement program activities for students enrolled in the WIL program.

Table 1. Schedule of the Placement Program

Dates	Activities	Assessment Due Date
Semester 1	○ University approaches companies, and obtains profiles. ○ Preparation of resume / CV with drafting of letters of application. ○ Preparation for interview. ○ Send applications to potential employers / providers. ○ Selection by Placement Provider; confirmation of work placement; appointment of Academic Supervisor.	
Mid-year Break	Health and Safety Day Lectures	
Semester 2		
Week 1-2	Start of Work Experience	
Weeks 3-6	First visit by academic supervisor	
Week 7-9		Report 1
Week 10	Intellectual Property Day Lectures:	
Weeks 11-13	Second visit by academic supervisor	
Week 14	Professional Day Lectures	
Week 15	Conclusion of 15 week placement	Seminar synopsis
Week 16	Seminars Industry supervisors are invited to attend	Report 2

Adapted from Jimmy,J and Randhawa, S. **Work Integrated Learning for Engineering Students at Flinders University.** Proceedings of 20th Australasian Association for Engineering Education Conference, University of Adelaide. (2009).

The three major components of supported WIL for Engineering Students at Vongchavalitkul University employment are: (1) prerequisites, (2) application for job placement, and (3) placement visits.

1. Prerequisites. Vongchavalitkul engineering students undertake work placement in the second semester of the third year. They must have passed most of the topics up to the level of the second semester of their third year before they are eligible to enroll in the WIL program. This is to ensure that the students have the minimum academic knowledge to cope with the challenges of the industrial placement. In addition, students must prepare for (1) skill development (professional engagement skill and technical skills), and (2) attitude and maturity development for WIL program.

Skill development	Professional engagement skills: o Communication skill development o Project management processes o Team-building processes o Workplace literacy o Independence, assertiveness and confidence training o Workplace health and safety knowledge o Legal and ethical obligations and responsibilities o Cultural awareness and values o Problem-solving approaches Technical skills such as: o Math and Drawing o Computer Programming o Materials o Using hand and machine tools o Measuring and marking out o Fastening and jointing o Route Survey
Attitude and maturity development	Students need a mature work ethic to gain real value from the WIL experience.

Adapted from **Guidelines for good practice in work integrated learning for the integrated resource Sciences** (May 2008)
http://www.cirm.org.au/pdf/work_integrated_learning_guidelines.pdf

2. Application for Job Placement. Vongchavalitkul Engineering students have to apply for the placement positions themselves during around their second year of

coursework. Our school will assist the students by making inquiries as to potential work placement providers. Interested placement providers can fill in a placement form for students any time that these placement providers want.

Common documentation includes explicit information regarding [9]:

- Students' placement rights and responsibilities, entitlements and support available
- guidance on meeting learning outcomes
- assessment guidelines
- health and safety documentation
- a learning agreement form which the student should be guided in completing and must be approved by the Faculty Placement Officer and should include:
  - contact details of the student and the placement provider
  - general areas of work and tasks to be undertaken in the placement role
  - application of the learning outcomes to the role and how the student envisages they will be met

3. Placement Visits. There are at least three placement visits by academic supervisors over the duration of the work placement.

1) **The first visit** happens at the work placement. The purpose of the first visit is to examine the workplace in order to ensure that our students have appropriate induction provided by the placement providers. The academic supervisor meets with all parties first to discuss about general issues regarding work placement. Various issues of the work in which the students are involved are also discussed [10], for instance, issues regarding students involved in classified, their responsibilities or organization culture.

2) **The second visit** takes place during the work placement. The academic supervisor mentors the student during their tasks and everyday activities. The purpose of the second visit is to ensure that the student is working in a supportive environment and to ensure the student has a varied group of meaningful tasks to complete [10].

   For this second visit, the academic supervisors have to recheck the student's time sheet at the end of each day. If there are any problems, these supervisors resolve them easily with effective monitoring and communication during the placement. The academic should meet with the student alone to discuss what these students think of their own performance, as well as what they think of the placement supervisor's assessment of them. Moreover, the academic supervisors have to meet with the placement supervisor alone to discuss the student's performance, their progress, the work quality, and so on.

3) **The third visit** takes place at the end of the work placement. This is a chance for our academics to assess and examine the quality of work carried by the student over the period of the work placement [10].

At the end of the work placement, the placement supervisor and the academic supervisor carry out the following list [11]:

- Complete the *Workplace Supervisor's Evaluation Report* in the student's journal
- Ensure the student completes the *Student Work Placement Evaluation Report*
- Check that the student has
  - handed back all finished and complete work
  - tidied their work area
  - returned all badges, security tags, keys, etc
- Give the student any additional information they require regarding their career paths, job opportunities, etc.
- Document anything that the co-op should do differently next time. This could focus on:
  - staff preparation before the work placement begins
  - the induction/orientation
  - the supervision of the student
  - the jobs list

3 Rewards of Work-Integrated Learning (WIL) in 2010

The WIL program at Vongchavalitkul Engineering School is offered only once a year. During 2008-2010, least than 10 companies and organizations have supported our WIL program, we have received many positive comments from placement providers towards our program. We summary the projects that we can get from WIL program.

Field	Number of Students in WIL	Number of workplaces in WIL	Number of Projects in WIL	Number of Reward
Civil	10	6	6	1
Mechatronics	12	8	8	-
Mechanical	16	8	8	-
Electrical	22	15	15	-
Computer Sciences	16	14	14	-
Total	76	51	51	-

We report all 51 WIL projects to conferment by the Thai Association for Cooperation Education 2011 (TACE 2011). Unacceptable, the project name "Crushes Rock Compacted for the Improved Quality of Rice Plant at Suphanburi Province" can be announced at the excellent engineering project of TACE 2011.

From an academic point of view, our students come back more mature after work placements and usually perform better in their final year topics. In addition, it is of critical importance that work-integrated programs in the creative industries focus on more than just training students to become employees [12]. The output and outcomes of work-integrated learning for the Engineering School at Vongchavalitkul University are summarized as follows:

Input	**Outputs**		**Outcomes**		
	Activities	Participation	Short-term	Medium-term	Long-term
University/ Facilitator Personnel: **- Administration** **- Academic experience** **Industry Personnel:** - **Practical experience** - **Mentoring** **Students**	- Appoint project manager - Engage mentors - Scope project - Document activities - Plan process - Establish guidelines - Develop networks - Develop professional development for participants - Review and Evaluate	- Project Managers - Industry Professionals - Teachers - Students - Community groups - Business	- Curriculum development - Student Projects - Reports	- University uptake - Partnership - Engagement - Enhanced student experience - Improved student work skills	- Improved education to workplace transitions - - Improved recruitment and retention of employees
External Factors					
Availability of work placements. Other commitments. Timing and resources. Continuity of employment. Sustainability of program.					

Adapted from **Guidelines for good practice in work integrated learning for the integrated resource Sciences** (May 2008)
http://www.cirm.org.au/pdf/work_integrated_learning_guidelines.pdf

4 Conclusion

The study has developed the work-integrated curricula in engineering school programs. Through this study, the academic supervisor, placement supervisor and students should communicate and co-operate more effectively and efficiently. The improved contextual environment should facilitate the learning process of the students and help to improve their degree of satisfaction. These work-integrated programs offer the students a wealth of experience in both academia and industry, and thus this can significantly improve their employability for a wide array of jobs upon graduation.

References

1. Fadzil, D.A.K.S.: The Importance of Work-Integrated Learning, http://www.jobstreet.com.my/learning/hr2.htm
2. Atkinson, L., Rizzetti, J., Smith, S.: Online Resources for Work-integrated Learning: A Case Study in Re- usability and Flexibility. Balance, Fidelity, Mobility: Maintaining the Momentum? In: Australasian Society for Computers in Learning in Tertiary Education (ASCILITE), QLD, Brisbane, pp. 39–45 (2005)
3. Atchison, M., Pollock, S., Reeders, E., Rizzetti, J.: Guide to WIL. RMIT, Melbourne (1999)
4. Heron, J.: The politics of facilitation: Balancing facilitator authority and learner autonomy. In: Mulligan, J., Griffin, C. (eds.) Empowerment Through Experiential Learning, pp. 66–75. Kogan, London (1992)
5. Dressler, S., Keeling, A.E.: Student benefits of co-operative education. In: Coll, R., Eames, C. (eds.) International Handbook for Co-Operative Education: An International Perspective of the Theory, Research and Practice of Work-Integrated-Learning, World Association for Co-operative Education, Boston (2004)
6. Fraser, S., Bosanquet, A.: The curriculum? That's just a unit outline, isn't it? Studies in Higher Education 31(3), 269–284 (2006)
7. Biggs, J.: Teaching for quality learning at university: What the student does. Open University Press & Society for Research into Higher Education, Buckingham (1999)
8. Bates, M.: Work-integrated curricula in university programs. Higher Education Research & Development 27(4), 305–317 (2008)
9. London Metropolitan University: Staff Guide and Resource Pack for Assessed Work Placements, https://intranet.londonmet.ac.uk/studentservices/careers/university-staff/placementsguide/placementsguide_home.cfm
10. Jimmy, J., Randhawa, S.: Work Integrated Learning for Engineering Students at Flinders University. In: Proceeding of 20th Australasian Association for Engineering Education Conference, University of Adelaide (2009)
11. A Supervisor's Road Map for Work Placement in NSW Government Agencies), http://www.docstoc.com/docs/74925036/Student-Work-Placement-Journal
12. Collis, C.: Developing Work-integrated Learning Curricula for the Creative Industries: embedding stakeholder perspectives. Learning and Teaching in Higher Education 4(1) (2010)

An Analysis on the Forming Reason of Dominant Firms of China

Hongli Ji[1,2,*], Yanqiu Liu[3], and Fanwen Wu[3]

[1] Renmin University of China, Beijing, 100872
[2] Inner Mongolia Finance and Economics College, Hohhot, 010051
[3] Inner Mongolia University of Technology, Hohhot, 010051
liuyanqiu2002@163.com

Abstract. China has gained outstanding achievements in economic development since she began opening up 30 years ago. Many dominant firms rise in the Chinese market and overseas and accelerate the economic development. But compared with that in the developed countries, there is still a long way to go. Dominant firms both home and abroad have common feature in the process of growing, so based on some theories and proceeded from China's actual conditions, the analysis on the forming reason of dominant firm will provide scientific basis for the development of dominant firms.

Keywords: dominant firm, large firm, leading firm.

1 Introduction

We are always concerned about how to make firms bigger and stronger as the strength of firms stand for the economic development of the country. The large firms are the dominant driving force of the economy. While the economy of our country has leapt forward in recent years, the result of economic development is unsatisfactory. As the processing factory of the world, there are many champions of processing in China, which is about the goods of costumes, toys, bags, umbrellas, lighters, stationery, and buttons and so on. China is second to none in outputs of PC, laptop computer and a number of components and devices, but we are working for others. Why corporations such as Kodak, IBM, Microsoft, Intel, and Google are not native to China? Why are the dominant firms so few in China? So, my question is how the dominant firms get the market power?

1.1 The Definition of Dominant Firm

It has long existed analysis for such a market structure of dominant firm in industrial organization , William • G • Shepherd (2007) depending on the state of competition and

* Hongli Ji is with the School of economics, Renmin University,Beijing,100872, China, and Inner Mongolia Finance and Economics college, Hohhot, 010051 (phone: 0471-3661206; e-mail: jhl406@ruc.edu.cn); Yanqiu Liu is with the School of Management, Inner Mongolia University of Technology, Hohhot, 010051 China.

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 37–43.
springerlink.com

monopoly, divided market structure into four main categories: pure monopoly, the dominant firm, strict oligopoly market, effective competition.[1] The dominate firm refers to one firm holds more than 40% of market share, and there is no similar competitors which are small firms and have no effective threat to the dominant firm. The higher degree of market dominance, the market structure is more close to a monopoly, but its market power is not complete. Dennis • W • Carlton (2009) considered this market structure as the dominant firm - the marginal competitors structure, A dominant firm which holds a large market share of large businesses, and many marginal firms which hold small market share are all existing in this market structure .[2]The dominant firm is also known as a leader. Domestic scholars think the dominant firm as the leading enterprises.

This paper mainly analyzes on the forming reason of the dominant large firm. Dominant firm as a special oligopolistic market structure is the most important expression of strong monopoly power in the real world. The dominant firm in the market is like Monopoly which can control its own price of products independently and make choice of production. The market share and dominant position express the efficiency of the dominant firm, so there is no doubt that the dominant firm is most efficient firm in this market structure.

2 Dominant Firms of China

There are many changes in our country after reforming and opening, we have changed from stagnation and a closed-door policy to reform and a policy of opening to the outside world, and we are carrying out all sorts of reforms. And we have been changing economic system from planned-economy to market-economy. And company's ownership has changed from in only state-owned to in different forms that may be stated-owned, collective-owned, individual-owned and privately-owned. Most of all is that Chinese firms experienced a progressive government-led top-down reform, so I agree the issue that firms are "embedded" in society, and firms must continue to understand, adapt, and integrate into the environment in order to avoid the effect of external environment, thus they could independently focus on their core business. Although it is still existed the low industry concentration, non-standard behavior, imperfect mechanism and so on, firms have more room to grow and have more opportunities to win the dominant position. But it is lack of the giant multinational firms which can coordinate and control of worldwide production, investment and marketing, which is very important to enhance the allocation efficiency of resource and scale advantages, promote the industrialization, establish and improve market economic system.

2.1 Dominant Firms and the Industry in China

As collection of data in China is more difficult, this is just an experiment to make inference which the dominant firm is in different industries, and it is also based on the existing literature and personal experience. There are the dominant firms of the domestic market of China, and they are all hold the larger market share. such as Wanglaoji holds 60-80% market share in the herbal tea market, Glanz 50-60% in the microwave oven market, Yadu 50-80% in the air purification market, Good boy 80% in the Baby carriage market, and CIMC in the reefer container market and so on.

Details are as follows table 1.

Table 1. China's major dominant firm on 2010

company name	Which market	market share approximation (income%),
Lenovo	PC	32%
Huawei	Network equipment manufacturing	60%
Haier	Washing machine	30%
Glanz	Microwave	50-60%
Yadu	Air purification products	50-80%
Jiuyang	Soybean milk machine	68%
Wanxiang group	Universal joint	60-70%
Good boy	Baby carriage	80%
Julonggroup	Nail clamp	60-70%
BYD	Nickel-cadmium batteries	40-50%
CIMC	Reefer container	50%
DJB(Wanglaoji)	Herbal Tea	60-80%
Chengde Lulu	Almond milk beverage	90%
Huiyuan	Pure fruit juice	56%
Nanfang	Paste products	68%
Xizhilang	Jelly	70-80%
Wangwang	Rice cracker food	68%
Tencent	Instant messaging software	76%
Baidu	Search engine	64%
Taobao	C2C	85%
Alibaba	B2B	80%

Note: Data comes from the network data and personal calculation

This table is not including all the dominant firms, and do not take into account others hidden in a large diverse enterprise. These dominant firms in China now are all in an absolutely dominant position without strong competitors, which is a leader in their markets. Although they have not all market shares, they are able to act like a monopoly. But the dominant firm is rare, and maybe also exists in some local places, such as the largest local hospitals and schools and so on.

2.2 The Main Features of Dominate Firms of China

The international dominant companies such as Standard Oil company during1870 to 1920, General Motors during the 1930s to the 1980s, IBM, and Kodak, Microsoft, Apple, Intel, Amazon, Google, etc., these companies is well known to the people, and attract attentions of many scholars. Compared to these international dominant companies the dominant firms of China have such features.

First, the top international dominant firms have developed generally in market-based economy for a long time, but ours developed in market-based economy for a relatively short time. The firm based market-based economy was emerged in the 1980s, according to the table, let see the established time of the dominant firms: the longest is Glanz for 32 years, the shortest is Taobao more than seven years, an average of

20 years. The international dominant companies operating time in addition to Nokia, Intel and Microsoft, other companies have all maintained at least 40 years, and the three exceptional companies have become the leaders in the early period of industry development. With short established time Chinese dominant firms have lost their first-mover advantage in the international market, have no significant accumulation process, and face fierce international competition. Summing up, we can state that the dominant firms is facing more difficult situation to gain a foothold in the world market.

Second, the dominant firms in different industries are just in the market of China, not in the world market, Few of them is known by the world, and most of them have no strength and willing to go to the world market. For example, the dominant firms in the 500 Fortune Companies in the world is only two, they are Huawei and Lenovo, two private enterprises. Chinese dominant firms is generally only in subdivided market, such as herbal tea, fruit juice of beverage industry; refrigerators, water heaters, microwave oven, soybean milk machine of household appliances industry; also batteries, universal joint, a clamp, etc. Despite few of dominant firms have gone to the world market and get more market share, the gap in the size and performance between China and advanced countries in terms of dominant firms is quite wide. All in all, the dominant firms of China are quite rare, and compared to the international dominant companies most of them are small in single kind market.

Third, according to Chinese industry classification, these dominant firms in addition to Tencent, Baidu, Taobao and Alibaba belong to computer services industry, the others are included in the secondary industry, especially the manufacturing industry, detail as including food manufacturing industry, beverage manufacturing industry, battery manufacturing industry, communications equipment manufacturing industry, computers and other electronic equipment manufacturing industry, daily chemical products manufacturing industry, textile, manufacturing, computer manufacturing industry, household electrical devices manufacturing industry and so on. Although these dominant firms are constantly innovating, international large companies are still controlling the market in the high-tech industry.

3 Reasons for the Formation of Chinese Dominant Firms

Many literary scholars have focused on the analysis of reasons for the formation of dominant firms, the reasons include low-cost, preemptive advantage, network effects, strategy, technique, professionalism, and learning from practice and so on, which is very important for development of firms. There are many factors caused the formation of dominant firms in the real world, and the factors interact between each other, so it is very complex to fully explain. In different countries and industries firms have different factors to get the significant market power. First of all, we agree that the ability in every aspect of dominant firms is the main reason. This ability is reflected on adaptation to the external environment, good operation including production, marketing, management, and coordination. If the firm can coordinate well, whether internal or external factors can work to the growth, and one of them maybe bring significant opportunities for the potential growth. From the characteristics of China's economic development the following sum up to the important reasons promoting the formation of dominant firms.

3.1 The Wisdom of Entrepreneurial

Good entrepreneurs can well coordinate production, marketing and management, and more important in China is to be able to fully understand and adapt to the external environment, thus to judge the trend of the future development of the industry. It is the key for the firm to win significant market power to adapt the government and its policies. The dominant firms of China are almost the private-based firms, so without the help of government they faced the lack of money at beginning. Firms through political participation, communication with government, political alliance, material incentives and other strategies create their own favorable conditions and get the support from government.[3] This shows that Local and national government protectionism is also one of the forming reasons of dominant firms. As Huawei and Lenovo are initial State-owned firms, their leaders must rational use disproportionate support of government and correctly understand the policy, and be able to not only break the shackles of the old system, but also to seize the opportunities of the new system, thus, the entrepreneurs successfully improved shareholding system reform and promoted the standardization and Systemization of management, and established effective corporate management structure.

3.2 The Continued Innovation

The key reason for the widening gap between Chinese and foreign large firms is technological innovation, which is closely related to the improvement from “made in China” to “innovated in China”. Innovation includes not only in high-tech research and development but also innovation to process of production. A firm that can creatively integrate technologies and applications and make better use of China's comparative advantages, will be able to achieve low-cost and win the significant market power. For example, Li-ion battery production of BYD is using Semi-automatic production line, which is the Combination of foreign high technology and domestic cheap labor. The production line is using dozens of patented technology altogether. This is the good example of Self-directed innovation. Besides, Huawei R4 softswitch technology and easily folding design of Good boy is also independent innovation, and Good boy has become the leader in the domestic baby carriage market and come into the U.S. market.

3.3 The First-Mover Advantage

The firm Early entered the market can get the advantages which later entrants can not get but must pay as a cost. For example, the first firm has already gained a significant market share, and has a large customer base, and the image of the product has become a part of customers, etc. which constitute the industry's barriers to entry. All the advantages become into disadvantages to new entrants, it is very difficult for new entrants to enter the market and share profit. Many of China's dominant firms have time to reach optimal size to gain market dominance just because of the first-mover advantage, such as: Wanglaoji, Chengde Lulu, Nanfang, Huiyuan juice, which have been the first brand in the subdivided market, and their products have been accepted by the people, Especially, Haier earned a reputation for the first to provide after-sales service to hold the market power.

3.4 Focus on New and Blind Market

With full understanding of the Chinese market, some of the dominant firms go the way of specialization of production. Because Chinese firms face competition not only from domestic firms as well as the large powerful foreign firms, most of them choose the market that have been ignored by large firms in the blind spot of competition and is small but with potential development, to start product and grow gradually, thus they can avoid fierce competition. So in this way they have time to gradually expand the scope of the market and eventually become the dominant firm of domestic market and the world champion of a single product.

In 1980 Wanxiang Group chose universal joints which is used to repair automotive, from then on it focus on this product and become first band in China in 1983. BYD select and specialize in the production of nickel-cadmium battery which is as breakthrough to market, because it can avoid the huge investments. It takes10 years to become the world battery king of power tools, cordless phones, electric toys, and market share of the three kind of products respectively is 45%, 72%, 30%.[4] Julong Group chose to product a nail clamp in 1998, has now become the NO.1 manufacturer of a clamp. Good Boy chose baby carriage, CIMC selected reefer container and semi-trailer, Yadu chose to enter a new air purifier industry, etc. They are opening up the industry and have become the dominant firms for focusing on small and new market.

3.5 Network Effect Causes the Formation of Dominant Firms and Make Them More Productive

Internet industry recovered in 2003, this new industry re-entered a period of rapid growth, Tencent, Baidu, Alibaba, Taobao, etc. as the dominant firms of networks benefit from the microchip and information technology. Another possible reason is increasing returns which are causes by network externalities. Network externalities means that the more users use the network, the greater value takes to users and firms. With a strong users base these dominate firms of networks can easily expand the business and reduce the cost and benefit from the learning by doing, which make them more efficient.

3.6 Merger and Acquisition, Joint Venture and Purchase

Enterprise merger is the main way to outside expansion for firm, the big foreign firms expand by merger with and acquisition of other enterprises. Because of the small firm size in a single production, mergers and acquisitions is rarely happens, but it will be a good way to expand in the future. Moreover, joint Venture and business affiliation are the main way to obtain capital and technology. Lastly, it is worth to mention that because It purchased the business of IBM's PC, Lenovo have become the largest of China and world's third largest personal computer maker, so to purchase the business is significant for its formation of dominant position.

4 Conclusion

There is a strong and potential market in China, and you can say that China has no small market in the future. Although the market is with high but low-level competition, instable pattern, low barriers to entry and low concentration, but for enterprises, it is both opportunities and challenges. In such conditions dominant firms could maintain a dominant position and there are many markets for firms to win. With the rapid economic development of China, enterprises with a Chinese symbol will grow gradually up and become the top dominant international enterprises.

References

1. Shepherd, W.G., Shepherd, J.M.: The economics of industrial organization. Renmin University of China Publishing house, Beijing (2007)
2. Carlton, D.W., Perloff, J.M.: Modern industrial organization. San Lien Book Store and Shanghai People's Publishing Agency, Shanghai (2009)
3. Zhang, X., Zhang, J.: Understanding Chinese firms from multiple perspectives. Beijing University Publishing house, Beijing (2010)
4. Lou, X.: To be the No.1, China Enterprise's historical opportunity. Beijing University Publishing house, Beijing (2009)

A Survey of Science Literacy Level for Senior High School Students in Taiwan

Chao Chih-Yang, Yang Shih-Chun,
Yen Chia-Sung, and Lin Yong Shun

National Chunghua University of Education, Taiwan
yaksbee@gmail.com, teljsy@ccu.edu.tw,
yslin_10@yahoo.com.tw

Abstract. The main purpose of the study was to explore the literacy of environmental science level for senior high school student in Taiwan. This study received 90.87 % return rate from the students. After the process of data analysis, the following results were obtained: (1) On the whole, senior high school students are only above average in scientific literacy. (2) Senior high school students have lower level cognition on the science applications.(3) Senior high school students have higher level cognition on the science ethics. (4) The public senior high school students generally have higher standard of scientific literacy than the private senior high school students. (5) Female senior high school students tend (tended) to pay attention to the science ethics, while the male high school students tend(tended) to have higher level cognition on the science applications. (6) The senior high school students whose parents have higher education have higher standard of scientific literacy. (7) The senior high school students whose parents have specialized jobs have higher standard of scientific literacy.

Keywords: Science, Scientific literacy, Senior high school.

1 Introduction

The development of science and technology is a sign of the competitive power of a state as well as the guarantee of the quality of life. In order to make science useful in people's daily life (Popli, 1999), all studies must be done earnestly since childhood, so is the cultivation of scientific literacy. In terms of promoting the planning of science education in Taiwan, nature study begins to be taught in grade 5 of national elementary schools, and biology, physics and chemistry are taught in secondary schools. This study also expects to explore the results of Taiwan's dedications to the promotion of science education in recent years, to know the students' cognition of scientific literacy, and to compare the differences of the cognition of scientific literacy of the students with different backgrounds. Whether the scientific literacy of the students in Taiwan can be improved or not under so many educational approaches deserves our long-term and further study and observation.

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 45–52.
springerlink.com

2 Discussion in Relevant Documents

2.1 Definition of Scientific Literacy

The wording "scientific literacy" is commonly used in the US, but referred to as the science of public understanding in the UK. Science is a noun familiar to us, but few people know its true meaning. Because to date there have been no a uniformed definition and common understanding of scientific literacy, each has his view (Durant & Thomas, 2006). World professional organizations and experts all have different definitions and views on scientific literacy (NRC1996; Eisenhart et al., 1996; Deboer, 2000). Scientific literacy is basically defined as two kinds, namely knowledge-centered and social sciences-centered (Bryan, John & Gregory, 2005).

2.2 Connotations of Scientific Literacy

Doris (2004) pointed out that it was not easy to define scientific literacy; in addition, there are many definitions and no definite and proper common understanding. Though experts and organizations of every country have different definitions, and there are no integral understandings on the implied meaning of scientific literacy from every side, it is undoubted that science should be one of the basic qualities of citizenry. It can be elaborated as below:

2.2.1 Science Nature

Rillero (1998) believes that if the public can have more understanding of the universe, the biosphere, and our body, they will have a quite different new horizon on health, the environment of the earth, and related issues. The most fundamental concept of scientific literacy is to have science nature (Pella, O'Hearn, & Gale, 1966; Showalter, 1974). If only one has a preliminary understanding of "science itself" and "science-related applications", he then has the so-called "scientific literacy" (Hurd, 1958; Durant, 1993; Rillero, 1998).

2.2.2 Science Process

The concept of science and the technical competence of implementation are the basic factors in scientific research, but scientific research is not only time consuming but strenuous. Therefore, in scientific research, whether the researchers' mental operating ability is sufficient for the practice of research is also an important factor for the success of the research. Quite a few scholars and experts believe that those with scientific literacy should possess the ability of science process (Pella, 1967; Agin, 1974; Showalter, 1974; Carin & Sund, 1989; Durant et al. 1989; Yager, 1996; Trowbridge & Bybee, 1996; Tao Xiandu 2008).

2.2.3 Science Attitude

Attitude refers to the evaluation and behavior disposition of an individual to a specific object in an environment, it is a disposition got after learning. Both scientific attitude

and value develop, and come into being before an individual grows into an adult. The people with scientific literacy may have the scientific spirit and attitude in terms of seeking truth, feeling the strength and beautification of science and loving exploration (Durant et al., 1989; Yager, 1996). In addition, the curious, open and skeptical attitudes to new scientific ideas are a must in scientific exploration, and many scholars also think a positive attitude should be held for science (Hurd, 1958; Durant et al., 1989; Durant, 1993). Li Ling (2006) believes science attitude refers to the basic understanding and consciousness of the influences of science and technology to society.

2.2.4 Science Ethics

Science and technology should have some contributions and limitations in promoting human well-being (Carin & Sund, 1989). Science's great contributions to human is undoubted, but it remains a big problem that how to properly use the scientific discoveries and applications. In exploring the universe, being able to observe the values of science is also what the researchers should strictly stick to (Pella, 1967; Agin, 1974; Showalter, 1974). Though there are not many scholars and experts advocating that science being included in scientific literacy, in the 21st Century, paying attention to science ethics may be an issue being unable to be removed, because this is recently a phenomenon of science's shock on society and also a hot issue emerging from scientific literacy.

2.2.5 Science Application

Human's scientific and technological inventions all initially came from scientific discoveries. As a citizen, we should have the ability of utilizing science concepts, process skills and values, and we also should be able to apply scientific research methods and knowledge to resolve daily problems in life and to make decisions (Showalter, 1974; Carin & Sund, 1989; Trowbridge & Bybee, 1996; Yager, 1996). The significance of scientific literacy lies in that possessing scientific literacy can make an individual "know the research process of science, know that these research processes are the crystallization of human's imagination and the rules of nature, and its purposes are to solve the problems encountered in life" (Hurd, 1958; Durant et al., 1989; Durant, 1993). An educated citizen should, on the basis of existing resources, methods and foundation, produce the ability of evaluating the quality of scientific knowledge.

3 Research Design

3.1 Research Structure

Questionnaire method is adopted in this study to know the cognition of scientific literacy of the students with different backgrounds. The variables of such backgrounds include: (1) school affiliation; (2) sex; (3) education of parents; and (4) parents' jobs. The research structure is as below (Fig. 1).

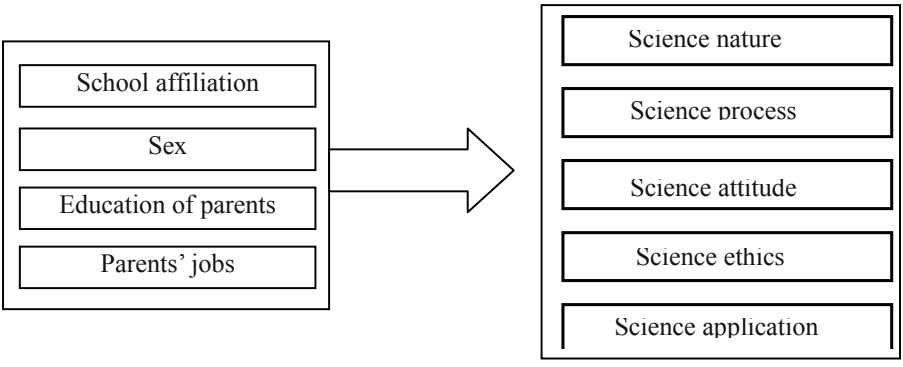


Fig. 1. Research structure

3.2 Sampling Method and Research Participants

The main purpose of the study was to explore the literacy of environmental science level for senior high school student in Taiwan. Cluster sampling was adopted and 30 public and private senior high schools in Taiwan were chosen as samples. Fifty questionnaires each school, 1500 in total, 1388 returned, of which 1363 are valid. This study received 90.87 % return rate from the students.

3.3 Reliability Test of Research Measures

Likert 5-point Scale was adopted in the pretest of the study, 200 questionnaires were given out, 196 returned, return rate 98%. Cronbach α coefficient was adopted for reliability test and the overall reliability of scientific literacy was 0.87. The analysis basing on the statistics shows that every aspect constituting the scientific literacy and the overall results reach the level of significance, of which the highest is science ethics 4.30, second is science attitude 4.04, then science process 3.93, science nature **3.85 and science application 3.78.**

The questionnaire consists of 5 aspects, i.e. science nature, science process, science attitude, science ethics and science application, and each is defined respectively as: (1) science nature contains 7 questions: dynamic, reproducible, material, social, objective, nonexclusive scientific methods and public; (2) science process contains 8 questions: research problems, data collecting, research purpose, research structure, research object, research tools, data analysis and research results; (3) science attitude contains 7 questions: objectivity, enlightening, initiative, trust, respect, interest and doubt; (4) science ethics contains 6 questions: unjust obtainment, injury evaluation, consent obtaining, informing after the event, protecting the participants and without being controlled; (5) science application contains 6 questions: being aware of, finding the problem, study the problem, analysis and judgment, solving the problem, and results application. There are 34 questions in total.

3.4 Analytical Procedure

t-test and ANOVA were adopted as statistical methods. t-test was adopted to analyze the level of significance of every aspect. t-test was adopted to test the different background variables such as school affiliation and sex, and the student gets to know the difference between the scientific literacy of themselves and others. One-way ANOVA was adopted to test the different background variables, and the student got to know the difference between the scientific literacy of themselves and others. If the value of F reached the level of significance, Scheffes' method was adopted to further carry out the comparison after the event, to know the significant differences of each group. The statistical test level of every aspect was $\alpha = 0.05$.

4 Research Results and Findings

4.1 T-testing of the Aspects of Scientific Literacy

Each aspect of the scientific literacy was analyzed through t-test, all reached the level of significance ($p<.01$); the average of the overall scientific literacy was 3.71, of which the average of science ethics the highest, 4.07, and average of the science application the lowest, 3.39.

4.2 Differential Analysis of the Cognition of the Scientific Literacy in the Schools of Different Affiliations

The scientific literacy for senior high school students from schools of different affiliations reaches (reached) the level of significance for both every aspect and the whole. The public senior high school students generally had (have) higher standard of scientific literacy than the private senior high school students.

4.3 Differential Analysis of the Cognition of the Scientific Literacy of the Senior High School Students of Different Sexes

The scientific literacy for senior high school students of different sexes had (have) significant difference in the aspect of science ethics. Female senior high school students tend (tended) to pay more attention to the science ethics. In the aspect of science application, the male high school students tend (tended) to have higher level cognition.

4.4 Differential Analysis of the Cognition of the Scientific Literacy of the Senior High School Students Whose Parents Have Different Education

In the differential analysis of the senior high school students whose parents have different education, the aspect of science process reach (reached) the level of significance. The level for the group of students whose parents have college/university or master's degree is higher than that for the groups of students whose parents have junior college education or lower education. The aspect of science attitude and the whole reach (reached) the level of significance, the college/university group is higher

than the groups of junior college education or lower. The aspect of science application reach (reached) the level of significance, the doctor's degree group is higher than the groups of junior college education or lower.

4.5 Differential Analysis of the Cognition of the Scientific Literacy of the Senior High School Students Whose Parents Have Different Specialized Jobs

In the differential analysis of the senior high school students whose parents have different specialized jobs, only the aspect of science ethics didn't reach the level of significance and all others did. In the aspects of science nature and science attitude, the group of the student whose parents have senior specialized jobs is higher than the group of students whose parents have semi-technical jobs; in the aspect of science process, the group of students whose parents have specialized or senior specialized jobs is higher than the group of students whose parents have semi-technical and technical jobs; in the aspect of science nature, the group of the student whose parents have senior specialized jobs is higher than the group of students whose parents have semi-technical jobs; in the aspect of science application, the group of students whose parents have senior specialized jobs is higher than the group of students whose parents have semi-technical and technical jobs; in the aspect of the whole, the group of students whose parents have semi-specialized and specialized jobs is higher than the group of students whose parents have semi-technical jobs, and the group of students whose parents have senior specialized jobs is higher than the group of students whose parents have semi-technical and technical jobs.

5 Discussion

The study was to explore the literacy of science level for senior high school students with different backgrounds in Taiwan, while there was a lack of the study on such subject through the studies of the scholars of every country in the world in modern times, and there was a serious lack of relevant information for comparison. Therefore, the comparison and discussion with relevant studies cannot be given wholly but the statement of the findings of this study. The discussion is as follows:

5.1 On the Whole, Senior High School Students Are Only above Average in Scientific Literacy

From the view of average number, the cognition of senior high school student in Taiwan on the overall scientific literacy is only 3.71, slightly higher than 3.5. Though in 2006 PISA's scientific literacy appraisal ranked highly fourth in the world, it is not high considering the student's overall cognition. Li Ling (2006) found in the survey of secondary school biology teachers' scientific literacy that, the level of cognition of secondary school biology teachers was higher on the whole.

5.2 Senior High School Students Have Lower Level Cognition on the Science Applications

From the view of average number, the cognition of senior high school student in Taiwan on science application is the lowest, only 3.39, and t-test reached the level of

significance. The average failing to reach 3.5 tends to be lower. There remains to be greatly improved for Taiwan senior high school student's ability of applying science concept, knowledge and method into daily life, and they have lower level cognition on the science applications

5.3 Senior High School Students Have Higher Level Cognition on the Science Ethics

From the view of average number, the average of senior high school student in Taiwan in the aspect of science ethics was as high as 4.07, and t-test reached the level of significance. Senior high school students have higher level cognition on the science ethics

5.4 The Public Senior High School Students Generally Have Higher Standard of Scientific Literacy Than the Private Senior High School Students

In Taiwan, the public senior high school students generally have higher standard of scientific literacy than the private senior high school students, and they reach(reached) the level of significance in every aspect of scientific literacy and on the whole. The public schools in Taiwan enjoy government subsidy, so their tuitions are much lower than that of the private schools. Students want to study in public senior high schools must pass scholastic aptitude test, and those scoring high will study in public schools. The public senior high school students generally having higher standard of scientific literacy than the private senior high school students perhaps is because their marks in scholastic aptitude test is higher than those of private school students.

5.5 Female Senior High School Students Tend (Tended) to Pay Attention to the Science Ethics, While the Male High School Students Tend (Tended) to Have Higher Level Cognition on the Science Applications

Analyzing from the data, though Taiwan senior high school students having a higher cognition in the aspect of science ethics, from the angle of sexes, female senior high school students tend (tended) to pay attention to the science ethics, and reach (reached) the level of significance; in addition, from the angle of science application, the male high school students tend (tended) to have higher level, and reach (reached) the level of significance.

5.6 The Senior High School Students Whose Parents Have Higher Education Have Higher Standard of Scientific Literacy

Parents' education is closely related to students' scientific literacy. The senior high school students whose parents have higher education have higher standard of scientific literacy on the whole, and reach the level of significance, of which only science nature and science ethics fail to show the level of significance, and the other four aspects all show the parents' education and their children's scientific literacy have positive correlation. This is possibly because those parents have higher education background pay more attention to or are able to the enlightening their children's scientific literacy.

5.7 The Senior High School Students Whose Parents Have Specialized Jobs Have Higher Standard of Scientific Literacy

The data show that, if the parents have relatively specialized jobs, their children have higher performance, and it reaches the level of significance. If the parents have relatively specialized jobs, their children have relatively positive cognition in every aspect. This is possibly because the parents have specialized jobs can better help or cultivate the children's scientific literacy.

References

1. Tao, X.: Mass media and the cultivation of the scientific literacy of public. Social Scientist (2008)
2. Li, L.: Survey and thoughts on the current situation of the scientific literacy of secondary school biology teachers. Journal of Zibo Normal College (2006)
3. Bybee, R.W.: Achieving scientific literacy. The Science Teacher 62, 7 (1995)
4. Agin, M.: Education for scientific literacy: A conceptual frame of reference and some application. Science Education 58, 3 (1974)
5. Bryan, A.B., John, M.R., Gregory, J.K.: Scientific literacy and discursive identity: A theoretical framework for understanding science learning. Wiley Inter Science (June 30, 2005)
6. Carin, A.A., Sund, R.B.: Teaching science through discovery, 6th edn. Merril Publishing Company, London (1989)
7. DeBoer, G.E.: Scientific literacy: Another look at its historical and contemporary meanings and its relationship to science education reform. Journal of Research in Science Teaching 37, 582–601 (2000)
8. Doris, A.: Reflective scientific sense-making dialogue in two languages: The science in the dialogue and the dialogue in the science. Wiley Inter Science (August 9, 2004)
9. Durant, J.R.: What is scientific literacy? In: Durant, J.R., Gregory (eds.) Science and Culture in Europe, pp. 129–137. Science Museum, London (1993)
10. Durant, J.R., Evans, G.A., Thomas, G.P.: The public understanding of science. Nature 340(6), 11–14 (1989)
11. Eisenhart, M., Finkel, E., Marion, S.F.: Creating the conditions for scientific literacy: A re-examination. American Educational Research Journal 33, 261–295 (1996)
12. Hurd, P.D.: Scientific literacy: Its meaning for American schools. Educational Leadership 16, 13–16 (1958)
13. NRC. National science education standard. National Academic Press, Alexandria (1996), http://www.nap.edu/catalog/4962.html
14. Pella, M.O.: Scientific literacy and the H.S. curriculum. School Science and Mathematics, LXVII 4, 346–357 (1967)
15. Popli, R.: Scientific literacy for all citizens: Different concepts and contents. Public Understanding of Science 8(2), 123–137 (1999)
16. Rillero, P.: Process skills and content knowledge. Science Activities 35(3), 3–4 (1998)
17. Showalter, V., et al.: Program objectives and scientific literacy, part 5 and 6, prism. Newsletter of the Center for United Science Education 2, 3–4 (1974)
18. Trowbridge, L.W., Bybee, R.W.: Teaching secondary school science— strategies for developing scientific literacy, 6th edn., Englewood Cliffs, New Jersey (1996)
19. Yager, R.E.: Workshop Science/Technology/Society As reform in science education. Science Education Center, National Taiwan Normal University, Taipei, Taiwan, R.O.C (1996)

Performance Evaluation of the Regional Innovation in the Integration of Production, Education and Research

Chaoyi Chen, Bo Wang, and Luo Chen

School of Management, Tianjin University, Tianjin, China

Abstract. In accordance with our current region situation in integration of production, education and research, the paper, combined with the qualitative & quantitive analysis approach, offers the reasonable performance assessment towards the regional innovation in integration of production, education and research in 31 provinces. The paper, after analyzing the innovation performance evaluation indexes influencing the integration of production, education and research, obtain 14 major branch index in order to set up the evaluation index system. The paper, after analyzing data from 31 provinces by factor-clustering statistical method, gain 4 major indexes and begin the mark sequence and cluster analysis before the advantages and disadvantages are discovered and suggestions are put forward.

Keywords: Integration of production, education and research, innovation, performance evaluation, factor-cluster analysis.

1 Introduction

At present, our country is in the process of transformation from the planned economy to the market economy system, with fast economic growth and the people's life standard raising. However, at the same time, the unbalanced development of regional economy is also exposed, with some industry excess capacity, the waste of resources, which weakens our national comprehensive competitive power with other developed countries. Based on the international and Chinese economic and science & technology trend, our central government pays high attention towards our science & technology development, and takes the independent innovation as our national strategy to propel our structure adjustment and improve our country's competitiveness, and takes the scientific and technological progress and innovation as the first driving force of our economic development. Takes the independent innovation as the central linkage to adjust economic structure, transform our growth mode and improve our competitiveness, and takes the construction of national innovation state as an important strategy in the future.

In the background of the national innovation strategy, how to create an innovative system among the government, enterprises, universities, research institutes, and intermediary institutes of various resources for the harmonious development is urgent. The integration of production, education and research is the main way and the direction of national knowledge and technology innovation.

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 53–60.
springerlink.com

At present, the issue of the integration of production, education and research, particularly at the focus on the model selection and combination of the effect the performance evaluation is still the barrier or the bottleneck to different countries to a certain degree. This article summarizes the existing problems and relative solutions towards different cities and provinces by 14 performance index analysis from 31 provinces and cities.

2 The Index Data and the Evaluation Method towards the Innovation Ability of Regional Integration of Production, Education and Research

2.1 The Index System Setting

Under the principles of operability, simplicity, comparability and scientificity, the performance evaluation index system on regional integration of production, education and research innovation is established respectively from the main body combined (enterprises, universities, research institutes), the index combined (input, output, combining effect) and other aspects concerned indexes to merge, conclude and cut, finally it obtains the following 14 evaluation index, and establishes the evaluation index system.

(1) The application for a patent license
(2) The contract number of the deal in technology market
(3) The proportion between industrial enterprise new product value and the total industrial output
(4) The science and technology services subject number of the higher school
(5) The number of institutes science and technology activities
(6) The raising total funds of the large and medium-sized industrial enterprise technology activities
(7) The proportion of the raising funds in science and technology activities from government and the large and medium-sized industrial enterprise
(8) The research and development personnel full-time equivalent of large and medium-sized industrial enterprises
(9) The raising total funds in science and technology from higher colleges
(10) The proportion in the science and technology raising total funds between the enterprises and higher college
(11) The research and development personnel full-time equivalent of universities
(12) The total raising fund from the institutes
(13) The enterprise fund proportion in the research institute science and technology activities
(14) The research and development personnel full-time equivalent of institutes

2.2 Acquisition of the Evaluation Data

This paper is mainly involved in the 14 evaluation index from data of 31 provinces in our country, which is mainly from the published yearbooks and official website.

(1) National Bureau of Statistics of China, The China statistical yearbook (2010), Beijing: China statistical publishing house.
(2) National Bureau of Statistics of China, the science and technology department. The China science and technology statistics yearbook (2010), Beijing: China statistical publishing house.

After getting the original data, the article has the synthesis of calculation to the parts of indexes, and has the establishment of a evaluation database.

2.3 Introduction of the Evaluation Method

2.3.1 Factor Analysis

Factor analysis is a technology which can simplify the multi-variations in order to decompose the original variables. And it can induce the potential categories, and takes the indexes strongly related as a class,Factor analysis is a statistical method used to describe variability among observed variables in terms of a potentially lower number of unobserved variables called factors. Factor analysis searches for such joint variations in response to unobserved latent variables which is the promotion of the main factor analysis. It can be applied into the reflection and the generation of some certain original variables.

2.3.2 Cluster Analysis

Cluster analysis begins its classification from the characteristics of numerals, is the result of the combination of numerical taxonomy and multivariate statistics. It is simple, and well-done in classification effect. The basic principal of the analysis is to observe the relative ties between the samples in accordance with the characteristics of numerals. And the relationship between samples decided by its distances, once the distance between the definitions has been offered, the closer samples will be classified one.

3 The Evaluation of the Regional Innovation in the Integration of Production, Education and Research Based on Factor-Cluster Analysis

3.1 The Selection and Analysis of Innovation Factor

In the light of factor analysis principle, after inputting the 31 group dates of 14 index and using standardization and factor analysis by software SPSS12, we can get the following results. From the Table 1, the value of KMO Test is 0.741 which is greater than 0.6, and the result of significance testing is less than 0.05 base on the Bartlett's Test, so it is fit for taking the factor analysis.

Table 1. The KMO and Bartlett's Test of Integration

KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.741
Bartlett's Test of Sphericity	Approx. Chi-Square	563.977
	df	91
	Sig.	.000

From the result of Table 2, we can see the eigenvalues of all the first four factors are greater than 0.9, their contribution rate of variance is respectively 50.465%,19.749%,8.684% and 6.57%, The contribution rate of accumulative total of variance is 85.468% which is greater than 85%. The first four factors include more than 85% information of all the variances, so we choose the first four factors instead of the original variances to make an analysis.

Table 2. The Variance Analysis of Integration

Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	7.065	50.465	50.465	7.065	50.465	50.465	5.559	39.707	39.707
2	2.765	19.749	70.214	2.765	19.749	70.214	3.840	27.429	67.137
3	1.216	8.684	78.898	1.216	8.684	78.898	1.432	10.230	77.367
4	.920	6.570	85.468	.920	6.570	85.468	1.134	8.101	85.468
5	.817	5.838	91.305						
6	.494	3.527	94.833						
7	.329	2.351	97.183						
8	.135	.965	98.149						
9	.121	.866	99.015						
10	.065	.464	99.479						
11	.030	.214	99.693						
12	.029	.210	99.904						
13	.010	.069	99.973						
14	.004	.027	100.000						

Extraction Method: Principal Component Analysis.

Table 3. Component Matrix of Integration Index

Rotated Component Matrix[a]

Index	Component			
	1	2	3	4
The application for a patent license	.230	.902	-.113	.166
The contract number of the deal in technology market	.911	.283	.021	.093
The proportion between industrial enterprise new product value and the total industrial output	.707	.063	-.304	.297
The science and technology services subject number of the higher school	.403	.680	.264	.232
The number of institutes science and technology activities	.950	.119	.015	-.085
The raising total funds of the large and medium-sized industrial enterprise technology activities	.128	.962	-.029	.042
The proportion of the raising funds in science and technology activities from government and the large and medium-sized industrial enterprise	.144	-.232	.831	-.074
The research and development personnel full-time equivalent of large and medium-sized industrial enterprises	.073	.959	-.026	.054
The raising total funds in science and technology from higher colleges	.869	.374	.229	.105
The proportion in the science and technology raising total funds between the enterprises and higher college	.146	.335	.557	.233
The research and development personnel full-time equivalent of universities	.729	.520	.243	.070
The total raising fund from the institutes	.953	.074	.198	-.062
The enterprise fund proportion in the research institute science and technology activities	.008	.192	.057	.929
The research and development personnel full-time equivalent of institutes	.925	.036	.315	-.030

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.
a. Rotation converged in 5 iterations.

From Table 3, the loads of the contract number of the deal in technology market, the proportion between industrial enterprise new product value and the total industrial output, the number of institutes science and technology activities, the raising total funds in science and technology from higher colleges, the research and development personnel full-time equivalent of universities, the research and development personnel full-time equivalent of institutes in factor 1 are larger, we can call factor 1 as output of integration and input of university's and institute's personnel factor; the loads of the application for a patent license, the raising total funds of the large and medium-sized industrial enterprise technology activities, the research and development personnel full-time equivalent of large and medium-sized industrial enterprises in factor 2 are larger, we can call factor 2 as the raising funds and personnel input of enterprise factor; the loads of the proportion of the raising funds in science and technology activities from

government and the large and medium-sized industrial enterprise, the proportion in the science and technology raising total funds between the enterprises and higher college in factor 3 are larger, we can call factor 3 as integration of enterprise, government and university factor; the loads of the enterprise fund proportion in the research institute science and technology activities in factor 4 are larger, we can call factor 4 as integration of institute and enterprise factor.

3.2 The Score and Sort of Total Factors

Table 4. The Integration Total Factors Score and Sort of Province Cities

Total Factors score	Region	Sort	Total Factors score	Region	Sort
1.9587	Beijing	1	-0.1913	Anhui	17
1.0412	Jiangsu	2	-0.2333	Heibei	18
0.8655	Shanghai	3	-0.2553	Fujian	19
0.7598	Guangdong	4	-0.2926	Jiangxi	20
0.5915	Zhejiang	5	-0.3016	Shanxi	21
0.3368	Liaoning	6	-0.3467	Yunnan	22
0.3217	Hubei	7	-0.3624	Guangxi	23
0.3207	Sichuan	8	-0.3789	Gansu	24
0.3111	Shan xi	9	-0.4190	Neimenggu	25
0.2982	Shandong	10	-0.4952	Guizhou	26
0.0531	Hunan	11	-0.5513	Xinjiang	27
-0.0407	Tianjin	12	-0.6152	Ningxia	28
-0.0639	Henan	13	-0.6548	Hainan	29
-0.0775	Heilongjiang	14	-0.6575	Qinghai	30
-0.1150	Jinlin	15	-0.6612	Xizang	31
-0.1447	Chongqing	16			

3.3 Cluster Analysis of Integration Innovation

Cluster Analysis can classify the dates automatically in accordance with the differences of their properties with no prior knowledge. In accordance with cluster analysis principle, we can get the tree diagram of the performance evaluation result of the integration.

The performance evaluation result of the integration can be classified into four types. The first type is Beijing; the second type includes Hubei, Sichuan, Shan xi, Liaoning, Shandong, Shanghai, Guangdong, Zhejiang, Jiangsu; the third type includes Jiangxi, Shanxi, Hebei, Fujian, Yunnan, Guangxi, Gansu, Neimenggu, Hennan, Heilongjiang, Tianjin, Jilin, Chongqing, Anhui, Hunan; the fourth type includes Hainan, Xizang, Ningxia, Guizhou, Xinjiang.

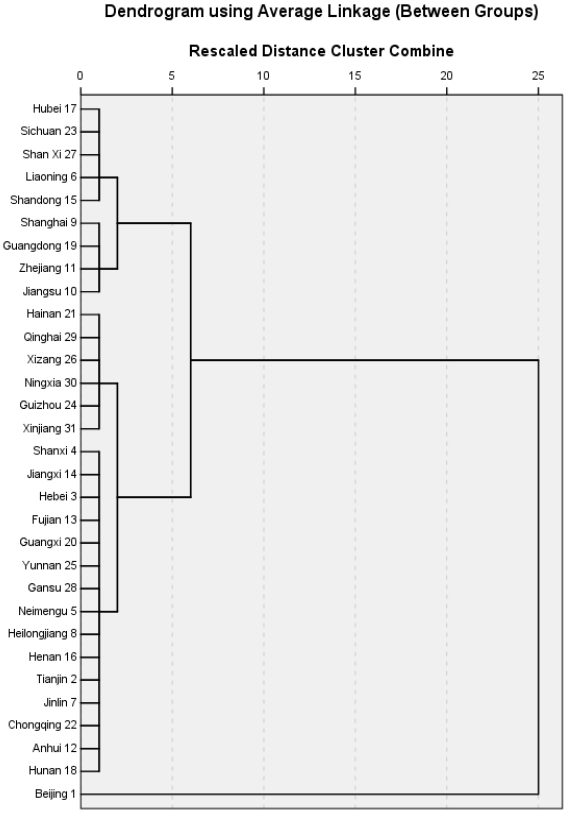


Fig. 1. Performance Evaluation Results Tree Diagram of Integration

4 Performance Evaluation Result Analysis of Regional Integration Innovation

(1) From the comprehensive factor score, Beijing, Jiangsu, Shanghai, Guangdong, Zhejiang province are top five in China. The reason of Beijing's high score is factor score 1st, output of integration and input of university's and institute's personnel factor; the reason of Jiangsu's and Guangdong is factor score 2nd, the raising funds and personnel input of enterprise factor; The reasons of Shanghai's are the factor score 1st and 2nd; The reasons of Shanghai's are the factor score 2nd and 4th. From all above, we can see most of the high comprehensive factor score provinces have a good achievement in the output and investment of integration of production, education and research. In addition, the achievement of cooperation between enterprise and other departments is much better.

(2) The second group fall behind Beijing mainly in produce of integration of production, education and research, it prove that the investments are not transformed to valid Technology Achievements very much. We need to find the reason by analyzing.

Meantime, we also can see the second group surpassing trend. The scores of the third group are all negative. There are some shortcomings in the investment, output and cooperation, we should work hard more.

(3) From the factor score, we could see there is a "three lags of the tripod" pattern appears in the integration of production, education and research. It is the economic community of cities around the Bohai Bay, led by Beijing; the economic community of cities around Yangtze River Delta, led by Shanghai and Jiangsu; the economic community of cities around Pearl River Delta, led by Guangdong. The cities have a great innovation capacity in the integration of production, education and research, giving impetus to the development of its adjacent regions.

(4) Each province and city should increase the input of fund and manpower, strengthen the cooperation between enterprise, institute and university, promote the commercialization of research findings. The governments of each province and city should build some communication platforms and establish a specialized division for organizing, cooperating and supervising so that guarantee the success of integration of production, education and research. The cooperation model of integration should have some innovation and breakout so as to arouse the enthusiasm of all sides. We should also strengthen innovation of the motivation and Interest Mechanism so that the investment of integration could get the reward in the highest degree.

References

1. He, J., Li, Y.: The Regional Innovation System of Research University, vol. (5), pp. 87–106. Tsinghua University Press (2008)
2. You, W., Zhou, S.: The Research on Mechanism optimization of Integration of Production, Education and Research. Science of Science and Management (10), 9–12 (2004)
3. Tang, X.: Research on the Technological Innovation of Regional Industry-University-Research Cooperation. Dissertation for the Degree of Harbin Engineering University D. Management, 103–138 (2009)
4. Lee, Y.: The Sustainability of University Industry Research Collaboration: An Empirical Assessment. Journal of Technology Transfer (25), 111–133 (2000)
5. Cai, B.: Innovation and Integration of Production, Education and Research, pp. 13–16. Guangdong Economic Press, Guangzhou (2008)
6. Wu, Y., Huang, Z.: A Study on the University Performance Evaluation of Integration of Production, Education and Research. Sichuan Province Government Agency College Journal (1), 39–41 (2008)
7. Li, W.: Multivariate Statistical Analysis. Beijing University Press, Beijing (2008)

Research on the Buffer Sizing Approach in Critical Chain Scheduling in Perspective of Flexible Management

Guofeng Ma*, Lei Li, and Zhiqin Chen

School of Economics and Management, Tongji Univ., Shanghai 20092, China
mgfsjtu738@sina.com

Abstract. An improved buffer sizing approach in critical chain scheduling was proposed based on the flexible management in projects. Through identifying the uncertainties which have influences on the duration of each activity and taking the project characteristics which include resource tightness, network complexity and risk preference of the project manager into consideration comprehensively, it decreases the subjectivity in the decision-making process effectively and makes the buffer sizing more reasonable.

Keywords: Critical chain, buffer, flexible management, uncertainties.

1 Introduction

Critical Chain Project Management (CCPM) is the application and development of the Theory of Constraints (TOC) in the areas of project management. In order to protect the critical chain and be able to implement the projects as planned, Goldratt introduced the concept of buffers [1]. Adding buffers into the schedule can shorten the construction period in a large scale without increasing the completion risk [2], and minimize the impact on implementation of the plan due to uncertainties. There have been several researches on determining the buffer sizes nowadays, and the most ones commonly used methods including two kinds: the Cut and Paste Method (C&PM) and the Roots Square Error Method (RSEM) [3]. Zinovy and Radovilsky applied the queuing theory to determine the buffer sizes [4]. Yongyi Shou and K T Yeo divided the uncertainty of activities into three levels: high, middle and low, and proposed conversion coefficients of buffer sizing [5]. In recent yeas, Oya I. Tukel has integrated two project characteristics, one of which is resource tightness and the other is network complexity to determine buffer sizes [1]. Behzad Ashtiani et al. addressed that duration of activities obeyed the lognormal distribution, and sized the buffers based on Root Square Error Method (RSEM) [6]. Luong Duc Long et al. [7] and Lili Chen et al. [3] determined the duration of actives and buffer sizes based on the fuzzy

* This research is supported by the National Natural Science Foundation of China for Young Scholars (No. 70802045), Ph.D. Programs Foundation of Ministry of Education of China (No. 20070247027) and Humanities and Social Science Foundation of Ministry of Education of China for Young Scholars (No. 07JC630020).

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 61–68.
springerlink.com

theory. In addition, in [2], the effects of project resource intensity, network structure complexity and managers' preference to risks on buffers were considered and integrated to determine the buffer sizes. Peng Gao proposed an approach based on the grey theory [8].

The references reviewed have proposed approaches to determine the buffer sizes at a different angle, and enriched the research of Critical Chain Project Management (CCPM). In most of these studies, however, due to the lack of determination of parameters in probability distributions and a large quantity of subjectivity existed in the process of activities duration estimation, many of these approaches are short of scientific decision-making, and it makes the project schedule management rigid.

In this paper, according to the recognition of flexible management, the source of uncertainty was first identified and analyzed to decrease the subjectivity in the estimation of the duration of each activity. And then three project characteristics including the resource tightness, network complexity and risk preference of the project manager are integrated comprehensively to optimize the buffer sizes.

2 The Buffer Sizing Approach in Perspective of Flexible Management

Flexible management derives from the recognition of flexibility. Flexibility is considered as the capability that the system can rapidly respond to the uncertain environment [9]. The capability and uncertainty are the key factors in the definition of flexibility [10]. To a certain degree, CCPM can decrease the influence of uncertainties over the project schedule. There shows flexibility existed in the CCPM. Project schedule management with much rigidity can cause the system not respond to the environment changes rapidly and effectively. Finally, the goals of project management: cost, schedule and quality are dramatically disjointed. CCPM is a new strategy where the behavior hypothesis of human, including Students Syndrome and Parkinson's Law are considered. However, human elements are not the unique factors leading to the rigidity. The reasons causing the rigidity should be a systematic composition where the things, environment, management and other uncertain factors must be taken into consideration. Consequently, in perspective of flexible management, the uncertainties having the influence on the duration of each activity should be first identified and analyzed in order to make the flexible mechanism of schedule management clear.

The uncertainties have two types: cognitive ones and accidental ones. Cognitive uncertainties are relevant to the lack of recognition and experience for projects, as well as accidental ones are relevant to the random or risk events when the projects carry out. In consideration of the two types of uncertainties, safety time is added into the duration of activities. In engineering projects, the executors always have high cognitive ability of their own tasks. If activities are executed by those people with less experience or lack of enough cognitive ability, the project schedule will be influenced considerably, as well as the cost, quality, and the stakeholders' requirement.

Although the accidental uncertainties have characteristics of randomness, they can always be predicted beforehand. Analyzing the comprehensive influence exerted by the accidental uncertainties, we can estimate the safety time added into the duration.

Based on above, the buffer sizing approach in perspective of flexible management mainly consists of two steps: the first one is determining the buffer sizes under the influence of uncertainties, and the second one is optimizing the result concluded by the first step considering resource tightness, network complexity and risk preference of the project manager comprehensively.

2.1 Determining the Buffer Sizes in Consideration of Uncertainties

The duration can be calculated from the construction norm. Theoretically, there is not much safety time added into the duration estimated by the norm. Therefore, it can be seen that the duration consists of two parts, one of which is fixation time for completion, and the other is safety time. Let d_i be the duration of activity i.

$$d_i = h_i + s_i \tag{1}$$

In (1), h_i is the duration calculated from the construction norm, while s_i is safety time added in comprehensive consideration of risk events caused by uncertainties. Because h_i is regarded as fixation time and its expectancy is $E(h_i) = h_i$, the probability distribution of d_i is determined by the distribution of s_i. We suppose the expectancy of s_i is $E(s_i) = \overline{s_i}$ and the variation of s_i is $D(s_i) = \delta_i$. The expectancy and variation of d_i is $E(d_i) = h_i + \overline{s_i}$, $D(d_i) = \delta_i$.

According to the transitional schedule methods such as CPM, PERT etc., the duration of activity i is d_i° ($d_i^\circ = h_i + s_i^\circ$), and s_i° is the safety time added in consideration of the uncertainties. The process of buffer size determining is constructed as follows:

1) Identify the uncertainties which may have impact on activities' duration estimation. We suppose that there are N kinds of risk events influencing on the duration of activity i, the influence exerted on activity i by risk event j is $\{s_{ij} \mid j = 1, 2...N\}$ (The influence is regarded as the safety time.), and those risk events are mutually independent.

2) Determine the weight of each risk event by the expert evaluation method. Let ω_{ij} is the weight of risk event j against activity i. And s_{ij} can be obtained as (2) shows:

$$s_{ij} = \omega_{ij}(d_i - h_i) = \omega_{ij} s_i \tag{2}$$

According to the transitional schedule methods to determine the duration, d_i° is the estimated duration of activity i. Therefore, s_{ij} is obtained by (3).

$$s_{ij} = \omega_{ij}(d_i^\circ - h_i) = \omega_{ij} s_i^\circ \tag{3}$$

Determine the probability distribution of each risk event referring to historical data or the expert investigation method. We suppose that the probability of risk event j is

p_{ij} and s_i^* is the safety time added into activity i based on the objective evaluation of risk events. s_i^* can be obtained by (4).

$$s_i^* = \sum_{j=1}^{N} \omega_{ij} s_i^{\circ} p_{ij} \tag{4}$$

According to the probability distributions of risk events and (4), we can evaluate the probability distribution of s_i^* by Monte-Carlo simulation method.

Consequently, in consideration of uncertainties, we propose that the calculation formula of buffer sizes is as the (5) shows.

$$\Delta B = \left[\sum (s_{i,0.9}^* - s_{i,0.5}^*)^2 \right]^{\frac{1}{2}} \tag{5}$$

2.2 Optimization of the Buffer Sizes in Consideration of the Three Project Characteristics

It is regarded in [1], [2] that resource tightness and network complexity can make a big difference in the buffer sizing process, and the schedule is much more likely to be delayed when the demand for resource is close to the supply or there are more predecessor activities. Under such condition, there should be bigger buffer sizes to decrease the influence due to uncertainties. The calculation formulas of conversion coefficients of the three project characteristics are showed as follows.

2.2.1 Resource Tightness

Let α_{iq} be resource usage for activity i for resource type q; α_i be the resource tightness of activity i; R_{qt} be the upper limit quantity of resource type q in the period of time t; r_{kqt} be total quantity of resource type q for activity k in t; m be the number of activities in t; ST_i and D_i be the starting time and duration of activity i respectively.

Then, for activity i

$$\alpha_{iq} = \sum_{k=1}^{m} r_{kqt} / R_{qt}, \quad t \in [ST_i, ST_i + D_i] \tag{6}$$

$$\alpha_i = \max\{\alpha_{iq}\} = \max\left\{ \sum_{k=1}^{m} r_{kqt} / R_{qt} \right\}, \quad t \in [ST_i, ST_i + D_i] \tag{7}$$

2.2.2 Network Complexity

The network complexity of activities can be represented by the complexity of the chain where the activities exist. Therefore, the network complexity of activity i can be showed as follows.

$$\beta_i = \frac{N_P}{N_T} \tag{8}$$

In (8), N_P is the total number of precedence relationships of activity i, while N_T is the total number of activities on the chain.

2.2.3 Risk Preference of the Project Manager

Although buffer sizes are determined to give a 95% chance of completing the project on time, from the angel of risk, the project manager still faces a 5% chance of being out of the buffer control. Let ε be the risk preference of the project manager, and the adjustment coefficient of buffer sizes at such level is calculated as (9) shows where $f_{1-\varepsilon}$ is considered as the multiple of standard deviation at 1-ε guarantee rate of completing the project on time.

$$\delta_i = \frac{f_{1-\varepsilon}}{2.0} \tag{9}$$

According to the analyses above, in the comprehensive consideration of resource tightness, network complexity and risk preference of the project manager, the optimized buffer sizes can be given by:

$$\Delta B = \left\{ \sum \left[(1+\alpha_i) \times (1+\beta_i) \times \delta_i \times (s^*_{i,0.9} - s^*_{i,0.5}) \right]^2 \right\}^{\frac{1}{2}} \tag{10}$$

3 Simulation Study

We suppose that d_i° is the duration of activity i obtained by transitional schedule methods, h_i is the fixation component of d_i°, and s_i° is the safety time added into d_i°. Table 1 shows the basic information of the project. The upper limit quantity of labor the project can supply is 10 persons. The level of risk preference of the project manager is 0.05.

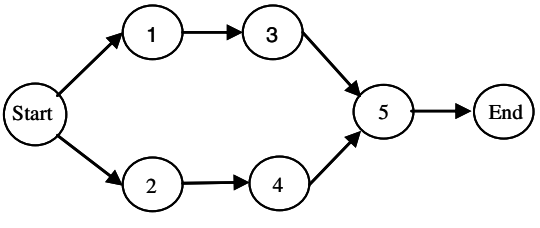


Fig. 1. Network Structure of The Project

Table 1. Information Table of Activities

Activity ID	Predecessor Activities	d_i° (wk)	h_i (wk)	s_i° (wk)	*Labor* (p)
1	/	5	3	2	1
2	/	10	7	3	3
3	1	10	8	2	3
4	2	8	7	1	4
5	3, 4	12	9	3	5

Identify and evaluate the uncertainties which may have impact on activity 1. Table 2 shows the results evaluated.

Table 2. The Results Evaluated of Activity 1

Risk event	Name	ω_{1j}	Probability distribution
1	Change of weather	0.1	N(60%, $4\%^2$)
2	Social influence	0.3	TR(30% ,40% ,60%)
3	Change of budget	0.3	U(30% ,55%)
4	Change of materials supplying	0.1	N(30%, $2\%^2$)
5	Change of requirement	0.2	N(40%, $2\%^2$)

According to the evaluation results and (4), the probability distribution of s_i^* added into activity 1 can be obtained after 1,000 times Monte-Carlo simulation by Crystal Ball. Fig.2 shows the simulation result. It illustrates that under the condition of completion with 90% and 50% guarantee rate respectively, the safety time is 0.92 and 0.84 respectively, in other words, $s_{1,0.9}^* = 0.92$, $s_{i,0.5}^* = 0.84$.

With the evaluation of probability distributions and weights of uncertainties influencing on other activities in the project, Table 3 illustrates the simulation results.

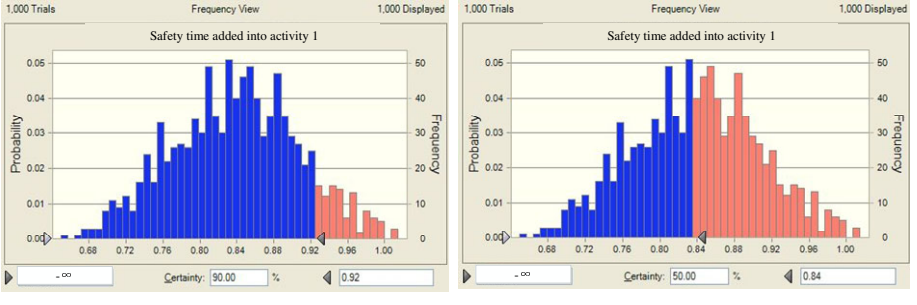


Fig. 2. The Probability Distribution of s_i^*

Table 3. The Simulation Results of Activities

Activity ID	$s^*_{i,0.9}$ (Safety time with 90% guarantee rate) (wk)	$s^*_{i,0.5}$ (Safety time with 50% guarantee rate) (wk)	$s^*_{i,0.9}-s^*_{i,0.5}$ (wk)
1	0.92	0.84	0.08
2	1.48	1.35	0.13
3	0.95	0.90	0.05
4	0.34	0.32	0.02
5	1.38	1.24	0.14

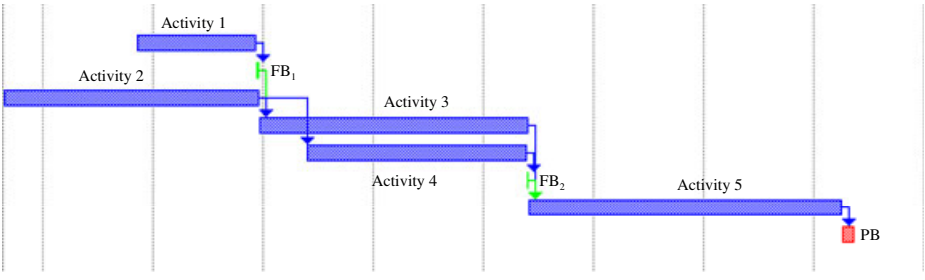


Fig. 3. The Schedule Plan after Adding Buffers

In the schedule plan after adding buffers, the critical chain is 2-3-5. In consideration of resource tightness, network complexity and risk preference of the project manager, adjustment coefficients showed in Table 4 are given to optimize buffer sizes. And Figure 3 is the schedule plan after adding project and feeding buffers.

Table 4. Adjustment Coefficients of Buffer Sizes

Activity ID	α_i	β_i	δ_i
1	0.4	0	1
2	0.4	0	1
3	0.7	1/3	1
4	0.7	1/3	1
5	0.5	2/3	1

According to (10), for the schedule plan

$$PB = 0.41(wk)$$

$$FB_1 = 0.11(wk)\ ;\quad FB_2 = 0.05(wk)$$

4 Conclusion

Buffer plays an important role in project schedule management. It can decrease the impact on the project exerted by uncertainties, prevent new resource conflicts and protect the critical chain. The project can respond to the uncertain environment rapidly in perspective of flexible management. And it is especially important to identify the uncertainties from the angle of the internal origin. The paper analyzed the influence on the construction period exerted by the risk events due to uncertainties and used the Monte-Carlo simulation method to determine the buffer sizes. In the other hand, it has taken the project characteristics into consideration in the process of calculation, and made the decision process more reasonable. However, the hypothesis that each risk event was mutually independent as premise of the paper may disagree with the facts. Buffer sizing on the premise of considering the interaction of each risk event is worth further studying. For all that, the paper presented a new idea to determining buffer sizes, and the constant perfection in theory or practice is hoped to provide vast development space for Critical Chain Project Management (CCPM).

References

1. Tukel, O.I., Rom, W.O., Eksioglu, S.D.: An investigation of buffer sizing techniques in critical chain scheduling. European Journal of Operational Research 172, 401–416 (2006)
2. Chu, C.-C.: Buffer sizing and critical chain project management. Computer Integrated Manufacturing Systems 14, 1029–1035 (2008)
3. Chen, L., Liang, F., Shi, X., Deng, Y., Wang, H.: Fuzzy-Safety-Buffer approach for project buffer sizing considering the requirements from project managers and customers. In: The 2nd IEEE International Conference on Information Management and Engineering (ICIME), pp. 482–486 (2010)
4. Zinovy, Radovilsky: A quantitative approach to estimate the size of the time buffer the theory of the constrains. International Journal of Production Economics 55, 113–119 (1998)
5. Shou, Y., Yeo, K.T.: Estimation of project buffers in critical chain project management. In: Proceedings of 2000 IEEE International Conference on Management of Innovation and Technology, Washington D.C., USA, pp. 162–167 (2000)
6. Ashtiani, B., Jalali, G.-R., Aryanezhad, M.-B., Makui, A.: A new approach for buffer sizing in Critical Chain scheduling. In: 2007 IEEE International Conference on Industrial Engineering and Engineering Management, Singapore, pp. 1037–1041 (2007)
7. Long, L.D., Ohsato, A.: Fuzzy critical chain method for project scheduling under resource constraints and uncertainty. International Journal of Project Management 26, 688–698 (2008)
8. Gao, P., Feng, J.-W.: A Grey Critical Chain Approach for Schedule Management and Its Application. Industrial Engineering and Management 14, 55–64 (2009)
9. Slack, N.D.: Flexibility as a manufacturing objective. International Journal of Operations and Production and Production Management 11, 3–13 (1983)
10. Karlsen, J.T.: Mestring av omgivelsesusikkerhet, Ph.D. thesis. The Norwegian University of Science and Technology (1998)
11. Intaver Institute Inc., Project Management Using Event Chain Methodology, http://www.intaver.com/Articles/Article_EventChainMethodology.pdf
12. Chen, W., Han, Z.: Study on application boundary conditions of Critical Chain Project Management (CCPM). Water Resources and Hydropower Engineering 38, 71–74 (2007)

Research on the Society Security of the Landless Peasants of Henan Province

Guo Xiaoli

Economics College of HeNan University, KaiFeng, 475004, China
bjguoxiaoli1978@126.com

Abstract. This paper studied the old-age insurance, medical insurance and job security of the landless peasants by empirical and theoretical analysis. The results indicate that the social security of the landless peasants is in the state of the serious deficiencies. With China's industrialization and urbanization accelerated, the scale of landless peasants is increasing. In order to protect the basic right to subsistence and development, We must take effective measures to strengthen social security of landless peasants.

Keywords: Henan, landless peasants, society security.

1 The Present Situation of the Social Security of the Landless Peasants in Henan Province

Social security is the foundation of social existence and development. With the industrialization and urbanization accelerated, the number of landless peasants is increasing year on year. It is predicated that by 2020 the landless peasants in our country will surpass 100 million. As the culture qualities and labor skills of landless peasants are generally not high, they are less competitive in non-agricultural industries. At present, the employment situation of the whole society is severe, and the social security system is not sound. Therefore, most of the landless peasants are in awkward "three no" circumstances: no land, no post, and no security. As a province with large agricultural population, Henan has witnessed a continually rising number of landless peasants in recent years. From the angle of developing trend, there will be tens of thousands of peasants losing their land in Henan every year in the future. In the process of landless peasants' acquiring citizenship, their subsistence allowances become a major worry. Old-age insurance, unemployment insurance, medical insurance and other social security issues are becoming increasingly prominent. According to the data from the Department of Land and Resources of Henan Province, from 2001 to 2004, there were 204,000-mu cultivated land taken over each year, and nearly 200,000 peasants lost their land every year. Therefore, it is necessary to pay close attention to the legitimate rights and interests protection of the landless peasants, explore effective protection strategies so as to build a new socialist countryside and harmonious society.

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 69–74.
springerlink.com

For a long time, the Party and the Government has been exhibiting great concern on the establishment of the social security system of landless peasants. Local governments at all levels have also been carrying through positive and effective explorations on this matter. However, in the actual practice, the ratio of the "three-no people" of the landless peasants stays high, even further expands.

1.1 Old-Age Security Situation of Landless Peasants

In China, the majority of peasants mainly rely on their offspring in the old age. In case of losing the land, their children have become landless peasants. With the unstable situation of re-employment and income, the traditional family pattern has been impacted severely. With the arrival of the aging society in rural areas, the supporting issue of the old age has becoming a new worry for landless peasants. It is investigated that there are only 15% of the landless peasants participating in various pension.

1.2 Medical Security Situation of Landless Peasants

In the aspect of medical security, the Province has given a certain amount of compensation in light of the specific condition of landless peasants. However, most landless peasants still cannot afford high medical expenses with limited compensation. They are still facing the issue of "poor access and high fee" in medical service. Although the New CMS has relieved the burden of landless peasants to a certain extent, the majority of landless peasants have acquired citizenship after losing their land. Therefore, they can't enjoy the rights of the New CMS, facing ever greater pressure on health care. The returning to poverty due to illness is more prominent.

1.3 Employment Security Situation of Landless Peasants

Employment is an important issue of subsistence and development for every member of society. Losing land means losing the most basic post. According to the findings of a survey into 105 landless peasant households made by the rural investigation team of Henan Province, the landless peasants of the Province have employment difficulties and few economic resources. Compared with urban workers, landless peasants lack non-agricultural employment knowledge, skills, and the supporting of favorable policies. Their ability to participate in social competition for jobs is very limited. Due to the low quality of rural labor force, poor information flow and other factors, the employment is difficult. Some landless peasants have few economic resources with decreasing income. Therefore, landless peasants can barely make the ends meet after entering cities. With the lost of the most basic means of production, i.e. land, no matter whether or not they are willing to acquire non-rural hukou, the life of landless peasants is basically urbanized. Before, they were self-sufficient, now they have to go to the market for most daily supplies. Therefore, their life cost increases. Although they gained a handsome compensation when the land is taken over, with the limitation of education background and labor skills, they could find no new employment opportunities and only rely on laurels.

2 Analysis of the Current Social Security Situation of Landless Peasants

2.1 Unreasonable Land Acquisition and Compensation Standards

Article 47 of Land Management Law stipulates that compensation for land requisitioned includes land compensation and resettlement subsidies and compensation for adjunct and young crops. The total sum of land compensation and resettlement subsidies shall not exceed 30 times the average annual output of the first 3 years after the corresponding land is taken over. This compensatory approach of setting only the highest price with no minimum protective price bears a clear planned economic characteristic, is contrary to the rules of market economy, and cannot accurately reflect the economic value of land acquired.

2.2 One-Way Resettling Pattern and Unfair Distribution

At present, the primary resettling pattern of the Province is money resettlement. Due to the one-way resettling pattern and unfair distribution between interest subjects, the legitimate rights and interests of landless peasants are severely violated. Governments at all levels take land use as an important inference, as a result, the compensation of the same land lot in the same village varies a lot due to different land uses, such as highway construction, enterprise sites, commodity house exploration. The difference can be several times or even more. A survey made by Zhejiang Province indicates that if the cost price of land acquisition is 100%, the profit distribution pattern of land taken over is: local government takes up 20-30%, enterprises 40-50%, organizations at the village level nearly 30%, peasants only 5-10%. Most of the substantial capital appreciation gains generated from the cost price to the sale price are obtained by developers or local governments.

2.3 Inadequate Social Security System, Landless Peasants Having No Old-Age, Medical and Employment Security

China's rural land bears the functions of ownership, economic benefits, employment and social security. After land is requisitioned, these four functions will be transferred or disappear. At present, the social security system of landless peasants in China is inadequate or basically not in place. After the state expropriate land, landless peasants should have been transferred to another security system from the old security system with compensation in hand. However, losing land does not mean the lost of their peasant identity. Urban-rural dual economy makes them different from urban citizens. Landless peasants themselves of Henan Province are a huge group of minimum standard of living. In this group, the people in distress are badly in need of the protection of social security, especially in the aspects of old-age security, medical security, employment security and subsistence allowances security. A survey into Pu Lizhai village, Kaifeng, Henan made by Yan Xing and Zhang Huiping indicates that in Pu Lizhai village, villagers cannot acquire any information about the subsistence allowances after losing their land, the government and the enterprises provide no clear information. In the process of urbanization, villagers lose the most basic means of

production and subsistence allowances. But they cannot enjoy the social security of urban dwellers to support their life in the future. It is even uncertain that whether or not they can receive the total sum of one-time compensation. How can they not complain?.

3 Thoughts on Addressing the Social Security Issue of Landless Peasants of the HeNan Province in the New Circumstances

3.1 Giving Reasonable Land Compensation to Landless Peasants in Light of the Rules of Market Economy

First, we must increase the compensation standards of land acquisition. The extremely low compensation standard of land acquisition is the most primary cause of peasants' suffering losses and appealing for help. Therefore, we have to as soon as possible revise the regulations on expropriating land of the existing land management law, standardize the acts of governments at all levels, and constrains public power in order to guarantee a long-term and reliable basic security for landless peasants. In the process of agricultural land being transferred to construction land, we must as fast as possible revise the administrative expropriating approaches in the planned economy, introduce market mechanism, and allow peasants directly involve in the whole process of land trade so as to ensure that their rights of using and disposing land being fully respected, and their interests being effectively protected. At the same time, we must as soon as possible improve land acquisition and compensation system and increase the compensation standards in order to benefit peasants.

Second, we must adhere to the principle of tilting the income distribution of land acquisition in favor of landless peasants. Landless peasants take the bulk, governments take the middle part, and the collectives take the small part. After land is requisitioned, peasants not only lose their land, but also a series of rights and interests relating to land. Therefore, new compensation items should be added, such as education subsidies, post-earnings of collective land in order to guarantee a long-term sustainable income for landless peasants.

3.2 Establishing Subsistence Allowances System for Landless Peasants

How to keep the living level under the condition of having no special skills is of great significance to maintain the stability of rural society. In case of not being able to fully establish a subsistence allowances system, we should as soon as possible build up subsistence allowances system for landless peasants and bring them into urban subsistence allowances system. It is especially important to make balancing urban-rural development our priority. We should set a basic living security benchmark for landless peasants in light of the economic development of the Province. Landless peasants whose living conditions are below the benchmark should be given a certain amount of compensation to secure their minimum living subsistence. Meanwhile, we should also pay attention to issues such as rationally define protection objects, scientifically set the subsistence allowances benchmark, and build up a plural and multi-channel financing mechanism. At the same time, other reform measures should also be taken to be matched up with the system.

3.3 Establishing Old-Age Insurance System for Landless Peasants

Landless peasants have made huge contribution to the urbanization of the Province. Therefore, we must establish old-age insurance system for them so as to guarantee their basic life and maintain society's stability. Specifically, there are two approaches: for those who have employment, basic old-age insurance of urban workers should be given to them; for those who have no employment, old-age insurance for landless peasants should be given to them. The fund should be provided by peasants, rural collective economic organizations and local governments. Generally speaking, individuals shall bear no more than 30% of the total, and the governments no less than 30%. The specific ratio shall be formulated in light of the actual local situation, adhering to the principle of landless peasants taking the small part.

3.4 Establishing Employment Security System for Landless Peasants

Employment security is the most direct and effective way to protecting landless peasants. While landless peasants enjoy the favorable policies of unemployment insurance of urban dwellers, we should actively realize their re-employment. First, governments should actively take effective employment promoting measures, improve employment training system, regulate that all landless peasants within the labor age can enter labor market, receive employment training, career choosing guidance, occupation introduction and other employment service. We should include the employment training of landless peasants into the important issues of government work and regard it as a significant link to assess governments' job performance. In the meantime, governments can also take a series of favorable policies such as tax reduction and exemption, social security subsidies, post subsidies so as to encourage enterprises to first employ landless peasants. Second, all quarters of society should give more attention to landless peasants. Peasants make huge contribution and sacrifice to urbanization. They lose their cultivated land and a series of rights and interests relevant to land. Therefore, society should show more care and attention to the disadvantaged group. For instance, enterprises should actively respond to the Party's policies and provide more re-employment posts for landless peasants. Third, landless peasants themselves should learn to change their employment perceptions, study and choose career by themselves. Last but not least, we should establish unemployment insurance system for landless peasants. For landless peasants who have acquired citizenship, the Province has included them into unemployment insurance according to Regulations on Unemployment Insurance of Henan Province. However, there are no protective measures for those who have no citizenship. Therefore, we should establish unemployment insurance system for landless peasants to secure their legitimate rights and interests. Landless peasants within labor age and having labor capacity and employment demand should be treated the same as urban laid-off workers, be granted Priority Certificate on Re-employment, and enjoy all the supporting policies of urban laid-off workers.

3.5 The System of Providing Legal Aid for Landless Peasants

The essence of legal aid system is to provide deducted or free legal service to some vulnerable groups. At present, landless peasants together with migrant workers and

urban laid-off workers belong to vulnerable groups in society. They have low status and limited economic power. When their legitimate rights and interests are violated in the process of their land being taken over, landless peasants usually cannot afford various costs necessary to start and further apply administrative rescue means (including money, time, legal knowledge, etc.). The principle of equality before law cannot be reflected in them, which requires that we should establish legal aid system for landless peasants in order to ensure the legitimate rights and interest landless.

4 Conclusion

Social security of landless peasants is a current complex and difficult social issue, urgently to be solved. As a Province with large agricultural population, with the urbanization accelerated, the issue of landless peasants will become more prominent in Henan. Therefore, we should protect the fundamental interests of the landless peasants and actively explore the scientific and reasonable social security system to protect the most basic right to subsistence and development for landless peasants.

References

1. Ling, W.-H.: On the Establishment of Social Security System for Farms without farmland. Journal of Economic Problem 2, 99–120 (2010)
2. Mao, M.: Lost-land peasants' social security system of China. Journal of XI'AN University of Post and Telecommunications 14, 139–142 (2009)
3. Ding, J.-Z.: Research on Land-Lost Farms' Social Security of Anhui Province. Journal of Anhui Vocational College of Police officers 9, 70–71 (2010)
4. Ouyang, L.: On the Government's Responsibility in China's Medical Care System. Journal of Xiangnan University 4, 60–63 (2006)
5. Chen, S.H.: Rural Social Security: Defect of System and Obligation of Government. Journal of Fujian Normal University 4, 30–33 (2004)

Nature of the Dynamics of Graphene

Zhang Yu-Qian, Li Yu-Chen, and Zhu Shi-Xi

National Lab. of Solid State Microstructures and Dept. of Physics,
Nanjing Univ., Nanjing 210093

Abstract. Graphene is a kind of two-dimensional carbon structure. Properties of graphene are very important for application of nano-technology. In "finite Temperature Lattice Properties of Grapheme has been studied with MC simulation and some anomalous predications have been given. Here I use dynamic simulation to restudy these properties and try to explain them.

Keywords: Graphene, dynamic simulation, MC simulation.

1 Instruction

1.1 Structure of Graphene

Graphene is a spin-off from the graphite layer of carbon atoms out of the surface material, Two-dimensional structure of the carbon. Graphite crystal film thickness of only 0.335 nm. The 200,000 added to the film together, and only a hair so thick. Graphene and graphite are the same as double hexagonal lattice, In the two-dimensional plane to each carbon atom sp2 hybrid orbital convergence,that is, each carbon atom with three nearest neighbor atoms to form three σ bonds. The remaining one p electron orbital plane perpendicular to the graphene with the surrounding atoms to form a π bond, the carbon atoms are each surrounded by hexagonal planar honeycomb-shaped structure, so that only two in the same spatial location of the surface atoms are different atoms.

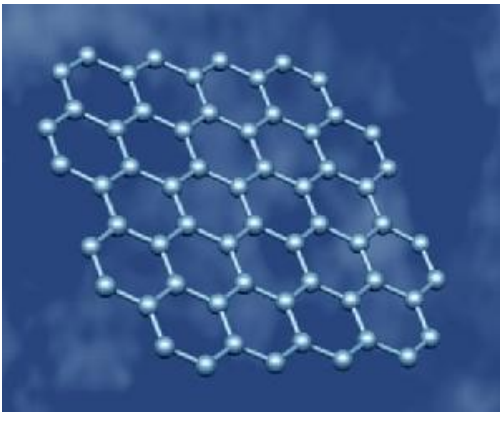


Fig. 1. Schematic diagram of graphene

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 75–82.
springerlink.com

1.2 The Nature and Application of Graphene

Graphene is one of the most important properties of electronic transport properties of graphene showed abnormal integer quantum Hall effect, the hall conductance of quantum conductance of odd times, and can be observed at room temperature. This phenomenon can be used to prove the opposite point of view of quantum mechanics, that graphene Dirac equation e is the electronic application.

Conductivity
Graphene is one of the largest electronic properties of the velocity of light has reached the 1 / 300, far more than an electronic conductor in the normal velocity. This makes graphene electronics, or more accurately, should be called "sub-load, and the relativistic nature of neutrinos is very similar.

Mechanical Properties
Graphene is the highest strength material known to man, hard than diamonds, stronger than steel but also the world's best high 100 times.

Electronic Interactions
Using the world's most powerful man-made radiation sources, the University of California, Columbia University and Lawrence Berkeley National Laboratory • physicists discovered a new secret features of graphene: Graphene between electrons and between electrons and exist in a honeycomb grid a strong interaction. Graphene honeycomb lattice and the electronic interaction between not only strong, but also electrical and electronic interaction between the strong.

1.3 Graphene Research

The current study graphene can be divided into two main directions: fundamental research and applied research, basic research in some reported that graphene can be attributed to the existence of stable micro-structure of the parent structure, but this understanding still a lack of evidence. Leaving the graphene may be applied to the transistor characteristics of the next generation - the quantum Hall effect has been confirmed. Some studies also focused on the identification of graphene research in this area are mainly intuitive and non-destructive AFM studies of Raman spectra, there is a way to use the traditional China-Africa theory of quantum Hall effect, and thus distinguish graphene (monolayer grapheme) and a few layer graphene (bilayer grapheme). In the preparation of graphene based on the macro to carry out graphene field-emission body, super capacitors, lithium-ion battery and the application of transparent conductive film to explore. As with the single-atom thickness, excellent electrical and mechanical properties of the boundary structure and rich features, it is an ideal graphene field emission materials, but the macro graphene lack of preparation and assembly technology constraints of its launch in the presence of applications. To take full advantage of graphene's structure and performance advantages, they have developed a method of electrophoretic deposition of uniform and dense surface was prepared and rich border protruding graphene films, the film and the substrate to achieve good contact between. Studies have shown that graphene films and comparable

to the carbon nanotube film field emission properties: low turn-on field and threshold values, good stability and uniformity of field emission, showing the graphene flat panel display and other aspects of applications.

2 Molecular Dynamics Simulation

2.1 Molecular Dynamics Simulation

The so-called classical molecular dynamics simulations, is the composition of the nucleus and the electron many-body systems, computer simulation of atomic nuclei with the movement, and thus the structure and nature of computing systems. Typically, the molecular dynamics simulation based on Born-Oppenheimer approximation (adiabatic approximation), each of which will be regarded as a nucleus provided by other nuclei and electrons . The potential field by Newton's equation, that is, by solving equations 2.1 by all the atomic details of the movement at different times.

$$\vec{a}_i = \frac{d\vec{v}_i}{dt} = \frac{d^2\vec{\gamma}_i}{dt^2} = \frac{\vec{f}_i}{m_i}, \qquad \text{i=1,2,…,N}$$

Where N is the total number of atoms in the system, $m_i, v_i, \vec{r}_i$ atom, respectively the quality, speed, position, and the force. In general, the case for no field, the force used by calculating the interaction potential function U of the gradient to be: $\vec{f}_i = -\nabla_i U$.

Molecular dynamics simulations can not only get the details of atomic motion, but also the same as doing a variety of experimental observations. For the balance system, you can do with molecular dynamics simulations to calculate the appropriate time average statistical average of physical quantities. For non-equilibrium systems, molecular dynamics simulations can also be time to observe the direct simulation of physical phenomena. In particular, some experiments can not be observed on the microscopic physical detail, molecular dynamics simulation can be easily observed. This advantage makes the molecular dynamics in physics, chemistry, materials science and other fields has been widely used.

Relative to the first-principles calculations, molecular dynamics simulation using empirical potential, although the accuracy worse, but because the program is simple, you can simulate a system with a large number of atoms and first-principles difficult or impossible to calculate the dynamical properties which has its unique advantages.

2.2 Periodic Boundary Conditions

For the macro-material, real samples usually contain a number of levels was 1023 atoms, which for the present calculations, it is impossible to simulate. If it is not going to study the surface properties of the sample, we can use periodic boundary conditions (PBC), which actually is the most current molecular dynamics simulation approach used. When using periodic boundary conditions, when the simulated atoms is considered to be in a box.

2.3 Ensemble Simulations and Thermodynamic Control

Molecular dynamics simulation of the actual set of external conditions is always carried out, the need for thermodynamic control, namely the need and the external heat exchange and mechanical balance. Equilibrium simulations, can usually be different ensemble classification. NVE ensemble is a natural molecular dynamics ensemble, only need to adjust the energy of the ensemble can be relatively easy to implement. In terms of molecular dynamics is required to adjust the temperature T, which is generally achieved through thermal bath.

3 Contents of the Study

3.1 Research Methods

This method of molecular dynamics simulations to study the structure of graphene bond length and lattice constant with temperature changes.First, the program generates a carbon tube, and then develop into a graphene structure, which is a bond length of 1.463Å of the hexagonal structure. Molecular dynamics simulation of the process that this state is the initial state, the use of periodic boundary conditions, tersoff potential and the Nose-Hoover heat bath for the evolution of computing systems to reach a certain temperature equilibrium. Calculated using the Lammps calculation software.

3.2 Description of the Stability

Simulation process from the initial state with the system dynamics simulations to make all the atoms in the Tersoff potential is under the influence of the time evolution of Ins in the process of evolution, every 1/20ns take a group of data, the 20 group data, by calculation, the standard size of 20 sets of data is less than 1% of the average data, so that the system has reached the temperature balance.

3.3 Results of the Calculation

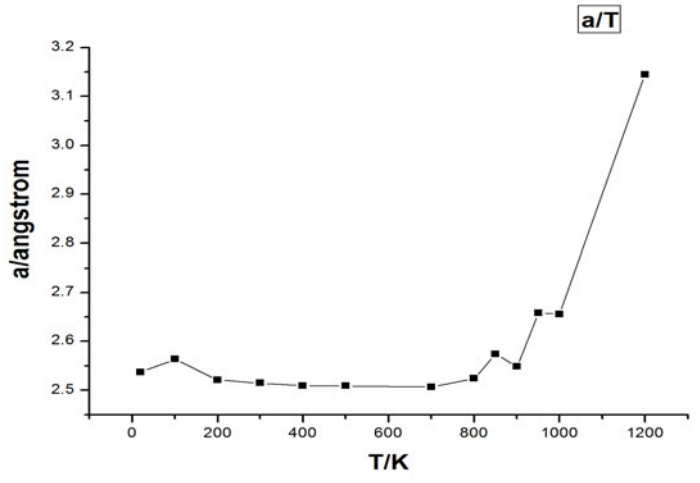
This includes the first of a 5,000 atoms wide 109.73Å, length 126.70Å size of graphene structure calculation. The results obtained are as follows:

(1) lattice constant

T/K	20	100	200	300	400	500
a/angstrom	2.5373572	2.5638205	2.5208515	2.5148373	2.5087957	2.508686

700	800	850	900	950	1000	1200
2.50702774 1	2.524038 7	2.573959 9	2.548358 3	2.658256 1	2.655549 4	3.144876 4

T is the temperature, a is the lattice constant, make the following image:

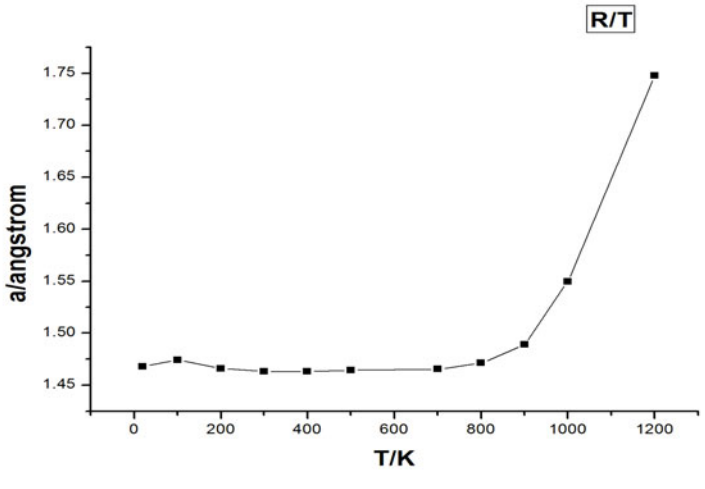


(2) Bond length

T/K	20	100	200	300	400	500
R/angstrom	1.46787	1.47421	1.46609	1.46339	1.46293	1.4646

700	800	850	900	950	1000	1200
1.46548	1.47138	1.50639	1.48882	1.55082	1.54958	1.74761

make the following image:



Found at a certain temperature range, graphene system there is a clear shrink, that decreases with increasing temperature the size of the phenomenon.

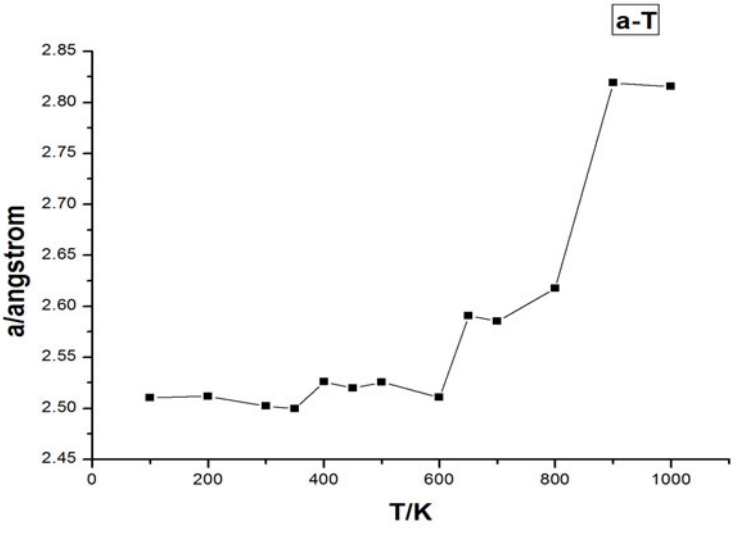
To study this phenomenon and the relationship between the size of the system, but also on a graphite containing 1800 atoms, length 65.84Å, width 76.02Å system similar simulation. The results obtained are as follows:

(1) Lattice constant

T/K	100	200	300	350	400	450
a/angstrom	2.5101725	2.5116177	2.5022305	2.499412	2.5257789	2.51974498

500	600	650	700	800	900	1000
2.525497375	2.510576	2.59060384	2.5854007	2.6175691	2.8191419	2.8156612

make the following image:

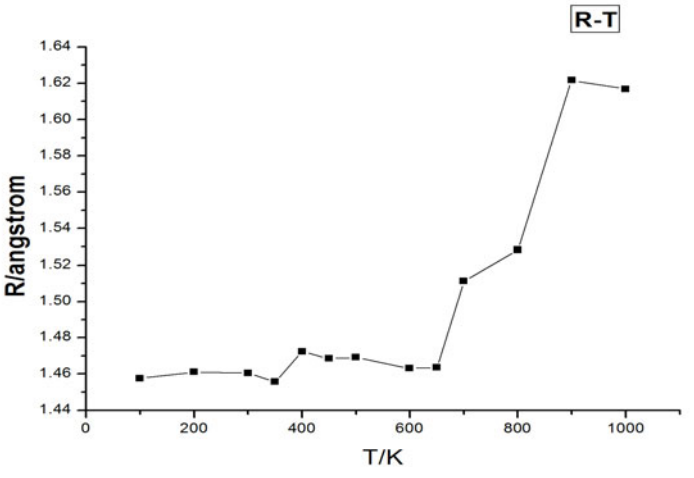


(2) bond length

T/K	100	200	300	350	400	450
a/angstrom	1.45748	1.46109	1.46049	1.45563	1.47226	1.46849

500	600	650	700	800	900	1000
1.46915	1.4631	1.46363	1.51113	1.52822	1.62173	1.61673

make the following image:



3.4 Results

Simulation results obtained by the above conclusion:

(1) at a certain temperature range, the system appeared with increasing temperature decreases the size of the phenomenon

(2) for the 5000 atom system, the lattice constant at about a minimum point occurs around 750K, ie before the system is a shrink, and in this, the system becomes thermal expansion;

(3) for the 1800 atom system, this effect is not so obvious shrink, but can still be observed in the temperature range of 650K on some before, there has been shrink;

(4) for two different sizes of atoms, the lattice constant and bond length with temperature trends are significantly different, this situation results obtained using tersoff and References [1] using empirical potential obtained LCBBOP Ⅱ significantly different results.

4 Conclusion

In this paper, molecular dynamics simulation method, the structure of the thermodynamic properties of graphene have been studied, the simulation results show that at a certain temperature range, the emergence of lattice constant and bond length decreases with increasing temperature anomaly. This phenomenon is because the system in the strong non-harmonic nature of this institution.

References

1. Zakharchenko, K.V., Katsnelson, M.I., Fasolino, A.: PRL 102, 046808 (2009)
2. Luo, C.: Low-dimensional molecular dynamics simulations of structural materials
3. Los, J.H., Ghiringhelli, L.M., Meijer, E.J., Fasolino, A.: PRB 72, 214102 (2005)
4. de. Heer, W.A., Berger, C., Wua, X., First, P.N., Conrada, E.H., Li, X., Li, T., Sprinklea, M., Joannna, Sadowski, M.L., Potemski, M., Martinez, G.: Solid State Communications 143 (2007)
5. Zhang, Y., et al.: Tersoff potential function of several amendments to the form

A L-stable Numerical Scheme for Option Pricing under Jump-Diffusion Models

Wei Li* and Shengwu Zhou

College of Sciences, China University of Mining and Technology,
Xuzhou 221116, China
liwei_123@126.com

Abstract. A L-stable and highly accurate method for option pricing under jump-diffusion models is developed in this paper. A semidiscretization scheme is performed on the partial integro-differential equation, and a numerical scheme is constructed based on Pade approximations of the matrix exponential. Due to the integral term, which cause the resulting system to be dense, an iteration to solve the equations in numerical scheme is present. Numerical examples for European option and barrier option with Merton's jump-diffusion model show that the algorithm is efficiently and avoid spurious oscillations.

Keywords: Option pricing, Jump-diffusion model, Pade scheme, L-stable method, Iterative method.

1 Introduction

It is widely recognized that classic option pricing model [1] proposed by Black and Scholes in 1973 does not ideally fit observed empirical market data. Two identified empirical features have been under much attention: (1) skewed distribution with higher peak and heavier tails (i.e. leptokurtic behavior) of the return distribution and (2) the volatility smile.

Many models have been proposed in literature to explain these phenomena. Here we are interested in the jump-diffusion model proposed by Merton [2]. For European option, Merton obtained analytical expressions but for most exotic options under jump-diffusion models, no closed form solutions exist. We thus need to consider numerical solutions for the partial integro-differential equation that arises. Khaliq et al. [3] proposed a L-stable method for pricing exotic options under diffusion models. In this paper, we expand the method to price option under jump-diffusion models. Due to the integral term, the discretization leads to a full matrix, which causes algorithm to be too expensive, and therefore we present an iterative method to solve the equations that arises. Numerical experiments for vanilla option and barrier option with Merton's jump-diffusion model show that the algorithm is efficiently and avoid spurious oscillations.

* Corresponding author.

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 83–88.
springerlink.com

2 Merton's Model and Semidiscretization

Under Merton's model [2], the value $V(S,t)$ of a option satisfies the final value problem defined by a partial integro-differential equation (PIDE)

$$V_t + \frac{1}{2}\sigma^2 S^2 V_{SS} + (r - \lambda\kappa)SV_S - (r+\lambda)V + \lambda\int_0^{+\infty} V(S\eta,t) f(\eta) d\eta = 0, \tag{1}$$

for all $(S,t) \in (0,\infty)\times[0,T]$. Above S and t denote the value of the underlying asset and the time, respectively. The risk free interest rate r, the jump intensity λ and the volatility σ are constants. $(\eta - 1)$ is the impulse function causing S to jump to $S\eta$, $\kappa = E[\eta - 1]$ and f is the density function. The maturity value is given by $V(S,T) = g(S)$. For an European call option, $g(S) = \max\{S - K, 0\}$; for an up-and-out call barrier option, it has the same payoff as that of an European call option except for an upper condition at the barrier B as $V(B,t) = 0$.

We first consider an European call option, taking the transformation

$$x = \ln(S/K),\ y = \ln(\eta),\ \tau = T - t,\ v(x,\tau) = V(Ke^x, T - \tau).$$

From (1) one gets

$$v_\tau = \frac{1}{2}\sigma^2 v_{xx} + (r - \lambda\kappa - \frac{1}{2}\sigma^2)v_x - (r+\lambda)v + \lambda\int_{-\infty}^{+\infty} v(x+y,\tau)\, p(y) dy, \tag{2}$$

where $p(y)$ is the density function of stochastic reject $y = \ln(\eta)$ and $v(x,0) = g(K\mathrm{e}^x)$. For Merton's model, $\ln(\eta) \sim N(\alpha, \gamma^2)$. Obviously, $\kappa = \mathrm{e}^{\alpha + 0.5\gamma^2} - 1$.

For constructing semidiscretization scheme, we need to truncate the infinite x- domain $(-\infty,\infty)$ to $\Omega_x = (x_{\min}, x_{\max})$. Considering an $N+1$ nodes grid $x_i = x_{\min} + ih$ with $0 \le i \le N$ and $h = (x_{\max} - x_{\min})/N$. We discretize the spatial derivatives using the second-order central differences. For the integral term, a second order discretisation is needed. Taking the transformation $z = x + y$, we split the integral into Ω_x and $\mathrm{R} \setminus \Omega_x$ and use the composite trapeizodal rule as

$$\int_{\mathrm{R}} v(z,\tau)\, p(z - x_i) dz \approx \frac{h}{2}[v(x_0,\tau)\, p(x_0 - x_i) + v(x_N,\tau)\, p(x_N - x_i) + 2\sum_{j=1}^{N-1} v(x_j,\tau)\, p(x_j - x_i)] + \varepsilon(x_i,\tau)$$

where using the fact that $v(x,\tau) \approx K\mathrm{e}^x - K\mathrm{e}^{-r\tau}$ for an European call option as $x \to \infty$, then the unbounded integral becomes

$$\varepsilon(x,\tau) = K\mathrm{e}^{x+\alpha+\frac{\gamma^2}{2}}\Phi(\frac{x - x_{\max} + \alpha + \gamma^2}{\gamma}) - K\mathrm{e}^{-r\tau}\Phi(\frac{x - x_{\max} + \alpha}{\gamma}),$$

where

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \mathrm{e}^{-\frac{z^2}{2}} dz .$$

Under the above discretisation, we can obtain a semi-discrete linear system

$$\boldsymbol{v}'(\tau) + \boldsymbol{A}\boldsymbol{v}(\tau) = \boldsymbol{b}(\tau), \quad \boldsymbol{A} = \boldsymbol{B} + \boldsymbol{C}, \quad 0 \le \tau \le T, \tag{3}$$

where $\boldsymbol{B}$ is a tridiagonal matrix, $\boldsymbol{C}$ is a dense matrix resulting from the integral term and $\boldsymbol{b}$ is non-homogeneous term resulting from the boundary conditions.

3 Time Stepping Schemes

Now, we focus on the linear system (3). Following [3], the exact solution of (3) satisfies the recurrence formula

$$\boldsymbol{v}(\tau_{m+1}) = \mathrm{e}^{-k\boldsymbol{A}}\boldsymbol{v}(\tau_m) + k\int_0^1 \mathrm{e}^{-k\boldsymbol{A}(1-\delta)}\boldsymbol{b}(\tau_m + k\delta)d\delta. \tag{4}$$

where $\tau_m = mk$ and k is the time step.

We use the $(p,q)-$Pade approximations $R_{p,q}(z) = P_{p,q}(z)/Q_{p,q}(z)$ to approximate the exponential function e^{-z} in (4), where

$$P_{p,q}(z) = \sum_{j=0}^{p} \frac{(p+q-j)!p!}{(p+q)!j!(p-j)!}(-z)^j,$$

$$Q_{p,q}(z) = \sum_{j=0}^{q} \frac{(p+q-j)!q!}{(p+q)!j!(q-j)!}(z)^j,$$

These schemes have property $R_{p,q}(z) = \mathrm{e}^{-z} + O(|z|^{p+q+1})$ as $z \to 0, z \to \in \mathbf{C}$. Similar to literature [3], the equation (4) is approximated by

$$\boldsymbol{v}_{m+1} = R_{p,q}(k\boldsymbol{A})\boldsymbol{v}_m + k\sum_{i=1}^{2} Q_{p,q}^{(i)}(k\boldsymbol{A})\boldsymbol{b}(\tau_m + \delta_i k), \tag{5}$$

where $\boldsymbol{v}_m$ is the approximation of $\boldsymbol{v}(\tau_m)$, $Q_{p,q}^{(i)}(z)$ are rational functions which share the same denominator as those of $R_{p,q}(z)$, and $\delta_1 = \frac{3-\sqrt{3}}{6}$ and $\delta_2 = \frac{3+\sqrt{3}}{6}$ are Gaussian quadrature points. The formula to obtain $Q_{p,q}^{(i)}(z)$ in [4] is

$$\sum_{i=1}^{2} \delta_i^l Q_{p,q}^{(i)}(z) = \frac{l!}{z^{l+1}}\left(R_{p,q}(z) - \sum_{j=0}^{l} \frac{z^j}{j!}\right), \quad l = 0,1. \tag{6}$$

It is well known [5] that numerical scheme (5) is L-stable as $p = 0$. However, the schemes mentioned above involve inverses of higher order matrix polynomials which can cause computational difficulties. We use the partial fraction form of the rational function $R_{0,q}(z)$ and $\{Q_{0,q}^{(i)}(z)\}_{i=1}^{2}$ proposed by Khaliq et al. [6]. For $q = 4$, following the notations in [3], we can write

$$R_{0,4}(z) = 2\sum_{j=1}^{2} \Re(\frac{w_j}{z - c_j}),\ Q_{0,4}^{(i)}(z) = 2\sum_{j=1}^{2} \Re(\frac{w_{ij}}{z - c_j}),\ i = 1, 2. \tag{7}$$

From (5) and (7), the algorithm becomes

$$\boldsymbol{v}_{m+1} = 2\Re(\boldsymbol{y}_1) + 2\Re(\boldsymbol{y}_2), \tag{8}$$

and

$$(k\boldsymbol{A} - c_1\boldsymbol{I})\boldsymbol{y}_1 = w_1\boldsymbol{v}_m + kw_{11}\boldsymbol{b}(\tau_m + \delta_1 k) + kw_{21}\boldsymbol{b}(\tau_m + \delta_2 k), \tag{9}$$

$$(k\boldsymbol{A} - c_2\boldsymbol{I})\boldsymbol{y}_2 = w_2\boldsymbol{v}_m + kw_{21}\boldsymbol{b}(\tau_m + \delta_1 k) + kw_{22}\boldsymbol{b}(\tau_m + \delta_2 k). \tag{10}$$

We write (9) and (10) as a generalization, $\boldsymbol{D}_i\boldsymbol{y}_i = \boldsymbol{d}_i, i = 1, 2$, where $\boldsymbol{D}_i = k\boldsymbol{A} - c_i\boldsymbol{I}$. Due to the matrix $\boldsymbol{A}$ is dense, similar to [7], we form a regular splitting for the coefficient matrix $\boldsymbol{D}_i$ as

$$\boldsymbol{D}_i = \boldsymbol{T}_i - \boldsymbol{J}, \quad \text{where } \boldsymbol{T}_i = k\boldsymbol{B} - c_i\boldsymbol{I} \text{ and } \boldsymbol{J} = -k\boldsymbol{C},\ i = 1, 2. \tag{11}$$

Then a stationary iterative method for the linear system $\boldsymbol{D}_i\boldsymbol{y}_i = \boldsymbol{d}_i$ reads

$$\boldsymbol{y}_i^{l+1} = \boldsymbol{T}_i^{-1}(\boldsymbol{d}_i + \boldsymbol{J}\boldsymbol{y}_i^l),\ l = 0, 1, \cdots,\ i = 1, 2, \tag{12}$$

where the initial guess $\boldsymbol{y}_i^0$ is taken to be the solution from the previous time step. Each iteration requires a solution with the tridiagonal matrix $\boldsymbol{T}_i$ and the multiplica- tion of a vector by $\boldsymbol{J}$.

Using the iterative scheme (12), we can obtain the solution of (9) and (10) more easily. Then taking into account (8) and using a cubic spline interpolation, we can get the option values.

4 Numerical Experiments

In this section, we price European call option and up-and-out-call barrier option using the following parameters

$$x_{\min} = -4, x_{\max} = 4, K = 50, r = 0.05, \sigma = 0.4,$$
$$T - t = 1, \alpha = 0, \gamma = 0.3, \lambda = 2, N = 400, k = 0.025,$$

For European call option, using analytical solution of Merton as the standard, the effectiveness of this algorithm is verified in following numerical simulation. The numerical results in figure 1 show that the errors resulting from the calculated results and analytical solution become smaller and smaller with nodes of difference increasing, and which also can be controlled within 0.02.

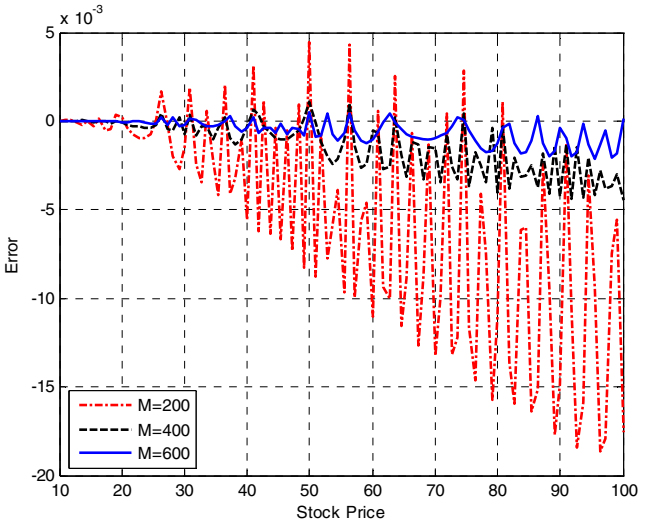


Fig. 1. Curve picture of errors

For up-and-out-call barrier option, the upper barrier $B = 70$, a comparison of our algorithm and Crank-Nicolson scheme is supplied in figure 2. From which, classic Crank-Nicolson method experiences spurious oscillations on near the strike price and barrier value, while our algorithm resolves the shortcoming and obtains reliable option value.

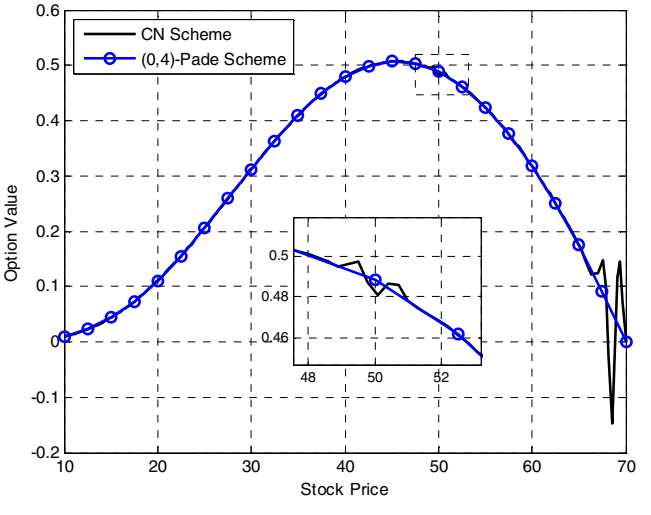


Fig. 2. Comparison of (0,4)-Pade scheme and Crank-Nicolson scheme

5 Conclusion

We have developed a L-stable scheme for pricing options under jump-diffusion model. Numerical examples have tested its effectiveness and it is also particularly efficient on multi-asset options as it has the same computational complexity as the fist-order implicit Euler method.

Acknowledgments. This work is supported by the Fundamental Research Funds for the Central Universities (2010LKSX03).

References

1. Black, F., Scholes, M.: The Pricing of Options and Corporate Liabilities. Political Economy 81, 637–659 (1973)
2. Merton, R.C.: Option pricing when underlying stock returns are discontinuous. Journal of Financial Economics 3, 125–144 (1976)
3. Khaliq, A.Q.M., Voss, D.A., Yousuf, M.: Pricing exotic options with L-stable Pade schemes. Journal of Banking and Finance 31, 3438–3461 (2007)
4. Wade, B.A., Khaliq, A.Q.M., Yousuf, M., et al.: High order smoothing schemes for inhomogeneous parabolic problems with applications to nonsmooth payoff in option pricing. Numer. Methods Part. Diff. Equat. 23(5), 249–1276 (2007)
5. Yousuf, M.: On the class of high order time stepping schemes based on Pade approximations for the numerical solution of Burgers' equation. Applied Mathematics and Computation 205, 442–453 (2008)
6. Khaliq, A.Q.M., Twizell, E.H., Voss, D.A.: On parallel algorithms for semidiscretized parabolic partial differential equations based on subdiagonal Pade approximations. Numer. Method PDE 9, 107–116 (1993)
7. Salmi, S., Toivanen, J.: An iterative method for pricing American options under jump diffusion models. Applied Numerical Mathematics 61, 821–831 (2011)

Empirical Research on the Level and Quality of Gains from ST Stocks in China

Wenxiu Hu*, Jing Niu, and Qingbang Mu

Yanxiang Road No.58, Xi'an University of Technology, 710054, Shaanxi, China
nimo19870315@126.com

Abstract. The paper investigated the question whether excess profits exist among ST stocks in China, under the circumstances of adjusting risk and not adjusting risk. Firstly, the paper made a group of t-test of significant difference between ST stocks and the financial market and normally listed stocks, on the holding income and accumulative total income ; Then, the article made and empirical study on the investing risk of ST stocks by estimating their Beta coefficients. The result shows that, only the ST stocks and *ST stocks outperformed the financial market and normally listed stocks in the short run by the geometry average earnings, but it is not always the case in the long run.

Keywords: ST stock, Beta coefficient, Excess profit, Empirical study.

1 Introduction

During the period of 1998-2001, the proportion of ST stocks among the whole stock market firstly increased from 2.82% to 4.74%, and a decreased of 3.79% followed. During the period of 2001-2011, this rate again climbed to 14.16%, meanwhile, affected by the conceptual speculation of recombination in the secondary market, individual investors scramble for them, and several funds ,now and then, add the holding of ST stocks even heavily hold them, so, ST stocks have made some difference in the financial market in China.

Because the ST (special treated) system exists in China alone, so the research result mainly centrally allocated several fields in China, such as ST institution[1-3], the causes of ST[4], the characters of ST company, measures of cap-picking [5], corporate governance[6], treatment of shell resource and earning management, and so it is with the research result about return and risk of ST stocks. Genxiang Shen[7](2003) described how the restriction of rise and decrease affects the volatility of ST stocks, using GARCH model based on the analysis of the structure of return; Jingjun Liu[8](2006) carried out an empirical research on the conversion between the status of normal, ST and stop listing using the order Logit regression model of 142 ST companies, whose result show that, the ST system is beneficial to the perfecting of finance and the survival of well-governed companies; Qiming Tang, Suxin Huang[9]

* Wenxiu Hu(1964-), male, Dean, School of Economy and Administration of Xi'an University of Technology, Professor, PhD, Research direction: M&A, Finance Engineering.

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 89–95.
springerlink.com

(2006) found that the hypothesis of half-efficient is not supported by Chinese market, using the event study methodology, which took the fluctuation into the GARCH model; Yuguang Shi[10] (2009) studied the abnormal return of ST companies during the period of 2007-2008, whose conclusion show that, the proportion and resource of abnormal return have something to do with the mechanism of delisting system and conversion; Yongxi Zhang, Wei Zhou[11](2011) discussed the relationship between the construction of ST institution and medium and small investors protection; Zhonglong Zhao[12](2011) found that the information asymmetry and self-traction are the leading factors affecting the development of capital market based on the analysis of actual effect after the quota system and supervision system put into practice.

But as to the level and quality of return from ST stocks, and the question whether the excess profit exist when investing ST stocks, which have been brought much special risk and system risk to the investors, have not been empirically tested, well, the result is also an important evidence to make decision, based on this point, the paper try to make an empirical research on the level and quality of investment earning from ST stocks, in order to offer an evidence for the improvement of return and the effective control of investment risk.

2 Empirical Research Plan of Abnormal Return in ST Stocks

2.1 The Research Hypotheses

According to the investors' characters in china, the investors, as a whole, are rational and risk-neutral in the long run, but limit rational in the short term, when risk evader and risk liker can be found everywhere. Actually, ST stocks are heavily held by funds, and individual investors are also interested in them, on one hand, the conceptual speculation of recombination in the secondary market play an important role, on the other hand ,the excessive profit from investing ST stocks may exist, based on this point, it is safe to design hypothesis as follows:

H1: The earnings' quality of ST stock is higher than that in the market
H2: Holding excess profits or Accumulative excess profits exist in ST stock
H3: Holding excess profits or Accumulative excess profits exist in *ST stock

2.2 Sample Selection

According to the difference in Income property, taking all A stocks as a general, the samples are divided into two parts. One is based on the holding period return data, the other is based on the accumulated return data; the time window of the former sample is 1998-2011,the time window of latter is 2003-2011,according to their life circle; the comparison datum include the finance market as a whole and the stocks normally listed.

When selecting, the test of normality was carried out using the statistic software SPSS18.0, and the data whose allocation does not accord with normal distribution is standardized, as to the data that is incomplete is replaced with the mean of other data in the same volume, and abnormal data is rejected through scatter diagram, effectively retention the original information, as shown in the table 1.

Table 1. Sample summary

Items	Holding period return data		Accumulative return data	
	Time window	Sample size	Time window	Sample size
ST	1999.4-2011.6	147	1999.1-2011.6	150
SST	2007.1-2011.3	51	1999.1-2011.6	150
*ST	2003.5-2011.3	98	1999.1-2011.6	150
S*ST	2006.11-2011.3	40	1999.1-2011.6	150
Normal listed	1999.1-2011.3	147	1999.1-2011.6	150
Market	1999.1-2011.6	150	1999.1-2011.6	150

Data sources: http://resset.cn/cn/; http://www.wind.com.cn/

As shown in the table 1, each sample size is bigger than 30, the time window is quite late, which can ensue the representativeness of data, and hunt the newest information.

2.3 Variables Setting

The style of data used in the paper is Monthly comprehensive income data, so does the risk free return data, the data process is carried out under the situation with adjustment of variance and without adjustment of variance, using holding period return data and accumulative return data, through the t-test between ST *ST and market and normally listed companies, at the same time, regression is also made to estimate the Beta coefficient to measure the investment risk of ST stocks, in order to make the anaphora latter clear, relative variables are set as shown in table 2.

Table 2. Variables setting

Variable meaning	name	expression
risk-free rate	r_f	$r_f = r_f$
Monthly comprehensive return rate	r_{xy}	$r_{xy} = \sum_{i=1}^{n} r_{ixy} w_i$
volatility	σ_{xy}	$\sigma_{xy} = \sqrt{\sum (r_{xy} - \overline{r_{xy}})^2 / (n-1)}$
Sharp ratio	SR_{xy}	$SR_{xy} = (r_{xy} - r_f) / \sigma_{xy}$
β coefficient	β_{xy}	$\beta_{xy} = \sum_{t=1}^{n} (r_{xy} - \overline{r_{xy}})(r_{ky} - \overline{r_{ky}}) \Big/ \sum_{t=1}^{n} (r_{xy} - r_{ky})$

In which, x=ST、SST、*ST、S*ST、m(market)、n (normally listed stocks); y=h (holding period return), y=a (accumulated return); i is the sequence number of xy; w_i is the circulation market value of ST stock divided by the circulation market value of the financial market as a whole; k=m (market), k=n (normally listed stocks).

2.4 Construction and Solving of Model

There are two models in the paper, one is the t statistics, and the other is a regression model, the former is used to check the existence of excess profit from ST stocks, the latter is used for the test of quality and level of investment risk in ST stocks further.

According to the variables setting, hypotheses and research target, the t statistics is constructed, shown as equation (1):

$$t = (\overline{SR_{xy}} - \overline{SR_{ky}}) \Big/ S_C \sqrt{n} \tag{1}$$

Where, $S_C = \sqrt{(n_{xy}\sigma_{xy}^2 + n_{ky}\sigma_{ky}^2) \Big/ (n_{xy} + n_{ky})}$ is the Composite standard deviation of SR_{xy} and SR_{ky}; n_{xy} is the sample size of ST stocks, n_{ky} is the sample size of comparison datum, n is the degrees of freedom.

Among the methods estimating the risk in investment, VaR is a common method to estimate investment risk in ST stocks but it is too complex for individual investors, and its feasibility is low, based on this point, this research take the method of Beta coefficient, which is simple and reliable, now ,the regression model can be set as equation (2):

$$r_{xy} = \alpha + \beta_{xy} \times r_{ky} + \varepsilon \tag{2}$$

Where, α stands for intercept, ε is a Random disturbance, the meaning of x、y、k shown in table 2.

In the course of solving the model, the significant difference t-test is carried out using the method of Paired-Sample t-test in the statistic software of SPSS18.0; the solving of regression model is carried out with the method of Multiple Liner Regression in the statistic software of SPSS18.0, in which the way to select variables is "enter", the significant level is set as default value, and modified the time window when there is conflict among different time windows by remain the shorter time window, in order to ensure the comparability of data.

3 Results and Interpretation

On the base of the method mentioned above, the t-test is made ,the result is shown respectively in table 3,where, *x=ST*、*SST*、**ST*、*S*ST, y=Market*; in table 4, where, *x=ST*、*SST*、**ST*、*S*ST, y=Normal;* in table 5,where, *x=ST*、*SST*、**ST*、*S*ST, y=Market*; in table 5, where, *x=ST*、*SST*、**ST*、*S*ST, y=Normal*, in table 6 ,where, *x=ST*、*SST*、**ST*、*S*ST, y=Normal.*

Table 3. T-test based on yields of holding period compared with the Market (%)

Item1	ST-Market	SST- Market	*ST- Market	S*ST- Market
Time window	1999.4-2011.6	2007.1-2011.3	2003.5-2011.3	2006.11-2011.3
GEOMEAN(x)	1.1200	3.6100	1.6600	1.9200
GEOMEAN(y)	1.4800	2.6200	1.6200	1.0400
Gains per std of x	1.1215	1.8256	1.2826	0.9054
Gains per std of y	1.8360	2.0882	1.8764	1.0770
T Statistics	-0.0684	0.4874	0.2568	0.5666
P(T<=t)	0.4728	0.3135	0.3988	0.2863

Table 4. T-test based on yields of holding period compared with the Normal (%)

Item 2	ST-Normal	SST-Normal	*ST-Normal	S*ST-Normal
Time window	1999.4-2011.3	2007.1-2011.3	2003.5-2011.3	2006.11-2011.3
GEOMEAN(x)	1.1600	3.6100	1.6600	1.9200
GEOMEAN(y)	1.3900	2.6600	1.7200	1.5400
Gains per std of x	1.1354	1.8256	1.2826	0.9054
Gains per std of y	1.8411	2.0182	1.9291	1.2893
T Statistics	0.0587	0.4526	0.1925	0.4209
P(T<=t)	0.4766	0.3259	0.4238	0.3375

Table 5. T-test based on accumulative return data compared with the Market (%)

Item 3	ST- Market	SST- Market	*ST- Market	S*ST- Market
Time window	1991.1-2011.6	1991.1-2011.6	1991.1-2011.6	1991.1-2011.6
GEOMEAN(x)	1.0700	0.8400	1.1800	0.6600
GEOMEAN(y)	1.5800	1.5800	1.5800	1.5800
Gains per std of x	1.3331	1.0124	1.3704	3.0635
Gains per std of y	1.8071	1.8071	1.8071	1.8071
T Statistics	-0.3587	-0.3900	-0.2378	-0.2098
P(T<=t)	0.3600	0.3484	0.4061	0.4170

Table 6. T-test based on accumulative return data compared with the Normal (%)

Item 4	ST- Normal	SST- Normal	*ST- Normal	S*ST Normal
Time window	1991.1-2011.6	1991.1-2011.6	1991.1-2011.6	1991.1-2011.6
GEOMEAN(x)	1.0700	0.8400	1.1800	0.6600
GEOMEAN(y)	1.4900	1.4900	1.4900	1.4900
Gains per std of x	1.3331	1.0124	1.3704	0.7362
Gains per std of y	1.9366	1.9366	1.9366	1.9366
T Statistics	-0.2412	-0.2803	-0.1179	-0.1093
P(T<=t)	0.4048	0.3897	0.4531	0.4565

At the same time, estimate the β Coefficients through Linear Regression Model, the result is shown as table 7.

Table 7. β Coefficient measurement results

Item	holding period yield		accumulated earnings	
	all	Normal listed	all	Normal listed
ST	1.0493	1.0556	0.9548	1.0650
SST	0.9895	0.9646	0.8872	1.0159
*ST	1.1121	1.0909	0.9818	1.0818
S*ST	1.3005	1.2995	1.1128	1.2861

Data Sources: Introduced from Regression Results According to Equation (2)

From the table 3 and table 4 we can see, compared with the market, based on the geometric mean return data, ST stocks that have not completed shareholder structure reform,*ST stocks that have risk of desisting,*ST stocks that have not finished shareholder structure reform beat the market; compared with the stocks that are normally listed based on the geometric mean return data, ST stocks as a whole, *ST stocks that have not completed shareholder structure reform beat the market.

Seen from the table 5 and table 6, ST stocks did non beat the market, nor the normally listed stocks, based on the accumulative return data without the adjustment of variance, but the *ST stocks that did not accomplish shareholder structure reform beat the market. Consequently, during all the time windows mentioned above, none of the results of t-test between them is significant.

It can be inferred from the table 7 that, the Beat coefficients estimated based on the holding period return data is more than 1 compared with the market and the normally listed stocks; when compared with the market, most Betas estimated based on the accumulative data are less than 1; when compared with normally listed stocks, most Betas are more than 1. As a result, the investment risk of ST stocks is greater than the market in the short run, in the long run; the investment risk of the ST stocks that have not finished shareholder structure reform is less than the market but more than the normally listed stocks.

4 Conclusion

According to the result of t-test, adjustment of variance and the estimate of Beta coefficients, it is safe to conclude as follows:

(1) Under the situation without the adjustment of variance, the ST and *ST stocks can beat the market and the normally listed stocks based on geometric mean return data in the short run, but they can beat neither the market nor the normally listed stocks in the long run.

(2) Under the situation with the adjustment of variance, the level and quality of ST and *ST stocks cannot beat the market or the normally listed stocks in the long run or in the short run.

(3) It is necessary to establish and perfect the retreat system, and regulate the conceptual speculation of recombination in the secondary market in China, in order to improve the efficiency of Chinese financial market.

Acknowledgments. This paper is funded by the doctoral fund project of ministry of education (20096118110010).

References

1. Yan, X.: The current problems and countermeasures to the ST companies listed in China. Commercial Times (21), 71–72 (2008)
2. Li, G.: Research on the Special treatment system and ST the company's corporate governance. Central South University (2007)
3. Jiang, G., Wang, H.: Should Firms with Two Consecutive Annual Losses be Specially Treated (ST)? Economic Research Journal 1(3), 100–107 (2005)
4. Qin, F.: Status and outlet of ST companies. Economic Management 1(1), 60–62 (2001)
5. Lv, C., Zhao, Y.-H.: Empirical Analysis on ST Firm Life Track. Economic Management (4), 44–51 (2006)
6. Cheng, B., Xu, X.: Research on the relationship between the nature of controller and the extent of supporting–based on the evidence of ST company experience in cap-picking. Macroeconomics 1(11), 46–51 (2010)
7. Shen, G.: Research on the influence of going up and dropping stop system to the income volatility of ST stock. Quantitative & Technica Economics (5), 135–139 (2003)
8. Liu, J.: Couldn't Special TR eatment System are as a Selection Mechanism? South China Journal of Economics (5), 112–119 (2006)
9. Tang, Q., Huang, S.: The Research of React of Market to ST Release and ST Withdraw: the Experimental Test from Shanghai and Shenzheng Stock Maarket. Statistical Research 1(11), 43–47 (2006)
10. Shi, Y.: Analysis of losses and retreat system of ST company irregular. Finance and Accounting for International Commerce (5), 12–16 (2009)
11. Zhang, Y., Zhou, W.: Capital market ST system and small investors' protection. Accountant (3), 39–41 (2011)
12. Zhao, Z.: On the Filling-Up of Stock Governing Policy to Written Law: Based on the Experience of Quota System and Special Treatment. Securities Market Herald (4), 33–39 (2011)

Research on the Impact of Work Heterogeneity on Human Resource Development Needs

Yan Qin and Dan Mu

Department of Information Technology and Business Management
Dalian Neusoft Institute of Information
Dalian, 116023, China
{qinyan,mudan}@neusoft.edu.cn

Abstract. Organizations have focused on heterogeneity, which has a significant impact on human resources development. By studying the correlations among heterogeneity and training development, career development, organizational development, management development, the impact dimensions of heterogeneity on human resource development were revealed. This study compared the differences of human resource development needs for different explicit heterogeneity by using one-way ANOVA and multiple comparisons analysis and found there were significant differences in human resource development needs for employees with different job positions, job responsibilities, organizational nature, organizational size and industry.

Keywords: Training Development, Career Development, Management Development, Organizational Development.

1 Literature Review

Heterogeneity of organizational members brought more severe challenges to the organizations trying to survive in the market competition and increase profit. Attention paid to heterogeneity by organization, accessibility and configuration of human resource and dynamic operational processes have a huge impact no matter from a strategic level or a practical level. Pelled (2003) based on personal characteristics and the relevance of the team task, classified team composition variables into "low-work-related characteristics" and "high work-related characteristics." "Low-work-related characteristics" refer to characteristics having a lower correlation with the team's task to be completed such as gender, age, etc., which are more of a social relationship with the team and not connected to the objective mission target; "High work-related characteristics" refer to characteristics directly related to task such as educational level, working years, etc., more reflecting task-related experience, differences of the view.

It is very important to discuss the professional disciplines of human resource development, which has been one of the hottest subjects in theorists. Many management experts and scholars focused on heterogeneity, resulting in different perspectives on heterogeneity studies. In addition, some scholars conducted some

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 97–104.
springerlink.com

research on heterogeneity regarding human resource management and development, such as heterogeneity training. However, empirical research on the impact dimensions of heterogeneity in human resource development activities is relatively lacking.

2 Research Framework and Research Hypotheses

This study elaborated from the impacts of employee's work heterogeneity on human resource development needs. The variables in the study include heterogeneity characteristic variables and human resource development variables.

Heterogeneity refers to differences of the research objects, including explicit heterogeneity and implicit heterogeneity. Explicit heterogeneity is obvious and can be observed easily. Based on the literature review, this study selected explicit heterogeneity mainly from work characteristics to elaborate, including five aspects of job positions, job responsibilities, organizational size, organizational nature and industry.

Human resource development is to improve individual, group and organizational efficiency by means of training development, career development, management development and organizational development. Human resource development activities include training development, career development, management development, organizational development.

Based on the literature review, we can see that empirical research of the impact of heterogeneity on human resource development needs is relatively lacking. Focusing on the framework of this paper, the following hypotheses are proposed:

Hypothesis 1: The training content which employees with work heterogeneity need is different.

Hypothesis 2: The teaching method which employees with work heterogeneity need is different.

Hypothesis 3: The training lecturer which employees with work heterogeneity need is different.

Hypothesis 4: The training form which employees with work heterogeneity need is different.

Hypothesis 5: The training participation which employees with work heterogeneity need is different.

Hypothesis 6: The career development which employees with work heterogeneity need is different.

Hypothesis 7: The management development which employees with work heterogeneity need is different.

Hypothesis 8: The organizational development which employees with work heterogeneity need is different.

3 Study Design

3.1 Questionnaire Design and Scale Development

This study used questionnaire survey as empirical research method.

Human resource development questionnaire included four dimensions of training development, career development, management development and organizational

development. The dimension of training development consisted of 5 questions: training content, teaching method, lecturer type, training form and training participation, to know the respondents' training development. There was one question in the dimension of career development, the respondents' career needs. There was one question in the dimension of management development to know management development needs. One question was included in the dimension of organizational development to know organizational development needs. All options were designed on basis of referring to human resource development review, relative questionnaires of human resource development and enterprises' human resource development applications.

3.2 Data Collection

In this study, 1,100 questionnaires were sent with 87% response rate ,75% valid questionnaires. In the data collection statistics, the questionnaires with some missing options or with the answers on the same scale to all questions due to the respondents' less seriousness were regarded as invalid questionnaires.

The questionnaires covered Dalian, Beijing, Shanghai, Shenzhen, Shijiazhuang, Guizhou, Panjin, Fuxin, Jinzhou, Taiyuan and other cities, involved HP China, Neusoft Group, Shijiazhuang Yiling Pharmaceutical Co., Ltd. Guangzhou Kingdee Software Co., Ltd, Agricultural Bank of China, Industrial and Commercial Bank of China Limited, GE Dalian, CITIC Securities Brokerage Limited, ING Capital Life Insurance Company Ltd., Dalian Broadcast Station, Hua Xia Bank Co., Ltd. Accenture, Dalian Bohan consultancy company, Citibank Shanghai Branch, Cisco System, Inc. and DHI·DCW Group Co., Ltd .

The questionnaires were collected in hardcopy and electronic form. Of all the questionnaires, 400 copies were sent by email and 700 copies were sent in hardcopy to be filled out on the site. Data were analyzed for 829 questionnaires.

4 Data Analysis

The paper compared the differences of human resource development needs for different dimensions of work heterogeneity including job positions, job responsibilities, organizational size, organizational nature, industry, mainly using one-way ANOVA and LSD multiple comparison methods.

4.1 Analysis of Job Positions Heterogeneity

The results showed that, in training content, first-line managers and ordinary employees were more likely to choose professional development and technical training; middle and senior managers, first-line managers were more likely to choose management skills than ordinary employees; while middle and senior managers were more likely to choose foreign language training than first-line manager.

In teaching method, middle and senior managers and ordinary employees were more likely to choose lecturer teaching than first-line managers; first-line managers were more likely to choose scenario simulation than middle and senior manager and ordinary employees. In lecturer type, ordinary employees were more likely to choose lecturer

teaching than middle and senior managers. In training form, ordinary employees were more likely to choose on job training than first-line managers. In training participation, ordinary employees were more likely to be good listeners in training than middle and senior managers and first-line managers.

In career development, middle and senior managers were more likely to choose administration path than first-line managers and ordinary employees; ordinary employees were more likely to choose professional technical development path than middle and senior managers and first-line managers; first-line managers were more likely to choose cross-sector path than ordinary employees.

In management development, middle and senior managers were more likely to choose management training than ordinary employees; first-line managers and ordinary employees were more likely to choose tutor training than middle and senior managers. In organizational development, ordinary employees were more likely to choose skills improving than middle and senior managers and first-line managers.

4.2 Analysis of Job Responsibilities Heterogeneity

The results showed that, in training content, IT staff were more likely to choose technical training than administrative staff and sales staff; administrative staff were more likely to choose management skills than IT staff and sales staff; sales staff were more likely to choose management skills than IT staff; sales staff were more likely to choose foreign language training than administrative staff.

In teaching method, IT staff and administrative staff were more likely to choose lecturer teaching than sales staff; while sales staff were more likely to choose audio and video teaching than IT staff.

In lecturer type, administrative staff were more likely to choose professors and scholars than IT staff; IT staff were more likely to choose professional trainer than IT staff and sales staff.

In training participation, administrative staff were more likely to choose to be good listeners in training than sales staff; IT staff were more likely to be active participants than sales staff and administrative staff; administrative staff and sales staff were more likely to be bold questioners than IT staff.

In career development, administrative staff and sales staff were more likely to choose administration path than IT staff; while IT staff were more likely to choose professional skills development path than sales staff and administrative staff.

In organizational development, administrative staff were more likely to choose to learn knowledge than IT staff and sales staff; IT staff were more likely to choose to improve skills than sales staff and administrative staff.

4.3 Analysis of Organizational Size Heterogeneity

The results showed , in training content, staff in organizations with 51-200 people were more likely to choose professional development than those in organizations with 0-50 people and over 5000 people , what's more, staff in organizations with 1001-5000 people were more likely to choose professional development than those in organizations with over 5000 people; staff in organizations with 201-1000 people

were more likely to choose computer training than those in organizations with 0-200 people and 1001-5000 people; staff in organizations with 0-50 people were more likely to choose career development than those in organizations with over 51 people; staff in organizations with 1001-5000 people were more likely to choose thinking development than those in organizations with 201-1000 people and over 5000 people.

In teaching method, staff in organizations with 0-50 people and 1001-5000 people were more likely to choose lecturer teaching than those in organization with over 5000 people; staff in organizations with 51-200 people were more likely to choose group discussion than those in organization with over 201 people.

In lecturer type, staff in organizations with 51-200 people were more likely to choose senior management than those in organization with 1001-5000 people ;while staff in organizations with 1001-5000 people were more likely to choose professional trainer than those in organization with 0-50 people and over 5000 people.

In training form, staff in organizations with less than 5000 people were more likely to choose on job training than those in organization with over 5000 people ; staff in organizations with over 1001 people were more likely to choose off the job training than those in organization with 0-50 people, in the meanwhile, staff in organizations with over 5000 people were more likely to choose off the job training than those in organization with 201-1000 people ; staff in organizations with 0-50 compared with staff in organization with 201-5000 people and staff in organization with 51-200 people and over 5000 people compared with those in organization with 0-200,1001-5000 people were more likely to choose part time self- study; staff in organization with over 5000 people compared with those in organization with 0-200 people,1001-5000 people, and staff in organization with 201-1000 people compared with those in organization with 51-200 people were more likely to choose online learning.

In training participation, staff in organizations with 1001-5000 people were more likely to be good listeners in training than those in organization with 0-200 people; staff in organizations with 51-1000 people were more likely to be active participants in training than those in organization with over 5000 people; staff in organizations with 0-50 people compared with those in organization with 51-200 people and staff in organization with over 5000 people compared with those in organization with 1001-5000 people were more likely to be bold questioners.

In career development, staff in organization with 0-50 people were more likely to choose administration path than those in organization with 51-200 people; staff in organization with 51-5000 people were more likely to choose professional skills development path than those in organization with over 5000 people; staff in organization with over 5000 people were more likely to choose cross-sector path than those in organization with1001-5000 people.

In organizational development, staff in organization with over 5000 people were more likely to choose to learn knowledge than those in organization with 0-200 people; staff in organization with 201-5000 people were more likely to improve ability than those in organization with over 5000 people; staff in organization with 0-50 people were more likely to change attitudes. than those in organization with 201-1000 people.

4.4 Analysis of Organizational Nature Heterogeneity

The results showed , in training content, staff in wholly foreign owned enterprises and joint ventures were more likely to choose professional development and foreign

language training than those in state-owned enterprises and private enterprises; staff in private enterprises were more likely to choose computer training than those in wholly foreign owned enterprises , joint ventures and state-owned enterprises; staff in state-owned and private enterprises compared with those in wholly foreign owned enterprises and joint ventures were more likely to choose thinking development; furthermore, staff in state-owned enterprises were more likely to choose thinking development than those in private enterprises.

In teaching method, staff in wholly foreign owned enterprises and joint ventures were more likely to choose group discussion than those in state-owned enterprises; staff in state-owned enterprises were more likely to choose scenario simulation than those in wholly foreign owned enterprises and joint ventures.

In lecturer type, staff in wholly foreign owned enterprises and joint ventures were more likely to choose professional trainer than those in private enterprises.

In training form, staff in state-owned enterprises were more likely to choose off the job training than those in wholly foreign owned enterprises ,joint ventures and private enterprises; staff in private enterprises were more likely to choose part time self-study than those in state-owned enterprises; staff in wholly foreign owned enterprises ,joint ventures and private enterprises were more likely to choose online learning than those in state-owned enterprises.

In training participation, staff in state-owned enterprises were more likely to be good listeners than those in wholly foreign owned enterprises ,joint ventures and private enterprises ; staff in private enterprises were more likely to be bold questioners than those in state-owned enterprises.

In career development, staff in state-owned enterprises were more likely to choose administration development path than those in private enterprises ; staff in state-owned enterprises were more likely to choose professional skills development path than those in s wholly foreign owned enterprises and joint ventures; staff in wholly foreign owned enterprises, joint ventures and private enterprises were more likely to choose cross-sector path than those in state-owned enterprises .

In management development, staff in wholly foreign owned enterprises and joint ventures were more likely to choose management training than those in private enterprises; staff in state-owned enterprises were more likely to choose tutor training than those in private enterprises; staff in private enterprises were more likely to choose short-term rotation than those in state-owned enterprises and private enterprises.

In organizational development, staff in state-owned enterprises and wholly foreign owned enterprises and joint ventures were more likely to choose to improve ability than those in private enterprises; while staff in private enterprises were more likely to choose to change attitudes than those in state-owned enterprises, wholly foreign owned enterprises and joint ventures.

4.5 Analysis of Industry Heterogeneity

The results showed, in training content, staff in IT industry compared with those in trading service , finance and insurance industry were more likely to choose technical training; staff in IT industry compared with those in pharmaceutical industry were more likely to choose foreign language training; staff in finance and insurance industry were

more likely to choose career development than those in IT industry; staff in pharmaceutical industry were more likely to choose thinking development than those in IT industry.

In teaching method, staff in finance and insurance industry were more likely to choose lecturer teaching than those in IT industry; staff in IT industry were more likely to choose group discussion than those in finance and insurance industry; staff in pharmaceutical industry were more likely to choose scenario simulation than those in IT industry.

In lecturer type, staff in finance and insurance industry were more likely to choose senior management compared with those in pharmaceutical industry; staff in pharmaceutical industry were more likely to choose professors and scholars compared with those in IT industry.

In training form, staff in pharmaceutical industry were more likely to choose on job training than those in finance and insurance industry; staff in finance and insurance industry were more likely to choose off the job training than those in IT industry; staff in IT industry and trading service industry were more likely to choose part time self-study than those in pharmaceutical industry; staff in IT industry compared with those in finance and insurance industry were more likely to choose online learning.

In training participation , staff in pharmaceutical industry , finance and insurance industry compared with those in IT industry were more likely to be good listeners; staff in pharmaceutical industry were more likely to be active participants than those in finance and insurance industry; staff in IT industry compared with those in pharmaceutical industry were more likely to be bold questioners.

In career development, staff in finance and insurance industry were more likely to choose administration path than those in pharmaceutical industry; staff in pharmaceutical industry were more likely to choose cross-sector path than those in finance and insurance and IT industries.

In management development, staff in finance and insurance industry were more likely to choose tutor training than those in IT industry; staff in pharmaceutical industry compared with those in IT industry were more likely to choose short-term rotation.

In organizational development, staff in trading service, finance and insurance industry compared with those in IT industry were more likely to choose to improve ability; staff in IT industry were more likely to choose to improve skills than those in pharmaceutical industry; staff in pharmaceutical industry compared with those in IT industry, trading service industry were more likely to choose to change attitudes.

5 Conclusions

This paper studied the impacts of employees' work heterogeneity on human resource development needs and proposed research framework and research hypotheses. The results showed that the employees with different work heterogeneity, whose needs for training content, teaching method, lecturer type, training form, training participation, career development, management development, organizational development were different and all hypotheses were supported. This study provides a reference for human resource managers in decision making when they organize training development, career development, management development and organizational development activities

carefully catered to work heterogeneity of employees and implement human resource development strategies. Besides explicit heterogeneity characteristics, other implicit heterogeneity will be in-depth studied in the future.

References

1. Combs, G.M., Luthans, F.: Diversity Training: Analysis of the Impact of Self-Efficacy. Human Resource Development Quarterly 18(1), 91–120 (2007)
2. Putney, D.M.: Ph.D. Thesis, SWAT Team composition and effectiveness. The University of tennessee, Knoxville (2003)
3. Torraco, R.J.: Challenges and Choices for Theoretical Research in Human Resource Development. Human Resource Development Quarterly 2(15), 171–188 (2004)
4. Holton, E.F.: The Mandate for Theory in Human Resource Development. Human Resource Development Review 1(1), 3–8 (2002)
5. Kuchinke, K.P.: Debates over the Nature of HRD: An Institutional Theory Perspective. Human Resource Development International 3(3), 279–283 (2000)
6. Lynham, S.A.: Theory Building in the Human Resource Development Profession. Human Resource Development Quarterly 11(2), 158–178 (2000)
7. Swanson, R.A.: Human Resource Development and Its Underlying Theory. Human Resource Development International 4(3), 299–312 (2001)
8. Lombardo, M.M., Eichinger, R.W.: High potentials as high learners. Human Resource Management 4(39), 321–329 (2000)

A Research of an Undergraduate Course "Exploratory" Closed Book Examinations Method* — The "Intermodal" Courses as an Example

Liu Mingwu and Zeng Haoran

School of Management, Chongqing Jiaotong University, Chongqing Zip Code 400074
liumingwu2007@yahoo.cn, zhrwan@qq.com

Abstract. The current examinations have some shortcomings, so there is a confrontation between teachers and students. Students usually do not work hard. In order to pass the test, they only study on the eve of exam. In fact, this is the isolation of teaching and learning, which is not conducive for students' know-how. In response to this undesirable phenomenon, this paper analyzes the status of examinations and proposes a new test method—exploratory closed book examination. A practical operation examination is performed, which proves this method and gets a risk assessment and concludes the prospect.

Keywords: Reform experiments, examination methods, exploratory closed book exam.

1 Introduction

At the "twelve-five" and the long-term plan, Chongqing Jiaotong University makes a developing strategy, which is from the "teaching-oriented" universities to "teaching and research" universities changing. In such an environment, recently arrived as a young teacher, I am very lucky that I am on an ample stage. With the tremendous expectations, I am full of passion for the work, and want to do something for my school. In accordance with the college's commitment, I work actively to deploy the task of teaching and teaching practice, at the same time I observe various phenomena, I still have the courage to think about the problems of teaching practice, and actively applied for Chongqing teaching reform subjects, and obtained experimental teaching reform (container multi-modal transport experiments and practice of teaching reform) project. Through several examiners, I found a universal phenomenon that in course examinations, many students chatted.

The examination of the teaching process is a very important part, and testing is a means of education quality [1]. There are always some students who chose the online course, and they only go to a few classes, even worse, they are also absent from the examination, the rates of public exam absenteeism are increasing[2]. Wang Li proposed a "people-oriented" means, which will help to mobilize the enthusiasm, initiative, students' gift performing as well as the healthy growth [3]. In the

* Chongqing Jiaotong University, Experimental reform projects (SYJ201018).

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 105–108.
springerlink.com

examination content, it is a shift from knowledge to ability assessment [4]. Researchers have also proposed a method of online assessment, which is completed on the computer the volume, the whole process of examination [5].

To what extent can the current open-book or closed book exam investigate the actual level of the students? What kind of examinations can better guide students to active learning? The best way is to emancipate students from the spot to recite the liberation, so that students are no longer afraid of exams, cheating will not exist. Preventing misconduct examinations from the source also fully embodies the school care for students.

2 The Practical Operation of Exploring Closed-Book Examination

In the literature review, we learned that all the existing tests are the existence of such drawbacks, these phenomena will more or less adverse impact on student learning motivation, learning efficiency, and ultimately student' learning objectives will be shifted, which will affect student' learning outcomes. So we need to find a better test method to solve the above drawbacks.

This paper proposes a new method which may be able to solve the problem of the utility. Through an induction way, we can turn the students from rote, fraud or misconduct of such a bad way to the student who actively learn, learn to learn, learn the way up. Ultimately students learning ability, the ability to learn from each other, learn to grasp the essence of the course of study, which can really inspire students, and cultivate students.

In accordance with school requirements, this course covers the key knowledge points which are grouped into questions A, B, C three papers; three weeks before the end of the course, we arrange for students preparing for examinations. We demand students summarize the course of study according to their understanding. Encourage students to discuss the review and summarize the core of this course. Students will be required to summarize their experience records in an A4 sheet of white paper.

3 Assessment and Countermeasures of Examinations Risk

As the scores of university courses have only compared with the value of the same class. The test design approach, students are in the same competitive platform, there is no unfair, unjust phenomenon.

Even if the experiment did not get the desired results, such as student achievement is generally high, and there is no differentiation, which affects the national education assessment. This situation can be detected via pre-marking method ahead of schedule, using the appropriate scoring criteria can achieve controlling of the student achievement goals. In fact, as long as we have well-designed paper, these problems can be avoided.

4 After-Evaluation of Exploratory Closed Book Examination

(a) Classroom observation evaluation. At the entire examination process, the paper was not found on any entrainment phenomenon, and whispering among students is

very rare. This method can prove that it performed a suppression test. In addition, when the test was terminated about 1 / 3 of students did not carry out an assignment, It indicated that they can not get too much from the information summarized (it seems it can be expected 1 / 3 of the students will fail).

However, students do get some answers from the information or inspiration. Because of this, this method played the role of inducing students to prepare, not to think about how to cheat. You can harvest after you do. While cheating is very risky, it may cause serious consequences.

(b) Analysis and evaluation of papers. Generally speaking, a good paper and a successful test show that student' performance will generally follow a normal distribution, that is, the middle is large and two ends are small, along with differentiation.

If the student scores generally obey normal distribution (kurtosis values close to 3), which reflects the differentiation. Then we get a very interesting conclusion: this study does not affect the way students' test scores distinguishing degrees, students can reflect real differences in the situation. The conclusion has a potential significance, which may really change examinations method of some speculative courses in university, and release the students from pain from reciting. It is not only beneficial to students' physical and mental health, but also reduces the number of students brought cheating incidents which should not be handled, reflecting the school can really care for the students and protect them. Therefore, how to effectively reduce or discourage students to cheat, is worthy of all educators for a serious consideration.

If the result did not show similar characteristics of the normal distribution, the reason may be that paper is not perfectly designed, there are too many questions easy to find answers; or it may be the true situation. There are two ways to resolve this situation, first, direct that experiment fails, the second is to continue to sum up experience, re-design papers and experimental aspects, in the latter class to carry out comparative experiments (this is only one class, so we cannot do comparative experiments). 57 students participate in this examination, 57 examination papers are received. After we do a JB normal for students' scores test, results are as follows: (1)The mean and median scores are 60, meaning that all achieve the teaching requirements;(2)The volatility of the results is 9.3 (9.3%), indicating that overall performance is stability, there was no extreme polarization;(3) The highest score is 85, the lowest is 45, indicating that this test method does not have a significant impact on scores. So students can not get too much score from the information on the;(4) The distribution of scores peak values 3.1, which is close to the normal distribution of the kurtosis value of 3. The skewed results value is 0.65. Although it is higher than the normal distribution skewed value of 0, the results have shown a degree of distinction; (5) From the image of view, the left tail distribution is thicker, while right tail is thinner. The number of lowest score is greater than that of highest score, but the number of two-layer segment and accounted contains over 15% of the total number.

The scores reflect the distinction between the degrees of student' achievement, the number of people whose results are very good or very poor is small, which is in line with the college basic status quo. The sharing and value is 60, so the whole has passed. The number of people whose paper score is below 60 occupies about half, while the number of people who almost passed is about 10. So the number of people requiring make-up is about 20,occupying about 35%, is in line with the actual

situation. As many students have to prepare for CET4 and CET6 this semester, the students do not spend enough time on this course.

(c) Student evaluation

In the survey, 51 questionnaires were distributed and total 51 were recovered, that means a hundred percent recovery rate. Throughout the survey process, as a course instructor also the sponsor of exploratory examinations, I was not on the spot, so I will have no influence on the survey. According to the survey results, all students receive this test method, and most believe that such examination method is better than closed book and open-book exam on average. According to the student's answer for the exam, you can see that most of the students believe that the explore way can promote their positive attitude examination and mitigate the burden of preparing for examinations.

5 Conclusion

A closed book exam exploratory teaching reform is not contrary to basic principles. Although the specific operation and the school seems "closed book exam are not allowed to bring data" provided some conflict. However, careful analysis of this test method is not entrained material behavior. "Closed-book examinations are not allowed to bring data" entrainment information can not be interpreted as students, teachers can not acquiesce students smuggle data or teachers instructed students not to smuggle information.

References

1. Liu, H.M.: Adult education exam problems and countermeasures. China Adult Education (3), 40–41 (2008)
2. Wei, B.G., Liu, L.: From the public way analysis of college teaching elective exam problems. China Education Research (5), 79–80 (2006)
3. Li, Z.P.: The university thinking politics theory class teaching reform. China Adult Education 8, 3–4 (2006)
4. Jia, Z.Z., Wang, L.L., Huang, P.: The ideological and political theory course examination reform. Education and Occupation (6), 124–125 (2007)
5. Wang, L.: Ideological and political class to implement people-oriented examination method reform. Science and Technology Market Economy (1), 73–74 (2011)

Study on Strategic Choices of China's Commercial Banks Based on SWOT Analysis

Yan-jing Bai* and Zhi-feng Yin**

School of Economics and Management, Beijing Jiaotong University,
Beijing 100044, China
13488883583@163.com

Abstract. The trend of financial globalization and liberalization has changed the competitive situation of the financial industry significantly. In this paper, we use the SWOT analysis to analyze the strengths, weakness, opportunities and threats of China's commercial banks, and give some suggestions on the strategic choices for them to develop healthily and steadily.

Keywords: SWOT Analysis, commercial bank, strategy.

1 Introduction

For banks, development strategy is the major issue related to overall situation and long-term directions. From a development perspective, the strategy considers developing trend of the future. From a competitive point of view, the strategy solves problems about competing with whom, how to compete and how to make better use of competitive resources. This paper will analyze the internal conditions and external environment of China's commercial banks by using the SWOT analysis to better grasp the strategic trends of banks.

2 SWOT Theory

SWOT analysis is a basic method to investigate development strategy of enterprises or industry. It's the self-diagnosis and environmental analysis before the development strategy. S stands for strength, W stands for weakness, O stands for opportunity and T stands for threat. Among them, S and W are the internal factors, while O and T are external. Enterprises identify their strengths, weaknesses and core competitiveness according to the internal conditions and identify opportunities and threats in the external environment, and then maximize the use of their internal strengths and external environmental opportunities, while avoiding their weaknesses and threats of external environment to minimize adverse effects. According to the concept of

* Yan-jing BAI (1987-), female, Han nationality, Hebei Province, School of economics and management, Beijing Jiaotong University.

** Zhi-feng YIN (1986-), male, Han nationality, Beijing, School of economics and management, Beijing Jiaotong University.

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 109–115.
springerlink.com

competitive strategy, an enterprise's strategy should be an organic combination between "can do", the organization's strengths and weaknesses, and "may do", environmental opportunities and threats. SWOT analysis is objective and important theory for the enterprise's strategic choice.

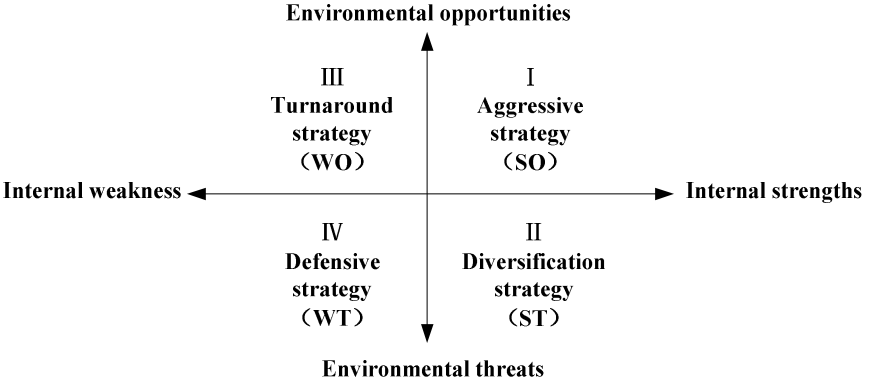


Chart 1. SWOT Analysis Diagram

3 Development of China's Commercial Banks Based on SWOT Analysis

3.1 Strengths of China's Commercial Banks

The internal strengths of China's commercial banks are mainly manifested in the following aspects:

1. Strong Financial Scale. China's commercial banks have significantly reduced the gap with big international banks. By the end of 2009, total assets of national commercial banks are 52.62 trillion yuan, up 29.5% on the previous year, total liabilities 4.986 billion yuan, up 30.4% on the previous year, equity 2.76 trillion yuan, up 15% on the previous year, deposits 42.48 trillion yuan and loans 27.35 trillion yuan , up 26.5% and 32.6% on the previous year. All capital adequacy ratios are appropriate to the standard. Asset quality improved significantly and profitability, risk-resisting ability and liquidity management has greatly improved.

2. As native China's commercial banks have a clear business and client resources advantages. It's not only helpful for enterprise culture to form internal cohesion and service brand, but also conducive to better communication, understanding and satisfying customers' needs.

3. After a series of reforms of institutions and mechanisms and competition with domestic foreign banks, China's commercial banks' competitiveness and sense of competition are greatly improved. By the end of 2009, China's commercial banks' non-performing loans is 497.3 billion yuan, 63 billion less than the beginning of this year, NPL ratio 1.58%, down 0.84 percentage points over the beginning. Asset structure and asset quality have improved significantly. The ability of offsetting has been enhanced.

3.2 Weakness of China's Commercial Banks

The internal weakness of China's commercial banks is mainly manifested in the following aspects:

1. Its profitability model is single, strongly depending on interest income. Business structure is in imbalance and development fatigues. International strength and experience is not enough and is still inward-oriented. Products and services lack innovation. In the interest rate market, it contains a great deal of potential market risks. Most domestic commercial banks' profitability model is obviously characterized by over-reliance on spread income.

2. The out-of-date human resource management mechanism results in staff surplus, high cost, low efficiency, and on the other hand lack of effective management, high management personnel, and the serious brain drain. Human resource management capabilities of China's commercial banks and that of foreign financial companies are still far apart.

3. With backward IT technology, China's commercial banks lack sufficient knowledge of the profound changes and lasting impact that IT technology may have on the banking sector. With the further development of network economy and the economic integration, digital banks began to emerge. Bank cards, online banking has become the key for China's commercial banks to gain the initiative in future competitions. How to integrate the advantages of traditional and electronic channels to build the IT channel model is a problem that China's commercial banks urgently need to solve.

3.3 Opportunities for China's Commercial Banks

1. Stable Politics and Good Macro-Economy. As one of the fastest growing economies in the world, China's economy maintained a sustained, stable and fast growth and was widely seen as power and hope for world economy to recover from the financial crisis, which is unique opportunities and advantages for China's commercial banks.

2. Government Support. Enjoying the national regulatory policies, China's commercial banks take the lead in seizing the domestic market after expanding domestic demand. The central government also increased the intensity and pace of commercial banks listed on restructuring in recent years and laid the foundation from the aspects of financial recapitalization and restructuring for the development and growth of China's commercial banks.

3. China's financial market is in its infancy and there is considerable room relative to the level of economic development and market demand. It provides an optimistic market expectation for innovation and development of banking business. Economic crisis in 2009 delayed the development of foreign banks in China to some extent. It provides a rare strategic buffer period for China's commercial banks and a business opportunity for China's banking industry.

3.4 Threats for China's Commercial Banks

1. Full liberalization of domestic financial markets makes an increasingly wider range of market competition and competition is increasingly fierce. HSBC, Citibank, Standard Chartered Bank and other foreign financial institutions have sighted on China's financial market to fiercely compete with China's commercial banks which will bring great impact on domestic commercial banks.

2. Integrated management is the main trend of today's international banking. International banks have accumulated a wealth of operational experience of operations on the consolidated. By contrast, domestic commercial banks is in an environment with regulatory constraints of the country's long-term business strictly separate policy, and lack comprehensive management capabilities and banking innovation.

3. Compared with the construction of the effective credit file and information sharing system in Western countries, which we often call "black list" system, China lack credit environment. Relevant laws and regulations supporting is obviously not perfect.

4 Strategic Choices of China's Commercial Banks

4.1 Aggressive Strategy

1. Actively Expand Overseas Business

With the domestic banks gradually going to the world, how to develop overseas business to truly become the leading international and globalization banks has become the focus of China's commercial banks and future business development point. To expand overseas and international business China's commercial banks must first choose a mature international financial product as a platform, and then gradually build an international financial products sequence. In addition, by making full use of the favorable opportunity that China's foreign exchange reserves are abundant and large multinational companies gradually enter the international financial markets, especially the debt markets, China's commercial banks should gradually carry out project financing, securities brokerage, derivatives securities, currency exchange, international mergers and acquisitions advisory and other new business to speed up the pace of international operations and enhance international competitiveness.

2. Carry Out Integrated Operation

With increasing competitions among modern commercial banks, the traditional business profit model whose profits rely on savings and loan interest rate is being increasingly challenged. China's commercial banks should adjust the structure of public and private business, changing from the "company bank" to "retail banking" to speed up the promotion of middle business, improve non-interest income, actively expand the overall business, and provide customers with a wide range of financial services to enable commercial banks to diversify business and revenue sources.

3. Create Product Branding

In the context of convergence of financial products and services, China's commercial banks should actively promote operating characteristic management structure, and build up a distinctive brand image to enhance the brand's social

reputation. Concentrate on brand marketing, closely combine brand value and culture of the bank to enhance brand meaning. Enrich brand with products, and drive products with brand.

4.2 Diversification Strategy

1. Product Diversification

Products of China's commercial banks are relatively single. To enhance competitive advantage and reverse market risk, they should adhere to follow the principle of Customer First, regard service quality improvement and customer focus as their own long-term strategy, constantly seeking product innovation, product diversification, the integration and improvement of existing products, innovation, and development of new financial products. Either the form-based development strategy or service China's commercial banks should continue to change in order to attract different customer groups.

2. Target Market Segments

China's commercial banks should put their customers in a target market position to form a rational market which would help build up a hierarchical business mode. According to different customers' needs, risk characteristics, commercial banks should establish an appropriate profit mode. Large-sized customers require banks to provide more professional, comprehensive financial services to enhance the capacity of the global financial services. For medium-sized customers, who are sensitive to price, we should enhance the ability to identify and control risk. Small-sized customers, who have a small amount of its business needs, always face many difficulties, such as high frequency, high risk and high transaction costs. Commercial banks should adopt complementary competitive strategy to meet different customers' needs and take different operations in response to intense competitive environment.

3. Attention Localization

When domestic commercial banks emphasis on international business and institutions, we must see the huge domestic market. China's economy is in a stage of rapid development and industrial structural adjustment phase. The domestic commercial banks should look for a long-term, strategic and effective demand from the country's industrial structure upgrade and development of emerging financial markets. Only when keeping a foothold on the fertile land in the country that is not yet fully competitive, China's commercial banks would be able to gain a foothold on the international market in the future.

4.3 Turnaround Strategy

1. Improve the Business Structure

For their own weaknesses and opportunities in external environment, China's commercial banks' business model should be located in the heart of the traditional business-oriented and develop core business around the traditional business. While avoiding risks in their vulnerable and the maximum advantage of market opportunities, China's commercial banks improve the business structure, increase efficiency and profitability, reduce operating costs, and focus on economies of scale

to create competitive advantage. By streamlining business processes and extending the business through online banking services, commercial banks should eliminate unnecessary products or upgrade and expand the business scope of traditional projects to promote the core brand and consolidate traditional business advantage.

2. Competition and Cooperation

Competition and cooperation is the only way for commercial banks to grow. Strategic cooperation among Chinese banks and between Chinese and foreign banks should be strengthened. This new partnership is a change from pure competition, confrontation into competition, cooperation, and ultimately a win-win. Share software development resources through complementary advantages in the business. Foreign banks can also provide us with and important ways of the management such as staff training and management services. We will achieve resource sharing in the credit card, settlement network marketing, information management systems, merchants and other aspects through the exchange of skills and experience.

3. Information Technology

Financial industry is a cutting-edge one that needs intensive high-tech applications. As a booster for banking reform and development, information technology plays a very important role. The level of information technology and the merits of network services determine success and failure in competition. China's commercial banks must proceed from reality and increase investment in information technology to promote the rapid development of electronic services. Further improve the internal financial information service system, bank management and speed up construction of decision-making information. By full use of advanced information technology, we realize deep data applications based on the data set and breakthrough technology development bottleneck of China's commercial banks.

4.4 Defensive Strategy

1. Maintain the Stability of Operation

China's commercial banks should pay attention to financial planning, maintain an appropriate capital adequacy ratio, adequate liquidity and leverage ratios and emphasize on the long-term stable source of funding to maintain the traditional banking business at a certain percentage, which is an important part to resist.

2. Improve Risk Management

China’s commercial banks should improve risk management comprehensively to build their own prevention and control system for risk management. First, set up a risk management committee to monitor portfolio risks. Secondly, establish an independent risk management department and grant certain rights to monitor business processes. Thirdly, improve risk management framework and system to further clarify the relationship between rights and responsibilities of relevant departments to form a risk management system, whose main body is risk management. Fourthly, improve risk measurement techniques to improve the level of valuation and risk pricing. Fifthly, improve the level of risk management information system, including the establishment of various departments, branches, all customer data, to enhance information sharing and risk control. Sixthly, establish clear reporting system to provide information for making decision.

3. Strengthen Human Resources Management

China's commercial banks should speed up the reform of human resources management system and establish a wide range of bank property rights legal system to take the corporate governance to the track. To achieve the bank's development strategy goal and enhance core competitiveness to meet the staff's overall interests and long-term interests, commercial banks should establish a modern commercial bank human resource management mechanism and enhance human resources development and management. It's also very important to enhance staff training and maximize the value of all the employees while maximizing the value of the companies.

5 Conclusions

Only if we formulate appropriate development strategy, China's commercial banks can survive even reach sustainable development in the increasingly fierce competition. If the commercial banks want to gain ground in the incentive competition, we should make comprehensive analysis of their advantages, disadvantages, as well as opportunities, threats of external conditions, thus choose and establish global and long-term strategies which are sustainable for the development. Keeping the inherent advantages, we should avoid risk to the hilt, seize the opportunity and improve the core competitiveness at the same time, so that the commercial banks will operate stably in the long run.

References

1. Hu, X.: Strategic Change and Enlightenment of Large International Commercial Banks during Post-crisis Era. China's Financial 15, 25–26 (2010)
2. Chen, H.: Study on Development Strategy of Banks in Financial Crisis. Association for Science Forum 8, 148 (2009)
3. Wei, J.: Analysis of the Commercial Banks' Development Strategy under the Influence of the Financial Crisis based on SWTO Analysis. Chinese Trade Tribune. 14, 71 (2010)
4. Jiang, Q.: Analysis and Countermeasures of China Commercial Banks' Strategic Development. Kong Economy 8, 89 (2010)
5. Wang, K.: Study on the Strategic Transformation and Product Innovation of Commercial Banks in Current Situation. Modern Business 20, 20 (2009)
6. Lian, Y.: Study on Accelerating the Business Strategic Transformation of Commercial Banks. Technology and Industry 7, 60–64 (2010)
7. Qi, S.: Study on Strategic Transformation of China's Commercial Banks after the Financial Crisis. North Finance 4, 43–44 (2010)

The Characteristics and Implementation Strategy of Modern Hospital Knowledge Management

Zhi-guo Liu and Hai-bo Yang

Library of HeBei United University, 063000 Tangshan, Hebei, China
qxlzg@163.com

Abstract. Hospital is a typical knowledge-intensive organization, knowledge is the core asset of the hospital. Implementing knowledge management in hospital is very suitable because hospital has the features including flat organizational structure, innovation is the vitality of hospital and medical staff focus on lifelong learning. Modern hospital should promote knowledge management through popularizing the concept of knowledge management, paying attention to development and utilization of the existing hospital information system, setting chief knowledge officer and establishing a knowledge-sharing mechanisms and other measures.

Keywords: knowledge management, knowledge asset, hospital administration.

Knowledge management is a new management concepts in knowledge-based economy era and a product of the wave of social informatization and Knowledgeable that has an important role in promoting the management level of various types of social organizations. As a typical knowledge-intensive organizations, hospital introduce knowledge management theories, methods and techniques has an important strategic role in improving hospital management, service level and core competitiveness.

1 Meaning of Hospital Knowledge Management

There has never been a widely accepted definition on knowledge management. Simply, knowledge management is the management that take knowledge as the core and a process of improving the innovative capacity and ability of create value to meet existing and future needs of enterprise to open up market opportunities through the recognition and effective using of existing and acquired enterprise knowledge. Therefore, the starting point of knowledge management is knowledge be regarded as the most important resource, maximize access and use of knowledge as the key to improving the competitiveness of enterprises.

For hospitals, knowledge management is the systems management on generation, collection, organization, dissemination, exchange and application and other related processes of medical knowledge, including subjective and objective medical knowledge management which the core is to create a mechanism and platform of subjective and objective medical knowledge converse into each other to achieve the

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 117–121.
springerlink.com

exchange and sharing of medical knowledge and promote the hospital's knowledge and technological innovation to improve the hospital's medical technology level and service quality, and finally to make the hospital to survive and develop in the increasingly fierce competition . With the rise of the knowledge economy and an increasingly wide range applications in hospital work of information technology, it has become a inevitable choice for modern hospital implement knowledge management to enhance its core competitiveness.

2 Knowledge Management Characteristics of Modern Hospital

Hospital has a good foundation and advantages to implement knowledge management, mainly in the following areas.

2.1 Knowledge Is the Hospital's Core Assets

Hospital is a typical knowledge-based organization, the hospital's most important asset is not fixed and financial assets, but the hospital's knowledge assets, knowledge and information is the hospital's survival and development foundation. The hospital's knowledge can be divided into two categories: one is the objective medical knowledge, also known as explicit medical knowledge that be divorced from the medical staff's brain and recorded in a variety of literature, network and contained in the products or services; the other is present in the medical staff's brain called subjective medical knowledge, also known as tacit medical knowledge, including capabilities, experience and skills of disease diagnosis and treatment. Tacit medical knowledge is not easy to express in words, it's the product of individual long-term accumulation and creation which is the key of medical knowledge innovation, medical technological innovation and application. For the modern hospital, the management ability of knowledge especially tacit knowledge directly determine the survival and development of hospital.

2.2 Hospital Has a "Flat" Organizational Structure

The hospital's organizational structure is relatively simple, and management layers is fewer. Even a medium-sized hospital with 1,500 employees and a 500-bed only has a middle management department director and by these officers reporting directly to the dean. This flattening of the hospital organization is very conducive to the implementation of knowledge management. The information transmit in traditional pyramid management structure is vertical that weakening the health information communicate between different sections, departments and employees, not only affected the reliability of information but also lack of interaction of knowledge production, information transmission delay, work low efficiency, organizational overlap, wastage of resources. Flat organizational structure overcome the above drawbacks and with the extensive application of information technology, everyone in the hospital easier to get comprehensive and accurate information. The characteristic of hospital organizational structure is necessary requirement and important features of technological innovation and knowledge management.

2.3 Innovation Is the Vitality of the Hospital

Today society develops and changes extremely rapid. The organization can not survive and develop if it have not adaptability and innovation capability. Only with adaptability and innovation can the organizations take advantage in the competition. Medical technology innovation is the inevitable choice for hospital to survive in the fierce market competition. Medical technology innovation are particularly urgent and important especially in the background that the rapidly changing of the new technological revolution representative by information technology and life sciences. Innovation is the core of knowledge management, the hospital knowledge management can maximum release the effect of the hospital's knowledge and skills to create infinite value for the hospital.

2.4 Medical Staff Focus on Lifelong Learning

The acquisition of medical knowledge and skills not completed only through formal education, a lot of knowledge and skills of the medical profession must through long-term medical practice and gradully get it, on the objective it necessary require the medical staff focus on continuous learning, focus on learning by doing. Especially with the rapid development of medical technology, medical staff need to continuously update their professional knowledge and skills to become a lifelong learner. Technological innovation activities require long-term knowledge accumulation, medical staff focus on continuous learning is just an important requirement of knowledge management and technological innovation.

3 Implementation Strategy of Modern Hospital Knowledge Management

Implementation of knowledge management is a necessary requirement for the development of a modern hospital. Although modern hospital already has the infrastructure and conditions of implementation of knowledge management, but in the following areas still need to do a lot of work to really play the role of knowledge management.

3.1 Popularize the Concept of Knowledge Management

Hospital knowledge management involves the hospital culture, hospital management system, application and innovation of hospital knowledge, hospital information resource management, construction of hospital information system and other aspects. Therefore, it is not a technology and can not solve the problems by only buying a computer system or software. In the process of hospital knowledge management, technology environment changes will make changes of the hospital organizational structure, communication patterns and operation way. Therefore, hospitals need to adjust their management philosophy timely to adapt the environmental changes. Meanwhile, in the process of knowledge management advancing, the biggest resistance

may come from those managers or medical staff who do not like or do not want to adapt to the new technological environment. Facing various obstacles, the implementater need to find good reasons to convince them and prove that knowledge management will indeed benefit to the hospital's development. In this process, education and training is very important, the hospital staff's consensus is the basis for the success of knowledge management.

3.2 Pay Attention to the Development and Application of Existing Hospital Information Systems

Knowledge management is the expand and deepen of the traditional information management methods and a more advanced form of information management. It is still depends on technology-based networks, databases or a specification document, the means is still the data's collection, storage, analysis and transmission that related with knowledge. The raw data of knowledge management hidden in hospital information systems or associated with some data of the hospital information system. Therefore, the construction of hospital knowledge management system definitely not to build a new system, but is the growth and expansion of the existing hospital information system. For example, build a platform in the hospital information system for published opinions to converse the objective and subjective medical knowledge into each other or try to make a valuable part of the electronic medical records can be retrieved to implement knowledge management.

3.3 Set Chief Knowledge Officer

The purpose of hospital knowledge management is based on innovation, but innovation is no precedent, often is the "crystallization" of the collective wisdom and an organized activity. Based on this, it is necessary for hospital implement knowledge management to set up an authority chief knowledge officer who responsible for the hospital's knowledge management activities. The chief knowledge officer should be equivalent to a vice dean. His main role is responsible for the unified management of hospital knowledge resources according to the hospital's development plans and strategies to ensure the sustainable development of the hospital. Chief knowledge officer must be able to combine the structured external knowledge and intuition to perceive tacit knowledge of organizational culture and behavior, ensure the hospital's intellectual assets ultimately translate into knowledge products that can bring profits for the hospital. This determines the chief knowledge officer is not only an important participant in the knowledge management, but also a managers in knowledge management activities.

3.4 Establish a Knowledge-Sharing Mechanisms

Establish knowledge-sharing mechanism is the most difficult part of hospital knowledge management. People have become accustomed to regarding the accumulated information and gained experience or skills as a part of personal wealth, refused to share with others. Modern hospital can solve this problem through these methods: first, gradually change the work environment and ways to build the network as a platform for the work environment, records and archive workflow and work files

online to increase the transparency of workflow. Second, maximize provide public information resources in the hospital local area network, such as hospitals guide, knowledge maps, a database of completed projects, status of project team collaboration, research project reports, evaluation of skills, using the Internet and groupware technologies to break barriers of communication between hospital departments to realize the hospital fast real-time communication and make the network became a actuator of hospital changing to learning organization. Third, create a team work mechanism in the R&D to allow more people to participate in the process of the tacit knowledge generation and use in the research and development activities in order to weaken the loss of tacit knowledge caused by the individual leaving.At the same time,take the network as a link and real-time information summary as management system to promote tacit knowledge explicit.

References

1. Gui, S.H., Lu, L.W.: Ten top management concepts that affect the future of China's enterprise. China Textile Press, Beijing (2004)
2. Ming, S.L.: Discussion on hospital medical knowledge management. Chinese Journal of Hospital Administration 18(9), 553 (2002)
3. Dian, Z., Bo, J.C.: Hospital knowledge management and technological innovation. Chinese Journal of Hospital Administration 19(4), 194 (2003)
4. Rui, X.: Knowledge management features of the Knowledge-based enterprise. Library and Information 1, 47 (2000)
5. Yin, C.G.: Introduction of a new hospital management - knowledge management. Soft Science of Health 15(4), 17–18 (2001)
6. Li, H., Dong, X., Zuo, M.: Knowledge Management in Theory and Practice. Huayi Press, Beijing (2002)

Competitive Advantage in e-Commerce: The Case of Database Marketing

Teresa Guarda[1], Maria Augusto[1], and Carlos Silva[2]

[1] Isla - Superior Institute of Languages and Administration of Leiria, Portugal
[2] CIEO – Centre for Spatial Research and Organizations,
Universidade do Algarve, Portugal
{tguarda,mfg.augusto}@gmail.com,
carlos.silva@unisla.pt

Abstract. In the contemporary competitive environment, the way that organizations generate, manipulate, select and make use of information about its customers has gained increasing importance for its potential to give the company an edge over competitors. Consumers demand a treatment increasingly personalized and products\services that better suit their needs, being receptive to actions that allow for greater interaction in the purchase process. The aim of this paper is to analyze the database marketing as a crescent usefulness tool, connecting businesses and consumers in the context of new market settings. With the appropriate structure can improve the integration and availability of data for customers, improving the company's performance, giving a more solid support at the level of decision making in business marketing activities and development strategies to acquire new customers, or ultimately helping to keep customers. These benefits become more significant in the digital economy.

Keywords: Competitive advantage, Database Marketing, e-Commerce.

1 Introduction

Currently, the terms marketing one-to-one personalization, loyalty and database marketing are often cited, as a means that enables organizations to differentiate themselves from the competition in the new scenarios of the current competitive environment. The growing use of online transactions is changing the way organizations interpret customer data. The active integration of data is vital, enabling the increase of market share online. Thus, the database appears as a marketing tool that is used properly, will increase the degree of customer satisfaction, leading the organization to better performance.

The database marketing (DBM) enables organizations to ensure their survival and gain competitive advantage in the digital economy. Organizations that do not have information with the features and characteristics of their clients have serious risk of succumbing to the action of competition.

This paper presents an overview of the contributions of several researchers, in order to discuss relevant issues of database marketing as a source of information for

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 123–130.
springerlink.com

the development of strategies to enable organizations to gain a sustainable competitive advantage, and analyze all aspects of e-commerce that justify the use of database marketing as a tool to support strategic decision making.

2 Database Marketing

DBM is an important strategic tool to support relationship marketing, as the main authors argue Gummesson [1], McKenna [2], Vavra [3]. Tucker and Brown [4] argue that, although significant benefits arising from their application, it is an underutilized resource, due to be circulated among some marketers. The DBM is focused primarily on the development of databases of consumer characteristics, including identification data, data relating to purchase history, demographics, and psychographics, and other useful information that enable the organization to build a foundation for the establishment of strategies targeted to an ever more specific to the development of new products, as is the case, communication, pricing, promotions. DBM aims to promote the integration of the profiles of consumers and their buying behavior [5], providing guidelines for the organization to adapt their activities to the profile of the target market in order to better meet their needs and gain a competitive position. DBM is an information-driven marketing process, managed by database technology that enables marketers to develop, test, implement measure and appropriately modify customized marketing programs and strategies [6]. DBM is being accepts as a fundamental marketing tool in a growing variety of business – from giant multinational corporations to local retailers offering a new product to their customer lists. The concept was introduced in the 80's, those who started using it made a wealth of money since then, but the real benefits of DBM remain to be discovered by the general public yet. The number of organizations utilizing DBM has been growing at an amazing rate lately. On the other hand a transition from simple accumulation of data to a serious analysis of this data just has started recently. We are going to see more and more stress put on the data analysis in the years to come. The technology to propel this process is denominated as data mining.

Some effects of a DBM can be summarized as follows[7]: it allows marketers to know more about various types of costumer and prospects, and to grade prospects by determining if a customer is willing to buy goods or services; it enables the organizations to reach customers with the right product and the right offer at the right time; it allows marketers to continually incorporate new information and results back into the database. So they can develop future strategies and executions from the collective results of previous efforts; DBM can be used by enterprising third-party marketers in order to bring together companies and customers; DBM can also develop special services to help easy customers to buy more, and offer the ability to establish a two-way communication with the customer through a variety of channels.

The DBM is a complex process that integrates the basic elements of strategic marketing, ie, targeting, targeting, positioning and communication [8]. We can use it in various ways, one being the analysis of information stored, in order to build detailed profiles of consumers, from a variety of characteristics. Appearing as a source of building customer loyalty, allowing the reduction of maintenance costs for the customer, enabling the company to build a sustainable competitive advantage in the area in which it operates.

According to Vavra [3], the objective of DBM is to create an electronic link between the organization and the consumer. A database of well designed can be very useful in implementing the activities of customer retention: enabling the marketer to identify the most important customers, to target promotions and offers of products and \ or services which are appropriate to your profile; allowing a clear monitoring of changes in terms of buying patterns; enabling marketers to interact in a more personalized way with customers; assisting in the establishment of programs to reward the most loyal customers, with the aim of strengthening their loyalty; being a good indicator for the development of new products from the characteristics of clients of the organization.

The DBM can be used in marketing activities [3]: the appropriateness of marketing activities to individual customers; in the segmentation of customers according to the frequency of purchase, recency, and purchasing preferences; in the service of customer support through the provision of the database (DB) employees responsible for care of these.

Hedgcock [9] considers the DBM as part of a system where marketing activities are interrelated, and the core client and there is a continuous interaction between the remaining elements of the system and customers, promoting the improvement of the service.

The greater the volume of data, the higher the quality of customer information, the greater the capacity of care, monitoring and recommending products \ services by the organization, the DBM is a resource that can be used strategically.

According to Seiler [10], there are several factors that promote the use of DBM by organizations to gain competitive advantage: a market orientation is the most important factor for the effectiveness of DBM, since it places the consumer as the centerpiece of the efforts of the organization. According to a survey by the National Center for Database Marketing, 86% of the differences in the levels of competitive advantage achieved through the use of database marketing are explained by the degree of business direction for the market; the placement of database marketing from a strategic perspective. Organizations prepared to make use of this tool should align the goals of the database with its strategic objectives and ensure that they are clear to the entire organization; the development of will and skill of cooperation between the departments and database marketing efforts; awareness that the use of external resources to build a marketing database does not remedy the lack of skills, capabilities and skills within the organization; the promotion of cultural and structural changes within the company to accommodate the database marketing; the encouragement of guidance for company information, which should enhance the manipulation of data in an effective manner and in the pursuit of results; the use of system data and allocating resources to programs for database marketing to gain competitive advantage. The well-established companies are able to extract information from data they have and turn them into marketing intelligence.

Shaw and Stone [11] suggest three methods of application of database marketing: creation of loyalty programs; creation of programs focused on prospecting for new customers; creation of an integrated approach to dealing with prospects and old customers.

The process of implementation and use of database marketing is dynamic. Thus, the company needs to continuously update and monitor the data to keep in line with the changing environment variables. If necessary the organization to assess their skills and resources, and take an orientation to the market.

3 Database Marketing and Competitive Advantage

Organizations seeking to develop their strategies in order to gain competitive advantage over their competitors. The concept of competitive advantage was treated first by Ansoff [12]. Competitive advantage can be understood as seeking unique opportunities that will give the company a strong competitive position. Based on this definition, several authors have proposed various devices for achieving a competitive advantage, since quality improvement to the construction of barriers to entry for new competitors in a given sector.

According to Porter [13], competitive advantage is achieved when a company can offer a superior value to customers by setting prices lower than their competitors, coupled with equivalent benefits, or the provision of benefits only at a price slightly higher than the average of competitors (differentiation). Efforts to search for an advantage are consolidated in the company's strategy.

To Lambin [14] the competitive advantage is the set of characteristics or attributes inherent in a product or brand, which gives it a certain superiority over direct competitors. These characteristics may refer to products, services that are aggregated, or the means of production, distribution or sale, own the product or the company. A multitude of factors can determine the superiority of the company relative to competitors priority. The value chain is a particularly useful mechanism to identify these factors, which can be organized into two categories based on the source of the advantage that they allow access: advantage foreign and domestic advantage.

The competitive advantage of foreign origin is based on the distinctive qualities of the product which is value for the buyer, producers and lower costs or improve performance within the company. In this case, the company may adopt a higher sale price than the competition. External competitive advantage gives the company a greater market power. The strategy based on competitive advantage foreign resembles the strategy of differentiation of Porter.

The international competitive advantage is sustained by a superior company in terms of matrix production costs, administration or distribution, providing a cost level lower than the leading competitor. The international competitive advantage is associated with increased productivity, allowing the company to achieve greater profitability and ability to resist a price drop imposed by the market or competition. A strategy based on international competitive advantage is similar to leadership in charge of Porter, and is supported by the technological know-how and organizational business.

Porter [13] believes that technology has a role in the structural change of a sector and the creation of new sectors, and may take the equalizer, neutralizer or balancer, annulling the competitive advantage of even healthy companies and leading others to take the lead.

Advanced information on the future state of the variables present in a scenario of a competitive environment have a high strategic value. The sooner a company can

anticipate situations favorable or unfavorable to a particular scenario, the sooner she can develop an appropriate strategy for action. Therefore, investments in competitive intelligence are always very timely [13].

For Schaars [15] there are two alternatives for obtaining competitive advantage. The first is to serve consumers better than competition. This advantage is based on the belief that the probability of repurchase is proportional to the satisfaction of consumers. The second is to be more agile than the competition, satisfying customers more quickly and creating barriers to the actions of competitors. The database marketing company can help in this task.

According to Albrecht and Bradford [16], the success of a company is closely linked to the sale of products that fit customers' needs. For the company to have access to these requirements, it is necessary to adopt a systematic approach to gain access to customers, so that the knowledge gained on them can lead to obtain a competitive advantage.

These considerations reinforce the idea of the importance of database marketing in the digital economy to obtain competitive advantage. Exploring e-commerce, Internet and database marketing for transporting to new contexts.

Turban [17] defines e-commerce from various perspectives: communications (e-Commerce is the delivery of information, products or services, or payment br mail); business process (it is the application of technology to automate business transactions and workflow); service (It is a tool that meets the desire of businesses and consumers to cut service costs while helping to improve the quality of goods and speed up service delivery); on-line (Electronic commerce provides the ability to buy and sell products and information on the Internet and other online services).

To Foo & Lim [18] the popularity of the Internet comes from a variety of factors: accessibility, ease of use, ability to obtain immediate global connectivity, quickly receiving and sending information, marketing opportunities and carrying out of business economies of scale and low entry barriers and equal opportunities it offers.

Peppers & Rogers [19] claimed that the global market and on-line has led to new attitudes of professionals and redefinition of the marketing activities. With the help of technology, nowadays there is the possibility to manage a database of consumers and competitors, promoting an interactive dialogue with the market, facilitating mass customization.

The emergence of electronic commerce has helped to see the creation of value from a new angle. The new and emerging business models are centered in intangible assets, such as relationships, knowledge and systems. For the company can gain competitive advantage, it is necessary that these resources are managed in some way, and database marketing figure as an indispensable tool in this task [20].

The advent of Internet and convergence with advanced database technologies allows organizations to remain competitive in the market and become more connected to consumers [21].

Evans and Wurster [22] analyzed the fundamental change in the information economy and its potential impact on strategy. Featuring two important features: reducing the capacity of organizations and individuals to monopolize control of information; and opportunities to improve access to information to reduce the time needed to implement a strategy.

Organizations continue to seek new ways to cut costs, increase productivity and improve customer service, technology becomes a facilitator rather efficient. The new technology will help not only to automate existing processes and improve resource utilization but also to redefine what work should be undertaken and how this should be done.

The growth of electronic commerce is changing the way organizations look at the data of their customers. The integration of databases has been an increasingly more important, the need to consolidate the company's image in the case online. Marketers belonging to the economy of traditional organizations still has problems to handle the complexities and new data are being incorporated with the wave of mergers and acquisitions and with the advent of the Internet [23].

In view of Cox [24], the increasing demand of databases in electronic commerce include the following requirements: unprecedented ability to handle a larger amount of data; the need to obtain a support increasingly sophisticated; the ability to work closely with software, other databases and legacy applications.

4 Electronic Commerce (e-Commerce)

E-commerce is to share information, maintain relationships and conduct business transactions via telecommunications networks inside and outside the enterprise. Thus, e-commerce is the way in which sales and purchases are generated, their achievements and activities of marketing-sales, all done online. The Internet is one tool of undeniable importance in electronic commerce, since it has a technology easy to understand, has publicly, and anyone, anywhere can access it just by a computer and a phone line. This facility started to advise companies on the sale to the final customer, without the need for investment. With the Internet, the person or company interested in buying or selling something, they can interact in finding information of interest, and need not be limited to what it initially appears. Following the opportunities opening up to mass customization, the Internet technology now enables the low cost, an increasingly more significant customization of services and care [25].

As observed by Albertin [26] e-commerce is the realization of the entire value chain of business processes in an electronic environment through the intensive application of communication technologies and information. The processes can be performed in whole or in part, including transactions business to business, business to consumer. Thus, electronic commerce covers any electronic business between two partners, or between a business and its customers.

According to Turban et al. [17] the Internet is promoting changes in the traditional physical market (marketplace), which is slowly being forced to coexist with the virtual market (marketspace). The latter is identified as an electronic marketplace, where goods are delivered directly to buyers to complete the purchase. The changes, which are reflected in greater efficiency of the marketing process.

The Internet is a valuable tool in marketing strategy because it allows a company will boost its presence and brand value in the market. In this context, the use of websites offers marketing managers the opportunity to communicate the mission of the company and brand, provide information about the attributes of products and services offered, as well as performance reports and future projects to a larger audience.

The database marketing organization is constantly updated with data from various sources, and the Internet plays a key role in this process. It is through this tool that marketing managers can monitor the behavior of customers, competitors and thus keep the information available within the company's up-date.

5 Conclusion

The database marketing not being a tool for the exclusive use of the area of marketing, is a strategic tool, which used properly can help effectively in achieving competitive advantage. The ideal scenario would be that all areas of the organization to work seamlessly, providing and using information and acting in a convergent way, promoting the improvement of organizational performance.

Large volumes of data does not necessarily mean useful information, since data alone may not be meaningful. The database is not the only marketing tool to support decision-making, updating the data is crucial and they must be integrated and converted into useful information when it is integrated, thus enabling support for strategic decision-making process. The role of database marketing has a strategic nature, to be able to approach their customers, the organization has the know, what is most evident in the digital marketplace, where there is a growing standardization of products, and that marketers have a notion increasingly clear that a good relationship with customers can lead to the emergence of competitive advantage and value creation.

The competitive advantages based on reliable information, provide the organization with knowledge of its customers. Technological advances and the emphasis given to help maintain customer relationships and customer loyalty, generating economic and financial benefits to the organization. The greater the knowledge of the organization to customers, the more likely that more efficient mechanisms for their loyalty and consequently to improve their performance.

The Internet is still a very recent phenomenon with implications for various segments of society. However, it is clear that the World Wide Web represents an important and integral tool of the marketing strategies of companies.

Marketing on the Internet, possibly, in some aspects, complement, can overcome traditional means, or replace the intermediaries more efficiently. In this context, some traditional distributors recognize the need to adjust their strategies to include online transactions in their operations. Internet marketing, as well as complement, may supplant traditional means, or replace the intermediaries more effectively. In this context, some traditional distributors recognize the need to adjust their strategies to include online transactions in their operations.

In the context of database marketing the Internet is essential as a source of primary and secondary data. The database marketing is a strategic tool that if used correctly, can assist the company in promoting and achieving strategies, since it provides relevant information about the market and customers.

The database marketing is not the only tool to support decision making, in this way, you should always update the data, integrate them and convert them into meaningful information to support decision making in operational and strategic levels.

References

1. Gummesson, E.: Total Relationship Marketing, Rethinking Marketing Management: From 4Ps to 30 Rs. Butterworth Heinemann, Oxford (1999)
2. Mckenna, R.: Marketing de Relacionamento. Ed. Campus, RJ (1995)
3. Vavra, T.G.: Marketing de Relacionamento (Aftermarketing). Editora Atlas, SP (1993)
4. Tucker, R., Brown, N.: Effective database marketing. Incentive 23(21) (December 1994)
5. Christopher, M., Payne, A., Ballantyne, D.: Relationship Marketing. Butterworth-Heinemann, London (1991)
6. Ananyan, S., Kiselev, M.: Marketing competence:: Contributions - DBM (2004), http://www.helmar.org/index.php?id=1&show=content&content_id=244
7. Batra, R.: The new direct marketing: How to implement a profit-driven DBM strategy. Irwin, Burr Ridge (1995)
8. Davies, M.J.: The Essential Guide to Database Marketing. McGraw-Hill Book Company, London (1992)
9. Hedgcock, R.: Automating the marketing environment & transforming data into insight: the datawarehouse & the database marketing. Direct Marketing, 56–58 (February 1998)
10. Seiler, M.: Organization key to database marketing. Marketing News, 18 (February 14, 2000)
11. Shaw, R., Stone, M.: Marketing com banco de dados: database marketing - entendendo e implantando o marketing com banco de dados em qualquer tipo de empresa. Editora Atlas, SP (1993)
12. Ansoff, I.: Corporate Strategy. McGraw-Hill, NY (1965)
13. Porter, M.E.: Cases in Competitive Strategy. Free Press, N.Y (1982)
14. Lambin, J.J.: Marketing Estratégico, 4ª edição. McGraw-Hill, Lisboa (2000)
15. Schaars, S.P.: Marketing Strategy: a customer-driven approach. The Free Press, NY (1991)
16. Albrecht, K.: Lawrence Bradford Service Advantage: How to Identify and Fulfill Customer Needs. McGraw-Hill (1989)
17. Turban, E., King, D., Lee, J.K., Viehland, D.: Electronic Commerce: A Managerial Perspective. Prentice Hall (2006)
18. Foos, S., Lim, E.: A hypermedia database to manage world-wide-web documents. Information & Management 31, 235–249 (1997)
19. Peppers, D., Rogers, M.: Enterprise One to One: tools for competing in the interactive age. Bantam Doubleday Dell Publishing Group Inc., NY (1997)
20. Boulton, R.E.S., Libert, B.D., Samek, S.M.: A business model for the new economy. Journal of Business Strategy, 29–35 (2000)
21. McKenna, R.: Leadership in the digital age Executive Excellence, p. 6 (2000)
22. Evans, P.B., Wurster, T.S.: Strategy and the New Economics of Information. Harvard Business Review, 70–82 (1997)
23. King, D.G.: Merging on to the information superhighway. Target Marketing, 73–74 (2000)
24. Cox, J.: E-commerce changing the face of databases. Network World 31, 38 (2000)
25. Chleba, M.: Marketing digital, novas tecnologias e novos modelos de negócio. Futura, São Paulo (1999)
26. Albertin, L.A.: Comércio electrónico: benefícios e aspectos de sua aplicação. ERA 38(1), São Paulo (1998)

The Impact of Credit Derivatives Issuance on Financial Institutions*

Yang Xing[1] and Zhong Yuqin[2]

[1] Institute of Finance, Department of Finance, School of Economics, Jinan University, 510632
[2] Finance Department, School of Economics, Jinan University, 510632

Abstract. This paper analyzes the impact of European Credit Derivatives issuance on the financial market stability. Based on the relationship between CDOs (Collateralized Debt Obligation) issuance volumes, market volatility, Europe 6-month risk-free interest rate and the stability of financial market, this paper uses coexceedances method and the Multiple Logit model. We find that CDOs issuance is positively related to the negative coexceedances, suggesting that the greater amount of CDOs issuance, the greater impact on financial stability. CDOs issuance is not related to the positive coexceedances, indicating that the risk management of banks through CDOs is limited. In addition, the asymmetric information and risk-free interest rate also affect the stability of financial market.

Keywords: Credit derivatives, coexceedances, Collateralized Debt Obligation.

1 Introduction and Background

When the U.S. subprime mortgage crisis happened in 2007, credit derivatives have been considered as a major cause for the crisis and questioned by many people. They consider credit risk spread rapidly and transfer among financial institutions, which lead to global financial crisis.

Commercial banks can transfer the credit risk to investors without transferring the ownership of credit assets through credit derivatives. In addition, it increases the depth and liquidity of credit markets, and accelerates the transferring of credit risk of the banking system to the capital market. Therefore, credit derivatives not only promote the financial institutions' efficiency, but also cause great instability of it.

Private equity companies, hedge funds, insurance companies and other new financial institutions can invest innovative products such as asset securitization and other derivatives. This can transfer credit risk and reduces the difficulty and cost of issuing bonds, which achieve risk transferring in different institutions and markets.

There is always controversial on whether credit derivatives enhance or undermine the financial stability. At first, scholars such as Duffie (2001) consider that financial

* The National Social Science Fund of China. "Study of counterparty default, liquidity risk impact and valuation of credit default swaps in international credit derivatives market."(11BGJ013).

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 131–139.
springerlink.com

innovation has a positive impact on financial stability. Through asset securitization and other financial innovations, banks transfer their long-term credit risk to many investors successfully and improve liquidity and promote bank system stability. Wagner and Marsh (2004) compare the cost of assets risk held by institutions and the social cost of risk assets, and illustrate that bank and non-bank financial institutions have sufficient incentives to transfer credit risk, which enhance financial stability. Jing Zhongxia and Zhang Tiantian (2007) point out credit derivatives allow market players to buy and sell credit risk protection more conveniently, and banks' operation and risk management have been greatly improved, which increase the pricing efficiency, market transparency and the financial system flexibility.

However, Instefjord (2005) and other scholars suggest that credit derivatives have a negative impact on financial stability. They believe that financial innovation tools allow banks tend to take risk, thereby increasing the risk level of banks and the volatility of banking system. Duffee and Zhou (2001) emphasized the internal financial system changes the traders' expectations about information asymmetry by credit risk derivative transactions, which leads to the collapse of loan market. Morrison (2005) constructed a model of corporate finance found credit derivatives has weakened the bank's motivation and supervision activities on borrowing ventures. Baur and Joossens (2005) suggest that securitization reduces the capital requirements of those who have transferred risk, if the risk-takers out of regulatory constraints. Allen and Carletti (2005) illustrate that credit derivatives not only diversify the spread of risk, but also increase the likelihood of crises contagion.

There are also many studies on the dual effect of credit derivatives on the financial stability. Static analysis of Santomer and Trester (1998) conclude that the reduction of asymmetric information makes the banks more likely to diversify the credit risk of derivatives transactions, and enhance the liquidity of bank assets, transfer risk to the non-bank financial institutions, which improve financial stability. Instefjord (2005) consider that the amount of credit derivatives is prone to sharply fluctuate and not conducive to financial stability.

In summary, from the views of credit risk transfer, improve cash flow, net assets yield rate, and the comparison of cost of risk assets in institutions and its social costs, scholars often conclude that credit derivatives improves the financial stability. From the views infectious among financial system, information asymmetry and change of liquidity requirement and price elasticity, absence of regulation about credit derivatives, scholars always suggest that credit derivatives undermines the financial stability.

Unlike other studies of credit derivatives and financial market stability, we consider a more specific angle: from a microscopic point of view, by calculating coexceedances (coex) and establishing Multiple Logit model, we analysis the coexceedances linkage between credit derivatives – CDOs' issue volume and banks (insurance companies) return. Then we discuss the impact of credit derivatives on financial stability.

This paper is arranged as follows: Section 2 describes the mechanism of CDOs' impact on financial market stability, and the measure of stability of financial market and the hypothesis; Section 3 describes the methods, models and estimation; Section 4 is the empirical analysis; Section 5 is the robust test; Section 6 is the conclusion.

2 CDOs Market Mechanism and Measure of Its Impact on Financial Institutions

CDOs reference entity is a portfolio of assets. In this portfolio, the assets include high credit rating (priority sector), intermediate credit assets (second priority sector) and poor credit rating plate (capital blocks). Among issued securities, kinds of securities can make credit rating higher than all the pool assets to enhance average credit rating. Because CDOs increase the second priority assets rating and issues higher rating securities, it causes financial risk.

Particularly, CDOs are not issued directly by the promoters, but by the Special Purpose Vehicle (SPV) who issues and fully assumes responsibility for CDOs debt service. Because the isolation of SPV and the sponsor risk, credit chain is lengthened, so it is more difficult to estimates the risk.

Different from the financial stability effect on the banking system or stock market caused by traditional financial products, CDOs are related to three financial sub-markets: credit market, bond market and stock market. Before credit risk transferred by CDOs, these three markets are separated which can avoid credit risk transmission. Credit risk is primarily borne by the banking system. Credit derivatives allows banks transfers risks to other financial institutions (e.g. insurance companies and hedge funds) which form a chain of risk between institutions; and a large number of asset securitization issued and traded in the stock market, transfer credit risk from credit markets to the securities market, formed a cross-market risks, so that the stability of the entire financial system may be affected.

Financial instability is usually assumed as the volatility of asset prices. Banks run on capital markets and its liquidity shortage caused by the accumulation and spread of risk. If the CDOs transaction exacerbated the volatility of financial assets equilibrium price, it destroys financial stability. On the contrary, if the CDOs transaction can stabilize the equilibrium price of assets fluctuations and make the financial markets more perfect, we say that it improved the stability of the financial system.

This study focuses on how CDOs affect the volatility of financial asset prices and then affect the stability of financial markets. We use the fluctuations of financial institutions' return as the dependent variables to measure financial stability. For the CDOs market, we choose the following measure as independent variables:

(1) We use European market monthly CDOs issuance as a measure of the credit derivatives market impact. Because CDOs have both positive and negative effects on financial institutions, we assume that H_1: the larger CDOs issuance, the higher probability of negative and positive coexceedances.

(2) We also use London Financial Times Index (FTSE100) condition monthly volatility as proxy variable of information asymmetry and heterogeneity calculated by GARCH (1, 1) model. King and Wandhwani (1990) show that when market volatility increases, coefficients between industries also increase. According to Bailie and DeGennaro (1990), who use GARCH-M model and the ARCH model of stock portfolio and conclude that return and volatility. It was assumed that H_2: the greater information asymmetry, the higher probability of negative coexceedances values and lower probability of positive coexceedances.

(3) The impact of short-term liquidity indicators of stock market includes short-term interest rates, M2, etc. Dynamic changes in interest rates can affect the coexceedances. In this paper, the European 6-month risk-free interest rate (Euribor) is regarded as proxy variable of shock on short-term market liquidity. Assuming H_3: the greater the liquidity shock, the higher probability of negative coexceedances and the lower probability of positive coexceedances.

3 Methods and Data

3.1 Methods

In this paper, we use "coexceedances" which has been introduced by Bae, Karolyi and Stulz (2003) to investigate the relationship between CDOs issuance and financial stability, which measures the linkage relationship of assets return which excess the given threshold. It is an extreme case rather than the linkage relationship under normal circumstances, which reflects the systemic risk better than the correlation. If the coefficient reflects the binary (or multiple) relations at a time point, the coexceedances measure relationship of N assets returns in a period of time.

The technical core of coexceedances is the threshold, which is defined as the joint occurrence of two exceedances (i.e. large absolute returns above a certain threshold) of two financial intermediaries' returns (Dirk Baur and Niels Schulze, 2003). Bae et.al.(2003) use the five-percent tails of the overall return distributions as a threshold that defines an exceedance. Positive and negative returns are treated separately.

The financial institutions asset returns are no more than 5% quantile (under the tail) of monthly average asset return as the negative values during the sample period. Financial institutions asset returns are no less than 95% the quantile (upper the tail) of monthly average asset return as the positive values during the sample period. The monthly asset returns less than 5% or over 95% means financial instability. This paper chooses 10 financial institutions monthly asset return, so the sum of positive and negative coexceedances is non-negative integers between 0 to 10, the greater coexceedances indicating more financial institutions returns are over the threshold, and larger financial systemic risk.

3.2 Data

Coexceedances data stems from the stock market monthly return of 10 financial institutions (including eight major banks, two insurance companies[1] in Europe) from March 2005 to March 2010. First, we calculate the monthly returns of financial institutions and the thresholds (below 5% tail and over 95% tail), the number of institutions which exceed the threshold (no more than 5% under the tail, no less than 95% upper the tail) as the current coexceedances, as the dependent variable in empirical analysis. Data stems from www.google.com / finance.

[1] They are Germany Allianz Insurance Group, France AXA Insurance Group, BBVA Bilbao Vizcaya Bank, BNP Paribas, Germany Dresdner Bank, Deutsche Bank, HSBC Bank, Intesa San Paolo Bank of Italy ,Societe Generale Bank, Germany Volksbank Bank.

European CDOs quarterly issuance data stems from European Securities Association (ESF (2010)), from March 2005 to March 2010. Because there are only quarterly issuance data, monthly data cannot be obtained. Therefore we calculate average quarterly data instead of monthly data which stems from http://www.europeansecuritisation.com/.

Market volatility of the London Financial Times Index (FTSE100) is the monthly return from March 2005 to March 2010. FTSE100 covers 9 major European countries, which is the most important indicator of European stock market by global investors. So it can be selected as the European stock market volatility indicator. Data stems from http://uk.finance.yahoo.com/.

European Interbank Offered Rate (Euribor) data stems from http://www.euribor-rates.eu/ from March 2005 to March 2010.

3.3 Model and Estimation

We use coexceedances of financial institutions' returns as object, establish Multiple Logit model. The regression equations are given as follows:

$$\Pr(coex_{it}=1 \mid \ln(CDO_t), \ln(vol_t), \ln(Euribor_t)) = \frac{1}{1+e^{-(\beta_0+\beta_{1i}\ln(CDO_t)+\beta_{2i}\ln(vol_t)+\beta_{3i}\ln(Euribor_t))}}$$

$coex_{it}$ are coexceedances which calculated by returns of financial institutions which is less than 5% or greater than 95% of the threshold in time t.

$i \in [0,10]$, $i \geq 2,3,4$. So, coexceedances have been divided into 3 groups.

The first group: $coex_{it} = \begin{cases} 1, i \geq 2 \\ 0, i < 2 \end{cases}$, shows that two or more banks (insurance companies) have reached to the threshold.

The second group: $coex_{it} = \begin{cases} 1, i \geq 3 \\ 0, i < 3 \end{cases}$, shows that three or more banks (insurance companies) have reached to the threshold.

The third group: $coex_{it} = \begin{cases} 1, i \geq 4 \\ 0, i < 4 \end{cases}$, shows that four or more banks (insurance companies) have reached to the threshold.

The larger i indicated that the greater the probability of banks and insurance companies encounter financial crisis,.

$\ln(CDO)$ $\ln(vol)$ $\ln(Euribor)$ represent the natural logarithm of CDOs issuance, FTSE100 stock index volatility and European inter-bank lending rate respectively.

$\beta_{ki}\ (k = 0,1,2,3, i = 2,3,4)$ are constants and coefficients. $k = 0,1,2,3$ represent the three independent variables of the regression respectively. While $i = 2,3,4$ represent the numbers of banks (insurance companies) have reached to the threshold respectively. The data and regression are processed by Stata11.1, and the result shows in Table 1.

Table 1. Constant and coefficients of the Multipl Logit regression

Independent variables	Coefficients of negative coexceedances				Coefficients of positive coexceedances			
	(1)	(2)	(3)	(4)	(5)	6)	(7)	(8)
$\beta_{02}[Const]$	−2.01[b]	−2.73[a]	−12.30[b]	−11.52[b]	−3.12[b]	−2.76[a]	0.61	3.24
β_{03}	−5.04[b]	−3.66[a]	−8.99[c]	−15.88[c]	−3.36[b]	−2.80[a]	−0.09	1.92
β_{04}	−12.94[b]	−4.51[a]	−5.41	−117.1	−6.96[c]	−3.84[a]	1.26	−2.84
$\beta_{12}[\ln(CDO)]$	−0.04			0.16	0.71			0.24
β_{13}	1.62			2.29	0.75			0.42
β_{14}	5.35[b]			27.25	2.12			2.04
$\beta_{22}[\ln(vol)]$		0.57[c]		0.46		0.60[c]		0.74[c]
β_{23}		0.90[b]		0.81[a]		0.50		0.56
β_{24}		1.02[b]		2.41		0.39		0.27
$\beta_{32}[\ln(Euribor)]$			2.89[b]	2.48[c]			−2.19[b]	−2.13[b]
β_{33}			1.88	2.46			−1.86[b]	−1.75[b]
β_{34}			0.71	15.53			−1.50	−1.39
Log-likelihood	-81.84	-66.0	-76.38	-58.70	-88.55	-88.06	-77.02	-75.90
Pseudo-R2	3.39%	22.09%	9.83%	30.71%	1.13%	1.67%	14%	15.2%

Notice: [a]、[b]、[c] indicate significance at the 1%,5% and 10% level respectively.

4 Empirical Analysis

4.1 Empirical Analysis of the Negative Coexceedances (5% tail)

Columns (1) - (3) in Table 1 are coefficients and significances of three single-factor models, column (4) for the coefficients and their significances of total-factor models.

From single-factor model perspective: CDOs issuance coefficient for the model of the first and the second are not significant, but in the third group, namely $coex_{it} = \begin{cases} 1, i \geq 4 \\ 0, i < 4 \end{cases}$, (hereinafter the same), coefficient is significant at the 5% level, and the coefficients of three model increase as coexceedances, which shows that with the increase of the CDOs issuance, the greater negative coexceedances, the greater the financial systemic risk, which is consistent with hypothesis H_1.

Coefficients of market volatility are significant at the 10%, 5% and 1% level in these three regression model. Significant level and the coefficient increase as the coexceedances increase. It is also significant for the second model in the total-factor models. That means the greater degree of information asymmetry, the greater probability of negative coexceedances.

Coeffecients of risk-free interest rate are only significant in the first group models in both the single and total-factor models. But it is reducing when the coexceedances increase.

We can see coefficients of total-factor models from the column (4): CDOs issuance in the three models are not significant, while the market volatility factors are significant in the second model, risk-free interest rate coefficients are only significant in the first

group models. However, all three variables coefficients increase as the coexceedances increase. The greater coexceedances mean greater probability that the financial intermediation return less than 5% quantile and greater probability of banks (or insurance companies) encounter financial instability.

4.2 Empirical Analysis of the Positive Coexceedances (95% tail)

Columns (5) - (7) in the Table 3 are single-factor models, (8) column are the total-factor models for the positive coexceedances. Not only in the single-factor models but also in the total-factor models, coefficients of CDOs issuance in the three groups models are not significant for the coexceedances,which is not consistent with hypothesis H_1 (explain in section 6).

Coefficients of market volatility are only significant in the first group models for both the single-factor and total-factor models. And the changes of coefficients indicate that smaller probability of coexceedances because of asymmetric information, which is consistent with hypothesis H_2.

Risk-free interest rate coefficients are significant in the first and the second models both in the single-factor and total-factor models. But they are negative correlative at the 5% level.

The greater impact of risk-free interest rate, the smaller probability of financial institutions returns exceed the quantile of 95%, which is consistent with hypothesis H_3. Greater liquidity shock means greater negative coexceedances linkages and the smaller positive coexceedances.

5 Robust Test

To further estimate the robustness of empirical results, we perform extend tests: take the return threshold 2.5% as the lower quantile, and 97.5% as the upper quantile; and then 10% as the lower quantile, and 90% as the upper percentile. Using these new thresholds, we perform robust test to estimate the significance of variable coefficients, and the results shown in Table 2 and Table 3.

Columns (1) - (8) in the Table 2 are negative coexceedances coefficients, which take the first four columns as the results of 2.5% quantile threshold. Columns (5) - (8) are the results of 10% quantile threshold.

Columns (1) - (16) in the Table 3 as positive coexceedances coefficients, in which columns (1) - (4) for the 97.5% quantile threshold, while columns (5) - (8) are the results of 90% quantile threshold.

In the Table 2 the results show that in the level of 2.5% quantile, the total-factor models of the Pseudo-R^2 is 46%.It is larger when compared with 30.71% of the Pseudo-R^2 in 5% quantile, indicating the model's explanatory power is stronger.The Pseudo-R^2 is 11.5% in the 10% percentile which is the minimum, indicating with the lower threshold tail, the more explanatory power of the models. And Pseudo-R^2 of CDOs issuance and market volatility increased at 2.5% level, indicating that these two explanatory variables impact to financial instability is increasing.

In Table 3, CDOs issuance and the coefficients of market volatility have increased a little at 97.5% and 90% quantile. Especially in total-factor models, they are still

insignificant. Other results are the same as results of 95% quantile. Robustness test results indicate that different quantile selection for the impact of explanatory variables was not significant.

Table 2. Robust test of negative coexceedances

Independent variables	2.5%				10%			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\beta_0(cons\tan t)$	-3.66[a]	-3.37[a]	-5.28[b]	-8.68[a]	0.30	-0.40	-3.10[a]	-3.80
$\beta_1(\ln(CDO_t))$	1.55[a]			1.54[b]	-0.13			-0.10
$\beta_2(\ln(vol_t))$		1.23[a]		0.99[a]		0.44[a]		0.39[a]
$\beta_3(\ln(Euribor_t))$			1.20[b]	0.95			0.94[a]	0.72[a]
$Log-likelihood$	-53.28	-35.37	-55.99	-31.45	-129.58	-119.77	-123.51	-114.93
$Pseudo-R^2$	9.26%	40%	4.6%	46%	0.24%	7.79%	4.91%	11.5%

Notice: [a]、[b]、[c] indicate significance at the 1%,5% and 10% level respectively.

Table 3. Robust test of positive coexceedances

Independent variables	97.5%				90%			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\beta_0(cons\tan t)$	-2.33[a]	-1.24	3.67[a]	2.35	-0.17	-0.12	3.50[a]	3.12[a]
$\beta_1(\ln(CDO_t))$	0.79[c]			0.69	0.21			-0.003
$\beta_2(\ln(vol_t))$		0.11[a]		-0.02		0.23[b]		0.24[b]
$\beta_3(\ln(Euribor_t))$			-1.56[a]	-1.47[a]			-1.07[a]	-1.03[a]
$Log-likelihood$	-58.48	-60.24	-52.86	-51.72	-114.94	-113.01	-102.98	-100.87
$Pseudo-R^2$	3.19%	0.27%	12.48%	14.38%	0.55%	2.21%	10.90%	12.7%

Notice: [a]、[b]、[c] indicate significance at the 1%,5% and 10% level respectively.

6 Conclusions

1. CDOs issuance is correlated with the probability of negative coexceedances significantly, which consistent with the H_1 hypothesis. Because credit derivatives issued by banks has much risk and may spread to credit markets, resulting credit expansion effect and causing negative impact to the stability of banks. These risks include: operational risk of using derivatives leading to excess speculative; credit risk caused by unique assets structure and duration mismatches arising from credit derivatives; transaction counterparty risk caused by either credit risk protection buyer or seller' default; liquidity risk caused by credit derivatives which cannot be traded in time and conveniently. CDOs issuance has insignificant impact to the probability of positive coexceedances. Because multinational banks in Europe developed countries always have highly modern corporate governance structure, customer demand-oriented, high efficient management and abundant diversity financial products, which ensure banks can prevent kinds of financial risks.

2. Volatility is a measure of uncertain return, which represents the assets at risk. In the case of lack of information, increased market volatility causes greater market risk, which causes individual investors to believe that other industries will face the same problem. Their investment behavior change which will lead to greater negative coexceedances, and also lead to greater positive coexceedances, but not significantly. Generally speaking, volatility increases the financial system risk.
3. The coefficients of risk-free interest rate show that the greater liquidity shock, the greater probability of negative coexceedances and less probability of positive coexceedances. It can be analyzed by the theory of liquidity constraints: when the impact of short-term interest rates (raise the interest rate) cause investors increase the cost of funds, arbitrage traders will suffer losses in industry transaction. Short selling, lending restrictions and wealth constraints will lead to shrinkage of wealth and reduction of ability to resist risks, which cause investors liquidate their assets, and aggravation of asset price volatility and declination of stock market. The market turmoil will increase systemic risk of financial markets and lead to further instability in the financial markets.

References

1. Alan, D.M.: Credit Derivatives. Disintermediation and Investment Decisions Journal of Business 2, 621–648 (2005)
2. Allen, F., Elena, C.: Credit risk transfer and contagion. Journal of Monetary Economics 53, 89–111 (2006)
3. Anthony, M.S., Trester, J.J.: Financial Innovation and Bank Risk Taking. Journal of Economic Behavior&Organization 1, 25–37 (1998)
4. Bae, K., Andrew, G., Rene, K., Stul, M.: A new Approach to Measuring Financial Contagion. Review of Financial Studies 3, 717–763 (2003)
5. Baur, D., Joossens, E.: The Effect of Credit Risk Transfer on Financial Stability. EUR Working Paper No. 21521 (2005)
6. Gregory, R.D., Zhou, C.: Credit derivatives in banking: Useful tools for managing risk? Journal of Monetary Economics 1, 25–54 (2001)
7. Risk, I.N., Hedging: Do Credit Derivatives Increase Bank Risk? Journal of Banking & Finance 2, 333–345 (2005)
8. Jin, Z., Zhang, T.: The significance of Credit derivatives market development to financial stability. China Finance 8, 15 (2007)
9. King, W.A., Wadhwani, S.: Transmission of volatility between stock markets. Review of Financial Studies 31, 5–33 (1990)
10. Richard, T.B., Ramon, P.D.: Stock Returns and Volatility. Journal of Financial and Quantitative Analysis 25, 203–214 (1990)
11. Wolf, W., Ian, W.M.: Credit Risk Transfer and Financial Sector Performance. CEPR Working Paper No. 4265 (2004)
12. Wang, Q.: Application of STATA in statistical and quantitative analysis. Nankai University Press (2007)
13. Wu, H.: Credit risk transformation of credit derivatives on financial stability. Economic Survey 6, 149–152 (2008)
14. Xu, R.: New Development of financial stability-The Theory of credit risk transfer review. Teaching and Research 7(7), 84–89 (2008)
15. Yang: China. Empirical Study of securities market industry coexceedances. Zhongnan University of Economics and Law 4, 87–92 (2009)

Empirical Analysis on Green Innovation of Chinese Paper Manufacturing*

Haiying Liu and Zhiyong Zou

School of Business
Shandong Polytechnic of University, Jinan Shandong Province, China
hjqlhyhb@163.com, zouzhiyong@live.cn

Abstract. In this paper, Moran index of spatial statistics, the spatial lag and spatial error model of spatial econometric economics, combined with Griliches &Jaffe knowledge production function(Knowledge Production Function, KPF), were used in spatial econometric analysis on China's provincial paper manufacturing R&D and innovation knowledge spillovers. The results show that there are significant global and local spatial autocorrelation in paper industry green innovation model; also find that China's 31 provincial contribution of green innovation in paper manufacturing sector was mainly achieved by enterprises R&D investment, while the combination of university and enterprise did not show significant effect on green innovation spillovers. The results show effect of knowledge spillovers are exist, but not significant. In the end, according to the conclusions, we draw some advices.

Keywords: Spatial econometric analysis, knowledge production function, green innovation spillovers, spatial lag model, spatial error model.

1 Introduction

The spatial characteristics of enterprise technological innovation behavior are an important field to be researched. The relevant foreign researches show that regional enterprises were easy to use adjacent academic institutions or business' knowledge and technical resources to make faster innovation. Thus the essentiality of regional enterprise innovation is coupling-based interaction innovation because its formation is the result of interactive learning, cooperation and competition among cluster actors. Actors' proximity may be decomposed into geographical, social and Indus- trial aspects in network for green innovation spillovers within regional enterprise cluster. Namely there exist interaction and radiation between green innovations of enterprises with the similar geographical position. Keller (2002) and Fu Sen (2009) proved the spatial correlation of R & D activities, and drew the conclusion that technology spillovers increase gradually weakened with geographic increasing distance. At the same time, geographical economists confirmed that spatial connection of economic things was not

* Shandong Social Science Research Projects "Paper Industry in Shandong Province of green technology innovation research (10CJGJ64)", Shandong Paper Industry Association commissioned the project "Technical Paper Industry in Shandong Province Economic Research".

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 141–147.
springerlink.com

only geographic proximity, more importantly was organization proximity (Based on a class of common or similar neighboring base). They believed that regional economic development, the quality of people's cultural, social environment, even customs, and many other forms of geographic adjacent factors in special circumstances will influence regional innovation. Such as, mainly due to areas with similar cultural backgrounds, enterprise are more easily achieved the dissemination and exchange of tacit knowledge, and the areas with similar economic level are better able to absorb and use of innovative resources. So what kind of relationship that exists between the geographical proximity of production and innovation performances? From the perspective of knowledge spillover, many researchers answered this question. They believe that the majority of innovation activities knowledge are implicit, and its marginal cost of transmission is increasing with increasing distance, while mostly of which can only be delivered through face to face contact. The geographic proximity promotes the overflow of green innovation knowledge between regional industrial enterprises within the same and different industries. This article will use the improved knowledge production function (KPF) to analyses the spatial structure of the green technical innovation of China's paper industry, to test the existence of spillover effects of green technological innovation of China's paper industry.

2 Model and Variables

2.1 Spatial Econometric Model

Scholars generally use Moran I, two Lagrange Multipliers (LMERR and LMLAG), R-LMERR, and R-LMLAG to determine whether there is spatial correlation of regional green innovation output or not, to measure Spatial Lag Model (SLM) and Spatial Error Model (SEM) which space model is more in line with objective reality. Anselin and Florax(1995) proposed the following criterion: If the LMLAG was more statistically significant than LMERR, at the same time R-LMLAG was significant and R-LMERR was not significant, it can be concluded that Spatial Lag Model is more suitable. On the contrary, if LMERR was more statistically significant than LMLAG, while R-LMERR was significant and R-LMLAG was not significant, you can conclude that spatial error model is the appropriate model. We also can use the natural log-likelihood function value (Log likelihood, LogL), likelihood ratio (Likelihood Ratio, LR), Akaike information criterion (AIC), Schwartz criterion (SC) to discriminate. Log-likelihood value of the greater, AIC and SC values of the smaller, the better the model fit.

2.2 Knowledge Production Function and Variables

Griliches (1979) first used knowledge production function (KPF) to estimate the impact of different elements of R & D output. Large numbers of empirical studies have shown that Jaffe knowledge production function expansion is a good statistical model on measure of knowledge spillovers and innovation performance.

Jaffe function model as:

$$\log P_{ikt} = \beta_{1k} \log(R_{ikt}) + \beta_{2k} \log(U_{ikt}) + \beta_{3k} [\log(U_{ikt}) \log(C_{ikt})] + \varepsilon_{ikt} \quad (1)$$

$$\begin{aligned} \log U_{ikt} &= \beta_{4k} \log(I_{ikt}) + \delta_{2k} Z_1 + \xi_{ikt} \\ \log I_{ikt} &= \beta_{5k} \log(U_{ikt}) + \delta_{2k} Z_2 + \mu_{ikt} \end{aligned} \tag{2}$$

Where: i – the unit of observation, k – the green innovation technology, t– time; P– the number of patents the company(the economic use of new knowledge); I – the company's R & D costs, U –the investment in higher education research, ε_{ikt} –random error term; C – the region's higher education institutions and R & D activities, blending geography index.

According the real situation of fund put in green innovation of home paper industry. In this paper, we use modified Jaffe knowledge production function model, and taking accounting into the introduction of the space dimension and university research, to measure green innovation spillovers of Chinese paper manufacturing.

Empirical models are as follows:

$$\begin{aligned} \ln NPS_{tj} = \ln A_{tj+} \beta_1 \ln ERD_{tj} + \beta_2 \ln HK_{tj} + \beta_3 \ln EURD_{tj} + \beta_4 \ln EGRD_{tj} \\ + \beta_5 \ln TRAD_{tj} + \beta_6 \ln FSN_{tj} + \varepsilon_{tj} \end{aligned} \tag{3}$$

Where: lnA – constant variable, t – time, j – on behalf of areas; NPS – paper products sales of green innovation (NPS) to be dependent variable; ERD– the paper companies green innovation funding, HK –the number of green innovation personnel, EURD –industry-academia cooperation green R & D variable, EGRD–government-industry cooperation in green R & D variable, TRADE – the introduction of innovative technology companies green fees variable, FSN – absorptive capacity of green innovation and enterprise (measured by average number of enterprises), which were used as explanatory variables in green innovation of China's paper industry spillover spatial econometric analysis model.

2.3 Sample Data

As time was needed for innovation from input to output, then this paper used two-year lag period to quantitative analysis. This paper uses two-year lag for empirical analysis. Taking account into the effect of geographical space, this paper use spatial econometrics spatial lag model or spatial error model to empirical analysis. Population for explained variables is selected from 2005 to 2009 , while the explaining variables are selected from 2003 to 2007 average data. All the sample data comes mainly from the 2001-2010 annual "Yearbook of China's paper industry" and "China Science and Technology Statistical Yearbook." Yearbook of the two parts of Tibet on the missing data, in order to ensure the integrity of the sample space and the estimated requirements for large sample, with reference to Jaffe's practice in the regression estimates of missing data values to -1 as an alternative.

3 Empirical Study

3.1 Spatial Statistical Test of Green Innovation Output

Moran's I was used to estimate the local spatial correlation of green innovation spillovers of Chinese paper manufacturing. Spatial autocorrelation coefficient Moran I

was 0.3567, while its normal statistic Z significantly. Thus there is a significant spatial autocorrelation (spatial dependence) in the spatial distribution of green innovation in China's provincial paper manufacturing. In the other word, the spatial distribution of the green innovation behavior of provincial paper industry was not completely random state, but showed similar values between the space clusters. This means paper enterprise green technological innovation has contact structure with space, area with higher green technological innovation close to each other in space. In all, there is spatial correlation in the province's green paper manufacturing's output of green technological innovation, meaning that there is a clear spatial clustering (Clustering) phenomenon. For example, relying on geographic area, the industry-university collaboration of paper enterprise and Shandong Institute of Light Industry has made an outstanding contribution to green technical innovation for the enterprise in Shandong province.

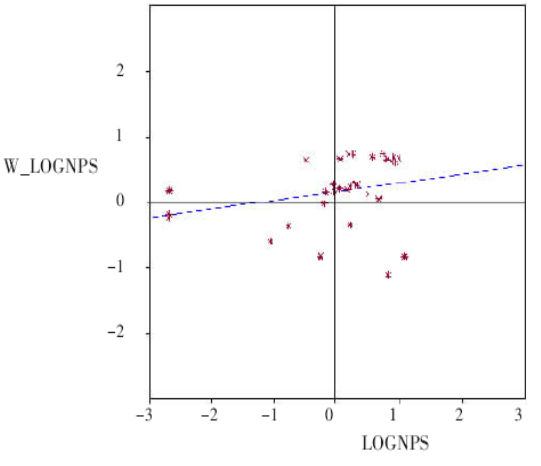


Fig. 1. Moran's I scatter of paper industry green innovation clusters

Local Moran's I scatter plot shows the spatial lag distribution of paper industry green technological innovation (Figure1. x-axis is W_LOGNPS, y-axis is LOGNPS). W_LOGNPS said the weighted average of neighboring values. As shown in figure1, the area as Shandong, Jiangsu, Zhejiang, Hunan, Guangdong and Anhui, all in the first quadrant, have stronger creativity capacity, and show positive correlation between self and clusters (HH). While in Heilongjiang, Fujian, Hubei, Liaoning, Hebei, Tianjin, Inner Mongolia, Xinjiang, Qinghai, and Ningxia province, are all in the third quadrant, with a relatively weaker green innovation capability, have a positive self-correlation and cluster (LL). Sichuan, Shaanxi, Henan province show a negative relationship between spatial autocorrelation. Provincial paper local green innovation and cluster differentiation of HH and LL, to some extent showed a cluster of green innovation the core of the paper industry - the edge spatial distribution pattern, namely provincial green innovation in geographical space exists on the distribution of dependence and heterogeneity. The performance of local clusters of high value coastal region is the high level of regional economic development; and local

low-value economic development in central and western regions is the low level areas. The results show that spatial pattern of the community species presented obvious collective pattern in China's paper manufacturing green technological innovation.

3.2 Analysis of Spatial Regression Results

The spatial lag model and spatial error model were estimated using the GEOSA software and were reported in Table 1. Shown in table 1, R square for our study is 0.8997; significant F statistics i-e 59.390 confirms the validity of the model at the level of 1%. Durbin-Watson Statistic is 1.9799, indicated that the phenomenon of residual serial correlation did not exist. Mainly because factors affect the paper manufacturing green technological were complex, and we use the full sample the noise is significant, thus the results are acceptable.

The regression results show that the coefficient of Provincial R & D investment in the green innovation of paper industry is significantly positive, indicating that more R & D expenditure, more output of paper industry green technological innovation. The coefficient of industry-university cooperation is significant positive shows that school-enterprise cooperation in research has a strong synergism on paper industry green innovation output. For example, the four largest listed companies are located in Shandong Province, relies on scientific research of paper manufacturing green technological innovation, with the bright regional technological characteristics. There is positive relationship between government-industry cooperation and paper manufacturing green technological innovation output, indicated that government policies can improve the green paper industry innovation output. The coefficient of absorption capacity (corporate average number) is significantly negative, indicating the paper-making enterprises scales are out of proportion to the efficiency of green technological innovation. That means paper-making enterprises in China could not use HR affectively. We also find the introduction of technology have positive impact on the green innovation output.

Table 1. Spatial Econometric Regression Results

Variable	OLS		SLM		SEM	
	β	t	β	t	β	t
C	0.0678	0.2957	0.1397	0.1791	0.1997	0.2851
lnHK	0.5367*	3.1812	0.5397	1.6173	0.6379*	3.3657
lnERD	1.2031*	5.7357	0.7892**	2.3572	1.5279*	6.2337
lnEGRD	0.3657**	2.6752	0.3027	1.1516	0.5738*	3.2352
lnEURD	0.1371	0.5537	0.0917	0.8627	0.0612	0.6357
lnTRAD	0.0627	0.5967	0.0653	0.6575	0.0517	0.3191
lnFSN	-0.0313	0.1751	0.5376***	1.8573	0.0657	0.5657
R^2	0.8997					
F	59.390*					
DW	1.9979					
ρ / λ			0.1789	1.3591	0.2575	0.9779
LogL	-23.0512		-36.2739		-99.3212	
LR			1.8796		0.5327	
AIC	62.0331		87.6752		101.97	
SC	72.5767		97.5297		123.7981	

Note: ***,**,* said that 10%, 5%, 1% significant

Taking into account the spatial correlation effect on the business of green paper the impact of innovation output, we use the SLM and SEM model for quantitative analysis (shown in Table 1). As shown in table 1; it is really alarming that ρ and λ for our study are statistically significant. This indicates that regional data with significant spatial dependence characteristics shown as the form of spatial lag dependent variable. The errors did not have phenomena of spatial dependence. Lagrange multipliers error, the lag and robustness test showed that LMERR, LMLAG, R-LMERR, and R-LMLAG are not significant at the level of 10%. By comparing the log-likelihood function value LOGL, AIC, SC, LR value of SLM and SEM model, the author find the effect of SLM model is relatively better. SLM and SEM model are clearly superior to OLS estimates.

SLM and SEM model regression results show that green technological innovation input elasticity coefficient are still stable positive, maintained at the level of 0.70 ~ 1.35, while controlling the average annual number. And this result is same with OLS regression results. The elasticity coefficient of enterprises invest on the human capital is positive, maintained at the level remained at 0.36 to 0.54. While OLS regression result are not identical. The results show, to a extent China's paper enterprises have got some green technological innovation capability. But the measures of innovation still stay at the low level and output is mainly thanks to increasing input. As for the regression results, the coefficients of import of advanced technology, industry-university cooperation, government-industry cooperation is positive. There is negative relationship between absorptive capacity of enterprise and green innovation output. This is indicating that recent green innovation in China's paper industry is mainly driven by the enterprise itself. And other R & D cooperation (government-industry-academia cooperation), technology transfer and other innovation means also show good results. In the terms of the composition of innovation, national and regional innovation output, the central regional innovation environment and eastern regional innovation input are great constraints on economic growth. For the reason of this, we should give effort to the construction of the central regional innovation environment and western regional innovation input. In addition, the promotion of cooperation of industry, academe and research institutes should be done to optimize the structure of the regional innovation output.

4 Conclusion and Recommendations

This paper utilizes special lag model, space error model in knowledge production function and special analysis to conduct the quantitative analysis for the function of China's paper industry green innovation spillovers. The results show that not only traditional economic variables determine the green innovation spillover, and there are significant global and local spatial autocorrelation in innovative output and its spillovers model. Thus the space spillover effects result of green innovation led to the surrounding areas innovation and innovation cluster. Compared to the traditional econometric models, this article uses the spatial lag model is more adapted to regional characteristics of the data, empirical results are more accurate and reliable. After having the detailed econometric analysis we have come to the conclusion that:

First, from the regression results, the author find that China's 31 provincial contribution of green innovation in the paper manufacturing sector mainly is enterprises R&D investment. Paper-making enterprise's own R & D dose capability

do have positive impact on the capital regional innovation competence, and while the spatial analysis results further confirmed this conclusion. In addition, the spatial quantitative model results showed that the cost of technology import and the absorptive capacity had no significant influence on green innovation performance. So each province should encourage paper enterprise to increase their capital input for this area, which is a long-term development strategy for China's paper industry.

Second, the spatial econometric results show that the main contribution to paper industry green innovation is enterprise R&D expenditure, while university R&D expenditure does not has obvious influence on green innovation, and the combination of university R&D expenditure and government-industry-academia collaboration does not show obvious influence on green innovation either. Mainly due to, one hand that our university's scientific and technological innovation resources are not fully utilized; the other hand that the business goals of our university's innovative research are not prominent. Then prompt the role of "official university-industry interaction", especially research universities, researchers and academic institutions in regional economic development is important. Therefore, in order to improve the ability of independent innovation of China's paper industry, active industry-university collaborative R & D knowledge spillovers, through policy measures the government to support the University of R & D and manufacturing enterprises to establish strategic alliances, to build industry-university cooperative research and development of institutional arrangements, to successfully achieve the University's new knowledge and new technologies to the business transfer and diffusion.

Third, the green innovation technical spillover effect of geospatial space in China's neighboring province paper industry does not exist. Studies by Jaffe (1996) showed that knowledge spillovers are more likely occurred in areas geographically close to each other, rather than cross-regional free flow. In other words, the exchange, sharing and application of technical knowledge with geographic space, all that need a local optimal environment and atmosphere for innovation. Therefore, the central and local governments must implement differentiated, rather than the average of regional public investment or R & D tax policy, subsidies, to guide the R & D investment, and thus form a good innovation system to improve the innovative capability of the paper manufacturing sector.

References

1. Jaffe, A.B.: Real effects of academic research. American Economic Review 79(5), 957–970 (1989)
2. Cohen, W.M., Levinthal, D.M., Klepper, S.: The Anatomy of Industry R&D Intensity Distributions. The American Economic Review 82, 773–799 (1992)
3. Cohen, W.M., Klepper, S.: Firm size and the nature of innovation within industries: The case of process and product R&D. The Review of Economics & Statistics 78, 232–243 (1996)
4. Kokko, A.: Technology, Market Characteristics, and Spillovers. Journal of Development Economics 43, 279–293 (1994)
5. Parent, O., Lesage, J.P.: Using the variance structure of the conditional autoregressive spatial specification to model knowledge spillovers. Journal of Applied Econometrics 23(2), 235–256 (2008)

Overview of Change Management and Its Implementation

Reza Malek[1] and Rashad Yazdanifard[2]

[1] Center of Post Graduate Studies, Limkokwing University, Cyberjaya, Malaysia
rezmalek7@gmail.com
[2] Faculty of Management, Multimedia University, Cyberjaya, Malaysia
rashadyazdanifard@yahoo.com

Abstract. The purpose of this paper is to provide an overview of the concept of change management and its objectives in today's fast-moving markets. You also can perceive why individuals resist change initially, and it is hard for them to leave the status quo, and shift to a new state. At the end the way of implementing the change management is examined.

Keywords: Change, Resistance, Process, and Objective.

1 Introduction

In the past, the majority of organizations had been designed for stability rather than change. Moreover they had been focused on creating a specific outcome, result or product, whereas these days, organizations focus on improving the process besides they need a new way to handle their job and maximize benefits. Having competitive advantage and profit maximization are the primary objectives of any company. One of the ideal approaches to achieve these objectives through the rapid change in technology, communication, and information is managing change within the organization. Change the values that organizations live by, values like awareness, teamwork, tolerance, responsibility and information are paramount – just as flexibility and change readiness [10]. It occurs by being knowledgeable about change management implementation and its procedures. Indeed change management is a significant component of any organizations.

Change management has typically been defined as a process involving unfreezing, moving, and refreezing values, practices, and procedures [5]. But people in an organization, as a human do not accept change simply, and see it as a negative thing. Therefore they initially resist change. Change management is an effective way to minimize this negative feeling and making the change process more efficient.

2 Change Management Definition

Change management is the process of planning and coordinating the implementation of all changes through individuals, teams, and organizations. This process basically is

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 149–153.
springerlink.com

defined as problem solving which is a shift from problem state to a desired future state [3]. It is the application of many different ideas from psychology and sociology fields, business administration, and system engineering [16]. Managing change is a process to make employee ready to accept change. In addition it is a way to make them comfortable about change and to carry it out with enthusiasm, and understand the role of their actions in the unfolding drama of the company's fortunes, and believe that is worthwhile for them to play a part [1]. To assess the impact of change and to monitor its implementation, it is essential to have a policy and procedure for Management of Change [11].

According to Kurt Lewin [3] the process of change three basic stages: Unfreezing, effecting change, and refreezing. Unfreezing is the creation of a difference realization between the status quo and ideal status of an organization that prepares people to change and decrease their resistance. Effecting change is moving from existing situation to ideal state that refers to various processes such as training, education, and restructuring that lead to the development of new behaviors, attitudes, and beliefs [5]. Refreezing regards reestablishing a new state within the organization by stabilizing the new patterns with a general support [5]. As a result change management is a strategic activity aimed at effectively implement new methods and systems in an ongoing organization and obtaining the best outcomes from the change process [4,6].

3 Change and Resistance

Organization's members often resist change, for the reason that habits are a normal part of every person's life. Humans are creatures of habit, are generally satisfied with the status quo, and will tend to initially resist change [8]. Initial resistance to change is caused by some reason. First, the uncertainty created by the announcement of impending change, it disrupts certainty and order of their work lives [3]. They feel a need to control their environment and resist anything that they perceive can threaten this control and it has an adverse impact on their job.

Second reason is that employees require modify their existing individual and organizational identities. Implementing organizational change initiatives invariably involves people change themselves [8]. The change cannot occur until people decide to change. This Process based on William Bridge [12] is expressed as transition. Transition is not the same as change. It is a psychological procedure while change is situational because of shift into the new state. Transition is an internal psychological re-orientation that people go through as they come to terms with a change [15]. The change will simply not work unless transition takes place. Workers' anxiety is another factor, they worry cannot conform and succeed in the post-change work environment. They fear they cannot adjust their practices or adopt new ones in line with their existing mind-sets in order to reach a new bottom-line target [1].

The last but not the least reason is caused by the organization, many organizations do not thoroughly strive to manage change. They simply announce what the changes will be and expect everyone to comply [8]. When changes are imposed without consulting, employees will be confused about why organization makes change, what the objective is, and how their job will be impacted. This can also result in denial, resentment, conflict, organizational rumor, and even increasing sickness absence [5].

Due to these reasons, it is understandable why individuals initially resist change, and see change as a negative thing.

4 Change Management Objectives

The primary purpose of change management is verifying changes in facilities, personnel, and operations, which are evaluated and managed to ensure that the environmental risks arising from these change are controlled [9]. Organizations that develop and promote strong change management obtain a significant return. Towers Watson [13] research shows that high-performance organizations perceived as highly effective in managing change are superior at responding to changes in the market, compared with high-performance organizations that are less effective change managers (65% to 41%).

A number of Objectives why organizations manage change are to:

- Minimize adverse impact of necessary change on system integrity, security, and the service level agreements [7].
- Maximize the opportunities presented by the proposed change [6].
- Take full advantage of employees' productivity.
- Make employees ready to change, and make sure they are engaged with change process.
- Prepare and support employees to ensure effective change to achieve strategic goals and vision [6].
- Allow the coordination and planning of changes in order to provide a stable production environment [7].
- Reduce the destructive consequences of change.

As a result, when organizations properly manage changes, they will obtain the best possible outcome, and they will achieve these above goals.

5 Change Management Process

The journey of change management starts with a request for change (RFC) [2]. A request for change comes into being by analyzing current situation and finding out a need for change. It also could be triggered by such activities as a customer request via service desk, the introduction or removal of a configuration item, or the output of a development project [7].

Second step is change prioritization by assessing impact and risk of change. The impact the change will have on the production environment and the services provided to the user and business reasons for the change are determined and approved [7]. Risk evaluating also is determined at this stage. Measuring risk can be defined as the actual risk associated with implementing the change versus the risk of possible failures if changes are not implemented.

The third station is creating readiness for change and overcoming resistance to change [6] by finding out what particular ideas or fears employees have about change. It is essential to take into account the organization's memories, norms and values.

Moreover, it is essential to have commitment and eagerness from top management to lower levels [16]. The next step is preparing employees to work in the new state by devising an effective education, training, and skills upgrading scheme for the organization [17]. Organizations must support employees to improve their change skills and make them ready for a return to enlargement. Another important step is effective communication via multiple ways, including speaking, writing, video, training and bulletin boards [14]. Give important amounts of time for people to ask questions and request explanation. Visibly communicate the vision, the mission, and the objectives of the change management effort therefore assist people to realize how these changes will influence them in person is important. Message should be positive and present opportunities for people to associate with each other, both formally and informally, to share ideas about change and change management [14].

The change coordination is the last step of this journey. This function controls day-to-day activities of the change management system with the goal of meeting the change management objectives [7]. The purpose of this meeting is review the appropriateness of changes scheduled for implementation and to confirm the coordination and cooperation of participating area. Change management is an iterative process that requires constant review and adjustment for continues improvement [2]. All of these steps are necessary to implement a successful change through the organization that needs change.

6 Discussion

In today's fast moving market change in organizations' environment has become a way of progress, and it is an inherent characteristic of every businesses. Advancement in technology, increasing customers' expectations, and the need to control costs and increasing efficiency mean that organizations need change to survive. This change requires a perfect management in order to obtain the best result.

Change management is a systematic approach that helps organizations and individuals cope with change and its effects. Indeed it is an essential process that facilitates organizations to apply new strategies. But there is a major problem to accept change and move to the new state. Initial resistance from employees, due to their human being and its characteristic that habit is a part of humans and it causes they refuse to accept change at first. In this situation a perfect change management is the key to overcome the resistance to change, kill the status quo, and protect the organization's environment. It requires several strategies to make the change successful such as analyzing the current situation and finding the reason for change, assessing the risk of change, training the employees, effective communication, and employee involvement. Another thing that is totally important to improvement is control the change by coordinating and reviewing the suitability of change.

At the end it is important to know that all of these strategy will work out and by building the sponsorship from the top to the lower level. Consequently change get result when organizations make a real effort to manage changes, and employees believe change is a positive process.

7 Conclusion

In conclusion, the change management approach should be fully integrated into organization's decision making, both informing and enabling strategic direction. It should be based on a realistic assessment of the organization's history, readiness, and capacity to change. Change management should minimize the adverse impact of necessary changes on system integrity, security and the service level agreements. Moreover, it is the coordination and planning of changes in order to provide a stable production environment which maximize the productivity of persons involved in the planning, coordinating, and implementation of quality changes.

References

1. Lawson, E., Price, C.: The psychology of change management, p. 30 (2003)
2. Doherty, P., Waterhouse, P.: Change management: A CA IT service management process map (2006)
3. Ash, P.: Fast and effective change management. Knowledge Solution (2009)
4. Nickols, F.: Change management 101, a premier. Distance Consulting LLC (2010)
5. CommonGap, Communication for governance and accountability program
6. Approaches to change management for flexible learning, Australian National training authority
7. Leopoldi, R.: IT service management. Change Management Method and Implementation Best Practices (2002)
8. Green, M.: Change management master class. Kogan Page (2007)
9. Cook, S., Macaulay, S.: Change management excellence. Kogan page
10. Beakdal, T., Hansen, K.: Change management handbook (2006)
11. Leal-Valias, M.: Management of change. NPC Training Program
12. Bridges, W.: Managing transition making the most of change (2005)
13. Yates, K.: Post-recession success depends on effective change management. Towers Watson (2010)
14. Paul, P.: Change management: communication, employee commitment, and leadership (2004)
15. JISC infonet, Change management
16. http://www.change-managemen.com
17. http://www.answers.com

Causes of American Subprime Mortgage Crisis and Its Enlightenment to China

Tao Song[1] and Shi-chen Fan[2,*]

[1] Finance Department of Hainan University, Haikou, Hainan, China 570228
[2] Applied Economics Postdoctoral Research Station of Southwestern University of Finance and Economics, Chengdu, 611130, China
{songtao7221,fansc1230}@163.com

Abstract. So far American Subprime Mortgage Crisis has lasted more than four years. It not only hit America's economy but also caused a global financial crisis. The study on its causes and its enlightenment to China undoubtedly is beneficial to the healthy development of China's financial market.

Keywords: subprime mortgage crisis, cause, enlightenment.

1 The Causes of American Subprime Mortgage Crisis

1.1 Double Roots: Low Savings Rates and Low Interest Rates

High consumption and low savings has always been the characteristics of American social economy. Especially in recent years, American people almost do not save. As is known to all, one of the important functions of savings is to prevent the risk and reduce the impact of uncertainty factors. Since 1980, the individual savings rate in USA had shown a downward trend. The average rate in 1980s was about 10%, and declined to 5% in 1990s. After 2000, the rate descended further to 2.5% and even in 2004 there was the negative savings [1]. With no savings, the ordinary American people, especially the borrowers of subprime mortgages, have no ability to resist even the mildest risk. The low individual savings rates in USA increased the threat of the subprime mortgage crisis.

As early as in 2001, with the dotcom crash and the September 11 attack, the American economy was faced with the occurrence of a recession. To stimulate the economic growth, Alan Greenspan, former chairman of the Federal Reserve, put forward the low-interest policy, cutting the interest rate for 13 times in succession. This policy greatly pushed the development of real estate industry, which became the major stimulus of the American economy then. But the low-interest policy also made the ordinary people stuck in the mortgage morass and caused the property bubbles, which objectively planted the seed of the subprime mortgage crisis. Joseph Stiglitz, the Nobel Economic Prize winner, pointed out in Les Echos in August 2007 that the subprime mortgage crisis was closely related to the policies of Greenspan and Greenspan should undertake the responsibility of the subprime mortgage crisis.

* Corresponding author.

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 155–159.
springerlink.com

1.2 Double Boosters: Lenders and Investors

Under the low-interest policy, some American financial institutions were carried away by the huge profits in the subprime mortgage market and forgot the management principle of risk avoidance. They put aside the lending standards for subprime mortgages and issued enormous amounts of loans. According to statistics, the subprime mortgages accounted for less than 5% of the total housing loans of America in 2001 and jumped to 23.6% then, amounting absolutely to $1.89 trillion [2]. Similarly, the investors who bought the subprime mortgage bonds —— banks, asset management companies, hedge funds, insurance companies and pension funds —— were obsessed with the short-termed high yield of the subprime mortgage bonds, ignoring the high risk of the subprime mortgage bonds and its nature of junk bonds. They not only bought a great amount of subprime bonds themselves, but they also sold the bonds globally through their branches, making the whole world the victim of this subprime mortgage crisis [2]. On one hand, the loan lenders and the investors, dizzy with the high profits of subprime mortgages, contributed to the temporary prosperity in the American real estate. On the other hand, they boosted the burst of the subprime mortgage crisis and enhanced its global impact, putting people in and out of America into the mortgage morass.

1.3 Double Incentives: Slowing Growth in Property Prices and Limited Rise in Mortgage Rates

According to the statistics from the American Federal Housing Enterprise Supervision Office, the highest growth rate in America's property prices in the period from 2000 to 2007, which was up to 4.33%, occurred the third quarter of 2004. Although the growth rate had stepped down since the second quarter of 2006, the property prices kept the positive growth. Since the slowing growth in property prices was predictable and acceptable to the market, it alone should not have caused a large number of defaults and the subprime mortgage crisis.

According to the statistics of the Federal Reserve, the American 10-year mortgage rates declined from the highest 8.55% to the lowest 5.25% from 2000 to 2003; but from 2004 to 2007, the rates presented the fluctuant improvement, rising to 6.75%. So after 2005, the American mortgage rates rebounded slightly, only 150 basis points overall [1]. Although the slight pick-up had certain effect on the interest costs of the borrowers, it should not be strong enough to cause the subprime mortgage defaults and disturb the mortgage market.

Neither the slowing growth in property prices nor the limited rise in mortgage rates can alone cause a large number of defaults of the subprime mortgages and then lead to the subprime mortgage crisis. But their joint impact was too heavy to American people who have little or no savings, especially the borrowers of subprime mortgages. Therefore, it is just the mild risks like the slowing growth in property prices and the limited rise in mortgage rates that stirred up the low savings and the mortgage morass, and even induced the subprime mortgage crisis.

1.4 Negligence: Duty Neglect of Credit Rating Agencies

In the subprime mortgage frauds, the credit rating agencies, such as Standard & Poor's and Moody's, were defeated by the allure of profits, forgetting their professional spirits and violating their professional ethics. They issued the rating certificates to the unchecked loans which concealed the possible risk, which damaged the interests of the defenseless people and investment institutions such as banks, asset management companies, hedge funds, insurance companies and pension funds. Therefore, the negligence of the credit rating agencies to some extent spread the influence of the subprime mortgage crisis [2].

1.5 Indulgence: Weak Supervision of Regulatory Agencies

The related regulatory agencies in America failed to realize in time the property bubbles caused by the low-interest policy, the mortgage morass that the public bogged in, and the weak risk resistance of ordinary people caused by the long-time high consumption and low savings. They ignored the high risk of delinquency in the mortgage market and mortgage securitization and did not take the timely measures to prevent and control the abnormal lending of financial institutions and the malicious concealment of credit rating agencies. Failing to effectively prevent and control the risk in the financial markets, the incompetent regulatory agencies were partly responsible for the subprime mortgage crisis [2].

2 The Enlightenment of American Subprime Mortgage Crisis

2.1 Strengthen the Awareness of Financial Risk and Construct the Double Defense at Home and Abroad

With the globalization of economy and finance, the international financial crisis occurs frequently and spreads quickly to other countries and regions, even to the whole world. Ten years ago, Asian Financial Crisis taught us an important lesson that the opening of the financial market must be careful and sound [3]; today, American Subprime Mortgage Crisis gives us a wake-up call that the financial risk cannot be ignored at any time. While opening up its financial market, China can ensure the safety of the state's finance only through strengthening the risk prevention. How to effectively resist the spread of the international financial crisis and defense its impact to the domestic finance is critical to the rapid and healthy development of China's economy. Tracking the latest development in the international and the domestic financial market and enhancing the cooperation with other countries, China should prevent and control the financial crisis at home and abroad [4].

2.2 Keep Vigilant against Various Financial Risks in the Development of the Domestic Real Estate

China now is going through a period of excess liquidity and asset price surge, the stock market and the housing market booming and the mortgage market growing rapidly. In the recent two years, the value of China's stock market has grown fivefold

and the real estate prices in China's large and medium-sized cities have risen unilaterally. Accordingly, Chinese banks keep increasing the loan balance: by the end of June, the national balance of commercial real estate loans is 4.3 trillion RMB, rising by 24.5% year-on-year; the loan balance of real estate development is 1.7 trillion RMB, rising by 25.6% year-on-year; the balance of commercial purchase loans is 2.6 trillion RMB, rising by 23.9% year-on-year [3]. According to the experience of the subprime mortgage crisis, the rapid rise in housing prices tends to cover up a great deal of credit risk and operational risk. Since the major means of financing in China's real estate market is the bank's indirect financing, the risk in the credit stocks of the commercial banks and the trend of the real estate markets are closely connected, and the concentration risk of development loans in the real estate markets and the potential credit risk in the individual housing mortgage loans should be fully concerned.

2.3 Carry Out Housing Mortgages Securitization Cautiously

The subprime mortgage crisis objectively raised the caution in importing the financial derivatives and instruments. As a double-edged sword, the financial derivatives and instruments has the risk of leverage ratio: under certain conditions they will develop the partial risk into the overall risk and turn the individual market risk into the systemic risk. The securitization of housing mortgages, as the typical one of the financial derivatives and instruments, diversifies the credit risk of banks and borrowers and spreads the mortgage risk into the stock and bond markets as well. At present, the housing prices have shown the bubbles and the trend of inflation, and the capacity of housing mortgage assets is being checked by all the risks in interest, exchange, market and price, etc. Therefore, China's securitization of housing mortgages must be carried out cautiously [4].

2.4 Reinforce the Government's Macro Adjustment and Control

The Chinese government should closely monitor the housing mortgage market, warning the lenders and investors before the outbreak of the crisis and reducing, even stopping, the dangerous housing mortgage loans. In the early days of the crisis, the government should shoulder the responsibility to stabilize the market to stop the crisis from extending to other financial markets. And the government should take measures to strengthen the confidence of lenders and stabilize the housing prices to prevent the crisis from endangering the whole society. The government should organize the capable financial institutions to set up the state-supported organizations, like American Federal National Mortgage Association and Federal Housing Mortgage Association, to purchase the risk assets and diversify the risk. Moreover, the government should make flexible and effective use of the international financial market, applying for the emergency loans from the international financial institutions such as IMF (International Monetary Fund) or selling the risk assets at packaged auction in the international market [5].

2.5 Enhance the Supervision over the Financial Market

China need promote the reform and opening up in finance, develop the corporate bonds market, build the financial market system that is multileveled, diversified,

capable and flexible, and improve the resilience and the risk resistance of China's financial system. Moreover, commercial banks need improve the measurement and management over credit risk, learn the foreign risk management technologies, select the most suitable credit-risk identifying model and combine it with the situation in China, and finally establish the credit-risk identifying model of China's commercial banks. Besides, the independent and credible rating agencies should be set up and the information disclosure system should be implemented. The supervision model should be transformed from separate supervision to functional supervision, resolving the conflict between the current supervision framework and the supervision idea from the system. The financial supervision should be coordinated around the globe [6]. The focus problem in the subprime mortgage crisis is the credit rating misalignment. So the supervision system over service agents in financial markets that is suitable for China's conditions must be set up promptly [7].

3 Conclusion

In conclusion, American Subprime Mortgage Crisis is the comprehensive economic crisis, rooting in the low savings rates and low interest rates, developing with the boost of the lenders and investors, bursting out with the stimulation of the slowing growth in property prices and limited rise in mortgage rates, and spreading around with negligence of the credit rating agencies and the indulgence of the regulatory agencies. The developing countries like China should have an objective and comprehensive understanding of the crisis and take the corresponding measures to prevent and control the domestic financial crisis.

Acknowledgement. Financial supports from: the National Natural Science Foundation Project of China in 2009 (NO.40961005), the National Soft Science Item of China in 2010(NO.2010GXS5D252), the National Education Ministry Planning Project(NO.09YJC840008), University Research Project of Hainan Ministry of Education in 2010(NO.Hjsk2011-28), Hainan University Education Project Fund in 2010(NO.HDJY1008).

References

1. Li, S.: Low Saving Rate Is the Root of American Subprime Mortgage Crisis. China Finance (21), 38–39 (2007)
2. Zheng, L.: Study on Cause and Influence of American Subprime Mortgage Crisis and Ten Inspirations. Review of Economic Research (6), 4–21 (2007)
3. Li, Z.: The Influence and Enlightenment of American Subprime Mortgage Crisis to China. The Teaching of Thought and Political Study (3), 4–6 (2008)
4. Zhen, B.: The Enlightenment of American Subprime Mortgage Crisis to China. Research on Economics and Management (11), 9–16 (2007)
5. Lu, P.: Study on American Subprime Mortgage Crisis and Its Enlightenment to China. Finance Observation (2), 102–104 (2008)
6. Liu, Y.: China's Countermeasures to the International Financial Crisis. China Finance (21), 56–57 (2008)
7. Huang, R.: Study on American Subprime Mortgage Crisis's Market Influence, Enlightenment and Countermeasures. China Finance (19), 60–61 (2007)

Research on Traffic Circle Capacity

Chaonan Jiang[1], Jing Yin[1], Jialei Gao[2], and Tao Huang[3]

[1] School of Economics and Management, Beijing Jiaotong University, Beijing, China
[2] School of Traffic and Transportation, Beijing Jiaotong University, Beijing, China
[3] Communication Signal Center of Operations Business Headquarters of Guangzhou Metro Corporation, Guangzhou, China

Abstract. Traffic circle is an important type of non-signalized intersection, and there is an extensive use of it. In order to improve traffic capacity and safety, different traffic circles adopt different traffic priority, for example, vehicle in-circle precedence method and the vehicle entry-precedence method. This paper studies the applicability of different traffic priority based on the traffic circle capacity. Assume that the time headway within the circle obeys Erlang distribution, and by using the gap-acceptance theory, the paper suggests the calculation formula of the saturation flow of the traffic circle. Based on the model solution, a critical value can be got to distinguish the applicability of different traffic priority. Therefore, the solution could help relevant departments make decision.

Keywords: Traffic circle, traffic capacity, gap-acceptance theory, vehicle in-circle precedence method, vehicle entry-precedence method.

1 Introduction

The social progress in many cities or communities gradually gives rise to serious problems of traffic jam. It has become the tendency to develop roads to three-dimension space----to construct intersection. But the high cost of constructing intersection and the high occupation rate in the space is the fundamental causes that traffic problems cannot be solved in time. So someone put forward a middle course and make research on plain traffic circle. The way can not only improve the present traffic status, but also reduce the construction cost greatly.

Traffic circles have been reported to be very useful in controlling vehicular traffic at road intersections (Brilon and Stuwe, 1990), Major advantages include the provision of adequate throughput and user safety, basically through slower vehicle speeds and reduced driving conflicts with the opposing traffic(M. Hossain, 1998). Many methods are used in the capacity of traffic circles. For example, Kimber (1980) brought forward an empirical formula for estimating capacity of traffic circles based on the data collected from the traffic circles of UK. Zhang Jian, Ma Jun and Liu Mo (2003) suggested a method to calculate the capacity of traffic circle based on probability theory in city. Shi quan and Wu zhong(2006) solve the problems by the gap-acceptance modeling.

The paper applies model of gap acceptance theory as well as concerning knowledge of probability theory and queuing theory, discusses the entrance traffic

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 161–168.
springerlink.com

capacity under the two kinds of situation of vehicle in traffic circle-precedence and vehicle entry-precedence. And by applying the model offering in the paper, the inside and outside factors that affects the traffic capacity in traffic circle can be analyzed to carry out real-time control and realize the optimum in a long term.

2 The Presentation of Questions

Many cities and communities have traffic circles—from large ones with many lanes in the circle to small ones with one or two lanes in the circle. Some of these traffic circles position a stop sign or a yield sign on every incoming road that gives priority to traffic already in the circle; some position a yield sign in the circle at each incoming road to give priority to incoming traffic; The goal of this problem is to use a method to determine how best to control traffic flow in, around, and out of a circle based on the traffic circle capacity.

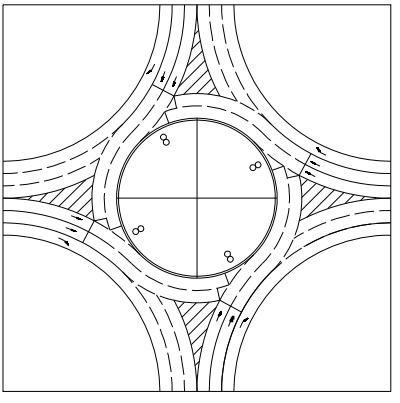


Fig. 1. The traffic circle

3 Traffic Circle Capacity Model

3.1 Prepare for the Traffic Circle Model

Gap and Time Headway Distribution. In accordance with the basic knowledge of traffic theory, gap distribution is in negative exponential. The function is:

$$f(t)=\lambda e^{-\lambda t} \tag{1}$$

The time headway within the circle obeys Erlang distribution and the probability density function of its distribution is:

$$p(t)=\frac{\lambda(\lambda t)^{r-1}}{(r-1)!}e^{-\lambda t}, \qquad r=1,2,3,\cdots \tag{2}$$

r means the order of Erlang distribution, Among them, $\frac{\lambda}{r}$ Defined as traffic flow amount in different R.

When r=1, it shows that traffic flow is in a state of complete freedom. With the increasing of R, the state of the traffic flow becomes a completely limited state.

Furthermore, on the bases of Cowan and Erlang distribution , We introduce the scale factor α (which is a proportion in vehicles with no touch in each other), and also assume that the least headway time of the unfreedom vehicles is t_c.We proposed that traffic headway time are M3 distribution, and its cumulative probability function is:

$$F(t)=\begin{cases}1-\alpha e^{-\lambda(t-t_C)} & t>t_C \\ 0 & t\le t_C\end{cases}, \tag{3}$$

Among it, $\lambda=\frac{q\alpha}{1-qt_C}$,

q ----traffic amount.

The simplest situation is that when α=1,t_C=0 , the same situation in Erlang distribution when r=1 .

The circle traffic flow. It is obvious that saturated flow of circle traffic is associated with architectural features and characteristics of the vehicles entering the roundabout. Without taking vehicles' delay into account, the circle saturation flow and inner-radius $R(m)$, the numbers of lane m , the average speed of limited ones $v(m/s)$ and also streets' average width $l(m)$, the formula of saturated traffic flow as below:

$$\lambda=\sum_{i=1}^{m}\frac{2\pi(R+\frac{il}{2})}{tv} \tag{4}$$

According related standards, m is less than 5, generally it is 2or3.And R is within 150m.generally, a smaller circle's r is within 50m, while it can be 150m at largest, say Paris Arc De Triomphe traffic circle.

3.2 Construction of the Traffic Circle Capacity Model

By the model of the traffic capacity of traffic circle, we can more clearly know of the state of the circle, and take measures by adjusting some factors to make the optimization of the capacity.

The entrance traffic capacity of the traffic circle respectively using vehicle in-circle precedence method and vehicle entry-precedence method

In the two methods, the vehicle gap-inserting theory is applied in the two methods with the probability theory and queuing theory.

The entrance traffic capacity of the traffic circle is represented by the number of vehicles entering the traffic circle per unit time. According to the assumed conditions, one of the entrance traffic capacities can represent it.

When the queue is enough long in the entrance-lane and there is only one lane in the traffic circle, the probability that there are k vehicles entry into one void of the traffic circle is $p(k)$, so the expectancy of the vehicles' number is:

$$C = \sum_{k=1}^{\infty} kp(k) \tag{5}$$

The Case of Single Lane. Specifically, The vehicle in-circle precedence method and vehicle entry-precedence method is more simple. Just one traffic flow insert into another traffic flow. There are different in the inserting flow and the inserted flow between the two methods. That is to say, for vehicle in-circle precedence method, the traffic flow entry into the roundabout in the right time and intersperse to the traffic flow in the circle. While in the vehicle entry-precedence method, the traffic flow in-circle in right time intersperse to the traffic flow entering the circle.

Assume that the time headway of the inserting flow is t_f, and the time headway of the inserted flow is T ,which Obeys Erlang distribution. Meanwhile, for guarantee the safety of vehicles there must exist a smallest time headway between vehicles in the inserted flow t_c must exist, which is the time head way in the case of no free traffic mentioned above.

When $t_c \le T < t_c + t_f$, one vehicle can intersperse to the inserted flow .so, when $t_c+(k-1)t_f \le T < t_c+kt_f$, there are k vehicles that intersperse into the inserted flow. Thus:

$$P(k) = P\{t_c + (k-1)t_f \le T < t_c + kt_f\} = P(T \ge t_c + (k-1)t_f) - P(T \ge t_c + kt_f)$$

According to the probability density function of the time headway, we can solve the above formula:

$$P(k) = P(T \ge t_c + (k-1)t_f) - P(T \ge t_c + kt_f)$$

$$= \sum_{i=0}^{r-1} \frac{\left\{\lambda\left[t_c+(k-1)t_f\right]\right\}^i}{i!} \times \exp\left\{-\lambda\left[t_c+(k-1)t_f\right]\right\} - \sum_{i=0}^{r-1} \frac{\left\{\lambda\left[t_c+kt_f\right]\right\}^i}{i!} \times \exp[-\lambda(t_c+kt_f)]$$

When the traffic flow is not much big• r=1, so:

$$P(k) = \exp\left\{-\lambda\left[t_c+(k-1)t_f\right]\right\} - \exp[-\lambda(t_c + kt_f)]$$

$$= e^{-\lambda(t_c+kt_f)}(e^{\lambda t_f} - 1)$$

In a certain period of time, $P(k)$ express the probability that there are k vehicles that intersperse into the inserted flow. In which k=1,2,3,⋯

Specifically, for the vehicle in-circle precedence method, the time headway of vehicle entry is the time headway of the inserting flow which is t_f, and the time headway of vehicle in circle is the time headway of the inserted flow which is T.

Meanwhile, there is smallest time headway between the vehicles in circle. When it comes to vehicle entry-precedence method, it is just opposite.

For the vehicle in-circle precedence method, every value of k directly express the number of vehicles entering the circle. So, the entrance traffic capacity is：

$$C_1 = \sum_{k=1}^{\infty} ke^{-\lambda(t_C+kt_f)}(e^{\lambda t_f}-1)=(e^{\lambda t_f}-1)\sum_{k=1}^{\infty} ke^{-\lambda(t_C+kt_f)} \tag{6}$$

For the vehicle entry-precedence method, the inserting flow is the traffic flow of the traffic circle. The number of vehicles entering the circle and the interspersed vehicles belong to the circle originally, the entrance traffic capacity equal to the velocity of the traffic flow multiplied by the no plug time. Thus:

$$C_2 = \lambda(g - \sum_{k=1}^{\infty} kt_f P(k)) = \lambda(g - t_f(e^{\lambda t_f}-1)\sum_{k=1}^{\infty} ke^{-\lambda(t_C+kt_f)}) \tag{7}$$

In which λ is the traffic flow in the circle, t_f is the time headway of the inserting flow, t_C is the smallest time headway of the inserted flow, g is a fixed period of time.

The case of multi-lane road (taking two lane roads as example). The difference between Multi-lane traffic circle and a single-lane one is that the entrance vehicles must be able to intersperse to the plugs of the all lanes, only in this condition the vehicles in the entrance can enter the traffic circle. In normal case, the critical gap and the flowing time are different in different lane of the traffic circle. So, the condition that just k vehicles enter the roundabout is:

$$P(k) = H\{t_{cout} + (k-1)t_{fout}, t_{cin} + (k-1)t_{fin}\}$$
$$- H\{t_{cout} + kt_{fout}, t_{cin} + kt_{fin}\}$$

Among them, t_{cout}, t_{fout} are the smallest time headway of the vehicles entering the outer lane of the traffic circle and the time headway of the inserted flow. And t_{cin}, t_{fin} are the smallest time headway of the vehicles entering the inner lane of the traffic circle and the time headway of the inserted flow.

And then, the main work is to get function expression of $H(t)$, for two lane roads, define the density function.

$$g_0(t) = q(1 - F(t))$$

Obtain the formula of $H(t)$ through calculation:

$$H(t)=1-\frac{Q(q_1\alpha_1 q_2\alpha_2)}{\Lambda(\lambda_1\lambda_2)}e^{-\Lambda(t-t_C)}$$

Among it, $Q=q_1+q_2, \Lambda=\lambda_1+\lambda_2$

The sum of q_1 and q_2 are replaced by Q, and the sum of λ_1 and λ_2 are replaced by Λ.

In the end,

$$C=\frac{Q(1-q_1t_f)(1-q_2t_f)}{1-e^{-Qt_f}}$$

If there are more lane roads, we can obtain results by flowing this above.

3.3 Generalization of the Traffic Capacity Model

There are three control methods that are vehicle in-circle precedence method, vehicle entry-precedence method and Signal control method. For a traffic circle, according its characteristics and the traffic flow of each entrance, the best method is determined which need combine three methods. So the specific model which offers some suggestion to the related personnel is as follow.

According the definition of traffic circle traffic flow, it's optimal when the sum of each entrance traffic capacity gets max.

$$C = \max \sum_{i=1}^{n} \{(1-z_i)[\sum_{i=1}^{n} y_i C_1 + (1-y_i)C_2] + z_i C_3\} \tag{8}$$

$$z_i = \begin{cases} 1 & \text{the signal control method is uesd} \\ 0 & \text{the signal control method is not uesd} \end{cases}$$

$$y_i = \begin{cases} 1 & \text{vehicle in-circleprecedence method is used} \\ 0 & \text{vehicle entry-precedence method is used} \end{cases}$$

$$st. \quad C_3 = f(t_{red}, t_{green})$$

$$t = t_{red} + t_{green} (among\ them,\ t\ is\ a\ cons\tan t)$$

We can solve the above formula:

$$\begin{aligned} C &= \max \sum_{i=1}^{n} \{z_i[\sum_{i=1}^{n} y_i C_1 + (1-y_i)C_2] + (1-z_i)C_3\} \\ &= \max \sum_{i=1}^{n} \{z_i[\sum_{i=1}^{n} y_i (e^{\lambda t_f} - 1) \sum_{k=1}^{\infty} k e^{-\lambda(t_c + kt_f)} + \lambda(1-y_i) \\ &\quad - \frac{\lambda(1-y_i)}{g}(e^{\lambda t_f} - 1) t_f \sum_{k=1}^{\infty} k e^{-\lambda(t_c+kt_f)}] \\ &\quad + (1-z_i) f(t_{red}, t_{green})\} \end{aligned} \tag{9}$$

Among them, λ is the traffic circle annular flow whose formula (4).

According the modeling, when the traffic capacity is maximum the results of t_{red}, t_{green} and y_i, z_i can be got.

4 The Calculation of the Model

According the Model, the formulas of C_1,C_2 are above as formulas (6) and (7) .There are three factors that are traffic circle annular flow λ, the minimum time headway of the inserted flow t_c, and the time headway of the inserting flow t_f.

4.1 The Case When t_c and t_f Are Fixed Values

By referring to plenty of material, we get the values of these variables, such as, l=3.75m , v=40km/h , 20m≤R≤150m , m=2 and the value of the time headway are replaced by the mean of t , here let t=1s It is also known by calculation of the traffic circle annular flow λ ranges from 400(pcu/h) to 1400(pcu/h) .

Under our assumptions, the values of t_c and t_f are fixed, so let t_c=3s • t_f=5s . According the formula (4), the figure as below shows the relationship between entrance traffic capacity and traffic circle annular flow.

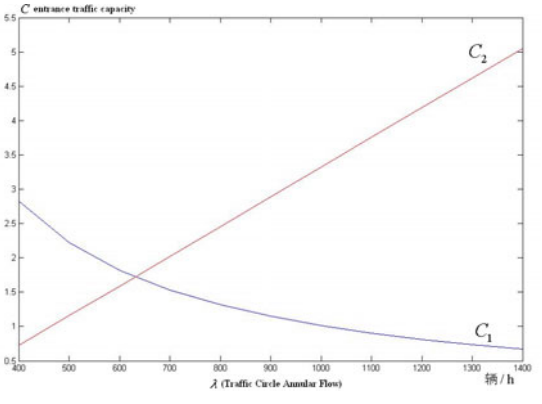


Fig. 2. The relationship between entrance traffic capacity and traffic circle annular flow

According the figure above, when t_c and t_f are fixed values ($t_c = 3s, t_f = 5s$ in this case), the traffic circle annular flow plays decisive role in making the entrance traffic capacity maximum in a certain time. By calculation, if the scale of the traffic circle is comparatively small (its radius is no more than 634 (pcu/s)) and before using Signal control method, the vehicle in-circle precedence method is better than the other one. Contrarily, the better one is vehicle entry-precedence method.

In this case, the conclusion can be applied in the traffic circles which are under construction, the radius can be determined based on the volume of traffic that will entry into the traffic circle. At the same time, it can offer the guidance suggestions whether adopting the vehicle in-circle precedence method or not.

4.2 The Case for One Fixed Traffic Circle

Let λ is equal to 400(pcu/h), 600(pcu/h), 800(pcu/h) 1000(pcu/h), 1200(pcu/h), so C is a binary function about t_c and t_f .making the figure as below by using Matlab.

We found that with the reducing of λ, the camber of the line becomes more and more obvious. At the same time, according to the value, if $t_c \geq t_f$, the vehicle in-circle precedence method is better than the other one. Contrarily, the better one is vehicle entry-precedence method.

In this case, the conclusion can be applied in one fixed traffic circles, it needs the signal light to control the entering flow and the inside traffic flow to make t_c and t_f reach a right relationship.

5 Conclusion

Traffic is related to people's everyday life, with the increasing traffic pressure, we should use scientific methods to organize the traffic flow especially at the traffic hub such as crossroads, traffic circles and cloverleaf junctions. This approach can make full use of transport capacity and tap the traffic potential adequately. Our model place particular emphasis on the situation that the traffic capacity is nearly saturated with a larger traffic flow under specific conditions. In the future, real-time surveillance system and Dynamic tuning system will be more viable, and reformation about Traffic Control System will appear sooner or later.

References

1. Hossain, M.: Capacity estimation of traffic circles under mixed traffic conditions using micro simulation technique. Transportation Research Part A 33, 47–61 (1999)
2. Brilon, W., Stuwe, B.: Capacity and design of traffic circles in Germany. Transportation Research Record 1398, 61–67 (1990)
3. Xiang, Q.J., Wang, W., Chen, B., Chang, Y.L., Gao, H.L.: A study of the model construction for roundabout's capacity. China Journal of Highway and Transport 12(4), 69–72 (1999) (in Chinese)
4. Shuangquan, Li, J., Du, P.: Roundabout Capacity Study with Simulation Technology. Journal of Highway and Transportation Research and Development 2, 95–98 (2004) (in Chinese)
5. Liu, H., Deng, W., Li, D.: Research on signal control method for complex roundabout. Road Traffic & Safety 8(2), 39–43 (2008) (in Chinese)
6. Chen, L., Guo, F.-X., Chen, C.: Study on the spatial organization optimum of annular intersection. Technology & Economy in Areas of Communications (2), 108–110 (2006) (in Chinese)
7. Wang, Z.: Research on Microcosmic Simulation of Traffic in City Road. Journal of Beijing Institute of Civil Engineering and Architecture 20(3), 54–58 (2004) (in Chinese)
8. Zhang, J., Ma, J., Liu, M.: Capacity of Roundabout Based on Probability Theory in City. Journal of Beihua University (Natural Science) 6(4), 541–543 (2003) (in Chinese)
9. Feng, Y., Chang, Y.: The Capacity Research of Roundabout. Journal of JiangSu University of Science and Technology 21(3), 91–94 (2000) (in Chinese)
10. Kimber, R.M.: The traffic capacity of roundabouts. TRRL report, LR942. Department of Environment. Department of Transport, UK (1980)

Characteristics of Heat Transfer and Resistance of Double Chevron Plate Heat Exchanges with Different Corrugation Pitch

ChunSheng Guo, WenJing Du, and Lin Cheng

Institute of Thermal Science and Technology, Shan Dong University, Jingshi road No 17923, 250061 Jinan city, Shandong province, China
csguo83@gmail.com, wjdu@sdu.edu.cn, cheng@sdu.edu.cn

Abstract. A mathematical model is developed in algorithmic form for the steady-state simulation of plate heat exchangers with generalized configurations. In present works, the thermal and hydraulic performances of working fluid in three corrugated channels have been investigated, numerically by a three-dimensional model with a new-type corrugation Plate Heat Exchanger (PHE)—double chevron PHE (DCPHE). The heat transfer character, the flow resistance and the Performance Evaluation Criteria (PEC) of the working fluid in double chevron PHE have been discussed; Results showed that the double chevron PHE with 9mm of corrugation pitch was better one.

Keywords: plate heat exchangers, numerical simulation, heat transfer, resistance.

1 Introduction

For being compact, easy to clean, efficient and very flexible, the plate heat exchanger (PHE) is widely employed in food, oil, chemical and paper industries, HVAC, heat recovery, refrigeration, etc.

The simplified thermal modeling of a PHE in steady state yields a linear system of first order ordinary differential equations, comprising the energy balance for each channel and the required boundary conditions. The main assumptions are as follows: plug-flow inside the channels, constant overall heat transfer coefficient throughout the exchanger, uniform distribution of flow in the channels, no heat loss and no heat exchange in the flow direction. Many reported studies include flow visualization by Focke and Knibbe [1], Kanaris et al. [2]; analytical models by Martin [3], Wang and Sunden [4], Srihari, et al. [5]; experimental data for actual models by Rao and Das [6], Rao et al. [7]; empirical correlations by Islamoglu and Parmaksizoglu [8],Wurfel and Ostrowski [9]; and numerical solutions by Croce and D'Agaro [10], Kanaris et al. [11], and Jain et al. [12].

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 169–174.
springerlink.com

Nomenclature

A	cross-section area of the cross-corrugated channel,m^2	L	Length of the channel
b	Gap between two consecutive plates, m	p	Pressure, Pa
C_p	Specific heat, J kg^{-1} K^{-1}	p_r	Prandtl number
D_h	Channel hydraulic diameter, m	Nu	Channel Nusselt number, $Nu=h_m D_h/k_m$
f	Darcy friction factor	Re	Channel flow Reynolds number, $Re=u_m D_h \rho_m/\mu_m$
h_m	Channel average heat transfer coefficient, W m^{-2} K^{-1}	T	Temperature, K
k	Thermal conductivity, W $m^{-1}K^{-1}$	u_m	Mean velocity of the fluid in a channel, m s^{-1}
ρ	Fluid density, kgm^{-3}	μ	Dynamic viscosity, Pa s

In recent years, the development of new plate-type heat exchanger has not been given enough attentions. The resistance factor of the new double corrugation plate heat exchanger developed by Zhijian Luan and Guanmin Zhang [13] are reduced by about 50% compared with the traditional chevron plate heat exchanger, but their heat transfer characteristics is also reduced by about 20%.

The present study aims to analyze a new type PHE—double chevron PHE with 3D CFD mode from their heat transfer characteristics and resistance characteristics.

2 Simulation

2.1 CFD Modeling

Preprocessing, the primary step of the CFD simulation was done by ICEM CFD13.0. It includes defining geometry, mesh generation etc. Geometry was defined by creating the required domain of known dimensions. Mesh generation was done by trial and error till the selected mesh could cover all the nodes of the domain. The fluids that are present in the flow geometry were defined. The boundary conditions and material properties of the geometry were specified. Then the domain was exported to ANSYS CFX 13.0. In solver, Finite elements method was employed to solve the set of governing equations numerically. Solver part solves the flow problem using numerical techniques like differencing schemes and discretization techniques. The final step in CFD, the post-processing was carried out to refine the simulation data for visualizing the vector plots, contour plots and particle tracking.

2.2 Defining the Geometry

Three different 4-plate PHE domains are created with the dimensions plate thickness = 0.61 mm, plate width = 170 mm, plate length = 390 mm, channel spacing = 2.5 mm; and double chevron plate structure is divided into big and small corrugations as shown in Fig.1.

2.3 Defining the Geometry Mesh Generation

Tetrahedral Meshing with spacing of 0.7 mm was applied to the plate volumes. Number of meshed volumes created on double chevron PHE were about 7,750,000.

2.4 Specifying Material Properties and Boundary Conditions

The material present is defined as a fluid. The hot and cold inlet boundary conditions were set as velocity inlets, and the outlets were set as pressure outlets. The metal plates were modeled as thin walls with the thermal resistance of a 0.61 mm stainless steel. All the exterior walls were modeled as adiabatic. The fluid domains were modeled with the properties of water.

The simulation was solved using RNG k–ε turbulence model. The problem was numerically solved using the finite elements method with the software ANSYS CFX 13.0. The simulations were carried in a Pentium 8 work station with 12 GB RAM. Each simulation took approximately 45 h to converge with a criterion of 1×10^{-5}. The simulations were done for sets of hot and cold flow rates for the given inlet temperature of 293 K for cold and 323 K for hot fluids.

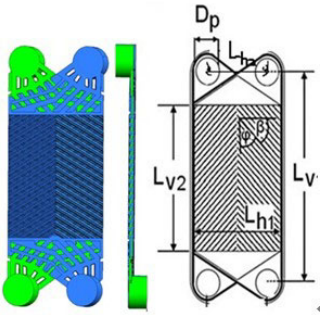


Fig. 1. Geometrical model of double chevron PHE

3 Results and Discussion

The Y-axis of Fig.2 is Nu/Nu_o, among which Nu_o =7.54, which is the Nusselt number of infinite plate of uniform wall temperature. Wherein the experimental result of chevron PHE is derived from the experimental formula put forward by Muley [14], which takes entrance section and fluid distribution into consideration. It can be seen from the figure that the Nu/Nuo of the double chevron PHE with β=60° is 5～20 times as that of the infinite smooth plate. And within the scope of Re=500～6000, it is about 15～25% higher than the experimental result of chevron PHE with β=60°conducted by Muley. At same times, the Nusselt number of DCPHE with 9mm of corrugation pitches was higher than three other.

The Y-axis of Fig.3 is f/f_o, among which f_o=96/Re is the flowing fictional resistance coefficient of infinite plate. As can be seen from the figure, the fictional resistance coefficient of the double chevron PHE with β=60° is 20～90 times as that of the infinite

smooth plate. Within the scope of Re=500～6000, it is about 10～15% lower than the experiment result of chevron PHE with β=60° conducted by Muley. At same times, the fictional resistance coefficient of DCPHE with 8mm of corrugation pitch was lower than three other.

Fig.4 shows the comprehensive performance evaluation number (PEC) of strengthening heat transfer of double chevron PHEe and traditional chevron PHE changes along with the change of Reynolds number. Among which PEC=$(Nu/Nu_o)/(f/f_o)^{1/3}$ can be seen from the figure, and with the same pumping power consumption, the heat transfer strengthening index of chevron PHE with β=60° is 20%～25% than the experiment result of chevron PHE with β=60°conducted by Muley. Therefore, compared with the widely used chevron PHE, the double chevron PHE has the strengthened heat transfer effect that is 20% higher. At same times, PEC coefficient of DCPHE with 9mm of corrugation pitch was higher than three other.

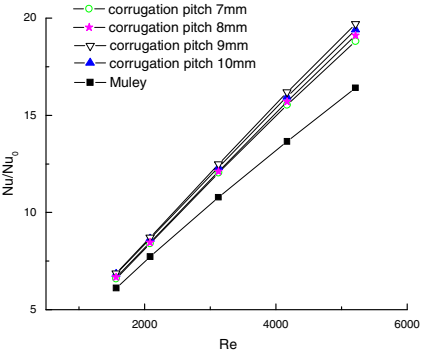


Fig. 2. Variation of Nu/Nu_o with Re

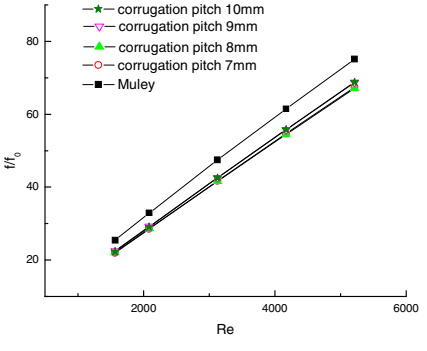


Fig. 3. Variation of f/f_o with Re

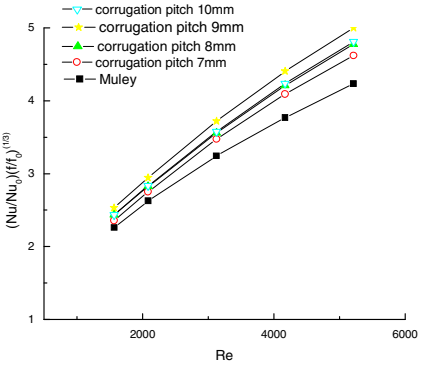


Fig. 4. Variation of PEC with Re

4 Conclusion

The application of the Nusselt number and the flowing fictional resistance coefficient evaluation number to a heat exchange process is considered, and the comprehensive performance evaluation number is used to quantify the deviation between the practical heat exchange process and the ideal one. This number is examined by the numerical simulation of the flow and the heat transfer in DCPHE. The results show that compared with the widely used chevron PHE, the double chevron PHE has the strengthened heat transfer effect that is 20% higher.

References

1. Focke, W.W., Knibbe, P.G.: Flow visualization in parallel-plate ducts with corrugated walls. J. Fluid Mech. 165, 73–77 (1986)
2. Kanaris, A.G., Mouza, A.A., Paras, S.V.: Flow and heat transfer in narrow channels with corrugated walls a CFD code application. Chem. Eng. Res. Des. 83(A5), 460–468 (2005)
3. Martin, H.: A theoretical approach to predict the performance of chevron plate heat exchangers. Chem. Eng. Process. 35, 301–310 (1996)
4. Wang, L., Sunden, B.: Optimal design of plate heat exchangers with and without pressure drop specifications. Appl. Therm. Eng. 23, 295–311 (2003)
5. Srihari, N., Rao, B.P., Sunden, B., Das, S.K.: Transient response of plate heat exchangers considering effect of flow maldistribution. Int. J. Heat Mass Transfer 48, 3231–3243 (2005)
6. Rao, B.P., Das, S.K.: An experimental study on the influence of flowmaldistribution on the pressure drop across a plate exchanger. J. Fluids Eng. ASME Trans. 126, 680–691 (2004)
7. Rao, B.P., Sunden, B., Das, S.K.: An experimental investigation of the port maldistribution in small and large plate package exchangers. Appl. Therm. Eng. 26, 1919–1926 (2006)
8. Islamoglu, Y., Parmaksizoglu, C.: The effect of channel height on the enhanced heat transfer characteristics in a corrugated heat exchanger channel. Appl. Therm. Eng. 23, 979–987 (2003)
9. Wurfel, R., Ostrowski, N.: Experimental investigations of heat transfer and pressure drop during the condensation process within the plate heat exchangers of the herringbone type. Int. J. Therm. Sci. 43, 59–68 (2004)

10. Croce, G., Agaro, P.D.: Numerical analysis of forced convection in plate and frame heat exchangers. Int. J. Numer. Methods Heat Fluid Flow 12(6), 756–771 (2002)
11. Kanaris, A.G., Mouza, A.A., Paras, S.V.: Flow and heat transfer prediction in a corrugated plate heat exchanger using a CFD code. Chem. Eng. Technol. 29(8), 923–930 (2006)
12. Jain, S., Joshi, A., Bansal, P.K.: A new approach to numerical simulation of small sized plate heat exchangers with chevron plates. J. Heat Transfer, ASME Trans. 129, 291–297 (2007)
13. Luan, Z.-J., Zhang, G.-M., Tian, M.-C., Fan, M.-X.: Flow resistance and heat transfer characteristics of a new-type plate heat exchanger. Journal of Hydrodynamics 20(4), 524–529 (2008)
14. Muley, A.: Heat Transfer and Pressure Drop in Plate Heat Exchangers [PH.D]. University of Cincinnati, Cincinnati (1997)

Perceived Risk in Marketing Strategy

Liu Tian-Que

The Australian National University, Canberra, ACT 0200, Australia
tianque0928@live.cn

Abstract. 'Perceived risk' in consumer behavior theory helps explain why often consumers do not move from the desire stage to the action stage, that is, make the actual purchase decision. Appropriate theories and practice can be used to identify what strategies marketers can employ to overcome this problem. However, there are still some gaps existing when applying these theories in practice. Future studies could be more related to psychology and focused on the risk perception of consumers who shop for goods at the virtual storefront.

1 Introduction

Perceived risk, the level of risk a consumer believes exists regarding the purchase of a specific product from a specific retailer, has been considered as influencing the consumer purchase decision process in marketing [1]. Proper marketing strategies such as positioning strategy and using the knowledge of consumer behavior to affect consumer decision making process can be used to overcome the problem that customers do not move from desire stage to action stage. The positioning strategy theories and practice illustrate employing a positioning strategy marker will help reduce perceived risk in high involvement purchase, and the practice and theories of consumer behavior indicate a marketer who is doing well in the study of consumer behavior can be employed to reduce the perceived risk in low involvement purchase. However, some theories related to the proposition that the smaller perceived risk is always preferred and the measure of the degrees of perceived risk may differ from the reality in applying such strategies. Future study in this area can be linked to psychology and focused more on the risk perception of consumers who shop for goods at the virtual storefront. Therefore, the essay will first focus on identifying the strategies marketers who can be employed to overcome this problem by using theories and practice in respects of low and high perceived risks, then pointing out the gaps between theories and practice, and finally giving some recommendations about how these theories should be improved and what direction academic research should be taken in future to more effectively reach the consumer.

2 Result and Discussion

The strategy marketers who focused on the area of developing positioning market strategy could be employed to reduce the high perceived risk, because they can educate consumers about why their product is the best choice in competing with other

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 175–178.
springerlink.com

brands in advance of his decision made. The ELM (elaboration likelihood model) indicates that before buying a high-involvement product such as a house or a car, consumers are likely to process all the available information and to have thought about the decisions carefully [6]. The brand personality, which is a distinctive image of a good's or service's character or benefits can not only attract customers' attentions faster but also leave a deep impression on them when they searching information [6]. Building a brand personality for a firm can make its product become unique and outstanding among all the products that are being considered by a customer, and thus reduce the perceived risk. For example, BMW Company launched the new generation car Mini Cooper in 2001, and created a personality of brave, active and full of vigor for this product by considering about its color, shape and performance [4]. [1]Results showed this product was Britain's seventh bestselling car in June 2010 [3], so it is clearly to see that the strategy marketers in this company successfully emphasize the characteristics of the car to customers and educate them why their product is the best comparing with other products.

Furthermore, positioning is a marketing method, through which positioning marketers try to create an image or identity for a product, brand or company in the perception of the target market [6]. They often create a perception of high quality, reasonable price and great credibility to reduce the perceived risk. For instance, Apple has established a unique reputation in the consumer electronics industry. One of the most important reasons is that the perception of the products created by Apple Company attracts a huge amount of loyal fans, and their devotion keeps this perception continues [10]. Therefore, developing a successful positioning market strategy make a great effect on reducing perceived risk in high-involvement purchase.

In respect of low perceived risk, the strategy marketers who focused on the area of consumer behavior can be employed since they can influence the decision makers by displaying the products at the time of purchase. The ELM shows that in low involvement situation, consumers make little or conscious effort in searching for information and purchases automatically, which is because they are not overly concerned about which option is chosen [6]. Through studying the internal influences such as psychological and personal influences on consumer behavior, strategy markers can develop appropriate marketing strategies to affect the customers' habits of shopping. The psychology of how consumers think, feel, reason, and select between different alternatives are important factors in developing such marketing strategy [7]. For example, by understanding that new products are usually initially adopted by a few consumers and only spread later, and then gradually, to the rest of the population, strategy marketers learn that companies that introduce new products must be well financed so that they can stay afloat until their products become a commercial success. Besides, it is important to please initial customers as they will in turn influence many subsequent customers' brand choices [7].

In addition to this, strategy markers can also use the external influences such as social and culture influences on consumers to reduce perceived risk. The effects of advertisement can be a good example in this case. Strategy marketers scheduled snack advertisements late in the afternoon [7] by understanding that consumers are more receptive to food advertising when they are hungry. Overall, through studying consumer behavior in the ways of internal and external influences, marketing

strategies can be developed to affect consumers' purchasing habits and thus reduce perceived risk.

However, there exist gaps between theories and practice. Firstly, in general, when using such strategies to overcome the problem that people do not move from desire stage to action stage, it is assumed that people prefer smaller risks to larger ones. But this statement is not always true in practice, it is considered to be true only in the case provided that other factors are constant such as expected value [5]. Additionally, when considering about if the product perceives a low or high perceived risk, the price of the product cannot be considered as the only measure. For example, a rich man may see a house as a low involvement product while a poor may see toilet paper as a high involvement product. Thirdly, the ELM explains the customer decision-making process in the situations of high and low involvement, but most decisions fall in the middle, and they cannot be simply determined to be solved by any specific marketing strategy. That is because for this level of involvement, customers often do some work to make a decision but not a great deal for them [6]. For instance, when people decide what movie they want to see, they often simply rely on "rules of thumb" instead of searching all the information or making habitual decisions. Therefore, as the degrees of the perceived risk are various appeared in practice while they are limited in theories, the problem of reducing perceived risk cannot always be well solved by positioning strategies marketers or consumer behavior marketers.

Some improvement should be made and some direction academic researches should be taken in future to make these theories more accurate and comprehensive. First of all, the study of reducing perceived risk in marketing can be combined with psychology. The psychology approach in the area of risk perception began with research in trying to understand how people process information, which is exactly consistent with the consumer decision-making process in consumer behavior [9]. Secondly, internet shopping has become the most popular way of shopping around the world. Although in past marketing literature, research has shown that the use of certain risk reduction strategies such as brand reputation, product trial, and warranty are successful in reducing the risk perception of consumers, these strategies cannot be effective the same way for virtual shoppers [8]. In fact, most of the existing literature concentrated on explaining the advantages and disadvantages of internet marketing, but only a few addresses the issues raised about the consumers' concerns in virtual shopping [8]. Therefore, more attention on the online perceived risk related to the online consumer behaviour should be paid and future study can be more combined with psychology.

3 Conclusion

In conclusion, consumer perceptions of risk have been widely dealt with in past literature as they accompany all purchases to varying degrees and influence buying behavior [2]. Developing successful positioning strategies and marketing strategies based on the theories of consumer behavior can help reduce perceived risk in the high and low involvement purchase respectively. However, there are still some gaps existing when applying these theories in practice. The study of combining perceived risk in marketing with it in psychology and consumers' perceptions of risk on

purchase intention in E-Shopping could be the future directions of this research. All in all, customers may perceive a certain degree of risk in most purchase decisions, but a successful marketing strategy in positioning or consumer behavior can help overcome this problem effectively.

References

1. Ashely: Understanding your customer's perceived risk. University of Connecticut (1995), http://www.hort.uconn.edu/agmarketing/distlrnarticles/percrisk.htm (viewed April 22)
2. Ben-Ur, Winfield: Perceived risk in the E-Commerce environment. University of Houston-Victoria (2006), http://www.sbaer.uca.edu/research/swma/2000/04.pdf (viewed April 22)
3. British Car Sales, the Sun News (2010), http://www.thesun.co.uk/sol/homepage/news/3043609/British-car-sales-up-20-in-a-year.html (viewed April 23)
4. Cooperate News, BWM Group (2001), http://www.bmwgroup.com/e/nav/index.html?, http://www.bmwgroup.com/e/0_0_www_bmwgroup_com/produktion/produktionsnetzwerk/produktionsstandorte/werk_swindon.html (viewed April 23)
5. Jia, Dyer, Butler: Measures of perceived risk. The Chinese University of Hong Kong (1997), http://www.mccombs.utexas.edu/faculty/jim.dyer/per_risk.PDF (viewed April 24)
6. Perner: Introduction to marketing. University of Southern California (1999), http://www.consumerpsychologist.com/marketing_introduction.html (viewed April 23)
7. Perner: Consumer behavior: the psychology of marketer, USC Marshall. University of Southern California (2010), http://www.consumerpsychologist.com/ (viewed April 23)
8. Samadi, Yaghoob-Nejadi: A survey of the effect of consumers' perceived risk on purchase intention in E-Shopping, Business Intelligence Journal (2009), http://www.saycocorporativo.com/saycoUK/BIJ/journal/Vol2No2/article1.pdf (viewed April 24)
9. Tversky, Kahneman: Judgment under uncertainty: heuristics and biases, New Series (2007), http://www.hss.caltech.edu/~camerer/Ec101/JudgementUncertainty.pdf (viewed April 24)
10. 800,000 iMacs Sold in First 139 Days, Apple Store (2008), http://www.apple.com/ca/press/1999/01/iMac_Sales.html (viewed April 23)

Discussion of EV Virtual Value Calculation Method for Light Environment Test System and Influence Factors of Calibration Experiment under Skylight

Yuan Zou[1,2,*] and Tianxing Shen[1]

[1] School of Architecture, Tianjin University, 300072 Tianjin, China
[2] School of Civil Engineering and Architecture, Wuhan University of Technology, 430070 Wuhan, China
yantzed@sina.com

Abstract. During calibration and measure process, to enhance scientificity and controllability of the light environment test system, taking the digital camera used for picking up images in the system as research object, we studied calculation and deduction steps of EV virtual value comprehensively, calculation method and image selection procedures are also proposed. 2 main influence factors of the calibration experiment under skylight, including dynamic range and alternation of sun's position are explored; meanwhile, the mechanism and prevent strategies of the factors are analyzed as well; the study has positive significance to improve stability and accuracy of the light environment test system.

Keywords: Dynamic range, color coordinates, light environment test system, calibration.

1 Introduction

As we know, physical environment of buildings and cities is closely associated with people's health and life qualities. As an important component of the physical environment, luminous environment impacts people's physical and psychological health greatly [1]. Therefore, based on the research of light environment physical properties, we can evaluate and design the light environment appropriately. At the moment, the digital measuring technologies in light environment testing field are particularly important. Blending with computer and digital imaging technologies, composed by light information collecting and processing modules, SM light environment test system is invented as a set of experiment devices which can provide services of overall testing indicators of light environment. Light information collecting module contains a calibrated digital camera and corresponding fisheye lens, these equipments take on the mission of image obtaining for measure and compute. During the process of system calibration, the variations of aperture and the shutter speed values are tightly related to the operation status of light environment test system [4].

* Corresponding author.

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 179–185.
springerlink.com

2 EV Virtual Value

Exposure value (EV) denotes all combinations of a camera's shutter speed and relative aperture that give the same exposure, as a comprehensive evaluation index of the object brightness and the camera sensibility, it is also used to indicate an interval on the photographic exposure scale. The theoretical formula of exposure value can be founded, as in

$$EV = \log_2\left(\frac{F^2}{T}\right) \tag{1}$$

"F" represents aperture value; "T" represents shutter speed. Meanwhile, EV virtual value can be defined as valid exposure value range of calibrated digital camera. Determining the value ranges of "F" and "T" can result in EV virtual value calculation.

First, a group of calibrating images were collected. In these images, a standard color board was shot as the object, the calibrating process is completed in the skylight situation. During this experiment, Canon EOS 5D was chosen as the photographic equipment. 12 images were recorded with different F-T value combinations, shown as Fig. 1. At the same time, a group of luminance values corresponding to the shooting time should be measured and recorded with a light meter, "L" was defined as the group of luminance values [5].

Fig. 1. Calibrating images exposed in skylight with various F-T value combinations

Then 144 groups of RGB values were extracted from the images in Fig. 1, a group of "D", the gray values, can be calculated with the conversion formula of RGB values and tristimulus values [1], as in

$$D = Y = 0.299R + 0.587G + 0.114B \tag{2}$$

According to the imaging principle of camera, when camera lens is fixed, the exposure "H", aperture values "F", shutter speed "T", and measured luminance values "L", will obey the following functional relationship:

$$H = \frac{T \cdot L}{F^2} \tag{3}$$

We put gray values "D" and logarithm of exposure "100lgH" into Cartesian rectangular coordinate, and fit a typical digital camera photosensitive curve named ABCD, which is comprised of discrete points.

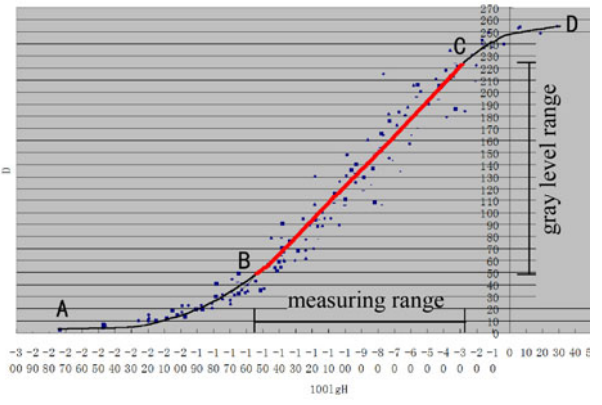


Fig. 2. Relationship between gray values "D" and logarithm of exposure "100lgH"

Curve AB represents underexposure range, Curve CD represents overexposure range, only BC, the straight line, is the valid range of digital camera. So the normal exposure values of light environment measuring system should be limited in this range. A conventional linear equation can be concluded by extracting Line BC, shown as the following expression.

$$y = ax + b \tag{4}$$

In the linear equation, logarithm of exposure "100lgH" is assigned to variable "x", gray value "D" is assigned to variable "y", The interval of vertical axis means the valid gray level range, while the interval of horizontal axis means the tolerance of digital camera, which can be regarded as equivalent luminance range of light environment measuring system. Now, we can deduce the measuring range of actual luminance from Formula 3 and Formula 4, "ΔL_{max}" is defined as this range, as in

$$\Delta L_{\max} = \frac{F^2}{T} 10^{\frac{(D_{\max} - b)}{100a}} - \frac{F^2}{T} 10^{\frac{(D_{\min} - b)}{100a}} \tag{5}$$

In this formula, "D_{max}" and "D_{min}" separately represent upper and lower limits of the gray level range, considering Formula 1, 3, and 4 simultaneously, exposure value of single point in the image can be concluded, as in

$$EV = \log_2\left(\frac{F^2}{T}\right) = 3.22(\lg L - \frac{D-b}{100a}) \tag{6}$$

During the actual calibration process, gray series ($D_1, D_2 \ldots D_m$) converted from RGB series should be sorted by order, and be corresponding to the series of luminance measuring values "L". We can take numerous points from gray level range, and

assign them to Formula. 6, then the EV average value is to be deduced, shown as Formula 7.

$$\overline{EV} = 3.22\sum_{i=1}^{n}(\lg L_i - \frac{D_i - b}{100a}) \Big/ n \tag{7}$$

EV average values can be transformed into EV virtual values by rounding method. Once EV virtual values are determined, appropriate camera shooting parameters will be easily chosen according to the corresponding relationship of aperture values "F", shutter speed "T" and EV values. By this way, value combinations of aperture, shutter speed and ISO can be more flexible. During shooting process of measure and calibration, a summation of EV corrections from 3 parameters is usually regarded as the calibrated EV virtual value.

Taking the calibrating images of Fig.1 as the research objects, to determine EV virtual value and eliminate invalid calibration photos, we figure out the calibrated linear equation which can be written as "y = 1.3774 x + 259.86" by fitting Line BC. According to the values from coordinates, gray scale range is limited from 50 to 225. Then we just choose one group of points, and assign them to the Formula 7 to calculate EV virtual value which equals to 9 in this situation. While the images are selected, by reading the EXIF information of the images to acquire F and T, exposure value of every image is calculated according to formula. 1. Based on EV corrections of F and T from photographic documents, we can also figure out EV valuations which are compared with EV virtual values to identify invalid images. The results of calculations and selections are shown in Table 1.

Table 1. Calculations and selections of various parameters

Image ID	T	F	T corrections	F corrections	EV valuations	EV virtual values
IMG_0091	1/10	5	3	5	8	9
IMG_0092	1/10	6.3	3	5	9	9
IMG_0093	1/10	10	3	7	10	9
IMG_0094	1/10	16	3	8	11	9
IMG_0095	1/10	22	3	9	12	9
IMG_0089	1/6	4	2	4	6	9
IMG_0090	1/10	4	3	4	7	9
IMG_0088	1/15	4	4	4	8	9
IMG_0087	1/25	4	5	4	9	9
IMG_0086	1/50	4	6	4	10	9
IMG_0085	1/100	4	7	4	11	9
IMG_0084	1/160	4	7	4	11	9

From the table, we can see IMG_0087 and IMG_0092 are valid images, if the rest of images are adopted; they may cause lower precision of the system, these invalid calibrating or measuring photos are suggested to get excluded. During calibration, even if the exposure value equals to EV virtual value, the photos are not always

recorded normally. When manual filming mode is adopted in dark environment, inappropriate low shutter speed will bring the images some negative effects including motion blur, jitter fuzzy, delay shooting, and some others which will have significant influence on testing data once they appear. In short, normative experiment operation steps are as important as EV virtual values, both of which are necessary conditions to obtain high quality test data.

3 Analysis of the Influence Factors during Calibration and Measure in Skylight

In the scenes of indoor test environment, intensity and angles of skylight often change quickly. Considering the variety of the building envelopes, asymmetrical indoor intensities of illumination distribution, large dynamic range and other similar situations, we boil down calibration interferences in skylight situation to 2 main influence factors, which are dynamic range including backlight, and sun positions of sunny days.

3.1 Dynamic Range and Backlight

Dynamic range is the ratio between the largest and smallest possible values of a changeable quantity, such as in sound and light. It is usually expressed in D-value or ratio. Photographers use "dynamic range" for the luminance range of a scene being photographed, or the limits of luminance range that a given digital camera or film can capture. In reality, dynamic ranges in many scenes are large, such as in backlight scene. The dynamic ranges are usually beyond the tolerance of digital camera in these situations. Therefore, for the purpose of calibration and data test in the skylight, testers should avoid shooting or filming photos in scenes with large dynamic ranges, and scenes with backlight or high brightness backgrounds [3].

In backlight scenes, objects are located between light sources and cameras. In this situation, the inadequate exposure of objects will be easily caused. While shooting in large dynamic ranges, light with high intensity goes through lens and projects onto CCD of a digital camera. Since the upper limit of photo sensitivity is determined by the brightest part of the scene, the upper limit increases naturally. The object with normal colors and light intensities is in backlighting status at this moment, the luminance is much lower than light with high intensity, if the gray maximum "D_{max}" representing the upper limit of photo sensitivity is taken as 255, gray levels of the object will be in the relatively lower positions of the whole gray range, and object gray scales will be compressed into a shorter interval than in normal dynamic ranges. So, calibrating or testing with light environment test system in backlighting status will lead to poor measure precisions and continual fiducial errors. As we know, fiducial error is one of important evaluation indexes of experimental facilities accuracy. In our case, the valid range of gray level "D" plays a significant role, if the range gets too narrow, it means that the measuring range, the denominator of fiducial error gets narrow correspondingly, the precisions of the system decrease automatically.

During calibration or measure process, if scenes get into large dynamic ranges and backlight situations, according to the importance of object divisions, photos should be

taken partially, sectionally, and separately. Of course, operations of experimental facilities and equipments layout should meet the following requirements. (1) Angle interval of the standard color board and horizontal plane should be 45°~ 90°, angle of vertical projection plane and window normal should not be more than 45° from both sides. (2) During filming process, testers should try to reduce unnecessary walking, and prevent the interference of skylight. Between the light source and color board, testers should keep the lighting path clear and unobstructed. (3) Angle of luminance meter and vertical projection plane normal should not be greater than 45°. (4) During the calibrating process of image collection, testers should keep the camera and luminance meter in the same direction. To ensure that the skylight path is clear and unobstructed, testers should appropriately locate and place test equipments as well.

3.2 Influences of Sun Position in Sunny Day

According to the past measured data, air transparency, sun's orientation and relative positions of calculation points affect luminance distribution of clear sky comprehensively. The brightest point is near the sun, the lowest luminance point is on solar meridian of celestial sphere, which has an angle of around 90° with sun [7].

From existing data, we know that the time range from 10:00 a.m. to 14:00 p.m. is the best calibration period. As the position of sun changes constantly, the orientation of the building used for measuring has significant effects on the calibrating results, window positions of calibration experiment can be calculated with formula of solar altitude and solar azimuth [6]. In this case, test location is Tianjin area (39°N 117°E), declinations are vernal equinox, summer solstice, autumnal equinox, and winter solstice, testing time is from10:00 a.m. to 14:00 p.m.. The calculations of solar altitude and solar azimuth are shown in Table 2.

Table 2. Solar orientations of various declinations in Tianjin area

Position	Tianjin (39°N 117°E)							
Solar terms	Vernal equinox		Summer solstice		Autumnal equinox		Winter solstice	
Declination	0°		23°27'		0°		-23°27'	
Local time	10:00	14:00	10:00	14:00	10:00	14:00	10:00	14:00
Hour angle	-30°	30°	-30°	30°	-30°	30°	-30°	30°
Solar altitude	42.3°	42.3°	60°	60°	42.3°	42.3°	21.5°	21.5°
Solar azimuth	-42.5°	42.5°	-67.9°	67.9°	-42.5°	42.5°	-29.5°	29.5°

From 10:00 a.m. to 14:00 p.m. in a year, solar azimuth changes in the range of -67.9° ~ 67.9°, which is symmetrical from south, it has the range of -45° ~ 45° in more than three-quarters of a year. From the historical data, it is known that he lowest luminance point appears at the corresponding place in north sky. Therefore, during the process of laboratory environment selection, northern angle interval of ±45° is confirmed to be the appropriate orientation, windows of the other directions will be easily affected by solar radiation without any efficient sun shading ways. Using three

parameters calculation method, we can also determine orientation angle intervals of any other locations during measure time range.

4 Conclusion

From the above presentation in the paper, we discussed the research on EV virtual value calculation process of light environment test system, proposed theoretical calculation method of EV virtual value, and established image extraction procedures scientifically. At the same time, 2 main influence factors of the calibration experiment under skylight including dynamic range and alternation of sun's position which reflect on their mechanism and prevent strategies are explored; the study has positive significance to improve stability and accuracy of the light environment test system.

References

1. Qicheng, J., Shulan, J., Bolin, Y.: Colorimetry. Science Press, Beijing (1979)
2. Jiaping, L.: Architectural Physics. China Architecture & Building Press, Beijing (2000)
3. Dingchuan, L.: Digital SLRS Imaging Principle. Image Technology 2, 29–32 (2007)
4. Bing, G., Qinxuan, Z.: Research on Experiment of Luminance Distribution Measurement by Using a Digital Camera. China Illuminating Engineering Journal 01, 15–18 (2003)
5. Tianxing, S., Jiuhong, Z.: Analysis of the Daylighting Influence Values of Outdoor Obstacles by Using Digital Image Processing Technologies. China Illuminating Engineering Journal 16(4), 1–4 (2005)
6. Manav, B.: An Experimental Study on the Appraisal of the Visual Environment at Offices in Relation to Colour Temperature and Illuminance. Building and Environment 42, 979–983 (2007)
7. Danny, H., Li, W.: Daylight and Energy Implications for CIE Standard Skies. Energy Conversion and Management 48, 745–755 (2007)

The Research on the Rectification System of Stereo Image Pairs

Qinyou Yang[1,2] and Xiaoyan Yu[1,2]

[1] Mudanjiang Normal University, Heilongjiang Province, China
[2] Harbin Normal University, Heilongjiang Province, China
yangqinyou@gmail.com

Abstract. Stereo image rectification is a preparation of the stereo image pair matching. It is a quick matching method, which can use to increase matching speed. Its result will directly affect the accuracy of the stereo matching. The rectification system of stereo pairs based on the small vision system (svs) is introduced in this paper .It uses the data from *.ini (produced by the svs in the procedure of rectification). For the rectification of stereo images, it adopts the undistorted points to obtain the counterpart by the arithmetic method. The method is also superior to implementation, low calculation complexity.

Keywords: Stereo image, the arithmetic of rectification, SVS.

1 Introduction

1.1 Stereo Image Correction Research Significance

Image correction is stereo image to match a preparation, fast stereo matching need through the high speed and high precision image correction of work to do. In theory, stereo image correction refers to the distortion image processing to the recovery from the sex that two images are projective transform, making the plane two images corresponding to a line at the same level of article,which be mapped to infinity pole. As in the very DuoZhong cases are easy to cause the distortion of the stereo image such as: imaging system of existence as poor, distortion, bandwidth limited caused by such as image distortion; Imaging devices shooting posture and scanning the image geometric distortion caused by nonlinear; Motion blur, radiation distortion, introducing the noise of the image distortion caused by, so stereo image correction of a lot of things in a wide range of applications. This is also the stereo image correction important meaning.

1.2 Stereo Image Correction the Present Situation of the Research

Now the correction methods are mostly based on image geometric projection values of the: the first for the image of the basic matrix and the camera, within the parameters, again through the transformation of the image pole mapping to infinity. For each image display device, benchmark images showing on the picture; According to the reference images and measurement show characters; By measuring parameters, calculating the parameters of the used for correct, using these parameters on image

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 187–193.
springerlink.com

data processing; After correction of image data will be provided to display device. This kind of research correction train of thought is very popular,and it make correction speed with a raise but ubiquitous computing capacity, correction algorithm is complex problems, and there needs to be further outer, raise the correction precision.

1.3 Research Goal

The goal of Stereo image correction study lies in to seek a kind of high efficient simple algorithm, after finishing the stereo image basic correction premise, it will reduce the existence of the outer correction, in order to improve the precision and speed of correction. This paper attempts to use Small Vision System (SVS) System produced by the corrected image, making interior parameters between the images corresponding points through the algorithm to calibration transform. Thought is trying to target the points on the image according to through some kind of algorithm to calculate the mirror of the original image of pixels (xd yd), the brightness of the value, and then assign values imaging. This line of thinking of the research objective is to ensure that corrected image, improve image as far as possible without outer calibration speed, and improve the quality of image matching and the speed of the match, so as to achieve the purpose of rapid matching.

2 Quick Stereo Image Correction Algorithm

2.1 The Existing Stereo Image Correction Algorithm

Based on the difference of the correction of ideas, the researchers used different correction method.

Hartley [1] puts forward a new method, with the change of the image point position is minimum constraint to optimize image distortion of photography. At the same time, he also puts forward the prospective affine transformation algorithm, preventing after correction of the image is divided. Gluckman [2] is using this line of thinking and photograph the determinant of the nature of the matrix changes to the distortion of the photography becomes small.

2.2 Improved Stereo Image Correction Algorithm

This article are given a cole overcome the disadvantages of two correction methods. That is to target image pixels (xu,yu), through the calibration algorithm to calculate the closest to the mirror of the original pixels (xd,yd) and finally complete work assignment. This algorithm is to ensure that the main advantages of the object image of each pixel is endowed with correspondingly brightness values, to eliminate the outer existence. It is based on the correction algorithm Small Vision System (SVS) image correction System provided reference algorithm. the basis are modified, and by using the SVS system calibration of produce ini file in the inside of the related parameters of the data. The method is simple and easy to implement, and the calibration speed, no outer existence, etc.

3 Based on VB Stereo Image Correction Algorithm Is Implemented

3.1 Program Design Process

Program flow chart is as follows:

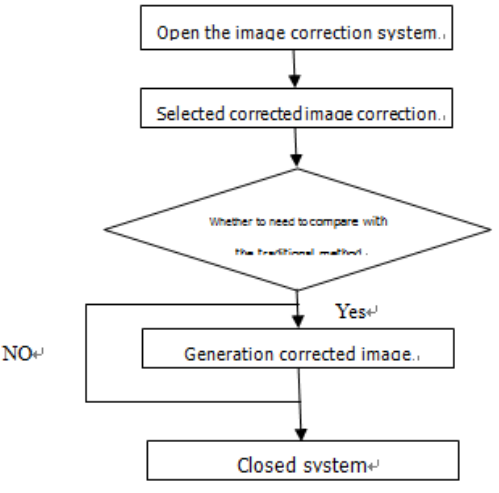


Fig. 1. Program flow chart

3.2 Interface Design

Interface image is as follows:

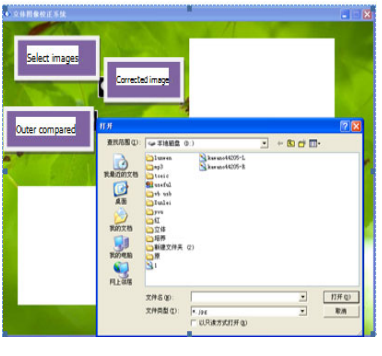


Fig. 2. Select file open

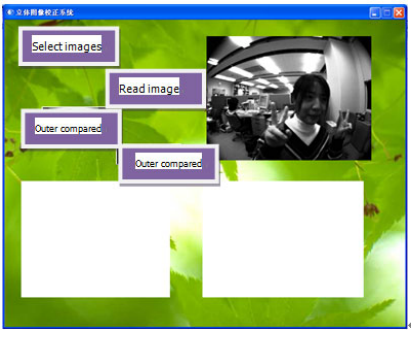


Fig. 3. Load the standard image

3.3 Program Realization

a) statement

Use the GetPixel () and SetPixel () two functions and methods, first declare functions and related use variables.

```
Option Explicit 'variables must statement
    Private Declare Function GetPixel Lib "gdi32" (ByVal hdc As Long, ByVal x As Long, ByVal y As Long) As Long    ' Declare functions
    Private Declare Function SetPixelV Lib "gdi32" (ByVal hdc As Long, ByVal x As Long, ByVal y As Long, ByVal crColor As Long) As Long   ' Declare functions
```

B) form in initial set. When

```
Private Sub Form Load()   ' Interface initialization
    Hide
    Picture1.BackColor = vbWhite
      Picture2.BackColor = vbWhite
      Picture3.BackColor=vbWhite
    Caption = " Stereo image correction system "
    Show
End Sub
```

C) use common dialog box open file

First need to add controls, the procedure is as follows: the choice of engineering → parts →Microsoft Common Dialog Control 6.0 选项，将 CommonDialogControls add to toolbox, and then create aCommonDialog1 control.

d) "Select image" button to trigger affairs

```
Private Sub command1_Click ()
On Error Go To no file   'error check
CommonDialog1.InitDir = "D:\"   ' The default path open for D disk
CommonDialog1.Filter = "*.jpgl*.bmp"   ' Open the file format for. JPG or. BMP
CommonDialog1.CancelError = True   ' Allow cancel
CommonDialog1.ShowOpen    ' Open the file dialog box
Text1.Text = CommonDialog1.FileName   ' Read path
Picture1.Picture = LoadPicture(Text1.Text) ' Into image
Picture1.AutoSize = True   'PictureBox1 According to the image size adjustment automatic height and width
no file: ‘error
If Err. Number = "32755" Then Exit Sub   ' exit
End Sub
```

e) "Corrected image" button to trigger affairs

```
Private Sub command2_Click ()
        Dim i As Long, j As Long, nSize As Long   ' Statement used variables
  Dim xu As Double, yu As Double, xd As Double, yd As Double, color As Long, ru As Double, fx As Double, fy As Double, cx As Long, cy As Double
 Dim t1 As Double, t2 As Double, k1 As Double, k2 As Double, k3 As Double, rd As Double
    t1 = -0.000876   ' According to SVS system to produce a *. Ini file set parameters values
  t2 = -0.002015
  k1 = -0.376997
  k2 = 0.199037
  k3 = -0.056334
  fx = 151.109345
```

```
        fy = 151.148059
        cx = 164.508426
        cy = 126.422274
     Picture2.AutoRedraw = True        ' According to load the image to adjust its size
  Picture2.Width = Picture1.Width   ' According to the initial image set goals image width
  Picture2.Height = Picture1.Height   'initial target image plans according to the image height
  Picture2.Cls   ' Screen clearing
          For xu = 0 To Picture2.Width   ' In Picture2 take pixel in turn on
                For yu = 0 To Picture2.Height
                ru = xu * xu + yu * yu   ' According to the correction algorithm is calculated in the
initial image of the corresponding points
          xd = xu * (1 + k1 * ru + k2 * ru * ru + k3 * ru * ru) + 2 * t1 * xu * yu + t2 * (ru + 2 * xu * xu)
          yd = yu * (1 + k1 * ru + k2 * ru * ru + k3 * ru * ru) + 2 * t2 * xu * yu + t1 * (ru + 2 * yu * yu)
         SetPixelV Picture2.hdc, xu, yu, GetPixel(Picture1.hdc, xd, yd)   ' Fu brightness values
       Next
     Picture2.Refresh   'Picture2 refresh
  Next
  End Sub
f) "Contrast" button trigger events outer
Private Sub Command3_Click ()
   For i = 0 To Picture1.Width   ' Records at the point correction
     For j = 0 To Picture1.Height
        xu = (i - cx) / fx
        yu = (j - cy) / fy
        ru = xu * xu + yu * yu
        xd = xu * (1 + k1 * ru + k2 * ru * ru + k3 * ru * ru * ru) + 2 * t1 * xu * yu + t2 * (ru + 2 * xu * xu)
        yd = yu * (1 + k1 * ru + k2 * ru * ru + k3 * ru * ru * ru) + 2 * t2 * xu * yu + t1 * (ru + 2 * yu * yu)
        xfd = cx + xd * fx
        yfd = cy + yd * fy
        udx(xfd, yfd) = i
        udy(xfd, yfd) = j
     Next
   Next
  For i = 0 To Picture3.Width   ' Pixel assignment
          For j = 0 to Picture3.Height
          Color = GetPixel(Picture1.hdc, i, j)
          SetPixelV Picture3.hdc, udx(i, j), udy(i, j), color
       Next
       Picture3.Refresh
    Next
    End Sub
```

3.4 Results

In VB6.0 environment engineering can be generated after from VB operation. Double click through the system, selection of need calibration, click on the button image correction, the image began to correction. After correction, the initial image system and corrected image as below.

Fig. 4. The initial image (adjusted before) The i image (adjusted before)

Fig. 5. Target image(After correction)

Press the button, the outer compared with the traditional algorithm improved algorithm trigger more events, through the generation picture can see clearly the existence of the outer.

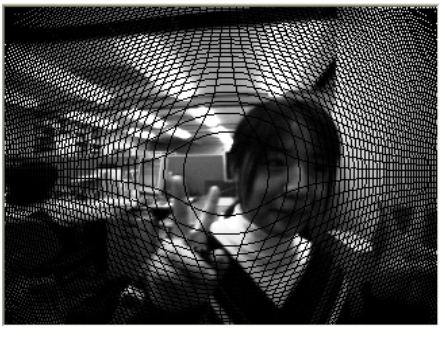


Fig. 6. The improved algorithm correction

Fig. 7. The traditional algorithm correction

3.5 Compared

Through to the before and after correction of wall beam bending degree and the overall image distortion degree contrast, can be bright to original observation of the wall beam bending and distortion part after correction has improved, make whole image the recovery. Also confirmed this paper provides stereo image correction system is stable and has a good experiment results.

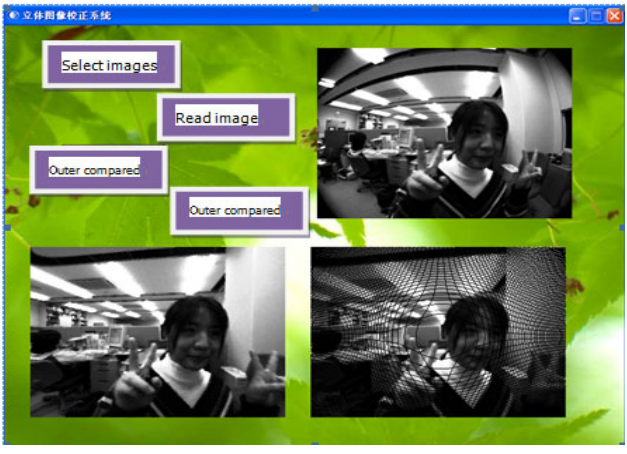


Fig. 8. Integral effect interface

4 Conclusion and Prospect

This study is based on image correction system, using SVS system to produce a *. Ini file in the inside of the data of the stereo image without outer correction. This kind of thinking makes the new correction method is simple and easy to implement, no outer existence, correction of high speed and precision, etc.

Stereo image correction is a very frontier research subject, the research content of it is extensive, the flexible methods. Especially in the help of stereo image matching is more to the majority of the staff of deepening the perfect space. Stereo image correction real practical and idealized and many problems remain to be solved. This paper provides ideas just stereo image correction in the field a new exploration, and a lot of work to carefully think we deepen study.

References

1. Hartley, R.: Theory and practice of Projective rectification. International Journal of Computer Vision 35(2), 115–127 (1993)
2. Gluckman, J., Nayar, S.: Rectifying transformations that minimize resampling effects. In: Jacobs, A., Baldwin, T. (eds.) Proceedings of International Conference on Computer Vision and Pattern Recognition, HaWaii, USA, vol. (I), pp. 111–117 (December 2001)

An Empirical Study on Industrial Homogeneity between Fujian and Taiwan Provinces by the Coefficient of Similarity and R/S*

Suqiong Wei[1,**], Yanhua Chen[2], Jinghuan Geng[1], and Liyun Zeng[1]

[1] Fujian Normal University, Fuzhou 350007, China
suqiongwei@126.com
[2] Hainan Normal University, Haikou 571158, China

Abstract. Industrial homogeneity is one of the bottlenecks in industrial upgrading with a region and in the cooperation among different regions. This paper tries contributing to the understanding of this issue, and Fujian and Taiwan are chosen as case studies. By using the coefficient of similarity, the paper first analyses the present situation and changes of the industrial structure(as a whole), the manufacturing industry and the tertiary industry. and then forecasts the trend of future development of industrial homogeneity of the two regions through R/S method in the fractal theory. The results shows that industrial homogeneity between these two regions is significant, and if the industrial development continues under the existing circumstances, the level of industrial homogeneity will continue to increase, that of the manufacturing industry will continue to rise in fluctuation, while that of the tertiary industry will decrease slowly. At the same time, the industrial homogeneity indicates that the division of labour between Fujian and Taiwan tends to transform from a vertical-division one to a horizontal-division one.

Keywords: industrial homogeneity, the coefficient of similarity, R/S, Fujian and Taiwan provinces.

1 Introduction

With the increasing trends of economic globalization and regional economic integration, the world economy has entered a period of the coexistence of cooperation and competition. The industrial homogeneity is the manifestation of high degree of similarity in industrial structures of different regions, which emerges and enhanced continuously in the process of changes in regional industrial structure[1-3]. To certain extent, such phenomenon will affect the efficiency of the allocation of regional economic resources and reduce regional comparative advantages. In recent years,

* Foundation: Science Foundation of Chinese Ministry of Education(09YJA790044); Science Foundation of Chinese Ministry of Education(10JJDZONGHE017).

** Author: Wei Suqiong(1965-), Ph.D. and Professor, specialized in economic geography and land resource and land use plan. Tel.:13003821102.

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 195–203.
springerlink.com

many scholars both at home and abroad have studied on status, mechanism and the impact of industrial homogeneity. In these studies, in terms of the price, labor productivity, inter-provincial trade and so forth, the development trend of Chinese regional industrial structure has been discussed[4] has been discussed, the degree of inter-regional industrial homogeneity has been revealed[5-8] through comparing the location quotient, Hoffman proportional coefficient, product weights, regional concentration index, the industrial structure similarity coefficient, the Gini coefficient and other indicators of various sectors and different regions with each other, the relationship between regional structure and these affecting factors on industrial distribution as well as the role of government protectionism and profit margin in these factors are investigated by using the classical theory and the core model in the new economic geography [9-12]. With regard to the impact of industrial homogeneity on regional economic development, though some scholars take negative views, there are some scholars believe that the industrial homogeneity can be classified into two cases, one is "satisfactory" and the other is "undesirable", which should be treated differently[12]. Currently, the study areas of industrial homogeneity are mainly focused on provinces and municipalities as well as the Yangtze River Delta, Pearl River Delta and other hot zones in China.

Fujian and Taiwan's inborn economic complementarity and difference in industry-evolution also create huge possibility for trade cooperation and industrial abutting joint. Currently the cross-strait cooperation of trade is increasingly close. In this context, some issues, like whether the industrial homogeneity exists between Fujian and Taiwan and what is their homogeneity's scope, extent, mechanism and the impact on the regional economic cooperation in the future and so on, are worthy of further exploration. This paper attempts to analyze the status and trends of industrial homogeneity between Fujian and Taiwan, statically and dynamically, in order to provide a basis for further optimizing the industrial structure and making Fujian a better role in connecting cross-strait.

2 Study Area Overview

From the geo-economic point of view, Fujian and Taiwan are at the heart of East Asian economic corridor, the world's most vibrant areas. Because of different economic development basis, the implementation of different social system, economic system and economic development strategies, economic development in Fujian and Taiwan are quite different. At present, Fujian is in the middle stage of industrialization, while Taiwan is in the stage of post-industrialization, the economic development of Fujian is trailed that of Taiwan by 20 years [13]. In 2005, Fujian's GDP(80351 million) was 23.2% (Table 1) of the size of Taiwan's GDP (346,389 million), equivalent to the level of Taiwan's late 80's (1986, Taiwan's GDP was the 80428 million). And the industrial structure, corresponding to the level of economic development, also reflects the development gap between the two regions (Table 1).

Tab1e 1. The survey of socio-economic development in Fujian and Taiwan provinces in 2005

Area	Land space (km^2)	population (ten thousand)	GDP(million dollar)	Proportion of three industries (%)
Fujian	1.24×10^5	3535	80351	36 : 43 : 21
Taiwan	3.60×10^4	22655	346389	2 : 25 : 73

data sources: "Fujian Statistical Yearbook 2005" and "Economic Yearbook of Taiwan 2005"

3 The Comparison on Fujian and Taiwan's Industrial Homogeneity on the Base of Similarity Coefficient

The study started from the internal structure of three industries, manufacturing industry and tertiary industry respectively. Then a system for industrial homogeneity comparison is came into being. Meanwhile, it adopts similarity coefficient to estimate inter-region industrial homogeneity, whose formulas is as follows:

$$S_m = \left(\sum_{k=1}^{n} X_{ik} \cdot X_{jk}\right) \Bigg/ \sqrt{\sum_{k=1}^{n} X_{ik}^2 \sum_{k=1}^{n} X_{jk}^2} \tag{1}$$

In the formula （1）, m=1,2,3 is on behalf of similarity coefficient of three industries, manufacturing industries and tertiary industry respectively, X_{ik} stands for sector k 's proportion in region i, and X_{jk} for k 's proportion in region j. S_{ij} = 1 suggests the same structure in the two regions, while 0 is completely different. $S_{ij} \geq 0.5$ indicates the high homogeneity level of the structure in the two regions, while $S_{ij} \leq 0.5$ indicates the low one.

3.1 Static Contrast of Industrial Homogeneity between Fujian and Taiwan Provinces

Table 2. The similar coefficient of industrial structures in Fujian and Taiwan provinces

Type	Fujian(2005)/ Taiwan(2005)	Fujian(2005)/ Taiwan(1980)
S_1 （Three Industries)	0.8270	0.9886
S_2 （The structure of manufacturing industry)	0.8594	0.8606
S_3 （The structure in tertiary industry）	0.8494	0.8137

Table 2 shows similarity coefficients of industrial structure in the same stage of industrialization development as well as on the same time in Fujian and Taiwan. On the whole, the similarity coefficient of three industries (S_1), internal structures of manufacturing industry(S_2)and tertiary industry (S_3), concerning Fujian in the year 2005 and Taiwan in the year 2005 and 1980, are all greater than the critical value 0.5 and approach 1, which means the high homogeneity level of the two places. Second, concerning the similarity coefficients of three industries and manufacturing industry,

Fujian(2005) and Taiwan(1980) reach the peak. Third, considering the tertiary industry, Fujian's and Taiwan's structural homogeneity level in the year 2005 is higher than that of Fujian in the year 2005 and Taiwan in the year 1980. That is to say, in the same stage of industrial development, Fujian and Taiwan's structural homogeneity level of three industries and manufacturing industry is higher, while Fujian and Taiwan's structural homology level in the aspect of tertiary industry is higher during the same period. As far as the year 2005 is concerned, Fujian and Taiwan's structural homogeneity of manufacturing industry and tertiary industry is still significantly obvious.

3.2 Dynamic Contrast of Industrial Homogeneity between Fujian and Taiwan Provinces

The relevant data of 1981-2005 years from the "Fujian Statistical Yearbook", "Taiwan Statistical Yearbook" and "China Statistical Yearbook"are adopted, and industrial similarity coefficients between Fujian and Taiwan have been worked out and showed in Figure 1.

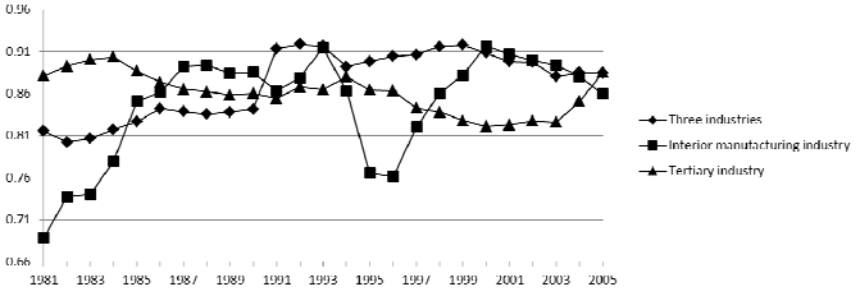


Fig. 1. The changing trend of the similar coefficient of industrial structures in Fujian and Taiwan provinces

Remark: Because of data problems, there are changes (increase and decrease) in the sectors of the internal manufacturing industry in Fujian and Taiwan.

Fig.1 characterizes as follows: firstly, from 1981 to 2005, all of Fujian and Taiwan's similarity coefficients of the structure in three industries, manufacturing industry and tertiary industry keep higher than 0.5, indicating a high homogeneity level, among which the structural homogeneity level of three industries and tertiary industry remains high, fluctuating between 0.8-0.92 over the years, but their trends are quite distinct from each other. The similarity coefficient of three industries fluctuates up while tertiary industry's fluctuates down. Secondly, concerning Fujian and Taiwan's manufacturing industry, the interior structural similarity coefficient curves a lot. The similarity coefficient goes up from 0.6890 in the year 1981 to the peak 0.9144 in the year 1993, then presents a steep rise after a nose dive with a turning point of the year 1996, and the homogeneity level somewhat eased from the year 2002 （from 2002 to 2005, it goes down by 0.04）.

3.3 Analysis of Industrial Homogeneity between Fujian and Taiwan Provinces

By analyzing the condition and background of industrial development in Fujian and Taiwan, we concluded that Fujian and Taiwan's industrial homogeneity changes rationally. Firstly, the similarity of industrial structure is related to the similarity of resource endowment. Fujian and Taiwan bears considerable similarity in natural condition, Humanities and history background as well as factor endowment, which gives birth to Fujian and Taiwan's high homogeneity level of industrial structure as a whole. Secondly, regional industrial homogeneity presents certain direct proportion to the degree of closeness of inter-regional ECOD[14],the closer ECOD is, the more similar industrial structure are. Different regions, at the similar level of development and in the similar stage of development, inevitably take on a similar supply structure and demand structure, thus form a similar resource structure, production function and demand preference. Therefore, the rising trend of homogeneity level of three industrial structure in Fujian and Taiwan indirectly indicates a fact that economic development gap between two regions is narrowing day by day, especially when Fujian has developed rapidly and stepped into the intermediate stage of industrialization since the year 1993, the growth rate of two regions' similarity coefficient seems more obvious than ever before. Thirdly, look from a dynamic angle, industrial homogeneity has different characteristics because of different development stages the industry stays in. At present, Fujian's manufacturing industry develops rapidly, staying in the intermediate stage of industrialization, while Taiwan in the post-industrialized society when tertiary industry highly developed, so that the different changing trend of the homogeneity level appears in terms of Fujian and Taiwan's manufacturing industry as well as tertiary industry. Fourthly, in recent years, the closeness of economic link between Fujian and Taiwan and factor flow brought about by it also enhance the similarity of industrial structure in the two regions. Up to the year 2006, the number of investment projects from Taiwan which approved by Fujian (including the reinvestment from third party) have reached 8930 with US$10.354 billion. Now the investment from Taiwan has become the second biggest offshore funds source [15] for Fujian, just behind HK's investment. The nearly 90% of Taiwanese business investment has entered into Fujian's manufacturing industry [16] and the investment priority was changed from wood & bamboo products manufacturing industry, leather & fur manufacturing industry, food & beverage manufacturing industry, metalworking machinery manufacturing industry and electrical & Electronics manufacturing industry in early time to high-tech and petrochemical industry. Especially the specialization and cooperation in the petrochemical, auto bile and IT industry between Fujian and Taiwan has converted gradually from labor intensive industry to capital and skill intensive industry, from vertical disintegration to de-verticalization, possessing stable degree of industrial relevance. Because of closer economic link and its resulting technology transformation and industry transformation, "coincidence" industry comes into being in related regions, which is called industry homogeneity.

From a negative point of view, Fujian and Taiwan's industrial homology would narrow the two regions' comparative advantages, which would lead to the loss of the benefit of both scale economy and division of labor, the block the sustainable development of both economic cooperation and win-win in the two regions in the

future. However, we just concluded the study above of industry homogeneity mainly by analyzing from three industries, together with interior manufacturing industry and tertiary industry. If we work on the study by classifying the industry more specifically and concretely, the problem would not be as serious as we preconceive. As it is showed in table 3, the first six industries in Fujian are all light industries with a proportion of 30.51% in Fujian's manufacturing industry, while that of Taiwan is 13.76%. At present, Fujian's proportion of leading industries such as petrochemical industry, IT industry and machinery industry and so on is far lower than Taiwan's. The difference of industrial structure between Fujian and Taiwan are fairly clear.

Table 3. The top 20 industrial sectors on the proportion of output value in Fujian and Taiwan provinces in 2005

Fujian		Taiwan	
The proportion of output value %	industrial sector	The proportion of output value %	industrial sector
6.59	Leather and Fur Products Manufacturing	12.83	Chemical Material Manufacturing
6.43	Nonmetal Mineral Products	9.62	Machinery and Equipment anufacturing
6.43	Textile Industries	9.55	Transportation manufacturing industries
6.02	Made-Up Textile Goods, Apparel & Accessories Manufacturing Industries	9.53	Petrochemicals Manufacturing
5.05	Food Processing	7.42	Basic Metal Manufacturing
4.76	Electrical and electronic communication equipment manufacturing	7.07	Metal Products Manufacturing

4 The Trend Forecasting of Industry Homogeneity Based on R / S in the Fujian and Taiwan Provinces

Figure 1 is a line graph taking time as the abscissa and similarity coefficient as the vertical axis and this non-smooth, non-differential fractal curve is suitable for a research with fractal theory. According to R/S mathematical model, we use similarity coefficients of industries between Fujian and Taiwan in 1981-2005, and with the help of Excel and VB programming software.

Overall, the similarity coefficient H_1(1981-2005) of three industries between Fujian and Taiwan is 0.9969> 0.5, C> 0, H_2 (1981-2005) of the secondary industry is 0.7358> 0.5, C> 0, H_3 (1981-2005) of the tertiary industry is 0.9230> 0.5, C> 0, showing that under the original regional development environment, the level of industrial homogeneity in Fujian and Taiwan in the coming years 2006-2030 will have the same trend as the level in 1981-2005. First, the degree of homogeneity on the three industries will increase and its value tends to 1, indicating a great momentum of the industrial upgrading and economic social development in Fujian in the next 25 years, to a certain extent, reducing the gap with Taiwan. Second, the similarity on the

internal structure of manufacturing industry in general will also show volatile changing trend, with its degree enhancing . Third, the degree of homogeneity on the tertiary industrial structure will also decrease slowly as it did in the past 25 years. In addition, H of three industry and tertiary industry index are closer to 1, indicating that compared with the manufacturing sector, their movements of similarity coefficients have stronger sustainability.

Of course, the concrete analysis in details in a short time period is also very necessary, for example, choosing the H_1, H_2 and H_3 when n=3 for analysis. The similarity coefficient H_1（1981-1983）of three industries between Fujian and Taiwan is | -0.1718 | <0.5, C<0, so the degree of industrial homogeneity in 2006-2008 and 1981-1983 will have an opposite trend, namely, in 2006-2008, slowly weak. Similarly, the H_2（1981-1983）=0.4408<0.5,C<0, as a sharply rising trend in 1981-1983 of Industrial homogeneity, so continuing weakening in 2006-2008. Then, the similarity coefficient H_3（1981-1983）of the tertiary industry is 0.3375<0.5, C<0, noting the degree of industrial homogeneity in 2006-2008 and in the past three years will have an opposite trend, weakening . Other values of n correspond to H are all greater than 0.5, closing to 1, basically the same as the overall analysis of the situation.

In addition, we get three figures (Fig2, Fig3, Fig4) about homogeneity of industry, the manufacturing industry and the tertiary industry by taking the logarithm for the

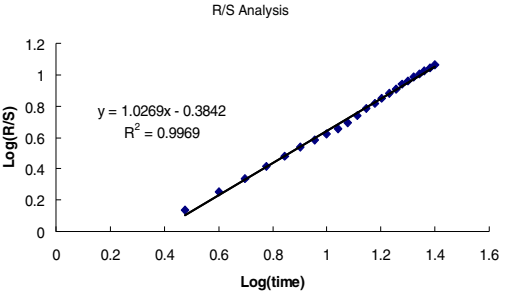


Fig. 2. The R/S of the similar coefficient of Industrial structure in Fujian and Taiwan provinces (three industries in 1981-2005)

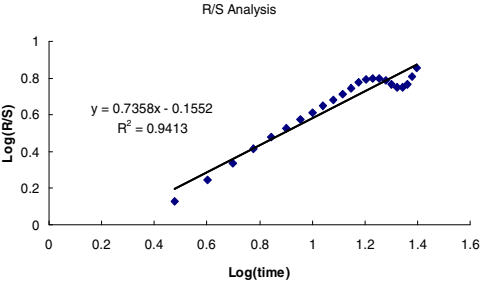


Fig. 3. The R/S of the similar coefficient of the industrial structure in Fujian and Taiwan provinces（1981-2005）

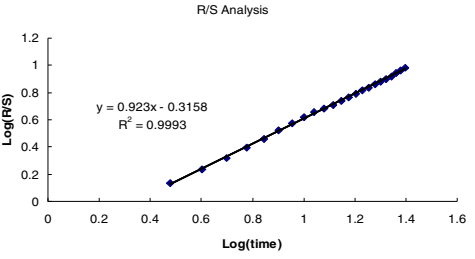


Fig. 4. The R/S of the similar coefficient of industrial inner structure in Fujian and Taiwan provinces (tertiary industries in 1981-2005)

log (R_n/S_n) and n. Their values of R2 were 0.9969, 0.9413, 0.9993, indicating that the return fitting effect is sound, and has the higher prediction accuracy.

5 Conclusions

The phenomenon of industrial homogeneity in these two regions is obvious. In the same stage of industrialization development, the level of the homogeneity of the internal structure in three industries and manufacturing industry between Fujian and Taiwan is high. The level of the tertiary industrial homogeneity in the same time between Fujian and Taiwan is also high. The R/S analysis of 25 years' dynamic data showed that: if the three industries continue to develop under the original circumstances, the level of industrial homogeneity will continue to increase with its value closer to 1, while there will be a fluctuant rising trend in the manufacturing industry and a slowly decreasing trend in tertiary industry. On the whole, they will all be growing in the future.

However, the appearance of industrial homogeneity between Fujian and Taiwan has its own inevitability. Because of the similarity of resource, closer economic development level and economic connection reflected by high-frequency flow of product and other factors, it exerts a great benefit to the occurrence of industrial homogeneity. At the same time, the industrial homogeneity reveals a trend that industrial division between Fujian and Taiwan has passed from vertical disintegration to de-verticalization.

References

1. Cui, G.: Regional Analysis and Planning. Higher Education Press (1999)
2. Hu, R., Zhang, X.: Analysis of Industrial Structure and Regional Interests. Journal of Economics and Management (2001)
3. Chen, Y., Wei, S.: An Empirical Study on the Industrial Isomorphism between the West of the Straits, Yangtze River Delta and Pearl River Delta. Tropical Geography 27(1), 54–59 (2007)

4. Naughton, B.: How Much Can Regional Integration Do to Unify China's Markets? Paper for Conference on Policy Reform in China, Center for Research on Economic Development and Policy Research, Stanford University, pp. 18–20 (November 1999)
5. Wang, Y., Hua, Y.: The Authentic Research on the Industry Structure Convergence in the Bohai Rim Region. Economy and Management 22(2), 30–33 (2008)
6. Dai, D.: To Research into Questions of Industrial Structure Convergence Based on the Profit Margin. Value Engineering (10), 28–30 (2007)
7. Zhang, T., Liang, J., Song, J.: Study on the Concentration and Dispersion of China's Manufacturing at Provincial Level. Economic Geography (3), 315–319 (2005)
8. Chen, Y.: Industrial Convergence of Metric and Consensual or Non-Consensual. China Industrial Economy (4), 37–43 (1998)
9. Kim, S.: Expansion of Markets and the Geographic Distribution of Economic Activities: The Trends in U. S Regional Manufacturing Structure, 1860-1987. The Quarterly Journal of Economics 110(4), 881–908 (1995)
10. Young, A.: The Razor's Edge: Distortion and Incremental Reform in the People's Republic of China. Quarterly Journal of Economics 115, 1091–1135 (2000)
11. Zhang, C., Li, M.: Protectionism and Cooperation: An Economic Analysis Based on Industrial Structure Convergence. Finance & Economics (3), 76–81 (2007)
12. Liu, C., Lv, L.: Industrial Isomorphism, Manufacturing Space Diffusion and Regional Economic Development of the Yangtze River Delta region. Management World (4), 35–39 (2005)
13. Wei, S., Chen, J.: A Comparative Research of Coupled Construction Land Change and Industrialization between Fujian and Taiwan Provinces. Geographical Research 25(1), 87–94 (2006)
14. Chen, J.: Study on Same Structure and Position of Industry in Delta Area of Changjiang River. China Industrial Economy 191(2), 19–26 (2004)
15. Shan, Y.: Economic and Trade Interaction Effect with the Future Development of the Strategy between Fujian and Taiwan. Fujian Tribune 9(10), 107–111 (2007)
16. Li, F., Xiong, J.: Docking of the Manufacturing Industry between Fujian and Taiwan. Fujian Finance (4), 15–18 (2007)

The Relationship between Financial Deepening and Economic Growth in Taiwan

Shu-Chen Chang and Cheng-Hsien Wu

Graduate Institute of Business and Management, National Formosa University, 64, Wunhua Rd, Huwei, Yunlin 632, Taiwan

Abstract. The goal of this paper is to investigate the threshold cointegration effect of financial deepening on economic growth in Taiwan over the period from 1981 to 2010. The results show that a threshold cointegration effect exists in relationship between financial deepening and economic growth. In short-run, economic growth has a significant and positive effect on financial deepening in the high-growth regime. This implies that economic growth can stimulate financial deepening in Taiwan in the high-growth regime. In addition, the impact of financial deepening on economic development has a significant and positive effect in the high- and low- growth regimes. Thus, financial deepening can increase economic growth in Taiwan.

Keywords: Financial deepening, economic growth, threshold cointegration model.

1 Introduction

The relationship between financial deepening and economic growth has been discussed in recent years. Some previous studies [1, 8, 13] point out that financial deepening can promote economic growth. Hasan, Wachtel, and Zhou's (2007) study uses panel data of Chinese provinces to examine the effects of financial deepening and political institutions on growth rates. Their results show that the depth of capital market has a significant influence on growth while the bank lending has insignificant effect. In addition, their study suggests that the greater development of financial sector as well as property rights, the legal environment, and political institutions have influence on economic growth. Abu-Bader and Abu-Qarn (2007) apply the vector error-correction model to examine the causal relationship between financial development and economic growth in Egypt during the period from 1960 to 2001. Their results show that financial development causes economic growth by increasing resources and enhancing efficiency for investment. Yang and Yi (2008) use the superexogeneity methodology to test whether financial development causes economic growth for Korea over the period from 1971 to 2002. Their results show that financial growth would causes economic growth while economic growth did not cause financial growth.

Khan et al. (2005), Ang (2008), Jalil and Ma (2008), and Anwur et al. (2011) use the Autoregressive Distributed Lag (ARDL) technique to estimate the short- and long-run effects of financial development on economic growth. Among these previous

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 205–210.
springerlink.com

studies, Ang (2008) using both quantitative and qualitative channels shows that financial development has a significant and positive impact on economic growth in Malaysia. Anwur et al. (2011) use Pakistan over the period from 1973 to 2007. Their study indicates that there is a significant and long-run stable relationship between financial development on economic growth. In addition, their results show that uni-directional causality exists in the external debt to exports ratio and the private sector credit to GDP ratio for Pakistan.

Regarding financial deepening and economic growth for Taiwan, most previous studies [5, 6, 12] have used the error-correction model to examine the relationship between these two variables. These studies show that financial deepening has a direct effect on economic growth in Taiwan. On the other hand, Chang and Lee (2003) use the ordinary least squares estimator (OLS) to examine the relationship between two variables. Their results show that financial development does not influence economic growth.

As above discussion, previous studies use a linear approach or error-correction model to investigate the relationship between financial deepening and economic growth. Their focus on whether the causal relationship between financial deepening and economic growth is uni-directional or bi-directional. In addition, these previous studies did not have a consistent result. Thus, this paper doubts of the linear cointegration relationship between financial deepening and economic growth. However, previous studies lack to consider the threshold cointegration relationship between financial deepening and economic growth. Thus, the goal of this paper is to investigate threshold cointegration relationship between financial deepening and economic growth. In addition, our approach uses Hansen and Seo's (2002) the threshold error-correction model to set the threshold relationship between financial deepening and economic growth.

This paper is organized as follows: Section 2 is to explain the empirical methodology. Section 3 describes the variables and our findings. Final section is conclusion.

2 Empirical Methodology

The long-run relationship between financial deepening and economic growth can be presented as follows:

$$\begin{bmatrix} \Delta GDP_t \\ \Delta PRIVY_t \end{bmatrix} = \mu + \alpha w_{t-1} + \begin{bmatrix} \Delta GDP_{t-1} \\ \Delta PRIVY_{t-1} \end{bmatrix} + \varepsilon_t \quad (1)$$

where μ is a two dimensional vector of intercepts, $w_{t-1} = GDP_{t-1} - \beta\ PRIVY_{t-1}$, α is a two dimensional vector of speed of adjustment coefficients, and ε_t is the error term vector. GDP_t denotes the economic growth; $PRIVY_t$ denotes financial deepening.

If the long-run relationship between financial deepening and economic growth has a nonlinear, the Eqn. (1) can be rewritten as follows:

$$\begin{bmatrix} \Delta GDP_t \\ \Delta PRIVY_t \end{bmatrix} = \mu_1 + \alpha_1 w_{t-1} + \Gamma_1 \begin{bmatrix} \Delta GDP_{t-1} \\ \Delta PRIVY_{t-1} \end{bmatrix} + \varepsilon_{1t}, w_{t-1} \le \gamma$$
$$\begin{bmatrix} \Delta GDP_t \\ \Delta PRIVY_t \end{bmatrix} = \mu_2 + \alpha_2 w_{t-1} + \Gamma_2 \begin{bmatrix} \Delta GDP_{t-1} \\ \Delta PRIVY_{t-1} \end{bmatrix} + \varepsilon_{2t}, w_{t-1} > \gamma \quad (2)$$

According to Hansen and Seo's (2002) study, all coefficients are allowed to vary with the regimes in the presence of a threshold effect in Eqn. (2). Thus, if the empirical model is nonlinear, cointegration effect from linear model may suffer from problems of misspecification and misleading conclusions. In order to avoid these problems, this paper uses Hansen and Seo's (2002) sup Lagrange Multiplier (LM) test statistic to determine whether the threshold cointegration effect is existed. Hansen and Seo (2002) suggest two bootstrap methods to calculate the *p*-values of the *sup* LM test and asymptotic critical values. When the error-correction model has a threshold effect, a two-regime threshold cointegration model is used as follows:

$$\Delta X_t = \begin{cases} A_1' X_{t-1}(\beta) + \mu_{1t} & if \ w_{t-1}(\beta) \le \gamma \\ A_2' X_{t-1}(\beta) + \mu_{2t} & if \ w_{t-1}(\beta) > \gamma \end{cases} . \quad (3)$$

with

$$X_{t-1}(\beta) = \begin{pmatrix} 1 \\ w_{t-1}(\beta) \\ \Delta X_{t-1} \\ \vdots \\ \Delta X_{t-p} \end{pmatrix} . \quad (4)$$

where x_t is a $p-$dimensional *I*(1) time series that is cointegrated with one $p \times 1$ cointegrating vector β; $w_t(\beta) = \beta' X_t$ is the *I*(0) error-correction term; μ_t is the error term; A_1 and A_2 are coefficient matrices which describe the dynamics in each of the regimes; and γ is the threshold value.

3 Empirical Result

The sample data comprises thirty years over the period from 1981 to 2010 in Taiwan. All data are obtained from Taiwan quarterly national economic trends databank. The financial deepening is measured by bank credit to the private sector. Many economic and financial time series could exhibit a trending behavior or non-stationary in mean. If variables are non-stationary, their results suffer from false results [11]. Therefore, this paper uses the Augmented Dickey Fuller (ADF) test and the Phillips-Perren (PP) test to determine whether all variables are stationary and the results are showed in Table 1. Table 1 shows that all variables can not be rejected at the 1 percent significance level. Thus, all variables are non-stationary in level.

However, all variables in the first difference can be rejected at the 1 percent significance level. This implies that all variables in the first difference are a stationary series.

Table 1. The results of unit root test

Test method	Variables	Level	First Difference
ADF	GDP	-4.024(0.0104)**	-7.993(0.0000)***
	PRIVY	-2.381(0.3882)	-4.401(0.0032)***
PP	GDP	-2.714(0.2329)	-8.077(0.0000)***
	PRIVY	-2.189(0.4909)	-4.455(0.0027)***

Notes: 1. The values in ( ) are *p*-values; the"**" denotes significance at the 5 percent significance level; and the "***" denotes significance at the 1 percent significance level. 2. Exogenous variables in the ADF and PP tests include individual effects and individual linear trends.

The goal of this study is to investigate the threshold cointegration effect of financial deepening on economic growth. The estimated cointegrating relationship is $w_t = GDP_t + 0.184 \times PRIVY_t$. Using Hansen and Seo's (2002) threshold cointegration test, the *p*-value is 0.000, implying that the null of linear cointegration can be rejected at the 5 percent significance level. The estimated threshold value γ is 6.29 where the error-correction term is used as the threshold variable. According to γ=6.29, our sample data can be divided into two regimes (e.g., $w_{t-1} \le 6.29$ and $w_{t-1} > 6.29$). About 0.51 percent and 0.49 percent of the observations fall into the first and the second regime, respectively. The first regime occurs as $GDP_{t-1} \le 0.184 \times PRIVY_{t-1} + 6.29$, implying that GDP is less than 6.29 percentage points below financial deepening. The second regime, on the other hand, the second regime occurs as $GDP_{t-1} > 0.684 \times PRIVY_{t-1} + 6.29$, implying that GDP is more than 6.29 percentage points above financial deepening. Thus, the first regime and second regime are denoted as "high-growth" and "low-growth" regimes, respectively.

The estimated threshold VAR for the effects of financial deepening on economic grwoth is expressed as follows:

$$\Delta PRIVY_t = \begin{cases} 2.194 - 0.042 w_{t-1} + 0.454 \Delta PRIVY_{t-1} - 0.042 \Delta GDP_{t-1} + \mu 1_t, w_{t-1} \le 6.29 \\ (1.943)\ (0.097) \qquad (0.210)^{**} \qquad (0.044) \\ -0.374 + 0.030 w_{t-1} - 0.04 \Delta PRIVY_{t-1} + 0.065 GDP_{t-1} + \mu 2_t, w_{t-1} > 6.29 \\ (2.272)\ \ (0.034) \qquad (0.107) \qquad (0.032)^{**} \end{cases} \tag{5}$$

$$\Delta GDP_t = \begin{cases} 9.315 + 0.466 w_{t-1} + 0.788 \Delta PRIVY_{t-1} + 0.512 \Delta GDP_{t-1} + \mu 1_t, w_{t-1} \le 6.29 \\ (2.699)^{***} (0.172)^{**} (0.294)^{***} \qquad (0.093)^{***} \\ 2.620 - 0.025 w_{t-1} + 1.106 \Delta PRIVY_{t-1} + 0.801 \Delta GDP_{t-1} + \mu 2_t, w_{t-1} > 6.29 \\ (4.032) \qquad (0.069) \qquad (0.561)^{**} \qquad (0.092)^{***} \end{cases} \tag{6}$$

where the parentheses in Eqns. (5) and (6) are Eicker-White standard errors; the"**" denotes significance at the 5 percent level; and the "***" denotes significance at the 1 percent level.

The estimation of the error-correction term in threshold VAR, w_{t-1}, allows for a straightforward investigation of the behavior of the gap between financial deepening and economic growth. The relationships between ΔGDP_t and $\Delta PRIVY_t$ have minimal error-correction effects and dynamics. The signs of these coefficients are used to explain the adjustment process by which the long-run equilibrium between financial deepening and economic growth is restored. Error-correction appears to occur only in the first regime in the GDP equation. The positive and significant error-correction effect exists in the GDP equation. The effect is strong. Furthermore, the dynamic coefficients of financial deepening for GDP equation are 0.788 and 1.106 in the first and second regimes, respectively. These effects have significant at the 1 percent significance level. It implies that financial deepening can stimulate economic growth. The dynamic coefficients of GDP for GDP equation are 0.512 and 0.801 in the first and second regimes, respectively. These effects also have significant at the 1 percent significance level.

On the other hand, the dynamic coefficients of financial deepening for financial deepening equation are 0.454 and -0.04 in the first and second regimes, respectively. The effect has significant at the 5 percent significance level in the first regime while it has insignificant at the same level in the second regime. The dynamic coefficients of GDP for financial deepening equation are 0.042 and 0.065 in the first and second regimes, respectively. The effect has significant at the 5 percent significance level in the second regime, but it has insignificant at the same level in the first regime. It implies that economic growth can increase the financial deepening in high-growth regime. In addition, the past financial deepening can influence current financial deepening in low-growth regime.

4 Conclusions

Our paper is to examine the threshold cointegration relationship between financial deepening and economic growth using Hansen and Seo's (2002) threshold cointegration approach. The finding shows that the null hypothesis of the linear cointegration would be rejected in favor of a two-regime threshold cointegration model. It implies that a discoutinuous or nonlinear adjustment of economic growth toward a long-run equilibrium. Therefore, there is a cointrgration relationship when the divergence between financial deepening and economic growth is below 6.29 percent. Furthermore, the adjustment response for economic growth is much greater than for financial deepening.

In the short run, the results show that the effects of economic growth on financial deepening are significant in the second regime. This implies that economic growth can stimulate financial deepening in high-growth regime for Taiwan. In addition, the impacts of financial deepening on economic growth are significant in both regimes. This result indicates that financial deepening can increase economic growth in high-growth regime for Taiwan.

References

1. Abu-Bader, S., Abu-Qarn, A.S.: Financial development and economic growth: The Egyptian experience. Journal of Policy Modeling 30, 887–898 (2008)
2. Ang, J.B.: What are the mechanisms linking financial development and economic growth in Malaysia? Economic Modelling 25, 38–53 (2008)
3. Anwar, S., Shabir, G., Hussain, Z.: Relationship between financial sector development and sustainable economic development: Time series analysis from Pakistan. Canadian Center of Science and Education 1(3), 262–270 (2011)
4. Chang, C.B., Lee, S.Y.: Financial development and economic growth in Taiwan and Japan. Chicago school might be right. Taiwan Zhongshan University Social Sciences Essays 1(1), 109–134 (2007)
5. Chong, C.F.: Financial development and growth: The case of Taiwan. Quarterly Journal of Taiwan Bank 50(4), 64–85 (1999)
6. Huang, R.D., Luo, S.W.: Financial development and Economic growth: The case of Taiwan. Quarterly Journal of Taiwan Bank 51(2), 50–76 (2000)
7. Hansen, B.E., Seo, B.: Testing for two-regime threshold cointegration in vector error-correction models. Journal of Econometrics 110, 293–318 (2002)
8. Hasan, I., Wachtel, P., Zhou, M.: Institutional development, financial deepening and economic growth: Evidence from China. Journal of Banking & Finance 33, 157–170 (2009)
9. Jalil, A., Ma, Y.: Financial development and economic growth: Time series evidence from Pakistan and China. Journal of Economic Coopration 29(2), 29–68 (2008)
10. Khan, M.A., Qayyum, A., Sheikh, S.A.: Financial development and economic growth: The case of Pakistan. The Pakistan Development Review 44(4), 819–837 (2005)
11. King, R., Levine, R.: Finance and growth: Schumpeter might be right. Quarterly Journal of Economics 108, 717–737 (1993)
12. Lee, C.C.: Banking development, stock market development and economic growth: The case of Taiwan. Institute of Economics, Academia Sinica 35(2), 79–105 (2005)
13. Yang, Y.Y., Yi, M.H.: Does financial development cause economic growth? Implication for policy in Korea. Journal of Policy Modeling 30, 827–840 (2008)

Residential Location, Commute and Urban Transportation

Xiaoping Yu

Beijing Jiaotong University, Beijing, China
yuxiaoping1983@sina.com

Abstract. This study considered the urban congestion in the view based on the relationship between residential location, commute and urban transportation. Jobs and housing location perform the trend as geographical imbalanced with the increasing in commuting distance and time, and as a result, it has been considered as the main reason to the congestion, and planners take jobs-housing balance as an effective solution to the transportation problems. Even though, there is little evidence shows that jobs-housing balance makes significant effect, for the factors affecting where to live and work are complex. The study indicated that despite the residential location choices and commuting time and distance are closed linked, the effective solutions for urban transportation problems such as congestion should take comprehensive considerations including main factors in both short-term and long-term to improve the regional accessibility.

Keywords: Residential Location, Commute, Jobs-Housing Balance, Congestion.

1 Introduction

The urban starts when economic activities clustered, while with the development of urban, urban sprawl has become a serious problem. The geographical imbalances in distributions of jobs and housing have been considered as the main course to the increasing commuting distance and time, leading to the urban congestion. Work and residential location choice has attracted particular attentions from the field of planning and transportation policy, and the jobs-housing balance at the same time has been taken as the effective solution to the congestions, with the consideration that the shorter journey to work, the less transport demand and correspondingly, the traffic condition improved.

Although the jobs-housing balance has won wide acceptance among the policy makers, there is no significant evidence showed that the improved jobs-housing balance can attribute to solving the urban congestion. Study on residential location choice indicated that there are complex factors affecting people`s decisions on where to work and live. People will consider housing cost, neighborhood quality, infrastructure, hospital, schools, accessibility to the urban center and available to park, etc.

Congestion and pollution are the main problems facing to the urban development, and attracting particular attention to consider the varieties of possible solutions. This study presented the trend of residential location, and the relationship between location

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 211–216.
springerlink.com

choice, commute and urban transportation. Based on the basic interactions, the paper indicated that effective solutions to solve transportation problem need the comprehensive improvement considering the affective factors in both long-terms and short-terms to improve the accessibility.

2 Relationship between Location Choice, Commute and Urban Transportation

The closed relationship existed between location choice, commute and urban transportation. People travel to work in different directions no longer the direct to the urban center that increase the transportation demand. Commuting at in the morning and after work make the worse traffic condition just like congestion and air pollutions. People change their jobs and the commuting distance, travel mode also have alternative situation, making the varieties of urban transportation.

2.1 Residential Location and Commute

Studies which show commuting intensified the congestion were most based on the assumptions that jobs are located in the urban centers and people prefer to work near home or to live as close to work as possible. But the factors impact the location choice are complicated, especially different people have variety utilities to affective factors.

Firstly, housing price is the main factor which is affected by land value. Reviewed the expand of urban spacial structure, people`s economic activities cluster in the urban center which became the central business district (CBD), and people preferred to live around the CBD area in the early location choice. Then, urban sprawls with economic development and commercial agglomeration closed to CBD, making the land value increased, and as a result, housing costs generally declined with the distance away from the urban center, and people have to live far away from the CBD. Alonso called it the "trade-off", that means people cannot get the low housing price and low commuting costs at the same time, and people have to choose from low price with high commuting costs or high land price with low commuting costs. Muth (1985) indicated that compared to housing costs, commuting costs are small. Even in modern world, living away from center cost extra commuting costs, but it is exactly the small amount compared with money you saved. Research from Credit Suisse took a family of four as an example from their survey, if the family live in Hergiswill with 10 minutes drive to Lucerne, their average disposable income is about $45,428 per year, while if they moved to live in Altdorf with about 35minutes drive, this number will increase to about $79.57. This survey also indicated that people moved to live in Fribourg have to pay 5,000 CHF, and at the same time they can get more than ten thousand saving from the housing costs.

Secondly, the accessibility to public transportation plays more important role in residential location, which affect directly on commuting to work or school, social activities and the convenience to hospital, in a world, accessibility influences the daily life. Generally, commuting costs refers to the financial cost, but now people pay more attention to the commuting time, which implicated the different value of time saving, which can be explained as the opportunity cost in economic sense. People with high

income level always prefer better transit condition, for their opportunity cost is comparably higher. In Japan, people take 1.5 hour in single trip as a limited to make the residential location choice, so the community closed to the public transportation station is always the best choice, of course, with high level rent and housing cost. The same to the situation in Beijing, China, that no matter the community locating in the second ring area or in the fifth ring area, the housing costs rely on the distance to the subway station.

Third, the environment of community is also a factor affect people who focus on children education, public facilities, neighborhood quality, or the other environment factors. Just this factor can explain the different location in some degree between China and US citizen. Compared New York with Beijing, people with low income level live in the center district in New York, and this trend attracted some sociologist to study on the race discrimination into the distribution in urban spatial structure, foe more and more black people live in the worse district in the center. While at the same time, people in Beijing chose the center community are mostly with higher income level. But the new trends both in New York City and Beijing, people move to the suburban or some community with great environment and large space.

Given the most three important factors, we can already get the conclusion that residential location choice cannot only depend on the distance to work, so the policy to lead people to live closed to work or find jobs near home, cannot get affective solution, in another worlds, there is no directly evidence show the policies aiming to jobs-housing balance can solve the transport problems, for the complex factors affecting location choice.

2.2 Commute and Urban Transportation

With urban economic development, there are more than one center in most urban, especially in some metropolitan even several employment centers, and at the same time, people have more chance to change jobs than to change house, so as a results, people can choose different transit tools to different area of the urban. The trend of commuting impact the urban transportation, which also made the most use of urban transit, it is also parts of the causes why people attribute the commute as the man reason to the congestion. Taking one line in Beijing, China, as example, the department of the government set a new line especially used to take the commuters in Tongzhou District, and as the Transportation Department of Beijing indicated, there are 67,000 people need to take the subway commute to work in the morning from 7 am to 9 am, which takes about 44 percent of all the passengers traveling to urban center during the daytime. Additionally, there are also more than 15,000 people taking bus and 12,000 people driving private cars should taken into consideration that these are the part of the transportation demand of commuting, and in order to get an equilibrium, improvement of transport capability supply is in need.

Brasee's Paradox (from Wikipedia), states that adding extra capacity to a network when the moving entities selfishly choose their route, can in some cases reduce overall performance. This is stated as follows, for each point of a road network, let there be given the number of cars starting from it, and the destination of the cars. Under these conditions on wishes to estimate the distribution of traffic flow. Whether one street is preferable to another depends not only on the quality of the road, but also on the density of the flow. If every driver takes the path that looks most favorable to him, the resultant

running times need not be minimal. Furthermore, it is indicated by an example that an extension of the road network may cause a redistribution of the traffic that results in longer individual running times.

Over the past decades, people commute farther than ever, resulting from the wave of suburbanization. Study of Fulton (1986) and Pisarski (1987) indicated that taking the nation as a whole, working trips made wholly within suburbs, the fastest growing commuting market, actually increased in length by around 15 percent during the 1970s. And the research from Klinger and Kusmyak (1986) showed that from 1977 to 1983, moreover, the mean journey to work for suburban Americans (defined as people residing outside a central city but within an urbanized area) increased from 10.6 miles in length to 11.1 miles, despite the mass migration of jobs to the suburbs during this period.

Commute characteristics shift with different situation, study on the commuting patterns and dual jobholder (Table 1) showed that the commute characteristics for the primary jobs of both single and dual jobholders are not quite different, with 75.8% for single jobholders and 79.6% for dual jobholders to drive alone to jobs. Both single and dual jobholders will choose carpool to work, with a slightly higher percentage from 14.9% with single jobholders to 12.2% with dual jobholders. Finally, there is the small percentage of people use public transit in U.S. To explain it, we can realize the relationship between commuting mode and transportation. Firstly, U.S. has the greatest highway system and most people will use private cars for daily travel, so the Dive take the largest percentage in the commuting mode, at the same time, it also indicates why there few people choose the public transit. Further, people who are dual jobholder may have double income and more travel need than single jobholders.

Table 1. Commute Characteristics

	Single	Dual Jobholder	
	Jobholder(%)	Primary Job(%)	Secondary Job(%)
Commute Mode			
Drive Alone	75.8	79.6	NA
Carpool	14.9	12.2	NA
Public Transit	4.3	3.5	NA
Other	5.0	4.7	NA
Commute Time			
0-4	20.8	21.0	NA
15-29	34.5	34.8	NA
30-59	33.9	33.9	NA
60 plus	10.8	10.4	NA
Commute Distance			
Same Tract	2.9	2.7	2.2
5 miles or less	32.1	31.8	28.9
5-15 miles	41.3	41.6	42.4
Over 15 miles	23.7	23.8	26.4

Sources: U. S Census Bureau, Census 2000; U. S. Census Bureau, LEHD Program 2005. Reported in Technical Paper No. TP-2007-01. Social, Economic, Spatial, and Commuting Patterns of Dual Jobholders. 2007.

Notes:

1.There is no made or time information for the secondary job of dual jobholders because that data is taken from the LF dataset, which contains only information on the commute to on job site and is assumed to be the primary job.

2.Certain values in this figure may not be significantly different from on another.
3.Data based on sample. For information on confidentiality protections, sampling error, and definitions, see <http://www.census.gov/prod/cen2000/doc/sf3.pdf>.
4.For further information on confidentiality protection and definitions with respect to UI data, see <http://www.lehd.did.census.gov/led/library/techpapers/tp-2006-01.pdf>.

3 Conclusions for Solving the Commuting Problems in Transportation

Given the analysis above, the paper indicates the relationship between residential location, commuting pattern and urban transportation, with the primary purpose to solve the transportation problems such as congestions and air pollutions, related to commute in some degree.

Commuting is exactly one of the most directly causes to the congestion, but there is no effective evidence show control commute help to improve the transit condition. First of all, the complex factors to location choice reduced the effect of policy aiming to guide people where to live and work. Additionally, population density makes great sense for the solution to congestion. Finally, geographical imbalance does not mean the sustainable urban structure.

The solutions to urban transportation problems should take into the long-term and short-term. With the view in long-term, transportation planning will work, and construction of infrastructure will improve the regional mobility, although the chance and the time to construct will make different sense. In short-term, tax policy, staggered office hours and some other policy can help to get improved transport conditions.

References

1. Bento, A.M., Cropper, M.L., Mobarak, A.M., Vinha, K.: The Effects of Urban Spatial Structure on Travel Demand in The United States the review of economics and statistics, August 2005, vol. 87(3), pp. 466–478 (2005)
2. Fulton, P.M.: Changing Journey-to-Work Patterns: The Increasing Prevalence of Commuting Within the Suburbs in Metropolitan Areas. Paper Prented at the 65th Annual Meeting of the Transportation Research Board, Washington, DC (January 1986)
3. Giuliano, G.: Is Jobs-Housing Balance a Transportation Issue? Transportation Research Record No. 1305, Finance, Planning, Programming, Economic Analysis, and Land Development. pp. 305–312 (1991)
4. Klinger, D., Kusmyak, J.R.: 1983-1984 National Personal Transportation Study. U. S. Department of Transportation, Federal Highway Administration, Washington, D. C. (1986)
5. White, M.J.: Firm Suburbanization and Urban Subcenters. Journal of Urban Economics 2, 323–343 (1976)
6. Cervero, R.: Jobs-Housing Balancing and Regional Mobility. Journal of the American Planning Association 55(2), 136–150 (1989)
7. Follain, J.R., Jimenez, E.: Estimation the Demand for Housing Characteristic; A Survey and Critique. Regional Science and Urban Economics 15, 77–107 (1985)
8. Qugley, J.: Consumer Choice of Dwelling, Neighborhood and Public Services. Regional Science and Urban Economics 15, 41–63 (1985)

9. Muth, R.: Models of Land Use, Housing and Rent:An Evaluation. Journal of Regional Science 25(4), 594–606 (1985)
10. Linneman, P.D.: The Demand for Residence Site Characteristics. Journal of Urban Economics 19, 129–148 (1981)
11. Pasarki, A.E.: Commuting in America. Westport, CT: Eno Foundation for Transportation, Inc. (1987)

An Online High-Speed Card Characters Recognition and Verification System Based on Multi-neural Network

Yongle Wang and Shuxin Xu

Xuchang ploughs the recent information science research institute, 461000, Xuchang, China

Abstract. In card manufacturing industry, it is very important to certify the characters printed on the cards. In this paper, we introduce a high-speed character recognition and verification system based on multi-neural network for verifying characters on plastic cards. The result based on real products shows the good performance of the system.

Keywords: OCR, Image Segmentation, Text Extraction, Neural Network.

1 Introduction

In modern society, a variety of cards, like mobile phone recharge cards and membership cards have been widely used. Plastic card plays an increasingly important role in people's daily life and its quality needs to be increased faster. Factories of card production one hour can produce over one million cards of the same type and it is also capable of producing more than thirty thousand kinds of different types of cards at the same time. However, in the process of card production, many factors will lead to the failure cards, including cards flip, cards of the same characters overlap and characters being incorrectly sprayed, and once an error occurs, the manufacturer must compensate based on the card face value. Therefore, the accuracy of the card test is a must and a very important step. However, manual inspection can't meet the needs of high-speed automated card production and for those cards with password, the password should be encapsulated in the tags, so the password must be verified and examined before the package. Checking through the manual can't achieve efficiency, accuracy and safety. In fact, the examination and verification has become the bottleneck of card production, therefore an online character recognition and verification system is urgently needed.

The document image recognized by traditional Optical character knowledge system is obtained by the flatbed scanner, and many commercial systems have been widely used[1,2,3]. But when using the camera to capture these images, there will be a lot of new problems[4], and the traditional optical character recognition system is not able to directly use the camera-based identification system.

In response to these problems, this treatise proposes a hybrid neural network-based high-speed online card character recognition and verification system. It consists of the following components: systematic plan design and the difficulties different with traditional OCR system, the positioning of identification character string, extraction, and analysis of character segmentation algorithm; the detailed

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 217–226.
springerlink.com

design of the mix neural network recognition engine integrating BPNN and CNN technology.

2 System Design and Difficulty

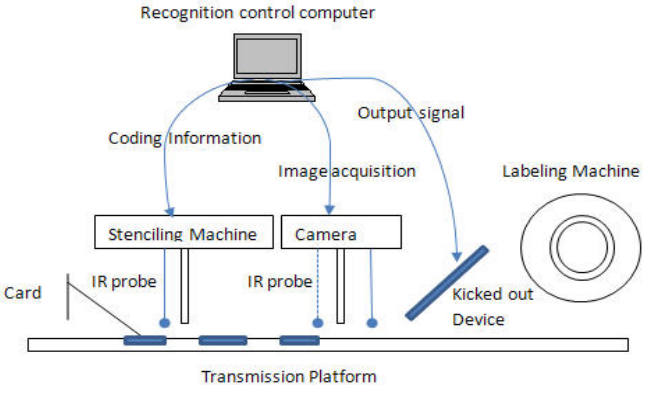


Fig. 1. Identification and verification system structure diagram

2.1 System Design

The process of card production is shown in Figure 1. Firstly, the blank cards are evenly distributed in machine drive platform transmitted from left to right with a certain time interval. When the infrared detector on the inkjet printer detects the card, the inkjet printer will coat pre-Set encoding on the card. In this stage, it will produce a lot of mistakes, so the unqualified card must be found and singled out before labelling. So we added a camera-based identification and verification system between inkjet printer and labeling machine, when the second infrared detector detects the card, this camera will capture the image and pass the card to identification Computer, the computer will recognize the characters having been sprayed and compare the results of recognition with the default encoding to determine whether the card is valid or not. The invalid card will be kicked out of the filter device and the effective card will be transferred to the labeling machine for labeling.

The system can be divided into three logical subsystems: image acquisition, image recognition, and output response.

2.2 Difficult Point of Achieving

Although the optical character recognition technology has been widely used, in the above-mentioned system, there are still many new and difficult points to address.

• High-speed requirements. The speed of card production line speed is close to 10 per second, so the response speed of detection system must meet requirements of production line.

• Low quality images, the image the quality of captured images will directly affect the results of subsequent processing. Since the card is moving at a high speed, so a clear shot with the camera is more difficult than using the scanner. So choosing what kind of camera is a must need considering. The property of camera directly impact on the transaction results behind. But the price is often the most important factor, in fact, it is difficult to get a good image in practice, so many post-processing need doing.

• Low-resolution images. Although industrial camera has many specialized features, the resolution is much lower than that of the household. To improve the speed, we set the camera's resolution 648x492 here, so the resolution of character is much lower than the resolution of images obtained by flatbed scanner with 150DPI.

• Variable number of characters and location. Because different cards have different layout, the amount of text which can be identified and their position are not constant, so the system should be able to handle a variety of cards.

• Complex background. For most cards, the characters are painted on the clear area, but there are a lot of cards are not the case, background textures and images increase the degree of difficulty of identification in large part.

3 Positioning and Segmentation of Text

After capturing the image, the next step is to extract the text information that need to be identified, and then transfer them to identify computer for recognition. This process can be divided into the following steps.

3.1 The Positioning of the Character String

For different cards, the location of the character string of sprayed words is different, because the card speed is too fast, it is impossible to automatically identify all the characters on the card. In fact there is no need to do so, for the same number of cards, the location of the sprayed character string is unchanged, if the picture of the cards can be gotten, the location of the sprayed character string can be determined. However, in practice, to obtain the card picture is not only contains the exact card picture but also includes some other background information (Figure 2). So, to get the card image, we must first extract the image of card, which requires the operation of edge detection, transformation and the Huffman binary methods will be used, this method of spray text detection is available, but the time complexity is high which influences the rate of detection.

Another method doesn't need to extract the image of card, just locating and recognizing the text character string directly from the correspondent position of captured images (Figure 3). But it must correctly set the capture trigger mode. Figure 1, when the infrared detector is placed before the camera, a delay needs to be set to capture the image. As the speed of conveyor is not completely uniform, so the corresponding position of the captured image is changeable. Therefore, the infrared detectors should be placed on the back of the camera in order to ensure that the trigger and there is no delay between the trigger and the capture operation. This approach can significantly reduce positioning time.

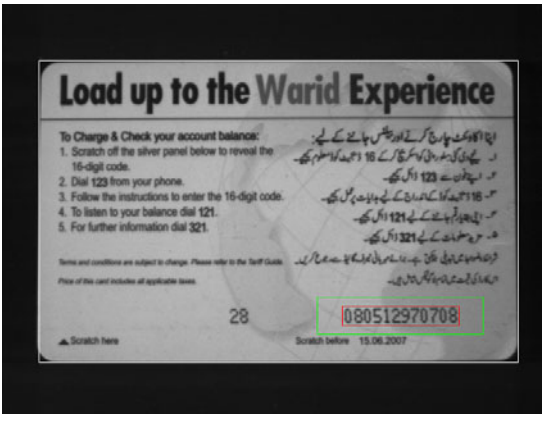


Fig. 2. To locate the relative position according to characters on the card

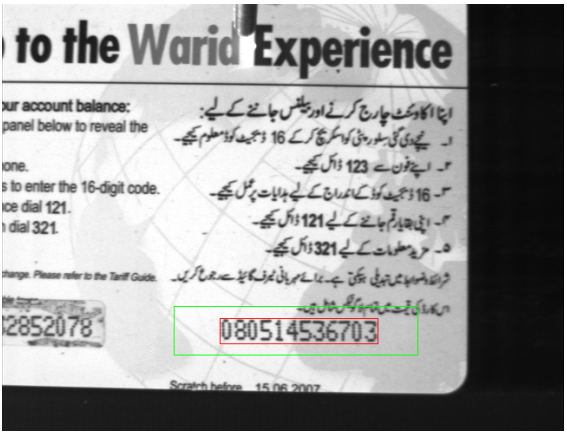


Fig. 3. To locate the position according to the string in the acquisition of the image

3.2 Extraction of Text Character Strings

The positioning setting discussed earlier only provide the general position of sprayed character string, while the actual location has some change to certain extent. It should first obtain the general area of the sprayed characters and then find the exact location of the string in this region. Figure 4 shows an image of approximate area that contains the painting text.

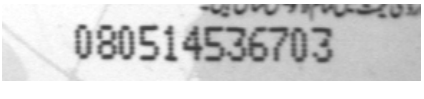


Fig. 4. Coding regions containing the general character

Accurately positioning the location of the spraying text can use the horizontal and vertical projection. In order to reduce the impact of textures and background text, it needs to conduct the pre-denoising in advance. After the results of projection generated, further analysis of the situation remove some errors, such as coding character strings are beyond this general area. In order to reduce the impact of lighting and texture background, it needs use a nonlinear processing method and local adaptive threshold method in preprocessing. Figure 5 shows the result of spraying character strings after binaryzation.

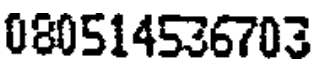


Fig. 5. After a string of binarization

3.3 Segmentation of Characters

Character segmentation processing plays a decisive position in optical character recognition. Richard made a review on the character segmentation methods and strategies and gave the level of segmentation and the space of segmentation strategy[5].

The initial segmentation method is based on the white space and pitch, this method is limited by the restricted print technology. Projection method is first used in the high-quality printing machines in which the next characters are separated according to line.

However, when the character width is variable, the pitch-based approach can't be used. Similarly, when the characters are tilted or there are siamese characters and similar strokes, the projection analysis method is also not available.

Connected component analysis provides a correct decomposition method for non-cursive characters picture. Therefore, connected component analysis method can be used to decompose the characters. Figure 6 is another deputy binarization results of picture text.

Fig. 6. Coding a binary character image

After the connected component analysis, the connection area are separated or connected based on the rules of height and width of their boundaries. Figure 7 shows the connection area after combination and it can be seen that the connection between elements should be further apart.

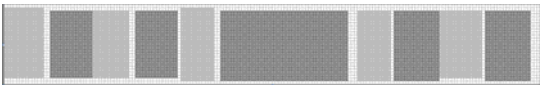


Fig. 7. Connected to do after the results of domain analysis

Because a character has a predetermined average width, so it should be dealt separately according to the width and the results of separation should be checked over the elimination of noise (see Figure 8), because of the low resolution of character (single character resolution is only 13x25), Direct Noise removal undermine the character itself, so first of all the resolution is increased to 26X50.

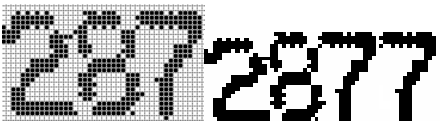


Fig. 8. Segmentation of characters on the adhesion and de-noising effect

3.4 Uniformization

Because of for recognition, neural network use the examples of normalization (36 x36) to practice, the extracted characters should be magnified to the normal size. Here, we adopt a method of normalization of middle.

4 Hybrid Neural Network

The system ultimately is used for practical applications, so it requires a relatively low recognition rate: the ideal state is close to 0. If the error recognition rate is over 1%, the system can't be used for commercial software. Because the single neural network has low efficiency and relatively high error recognition rate, reliability and robustness are not enough. Therefore, a hybrid neural network is designed.

4.1 Backward Propagation Neural Network （BPNN）

BPNN [6] is widely used in pattern recognition. This system is one kind of a neural network. In the BPNN of the system, there is an input layer, a hidden layer and an output layer. Input layer is a pixel group OF single character. The size of the input character is a36X36 pixel matrix, therefore, the number of neurons in input layer is 36X36 and the number of Output layer is 10, which represents 10 digits. In order to improve the stability of BPNN, it needs to maintain that the number of neurons is an inverted pyramid from the input layer to output layer. The number of neurons in hidden layer is from 100 to 200 and uses a gradient-based learning algorithm to update the weights of neurons.

4.2 Convolutional Neural Network (CNN)

The identification essence of neural networks is feature extraction. The More traditional method is manually extracting the whole process. The untreated input learning method is to find the characters in this area. Each method has its advantages and disadvantages. For the previous method, the main difficulty is to determine an appropriate feature class to extract, and using a reliable and robust method to extract features. Advantage of this approach is can be designed to extract the interesting features of the

extractor and ignore the noise. In addition, the automatic features learning methods such as BPNN is particularly sensitive to noise, it can wrongly think that the noise is some characteristics of some characters. Due to the interference of some noise, the recognition of BNPP is reduced [7].

To complement the defects of BPNN, a dual-channel convolution neural network (CNN) [8]is used. Figure 9 are the untreated eleven binary image, I+ is a copy of I, I- is the Inverted map of I+, Q+ and Q- is a convolutional layer and marked the feature extraction layer, the core of convolution is 16 hand- features, F + and F-test exist in each window test features, they like the feature class, BP_H and BP_O is the classifier.

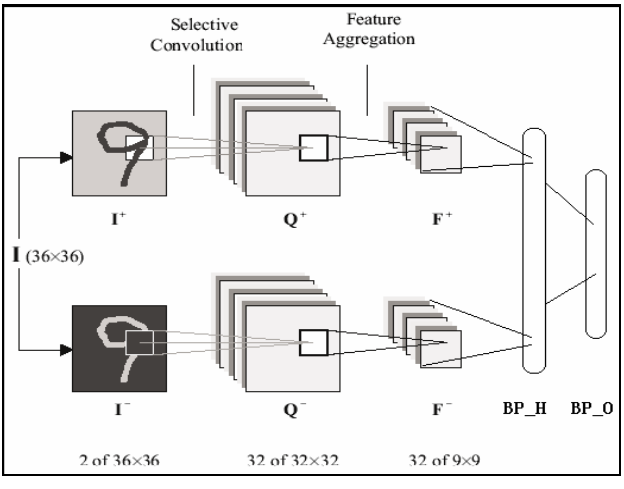


Fig. 9. Convolutional neural network processing

In the initial system of CNN, a layer of the perceived network is used as discrimination system [9]. But after testing, a layer perception network may have a high score in the case of no characters or destroyed characters, but when the character is recognized, it will lead to error identification. To avoid this from happening, a hidden layer is added in the identification system. Through the practical application, it is very effective. The CNN method layer with a hidden identification layer can recognize the non-character lowly that is to say it is the characters that can't be recognized. Therefore, this model will be more stable and robustness is relatively strong. BP_H is an added hidden layer, BP_O is the output layer.

4.3 Hybrid Neural Network System

For the problems encountered, a single neural network can't be effectively addressed. Taking into account the advantages and disadvantages of BPNN and CNN, we designed a neural network based hybrid system (Figure 10). In the system, three kinds of neural networks are used. A BPNN and two CNN. In the 4 neural networks of OCR1, a BPNN, and a CNN are used. In BPNN1, the number of neurons in hidden layer is 150, when the BPNN is no training set, CNN1 training samples are switched

from the upper left to lower right corner to produce more training image to accommodate the test set of conversion. Because convolution and CNN have more convergence than the BNPP, When the BPNN can't converge, and when to switch sample, CNN can converge. This is why CNN has a higher recognition rate than the BPNN. After OCR1, the results of the adjustment will be an integer. Y is marked as the largest maximum output of the neural network. The second Y is the second largest neural network output.

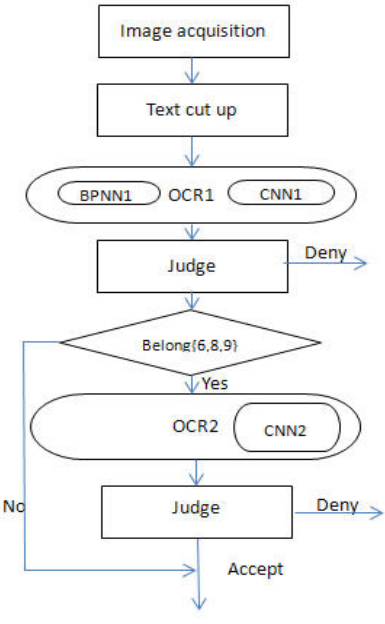


Fig. 10. Hybrid neural network processing

We accept the results as corresponding class of the largest neural network output, on the contrary, the result will be discarded. In the two results of neural network OCR1, if all are accepted and acceptable results are equal, then, the results of OCR1 are reliable. Opposite result will be discarded. The 6, 8 and 9 characters in OCR1 seem the same, as the same class. This means that there are two consequences, 6, 8 or 8, 9, we accept the 689 class as a temporary OCR1 results. So, 6 or 8 or 9 are not reliable. In OCR2 will be more carefully classified. OCR2 has only a neural network CNN2. CNN2 is the classifier designed specifically for the 6,8,9. The training set of CNN2 contains only three classes: 6, 8 and 9. Compared with the CNN1, CNN2 have fewer classes to be identified. Therefore, the ability CNN2 classification has been improved. By using CNN2, the efficiency of identification of 6,8,9 has a lot of improvement.

In OCR1, it needs the classification of 10 numbers (0-9), so in OCR1 using the neural network to ensure more reliable identification results. In fact, in OCR1 the more the neural networks are used, the more reliable the result will be. However, the rate of identification will be reduced. Considering the time and speed, we used two

neural networks in OCR1. In OCR1 the contribution to the hybrid neural network, the majority of non-characters will be discarded. In OCR1, in addition to 6, 8, 9 which are considered to be a class, the output is basically correct. Based on the correct classification of the OCR1, using a neural network in OCR2 is enough, because it only needs to classify 3 digits (6, 8, 9). So the goal of OCR1 is to reduce the error rate, at the same time, the goal of OCR2 is to increase the recognition rate....

5 Experiments and Results

The following experiment is used to test the system performance. First, establishing a number neural network and the network is trained by 2270 digital samples in the 220 cards (many samples were removed). 143 cards are tested and table 1 is the test results. For the method we proposed, we suggest the card's processing speed is 6 per second.

Table 1. Identification and verification results

	Correct rate of identification and verification	Validation failure rate	Rejection rate
BPNN	89.51%	6.99%	3.50%
CNN	93.70%	2.10%	4.20%
Method	97.21%	0.69%	2.10%

This treatise introduces a high-speed character recognition and verification system card. This system can reduce waste card rate card can also greatly improve production efficiency. Through further work, we can improve segmentation algorithm and neural network and better improve the system performance.

References

1. Mori, S., Suen, C.Y., Yamamoto, K.: Historical review of OCR research and development. Proceedings of the IEEE 80(7), 1029–1058 (2005)
2. Wesolkowski, S., Bowman, P., Tunnah, D.: A high speed colour cheque image capture and processing system. In: IEEE Canadian Conference on Electrical and Computer Engineering, vol. 1(2), pp. 457–460 (2006)
3. Zhao, S., Wang, Z.: A High Accuracy Rate Commercial Flight Coupon Recognition System. In: Proceedings of the 7th International Conference on Document Analysis and Recognition, vol. 8(86), pp. 3–6 (2003)
4. Liang, J., Doermann, D., Li, H.: Camera-based analysis of text and documents: a survey. IJDAR 7(1), 84–104 (2005)
5. Mo, S., Mathews, J.: Adaptive quadratic preprocessing of document images for binarization. IEEE Transactions on Image Processing 7(7), 992–999 (2007)
6. Niblack, W.: An Introduction to Digital Image Processing, pp. 115–116. Prentice Hall (2004)

7. Richard, G., Casey, Lecolinet, E.: A survey of methods and strategies in character segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence 18(7), 690–706 (2006)
8. LeCun, Y., Bengio, Y.: Word-level training of a handwritten word recognizer based on convolutional neural networks. In: Proc. of the International Conference on Pattern Recognition, Jerusalem, pp. 88–92 (2004)
9. Teow, L.-N., Loe, K.-F.: Robust Vision based Features and Classification Schemes for Off-line Handwritten Digit Recognition. The Journal of the Pattern Recognition Society 35, 2355–2364 (2002)

About the Author

Wang Yongle (1972 -), male, Henan Biyang, associate professor, engineer, master, mainly engaged in information security, Web design research.

Xu Shuxin (1971 -), female, Xuchang, Henan, associate professor, engineer, master, mainly engaged in data mining, program design research.

Tel: 13849879289
E_mail: wangyongle316@163.com
Address: New Road, Xuchang City, No. 4336
Zip Code: 461000

Impact of Macroeconomic Variables on Stock Prices: Empirical Evidence from Karachi Stock Exchange, Pakistan

Muhammad Nauman Khan[1] and Sharif Zaman[2]

[1] Institute of Management Sciences,1-A, sector E-5, Phase VII, Hayatabad, Peshawar, 25000, Pakistan
mohdnaumankhan@gmail.com
[2] Institute of Management Sciences,1-A, sector E-5, Phase VII, Hayatabad, Peshawar, 25000, Pakistan

Abstract. Movement of stock indices is responsive to changes in macroeconomic fundamentals. The objective of this study is to find the relationship between macroeconomic variables and stock prices in Karachi Stock Exchange (KSE), Pakistan. The study considers annual data of several macroeconomic variables from 1998 to 2009: gross domestic product, exports, consumer price index, money supply M2, exchange rate, foreign direct investment and oil prices. The stationarity of data is checked through Augmented Dickey Fuller test. All variables are stationary at zero lag. Multiple regression analysis with Fixed Effects Model is then used. Results show that gross domestic product and exchange rate positively affect stock prices while consumer price index negatively affects stock prices. The results of export, money supply M2, foreign direct investment and oil prices were insignificant. This study will facilitate investors in taking effective investment decisions by estimating the expected trends in macroeconomic variables.

Keywords: Stock prices, Gross Domestic Product (GDP), Consumer price index (CPI), exports, Money supply M2, Exchange Rate, FDI, Oil prices and Linear Multiple Regression and Correlation Model.

1 Introduction

It is believed that domestic economic fundamentals play determining role in the performance of stock market. Knowledge of stock market sensitivity to macro economic behavior of key variables and vice-versa is important in many areas of investments and finance. This study will be useful to understand this relationship between macroeconomic variables and stock prices. The objectives of the study is to find out the nature of relationship between different macroeconomic variables and stock price.

2 Literature Review

The efficient market hypothesis (EMH) was dignified by Fama (1970). In an efficient market, current as well as past information on the growth of macroeconomic variables

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 227–233.
springerlink.com

are fully reflected in stock prices so that investors are unable to make some profitable trading rule using the available information.

In a study conducted by Suliaman D. Mohammad, Adnan Hussain and Adnan Ali (2009) the relationship between macroeconomic variables and stock prices in Karachi Stock Exchange is found considering the quarterly data of economic variables. The result shows that foreign exchange rate and foreign exchange reserve greatly affect stock prices. Interest rate and money supply (M2) is also significant.

Arshad Hasan and M.Tariq Javeed (2009) explored the long term dynamic relationship between monetary variables and equity prices for the period June 1998 to June 2008. Results showed that inflation and stock prices were negatively related.

Anokya Muhammad Adam and George Twenebah (2008) examined the impact of macroeconomic variables on stock prices in Ghana. The study established that there was cointegration between macroeconomic variables and in stock prices.

Liaquat Ali and Nadeem Ahmed (2008) examined the relationship between stock market development and economic growth in case of developing economy Pakistan. The data set covers annual time series from1971 to 2006. The results of the study indicated that there is a long run relationship between economic growth and stock market development for Pakistan.

Robert D. gay (2008) examined the relationship between stock prices and macro economics variables effects on four emerging economies India, Russia, Brazil and China. The findings were insignificants which proves inefficiency in market.

Shahid Ahmed (2008) conducted a study on SENSEX index price effects by real and financial sector performance in economy. This study finds the nature of the causal relationships between stock prices and the key macro economic variables of the Indian economy using quarterly data for the period March, 1995 to March, 2007.

Raman K. Agerawalla (2008) studied the causal relationship between the share price index and industrial production in a multivariate vector error correction model for India. The results of the study revealed that stock market in India are demand driven and industry led.

Dr. Shahbaz Akmal (2007) examined the impact of inflation and black economy on stock market prices. This study inspect whether this hold for Pakistan, over the period 1971-2006. The findings support the hypothesis.

Mohsen Mehrara (2006) examined the causal relationship between stock prices and macroeconomic variables in Iran. It was argued that stock price variability was fundamentally linked to economic variables, though the change in stock price lags behind those economic activities.

Desislava Dimintrova (2005) studied the hypothesis that there is a link between the foreign exchange and stock markets. The empirical findings of the study were somewhat weak.

Dr. Nishat and Rozina Shaheen (2004) investigated long term relationship between a group of macroeconomic variables and stock price in the Karachi Stock Exchange index. The result showed a causal relationship between stock prices and the economy. Macroeconomic variables Granger-caused stock price movements.

Bahram Adrangi, Arjun Chatrath and Antonio Z. Sanvicente (2002) studied a widely reported negative relationship between real stock returns and inflation for a major emerging market, Brazil. The results of the study back the negative relationship between inflation and real stock returns.

Fazal Hussain and Tariq Mehmood (2001) examined the relationship between stock prices and macroeconomic variables. The correlation analysis suggested low correlations between stock prices and macroeconomic variables. The cointegration analysis showed the presence of a long run relationship between stock prices and macroeconomic variables.

G.P.Diacogiannis, E.D.Tsiritkas (2001) study examines the case of the emerging Greek stock market and economy. Hussain and Mahmood (2001) observed unidirectional causality from macro variables to stock prices. Nishat and Saghir (1991) examined unidirectional causality from stock prices to consumption expenditure.

3 Research Methodology

The following variables can be identified from the literature.

Table 1. Variable Description

Variable	Concept	Description	Units	Source
SP	Stock price	KSE 100-index	Share price in Rs.	Business Recorder
GDP	Economic activity	Market value of all final goods and services during a year	Percentage growth per annum	State bank of Pakistan annual reports
EXPORT	Exports	Goods which are sends to foreign countries	In Million of Rs.	State bank of Pakistan annual reports
CPI	Inflation	Consumer price index	Percentage per annum	State bank of Pakistan annual reports
M2	Money supply	Money supply M2	In Million of Rs.	State bank of Pakistan annual reports
ER	Exchange rate	Average Exchange Rate	US dollar/ Rs.	State bank of Pakistan annual reports
FDI	Inward FDI	Volume of foreign capital invested in the economy	In million of Rs.	State bank of Pakistan annual reports
OP	Oil prices	Average Brent crude oil prices	Rs. per barrel	World Bank Commodity prices

3.1 Data Collection and Sample Size

The data used for the study is annually from 1998 to 2009 (12 years). For share price data one company is selected randomly from each sector. As there are 34 sectors in KSE, therefore, 34 companies and one company from miscellaneous sector a total of 35 companies were selected.

3.2 Hypothesis

Null Hypothesis: H_0: no impact of macroeconomic variables on stock prices

Alternate Hypothesis: H_1: impact of macroeconomic variables on stock prices

H_{11}: GDP is positively related to stock prices.
H_{12}: Exports is positively related to stock prices.

H_{13}: inflation is negatively related to stock prices.
H_{14}: monetary growth is negatively related to stock prices.
H_{15}: exchange rate is negatively related to stock prices.
H_{16}: foreign direct investment is positively related to stock prices.
H_{17}: oil prices are negatively related to stock prices.

3.3 Econometric Model

Fixed effects regression model is used here to analyze the individual effect of each and every unit in the model on the dependant variable. Many researches in the literature have analyzed panel data to control for time-specific effects. A frequently used method in this study is to include dummy variables and differentiate each time period from the others.

$$\log SP = \alpha + \log\beta_1 GDP + \log\beta_2 EXPORT - \log\beta_3 CPI - \log\beta_4 M2 - \log\beta_5 ER + \log\beta_6 FDI - \log\beta_8 OP + Dfirm_1 + \ldots\ldots\ldots + Dfirm_{34} + DY_1 + \ldots\ldots\ldots + DY_{11} + \eta \quad (1)$$

Where:

η is normally distributed error term, α and β_i's are unknown population parameters, α is the intercept, β_1, β_2 …, β_k are regression coefficients for variables GDP, EXPORT ……, OP, DY are dummy variables for years and D Firm are dummy variables for companies respectively. 34 dummies were created for 35 companies and 11 dummies were created for 12 years.

4 Data Analysis

4.1 Stationarity of Data

For checking stationarity of variables augmented Dickey fuller test was used at various difference levels. The results of ADF test for GDP, exports, CPI, money supply M2, exchange rate, FDI and oil prices shows that the series is stationary at 1% level at zero lag.

4.2 Multiple Regression

Simple Multiple Regression: In simple multiple regressions the impact of macroeconomic variables is analyzed on the share prices. The following Model was used:

$$\log SP = \alpha + \log \beta_1 GDP + \log\beta_2 EXPORT - \log\beta_3 CPI - \log\beta_4 M2 - \log\beta_5 ER + \log\beta_6 FDI - \log\beta_7 OP + \eta \quad (2)$$

Table 4. Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	.424[a]	.180	.166	.542688423813850	.291

4.3 Autocorrelation and Multicollinearity

The sample size is 420, there are 7 regressors and there is an intercept term in the model. The Durbin- Watson test statistic value is 0.291. Null hypothesis is rejected of non-autocorrelated errors in favor of the hypothesis of positive first-order autocorrelation. R^2 value is 18% which is a small variation in the dependant variables but there are few significant t ratios. Multicollinearity exists but its small.

4.4 Coefficients and Interpretation

The relation between GDP and share price is positive. H_{11} is accepted. This result is consistent with the studies of Suliaman D. Mohammad, Adnan Hussain, Adnan Ali (2009) and Fazal Hussain and Tariq Mehmood (2001) that GDP positively affect stock prices. The relation between exports and stock prices is positive. H_{12} is rejected.

The relation between inflation and Stock prices is negative. H_{13} is accepted. The results of the study are consistent with the studies conducted by Arshad Hasan and M.Tariq Javeed (2009), Anokya Muhammad Adam and George Twenebah (2008), Mohammad Shahbaz Akmal (2006), and Dr. Nishat and Rozina Shaheen (2004). The coefficient of ER is positive Rs.5.274. The hypothesis comes out wrong and ER rate has strong positive relationship with share price.

The relation between FDI and stock prices is positive. H_{16} is rejected. The result of the study is consistent with the results of the study conducted by Shahid Ahmed (2003).The relation between oil prices and Stock prices is negative. H_{17} is rejected which suggests that there is no significant relationship between oil prices and stock prices. The result of the study is consistent with results of the study conducted by Robert D. gay (2008), and Anokya Muhammad Adam and George Twenebah (2008).

Table 5. Coefficients

Variables	Coefficients	Std. Error	t-Statistic	Prob.	Tolerance	VIF
(Constant)	-5.274	5.148	-1.024	.306		
GDP	.886	.245	3.611	.000	.279	3.587
EXPORT	.699	1.543	.453	.651	.008	118.585
CPI	-.233	.246	-1.949	.145	.209	4.780
M2	.107	1.785	.060	.952	.005	210.905
ER	5.274	1.946	2.710	.050	.044	22.907
FDI	.639	.398	1.605	.184	.018	55.421
OP	-.425	.383	-1.110	.267	.050	19.892

4.5 Fixed Effect Model

The fixed effects model consists of 407 observations. It was a 12 years data from 1998-09 of 35 companies and 7 macroeconomic variables. The results of F-test shows that the model is significant at 1% level. The value of R square is 0.829 which means that 82.9% of the variation caused in the dependant variable is explained by these independent variables. The coefficient of constant is insignificant. The coefficient of GDP is positive Rs.1.37. This value is insignificant. The coefficient of export is negative Rs. -0.257664 but this value is insignificant. The coefficient of CPI

is negative Rs.-0.652359 and insignificant. The coefficient of money supply M2 is negative Rs.-0.442082. This is also insignificant. The coefficient of exchange rate is positive Rs. 6.69604. This value is insignificant. The coefficient of FDI is positive Rs. 0.341060 and this value is insignificant. The coefficient of oil prices is positive Rs.0.273065 and this value is insignificant. Out of the 34 firms only 15 firms has negative coefficient which means that they have a negative impact on the share prices. Eight of them have a significance of 1% two have 5% and the other five have an insignificant effect on stock prices. Nineteen firms have a positive effect on share prices. Ten of them have a significance of 1% two have 5% and the other seven have an insignificant effect on stock prices. In years the results of year1998, year1999, year2000 and year2001 have a positive effect on stock prices but they are insignificant. The other year's data is omitted due to collinearity problems.

Table 7. Selective Fixed Effects Model Results

const	-7.52146	28.7569	-0.2616	0.7938
GDP_growth_rate	1.37587	1.09165	1.260	0.2084
Exports	-0.257664	4.09741	-0.06288	0.9499
CPI	-0.652359	2.33962	-0.2788	0.7805
Money_supply_M2	-0.442082	4.26258	-0.1037	0.9175
Exchange_rate	6.69604	12.0544	0.5555	0.5789
FDI	0.341060	0.912195	0.3739	0.7087
average_crude_o	0.273065	4.05045	0.06742	0.9463
D1998	0.862943	4.29823	0.2008	0.8410
D2000	0.136514	0.783053	0.1743	0.8617
D2001	0.341235	0.515202	0.6623	0.5082
DAl_Meezan	-0.625706	0.108136	-5.786	1.56e-08***
DYousaf_Weaving	-0.962282	0.105690	-9.105	5.93e-018***
DNishat_mills_l	0.0223045	0.105690	0.2110	0.8330
DThal_jute_mill	0.174702	0.105690	1.653	0.0992 *
DHabib_sugar__m	-0.218664	0.105690	-2.069	0.0393 **
Dpakistan_natio	-0.308738	0.105690	-2.921	0.0037 ***
Dcentury_paper	- 0.0300241	0.105690	-0.2841	0.7765
DWazir_ali	- 0.290508	0.105690	-2.749	0.0063 ***

5 Conclusion

The result of the study showed that GDP, ER positively affect stock prices while Consumer price index negatively affect stock prices. The results of Export, Money supply M2, FDI and oil prices were insignificant which shows that these variables do not have significant relationship with stock prices. The correlation analysis shows strong correlations between stock prices and macro variables. Share price has a weak correlation with CPI while with other variables it has relatively stronger which suggests that these variables does influences stock prices and investors should aware themselves on information regarding these variables.

Affect of macroeconomic variables on stock prices is an international phenomena and it cannot be forecasted easily. Stock exchange performance is affected by the constant tightening of the monetary policy by the government. These policies should be revised to make them effective. Rise in oil prices and inflation plays an important part in determining the stock market trends. Inflation stays the biggest danger to the

economy jumping to more than 20% in 2009. The consistent rise in inflation and oil prices badly affect stock prices. When inflation increases it results an unexpected increase in the interest rate by the state bank of Pakistan which ultimately resulted in sharp decrease in Karachi stock exchange.

Foreign direct investments boost up the secondary market. In Pakistan the foreign investment is continuously falling and it needs to be increase because of stable political condition and effective law and order situation. Investors will invest in those countries where they feel they are safe and secure to invest. The current situation of Pakistan is not appropriate for the foreign investors. The law and order situation is worsening day by day which discourage for the foreign investors. The government should provide protection to the foreign investor in order to increase foreign investment in country.

References

1. Mohammad, A.A., George, T.: Do Macroeconomic Variables Play any role in the Stock Market in Ghana? MPRA Paper No 9368, pp. 20–32 (June 24, 2008)
2. Adrangi, B., Chatrath, A., Sanvicente, A.Z.: Inflation, output, and stock prices: Evidence from Brazil. Journal of Applied Business Research 18(1), 61–76 (2002)
3. Diacogiannis, G.P., Tsiritakis, E.D., Manolas, G.A.: Macroeconomic factors and stock returns in a changing economic framework: The case of the Athens stock exchange. Managerial Finance 27(6), 23–41 (2001)
4. Dimitova, D.: The Relationship between Exchange Rates and Stock Prices: Studied in a Multivariate Model Retrieved in Issue in Political Economy (2005)
5. Akmal, S.: Stock return and inflation in case of Pakistan. Pakistan Social and Development Review (2007)
6. Gay Robert, Jr., D.: Effect of Macroeconomic Variables on Stock Market Returns for Four Emerging Economies: Brazil, Russia, India, And China. International Finance and Economic Journal (2008)
7. Hasan, A., Tariq Javid, M.: An Empirical Investigation of the Causal Relationship Among Monetary Variables and Equity Market Returns. The Lahore Journal of Economics 14, 115–137 (2009)
8. Hussain, F., Mahmood, T.: The stock Market and the Economy of Pakistan. The Pakistan Development Review 40(2), 107–114 (2001)
9. Mehrara, M.: The relationship between stock market and macroeconomic variables: A case study for Iran. Iranian Economic Review 10(17) (Fall 2006)
10. Nishat, M., Shaheen, R.: Macroeconomic Factors and Pakistani Equity Market. The Pakistani Development Review 43(4) (2004)
11. Agrawalla, R.K.: Share prices and macroeconomic variables in India: An approach to investigate relationship between stock markets and economic growth, pp. 136–146 (2008)
12. Shahbaz, M., Nadeem, A., Liaquat, A.: Stock market development and economic growth: ARDL causality in Pakistan. International Research Journal of Finance and Economics (14) (2008) ISSN 1450-2887
13. Ahmed, S.: Aggregate economic variables in stock markets in india. International Research Journal of Finance and Economics (14), 141–164 (2008) ISSN 1450-2887
14. Mohammad, S.D., Hussain, A., Ali, A.: Impact of macroeconomic variables on stock prices: empirical evidence in case of KSE. European Journal of Scientific Research 38(1), 96–103 (2009) ISSN 1450-216X

Research in Transaction Model of ERT Database Management System

Yongle Wang and Shuxin Xu

Xuchang ploughs the recent information science research institute, 461000, Xuchang, China

Abstract. This paper gave a bird's-eye view of the Embedded Real-Time Database Management System（ERTDBMS） and its transaction model of the characteristics of the status quo, Analysis Services and Analysis, alternative /compensation model, based on function and value of the object-oriented framework Panel model Analysis Servicese and Analysis of the ERTDBMS, and discussed the scheduling strategy ERTDBMS Affairs and concurrency control, the Panel stated ERTDBMS model of significance.

Keywords: Embedded database, real-time database, transaction mode.

1 Introduction

At present, the computer has entered the post-PC era. Following the PC products, the digital information processing has become a major tool. In these embedded devices, it requires sophisticated data management capabilities and the required functionality are often very complex, so that the flat file system is insufficient to handle and manipulate the data, which contributed to the demand for embedded database system. The designed aim of embedded database system is to conducting data storage and recovery under the circumstance of minimum interference and the minimum influence of system. As it is often necessary to make real-time response to the environment, this concept is based on real-time or near real-time embedded computer, more precisely, the more demand are embedded real-time database systems. Although Embedded real-time database systems are research focus, the research is still in its infancy, no mature products are made currently, the market has some embedded database systems and real-time database system products, but the embedded database system is basically the cut of traditional commercial database system, real-time database system model and the concurrency control mechanism is not new, they follow some control mechanism of common database system, or be partly transformed.

2 Embedded Real-Time Database System

Embedded real-time database management system (ERTDBMS) is a new data management techniques rising in only recent years. It is currently a sophisticated database technology, directing at the characteristics of embedded devices and realizing data storage, organization and management [1] of mobile devices and embedded device.

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 235–242.
springerlink.com

Embedded systems, desktop and corresponding parts of server version have much in common. However, based on its unique needs and limitations, to support such applications in embedded real-time database system should be with the following features. (1) transaction complexity. (2) transaction compensatory and implementation . (3) determining the correctness of transactions. (4) access to external databases. (5) higher requirements of the analysis of the Ability. (6) Time and space trade-offs. (7) a brief recovery mechanism [2].

In short, the embedded real-time embedded database system application requires the embedment, initiative, real-time and memory database, particularly it needs more in-depth study and discussion in the transaction model, features, correctness criteria, transaction processing and the storage structure of the system.

3 Characteristics and the Research Status of ERTDBMS

3.1 ERTDBMS Transaction Characteristics

Embedded real-time database system model is similar to traditional database and its operation is also based on the form of transaction, the system as a unit assigns transaction CPU, data and other resources, conducting priority allocation and scheduling processing. At the same time, ERTDBMS the transaction is very different with the traditional transaction, the transaction can have time constraints, the system throughput is not the goal, but a single transaction to meet the timing constraints , in order to satisfy the largest proportion of the transaction of timing constraints; the atomicity, consistency, isolation, and permanent of traditional transaction seems too strict or even impossible in embedded real-time environment; require the use of "know when" mechanism to deal with scheduling or concurrency control of matters, rather than the traditional first-come, first Service mode.

The affairs in ERTDBMS and in the traditional have fundamental distinction, mainly in:

(1) the transaction correctness is not only in the correctness of logical results, and time requirement, the transaction must be given within the deadline of completion;

(2) the predictability of services;(3)restorability(Compensation).

3.2 ERTDBMS and Services Model Research Status

The current embedded real-time database system can be divided into two categories: one is commercial-grade embedded real-time database systems, such as the eXtremeDB memory real-time database provided by United States McObject company. The other is the embedded real-time database system which is the user designed and developed according to the specific application objects.

Kim created a real-time database transaction model. Braoudakis take a different approach to associate matters with a value function. Zhou, Rundensteiner, and Shin will integrate object-oriented view into of real-time database system, the proposed ROMPP, using object-oriented framework to discuss the consistency and correctness of temporal and logical. Currently, among the transaction model of embedded real-time database system, the real-time transaction model of function-based alternative/compensation

designed by Professor Liu Yunsheng of Huazhong University of Science and PhD Professor Xia Jiali is the most sophisticated[3].

Real-time transaction-based model based on functional alternative replace the concurrency control and scheduling as the basic unit can improve the success rate of transactions. In the transaction model of support real-time compensation, the alternative improve the success rate of the transaction, but can not guarantee absolute success matters, when all alternatives have failed or can not be successfully replaced in its execution before the deadline, it needs compensatory measures, in particular, Some which do not have the alternative real-time transaction need compensation for the failure.

4 Transaction Models and Processing Techniques of ERTDBMS

4.1 ERTDBMS Transaction Pre-analysis

The real-time application in ERTDBMS contains a wealth of data, transactional semantics and complex structure, with many features different to traditional applications, in the transaction we must take into account the conflicts between data and resource, the interdependence between transactions, the dynamic "production" of new activities (services) and the factors of affairs mortality[4]. Thus, the error of actual transaction execution time and estimated the worst case execution time is big, so the system must have a certain ability to predict and handle, can carry out " schedulability "analysis and prediction, know a transaction to meet their time limits in particular the possibility of cut-off time in advance, whether there are "danger" and to take special measures to ensure the works are properly completed.

(1) When transaction reaches system or the system is produced, it needs conducting preanalysis. Extracting the alternative set, data sets, logical operations(type and order),timing requirements,urgency, and the key, running time estimates, possibly triggered activities/affairs, information, and the structure, behavior, and timing and other related information of other aspects between the various panels.

(2) When the system is running at the right time, using certain algorithms conducts correlation analysis on the current active affairs to support and assist scheduling algorithms and the implementation of concurrency control strategies.

(3) Before the implementation of scheduling and concurrency control, conducting "Schedulability" forecast dynamically, determining the likelihood that they can complete normally or the degree in order to take appropriate measures to ensure its implementation as possible.

(4) If necessary, take appropriate measures, such as the dynamic adjustment of priority and scheduling algorithm, the implementation of "functional replacement"or"compensation" and other activities. The pre-analysis can also be a static or dynamics implementation, which does not affect system performance. The steps pre-analysis algorithm is generally as follow:

Pre-analyse(T)
Input: real-time transaction T
Output: alternative set of T

Steps:①generate task tree TK-Tree, and to extract information about structure, behavior, resource requests, knowledge of time limits;②decompose TK-Tree and generate the scheduling tree; ③to conduct a schedulability analysis of scheduling tree; ④produce weak scheduling trees and strong scheduling trees.

Deadline is an important feature of ERTDBMS affairs, to determine the cut-off is the primary work of pre-analysis of transaction. From the identification of ways, the main are assignment deadline, deduct deadline and deadline of implicit interpretation and so on. In ERTDBMS, the strategy of precise transaction scheduling involves deadlines and execution time. Like deadline, the program execution time becomes another important problem expected to solve [5]· Literature 5 presents a method for estimating program execution time and it will be divided into two steps: First, the static WCET (Worst Case Execution Time) estimate, explaining that the implementation of the program is finished in X seconds, the accuracy of scheduling dependents on the tightness of this interval; the second is the gray system theory, by analyzing the gray information to find out the law and to predict the implementation of transaction events and get dynamic time estimates.

4.2 ERTDBMS Alternative/Compensation Services

A application with a time limit is a real-time transaction, it has a number of tasks, some tasks can finish real-time transaction capabilities. Real-time transaction possesses the feature of functional alternative, a real-time transaction set contains more than one task and they are functional equivalent, known as the alternative task set. Each task also contains a set of child transaction of functional equivalence. In A task alternative set of real-time transaction, the set constituted by each member of a mission is called functional alternative set. Real-time transaction set is composed of multiple functional alternative, when the transaction is scheduled, a suitable capabilities replacement set is selected from the real-time transaction, if this feature alternative set fails and the deadline of services is not yet reached, the system will attempt to select another feature alternative set of this real-time transactions to continue investment in schedule, if the system can consciously choose functional alternative set with high success rate, the efficiency of the system can be improved.

The establishment of real-time transaction model based on function alternative (Function Alternative Transaction Model, FATM) fully demonstrates that a real-time transaction actually possesses complex three-dimensional structure. To take three steps in the pre-analysis method: obtaining all the functional alternative set by decomposing affairs; extraction and analysis of information and knowledge related to the operation one by one; study the function of schedulability and order to provide direct information for scheduling. The pre-analysis process of functional alternative set is shown in Figure 1. When conducting system scheduling, firstly selecting an object from the tree of emphasis degree, if there is no emphasis degree tree, selecting an object from weak emphasis degree tree, in the soft real-time environment, the object may be selected from the general operation trees.

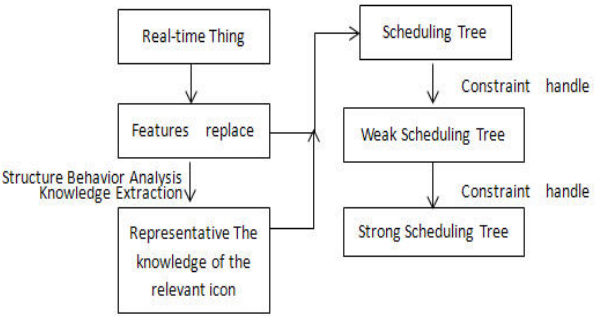


Fig. 1. Alternative set of real-time transaction capabilities pre-analysis diagram

ERTDBMS transaction model based on functional alternative is in favor of the transaction pre-analysis, as the early scheduling of scheduling, it is directly related to the implementation success rate, thus saving the system run time. Meanwhile, the substitutability of affairs make implementation possess multiple paths to improve the success rate of the transaction. However, if all alternative sets(or time does not allow) can't be successful implemented, then it still has to die. Therefore, in order to avoid a system disaster caused by this, it is necessary to expand FATM to make it support compensatory, that is, transactions are composed of compensation tasks and the main tasks, when the main task can't be successfully executed, the compensation task to make it finished safely(of course not every services are available for compensation.)

4.3 The Real-Time Transaction Model of Value Function-Based

The transaction model with functional alternative possesses regularity, functional equivalence, isomorphism and so on[6].Using the consensus view to treat real-time and non-real-time transactions in system, directing the value function of transaction into the transaction model and enabling real-time transaction model to meet the better features. The value function of real-time transaction is shown in Figure 2.

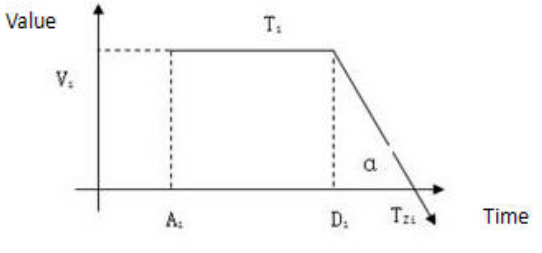


Fig. 2. Real-time transaction value function

Non-real-time transactions can be considered as a special real-time transactions column (not Di's requirements), and its value does not change over time, is constant Vi. The value function of real-time transaction is discussed as below according to different situation:

(1) t<Di Time,VTi∈RT, with V (Ti) = Vi (Vi is a constant);
(2) t>Di Time,VTi∈HT, with V (Ti) =- P(P is a constant);
VTi∈FT, with V(Ti)=0; VTi∈ST, with V(Ti) = 0 (TZi-t) tanα.

4.4 The Transaction-Based Model of Object-Oriented Framework

Currently, most of the works of ERTDBMS is based on the relational model. However, the area of object-oriented database system ERTDBMS (Object-Oriented Database Management System, OODBMS) attracted the attention of researchers. Usually, ERTDBMS research is motivated by the advantages of borrowing OODBMS technology to solve real-time system (RTS) data management issues.

RT-OB：=<O_ID，O_ATTR，O_M，O_C>
A real-time data object of (Object-Oriented OO) data model is defined as quad:
RT-OB: = <O_ID,O_ATTR,O_M, O_C>
RT-TRAN：=<T_ID，T_OP_Object，T_OP_Relation，TTP，TT，Priority，RR>；

With four parts of the structure of the object model, each element is extended to specify real-time constraints of different situation and meet the time limit in the operation of data manipulation. Consistency characteristics of the data (especially time consistency) will inevitably lead to real-time characteristic of ERTDBMS affairs, the characteristic of timing. In traditional database system, a transaction is defined as a data submission and died at the sequence of operations. Transaction model is defined as the following six groups by the call feature of one or more objects or relations:

RT-TRAN：=<T_ID，T_OP_Object，T_OP_Relation，TTP，TT，Priority，RR>；

①T_ID: affairs, the only identifier; ②T_OP_Object: matters set of related objects; ③ T_OP_Relation: relation sets of association affairs; ④TTP: transaction type; ⑤ TT: transaction time information; ⑥ Priority: database manager give priority to services; ⑦ RR: Services Resources Needs.

4.5 ERTDBMS Transaction Scheduling Strategy and Concurrency Control

In ERTDBMS transaction scheduling, it mainly take advantage of two factors: transaction arrival time and transaction deadline. Scheduling strategy mostly used the priority scheduling method, such as the first release of the highest priority (Earliest Release First, ERF), the first deadline priority (Earliest Deadline First, EDF), up to the earliest deadline first (Earliest Feasible-Deadline First, EFDF), the minimum relaxation time priority (Minimum Slack First, MSF), the value of the highest priority (Highest-Value First, HVF), etc.[7]. And thus derive some new strategies, such as adaptive earliest end time (AED) protocol, which is based on the EDF allocation strategy, the

transaction is divided into two groups according to predetermined control variables, the sorting of critical value, according to the key value of the order list .

Concurrency control is to control the interaction between concurrent transactions (or coordinated scheduling), so that the database consistency (ERTDBMS in particular, transaction time consistency) is not destroyed. ERTDBMS concurrency control is based on the traditional database, consider the time factor, propose a pessimistic concurrency control, optimistic concurrency control and multi-version concurrency control protocols. Phase locking is the most popular pessimistic concurrency control protocol, using this method, the transaction is prohibited to access the database target before obtaining the lock, the transaction execution is constituted by two parts, the lock and release the lock, which Includes the suspension of priority, priority inheritance and conditional priority inheritance and so on. Optimistic concurrency control protocol allows the transaction barrier-free operation until all the implementation are completed, and then verify when submission, if it is passed, submitted or restart. In the multi-version concurrency control protocol, each data has multiple versions, reading the transaction can read the old version had not been covered to ensure serializability, thus increasing the concurrency and reduce the reject operation.

5 Conclusion

This paper mainly discusses the characteristics of embedded real-time database and its transaction model and its treatment of the theory, technology and implementation methods, including ERTDBMS features, ERTDBMS transaction model characteristics and Research, the transaction model based on the replacement / compensation, the transaction model based on value function, transaction model based on object-oriented scheduling strategy and ERTDBMS affairs and concurrency control. Embedded real-time database system as an user process of embedded real-time operating system, , the implementation of its affairs in time and space must have a major conflict, how to resolve the conflict between them can be resolved partially gradually from the improvement of transaction model and its scheduling policy and the complicated mechanism. Therefore, the more in-depth study and treatment of the transaction model ERTDBMS become a research hotspot and also has high research value and practical significance.

References

1. He, X.-G., Liu, Y.-S.: Special Database Technology, pp. 203–230. Science and Technology Press, Beijing (2007)
2. Pan, Y., Lu, Y.-S.: Nested transaction concurrency control in parallel real-time databases. J. of Donghua Univ. 22(2), 114 (2005)
3. Liu, Y.-S., Ding, L.: Design of embedded real time database manage system. Application Research of Computers 8, 230–231 (2006)
4. Abhott, R.K., Garcir-Molina, H.: Scheduling real-time transactions: a performance evaluation. ACM Trans. on Database Syst. 17(3), 513 (1992)
5. Xia, J.-L.: Real_Time Transaction Model Based on Alternative/Compensate. Computer Engineering and Applications 39(34), 25–27 (2003)

6. Xia, J.-L., Liu, Y.-S.: Sitisfying Predictability for Embedded Real-Time Database Systems. Mini-micro Systems 24(2), 234–237 (2003)
7. Braoudakis: Timeliness Via Speculation for Real-time Database. In: Pro.of IEEE Real-time System Symposium (1994)

About the Author

Wang Yongle (1972 -), male, Henan Biyang, associate professor, engineer, master, mainly engaged in information security, Web design research.
Xu Shuxin (1971 -), female, Xuchang, Henan, associate professor, engineer, master, mainly engaged in data mining, program design research.
Tel: 13849879289
E_mail: wangyongle316@163.com
Address: New Road, Xuchang City, No. 4336
Zip Code: 461000

Study on the Learning Evaluation of Web-Based Learning Platform*

Kejin Xu[1] and Yang Du[2,**]

[1] School of Pharm Ceutical, Changchun College of Traditional Chinese Medicine, Changchun, P.R. China
[2] School of Chemistry and Chemical Engineering, Southwest University, Beibei, Chongqing, P.R. China
duy360@swu.edu.cn

Abstract. Network bring further study and further increase opportunities for people to build up a platform for learning. Learning platform is an open, equitable and other characteristics; learners learn to use e-learning platform, showing the individual, collaboration and unstructured features. I start from these features, combined with study and evaluation made of the current web-based learning platform for learning evaluation there are some problems. This learning platform for learning to re-evaluate the content of the building, including the evaluation of the learning environment, learning methods and assessment of learning styles, learning attitude and emotional evaluation of network resources utilization of the evaluation.

Keywords: Online learning, learning evaluation, unstructured learning, academic abilities.

Introduction

With the rapid social development, information technology in people's lives increasingly dominant, and the network learning platform for people has been build up. However, due to the special nature of learning, the teacher's control weakened, to make learning to be successful, in large part dependent on the learners to manage their own learning activities. In this regard, how to have a general grasp of the situation of self learning, and guide teachers on how to learn more about teaching has become an urgent problem to be solved for learners and teachers. Web-based learning platform for learning evaluation of the learning process and learning effect and to provide learners with all aspects of learning support is a very good answer to this question. Therefore, how to use Web-based learning platform for learning evaluation to improve the quality of network learning has an important and practical significance.

* This study was supported by the Southwest University Teacher Education Innovation Platform projects.
** Corresponding author.

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 243–248.
springerlink.com

1 A Web-Based Learning Platform Evaluation Overview of the Learning

Learning evaluation Web-based of learning platform is a network learning tool of teachers and learners for mutual understanding, is a powerful tool to improve the quality of online teaching. To evaluate the implementation of e-learning platform, learning platform for understanding the characteristics of the learner's own characteristics and use this platform to the characteristics of learning activities is very important, these features are the basis for the evaluation of learning, but also help smooth correct evaluation implementation.

1.1 E-learning Platform

Online learning is defined as "a variety of courses to users to learn, and make the learning process records online platform." Familiar with its features, it is correct to ask questions and make the basis for evaluation. So that learning is a fundamental change in the learning environment provided, it is this transformation of the learning environment in which learning has become a new way of learning, showing openness, equality, independent of other characteristics.

1.2 Evaluation Based on Learning Platform

Web-based learning platform for learning assessment, online learning is based on the characteristics and features to make learning presents a reasonable assessment. Many studies have shown that learner self-evaluation can play an active and creative role, learning to improve the quality of the individual learner. Learner interaction in the process, they evaluate each other's role is more important when operating in an open, students can be shared between test operations, while increasing the quality of the learner's sense of responsibility for the job is complete, collaborative learning to improve the ability of learners .

A wide range of evaluation methods for the assessment of this study is the most appropriate, and the network itself provides a learning platform to implement the possibility of multiple assessment techniques, a variety of communication tools, collaboration tools, tools, and published faster feedback, virtual reality technology are toward the support of multiple assessment of direction. Compared to traditional evaluation methods, evaluation of online learning is more focused on the learner's learning style, learning style assessment.

2 Learning Evaluation Problems of Web-Based Learning Platform

Web-based learning platform to learn the special nature of the activities of teaching and learning activities is relatively isolated, traditional teaching and social impact of the high demand for qualified personnel to such learning has brought some problems. Among them, the learning evaluation is particularly prominent.

2.1 Learning Evaluation Only on Results, the Lack of Procedural

Many of China's network of educational institutions for the evaluation of the learner continues to follow the traditional "homework + examination" of the evaluation methods. In this way emphasis on evaluation of learning outcomes, the learning process of learners cared little about the learner in the learning process demonstrated the ability, learning attitude, learning styles, and the accumulation of prior knowledge of learners and other little attention , the lack of prediction, diagnosis and guide the development of learner information. Lack of evaluation of the entire learning process, making the lack of systematic study and evaluation, can not truly reflect the web-based learning platform for learning evaluation system characteristics and, therefore, can not fully reflect the learning of facts, lack of guidance on teaching.

2.2 The Evaluation Too Much Emphasis on Theory, Assessment Forms the Lack of Diversity

Currently, the online education course examination is still the intellectual, academic content-based, rote learning scores share a large proportion of the kinds of questions, not only suited to the characteristics of open education, more and contrary to the modern concept of online education, and thus fail to realize open education self-study, is not conducive to learners' professional skills training, innovation and the main spirit. In addition, too much time on the Internet to conduct an objective evaluation of the test, for example, the results of subjective questions to assess, due to computer software evaluation criteria in the test results there are some technical difficulties, so many web-based learning tests using objective evaluation methods. This form only applies to certain disciplines, for strong liberal arts and applied disciplines showed a significant discomfort.

2.3 Learning Evaluation Is Not Flexible

In strong support of modern information technology, in order to provide all learners are not time-limited learning opportunities, online education has moved from "the object of open time, open space," has taken a solid step, and is exploring professional open, open the course, teaching the open and so on. But the openness with the teaching, online education learning evaluation is still closed, it is mainly reflected in: evaluation of the main single, ignored by critics in the evaluation of the status and role; examination there is "one size fits all" phenomenon.

2.4 Learning Evaluation Delay

Although Web-based learning to overcome the delay in the evaluation of transfer letter, but the separation of its students and individual learning characteristics, inevitably leads to the feedback time delay. This will enable learners to the knowledge and skills, the knowledge content of the questions and problems on the job requirements is often not timely answer, which sometimes is not well understood learning content, coupled with space-time distance students, teachers in the learning the virtual nature of the minds, the learner's own self-confidence and other reasons, tend to affect the learners'

interest in learning and learning motivation, and can not be adjusted so that students learn the next step.

3 Re-build Web-Based Learning Platform to Evaluate the Content of Learning

Prior learning assessment often used summative evaluation. More attention to the summative evaluation phase of the learning outcomes, and learning itself is a dynamic process of development, not only to conduct summative evaluation, more importantly, the whole process of evaluation of learners learning --- the process of evaluation. The web-based learning platform of learning evaluation is learner-centered learning to study the background, means, methods and results achieved by all the learning process in order to find ways for learners to study or test the quality of learning of learners. Evaluation follows the guidance, integrity, objectivity and scientific principles, based on these principles, Web-based learning platform, learning evaluation from the following aspects.

3.1 Evaluation of the Learning Environment

Evaluation of the learning environment is in order to examine the status, characteristics, problems and learners' actual and potential to produce various effects, and these information to teachers and administrators, the environment to maintain the advantages and disadvantages can be excluded. We will be learning environment evaluation is divided into three categories: environmental indicators hardware, software, environmental indicators and indicators of psychological environment. Hardware refers to the school principal environmental indicators hardware; software environment indicators such as primary school teachers; psychological environmental indicators, including the learner's own personality traits, temperament, achievement motivation. Understand the learning environment will enable more targeted teaching, develop in line with the actual situation of the learner, teacher training measures, individualized, but also give full play to people's initiative, to track the learning process of development and change.

3.2 Evaluation of Utilization of Network Resources

Utilization of network resources is the learner's understanding of learning content, digestion, that is, knowledge of the process of constructing meaning and can always upload your own opinion on the content of learning resources, reflection notes. Network resources include their own courses resources and Internet resources. How can grasp the learner a more specific level of utilization of resources, we can login by user or IP address tracking and monitoring real-time tracking monitoring monitor tracking to solve.

3.3 Learning Style and Ability to Evaluate

Learners in the use of some information about self-learning methods and learning strategies, is based on their own learning style set. Learning style assessment can help

learners understand their own characteristics. Under the guidance of teachers choose appropriate learning methods and strategies to regulate their learning behavior, from the truly achieve individual learning.

Web-based interactive learning platform and the characteristics of individual learning, learners participate in the exchange and collaboration and academic ability is the ability to develop basic learning ability, the ability of these two comprehensive evaluations of the quality of learning is a test, is also a learning platform to build a test case.

3.3.1 Evaluation of Learners' Communication and Collaboration Capabilities

Distance learning is a two-way interaction, two-way transmission of information, network with each other to share resources and ideas, everyone can participate in group discussions, so that learners can learn on the platform BBS, chat rooms, Email, video or audio conferencing systems and other interactive media network with teachers or other learners in real-time and non real-time interaction. In learning, so that communication and collaboration is to get learners to acquire knowledge and an important means and motivation for learning to stimulate and maintain the results is extremely important. Therefore, the learners participate in the exchange and collaborative learning ability is an important criterion for evaluation of learners.

3.3.2 Evaluation of Learners' Academic Abilities

Bloom's target the cognitive domain is divided into six categories, namely, knowledge, understanding, application, analysis, synthesis and evaluation. Level gradually increased the ability of the corresponding requirement will be increased gradually during the capacity evaluation, different situations can show different characteristics, which, in the evaluation of capacity combined with the discipline, the same learners face different disciplines, the ability evaluation results is not the same. Information age requires students to have the practical ability, innovation, provide more learners to play, the ability to display their own space, and arranged some of the development potential of the work of learners, while learners work on the evaluation of high-quality and timely feedback to learners, to promote more timely understanding of their learners and make further efforts.

3.3.3 Online Learning Outcomes Assessment

Knowledge of learners study, testing the improvement of thinking skills of learners, learner evaluation of online learning through the network achieved a measure of learning outcomes is important to the success of online teaching standards. Assignments and test scores is a measure of mastery of knowledge. But the purpose of education is more important is to improve the quality of learners to improve learners discover, analyze and solve problems. Therefore, online learners' learning outcomes reflect the usual manifestations. By evaluating the learner's learning, on the one hand can help learners to make better self-evaluation; the other hand, learners can help teachers according to the situation of teaching strategies and make appropriate adjustments.

Evaluation of learner achievement is a dynamic online learning process. In this process, there is no time limit, or the learner can always graduate school. So that learners can own some cases to make reasonable arrangements for their own learning,

so as to achieve the best learning results. Dynamic assessment of learning outcomes of learners for learners can provide a comfortable learning environment so that students practice reasonable arrangements for examinations or the time. In addition, online learning is not just the final results of the assessment, which also includes learner engagement in learning, learners complete the job as well as participate in discussion of the situation, etc., to achieve the diversification of evaluation. In this multi-factor, multi-level evaluation of complex issues, the use of fuzzy comprehensive evaluation method can often be more reasonable and comprehensive results. Fuzzy comprehensive evaluation method is applied to the evaluation of fuzzy mathematics to solve the multi-factor fuzzy data under the premise of a comprehensive problem. Using this method of evaluation can truly reflect the learning outcomes of students online.

4 Conclusion

Compared with the traditional evaluation of learning, learning platform, web-based learning evaluation has its own particularity, the learners learn and teachers have a great help to teachers and students can get in the shortest possible time and traditional learning almost the same feedback. Of course, Web-based learning platform to learn the specific implementation of continuous evaluation, with the continuous advancement of learning, technology continues to develop, will start to appear some problems, we continue to develop only constantly to explore, to be able to timely to solve these problems, so that learning can go on to better evaluate the system continue to grow and develop.

References

1. Carnevale, D.: Learning online to teach online. The Chronicle of Higher Education 50(10), A31–A32 (2003); Smith, T.F., Waterman, M.S.: Identification of Common Molecular Subsequences. J. Mol. Biol. 147, 195–197 (1981)
2. Easton, S.S.: Clarifying the instructor's role in online distance learning. Communication Education 52(2), 87 (2003)
3. Hassett, J.: The e-learning survival guide: team e-learning managers at financial service firms collectively representing more than $300 billion in revenue share words of wisdom on surviving the e-learning jungle. Training 39, 120 (2002-2003)

Research of Public Service Units Performance Appraisal Based on Weight-Variable Theory

Zhou Xiaoshi, Jin Minli, and Ling Ping

School of Economics and Management, Shenyang Ligong University, Shenyang City, Liaoning Province, P.R. China, (110159)

Abstract. This paper built a common performance appraisal system that based on comprehensive research of scholars at home and abroad to public service units performance appraisal, including assessment indicator framework, weight-variable of organizations performance appraisal. The weight-variable model of public service units performance appraisal can evaluate the result quality, help to reward the excellent and publish the inferior and overcome the disjoint between performance and strategy.

Keywords: Performance Appraisal, Weight-Variable Model, Evaluation Indicators, Public Service Units.

1 Introduction

In the planned economic system, Our country's public service units have very obvious government characteristics, because they belong to the government, and they are derivatives of the government. With the establishment of market mechanism, more and more contradictions and problems appeared just because of the excessive attachment of the government. Since 1980s, The Public Service Units have carried out various forms of reform. However many problems still haven't been solved, what's more many new problems appeared in the new situation. Among them the most important thing is how to improve the efficiency of the Public Service Units in the changing market environment, which is directly connected with the work efficiency and service quality of the Public Service Units, and the enhance and improvement of efficiency is the starting point and destination of performance evaluation.

In the background of this situation, Research of our country's Public Service Units performance appraisal has important meaning to improve the efficiency and promote the reform of our country's Public Service Units. First, We should make some theory and practice discussion in order to make our country's Public Service Units performance appraisal get into the scientific, standardized, institutionalized track. Second, with the background of the globalization, it can drive things as follows: Improving the management of the Public Service Units, changing the management concept, improving the work efficiency and management ability and the ability to respond to of Public Service Units.

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 249–258.
springerlink.com

2 Study Document Summarize of Internal and External

The name of "Public Service Units" exists only in our country, Its basic characteristics is hosted by country, public welfare and the special field of activity. There is no appellation of "Public Service Units" in foreign countries. However , the Public school, the Public hospital, the public library, the public museum, the public research institutions and so on also have the same basic characteristic of our country's Public Service Units. So our country's institution is equivalent to "Public institutions" or "Public Organization" in Overseas.

Through the study document of internal and external, we can summarize the progress and the theory degree of Public Service Units performance evaluation at present.

2.1 Research Degree

"New public management movement" of foreign countries, and the reform of Public Service Units since 1980s can be regarded as the beginning mark of theory and practice performance evaluation in the world. Trough 20 years development, Public Service Units performance appraisal of our country has accumulated a solid theoretical foundation and rich practical experience, but at present Public Service Units performance appraisal has not formed a complete and systematic theory. Even so, based on the Government and enterprise theory, we still can form the basic knowledge of the targets, standards, indicators and methods of the performance evaluation.

2.2 Development

1) values appraisal: Comprehensive development from "3E" "4E",to"Quality", "Service first", "Responsibility".

2) The diversification of evaluation subject: the development from evaluated by Government or institution it self, to evaluated by Intermediary organization(third party) or by the public.

3) Evaluation Framework: It has experienced a developing process from focusing on the results, to focusing on the process and efficiency of input and output, finally to the connection of performance appraisal strategy and quality.

4) Evaluation methods: In several common evaluation methods in practice at home and abroad. balanced scorecard has been widely used because of its following benefits: comprehensive, systematic, strategic, benefit to find the gap and so on. In particular, balanced scorecard is tissue performance evaluation method which has been widely used. What's more, With its simple and practical, rigorous scientific characteristics, Common public organization assessment indicator framework in Europe is taken seriously.

5) Evaluation index system: Public Service Units performance appraisal is divided into two categories: The first is special use type, evaluation index system is

established based on the difference between the industry and region of Public Service Units. The advantage is specific, pertinent, benefit to further study. But the disadvantage is that there is no general form, so it is not benefit to the promotion and lateral comparison. The second is universal type, such as common public organization performance appraisal in Europe. The advantage is it is benefit to sum up experience and promote comprehensively, but the disadvantage is that it can not reflect different situations in different Units, because it is too vague and general. So we should add new index which can reflect different situations in different Units when we use it.

3 The Contraction of Public Service Units Evaluation Indicators System

3.1 Public Service Units Performance Appraisal Dimension

With the aid of Balance Mark Card we divided Public Service Units performance appraisal into 4 dimensions.

1) The public interest of Public Service Units. The basic value of Public Service Units performance appraisal lies in the realization and promotion of social public interests, which is decided by the public interest. 2) Public satisfaction. Public Service Units is service provider, then citizen is the consumer of public service. 3). Internal process. The internal process of public service units is the foundation of providing service and realizing benefits, for the purpose of obtaining public interest success, we should satisfy the public demand by creating new public product and service ,So Public Service Units should optimize internal process. 4) Learning and Innovation. Learning and Innovation is the foundation of achieving the long-term public interest and achieving improvement. The other 3 dimension is the disparity in public benefits, public satisfaction and internal process, While Public Service Units can make up the disparity by learning and innovation.

3.2 Index System of Public Service Units Performance Appraisal

Every evaluation index is corresponding to the best performance standard In order to express the connotation of index, example is established under evaluation index(third grade indexes), example can list representative behavior or case which can express index connotation.

To different Public Service Units, Referring to same evaluation dimension, bur evaluation index is different. This is decided by complicated Management Objectives, special Product Form, multidimensional product form, price mechanism. So the establishment of evaluation index is the hardest and most difficult work in the establishment of evaluation system.

The following is Indicator System of Public Service Units Performance Appraisal Indicators in XX province as Table 1.

Table 1. Indicator system of Public Service Units Performance Appraisal

Dimensionalities (the first grade indictors)	Evaluation Indicators (the second grade indictors)	instances (the third grade indictors)
Public Benefits	economic Benefits	the extent of budget execution The extent of financial goal completion Ability of meeting and balancing all stakeholders′ financial interest Activities of increasing profits
	Social Benefits	The Units' impact on the local, national and international society Quality and frequency of the Units' communication with local or social partners moral behavior of organization mass participation to support social vulnerable groups to combine Social changes Organization's own functions
	Ecological Benefits	The extent of the integration in the principles of sustainable development and group decision-making process To accord with environmental protection standard To avoid making harassment and damage to the public and to guarantee public safety Considering environmental and natural protection indecision-making
the public	customers satisfaction degree	degree of satisfaction with overall image of their organization degree of satisfaction with customer's Participation degree of satisfaction with Service accessibility degree of satisfaction with products and services
	Employee satisfaction degree	degree of satisfaction with the whole organization degree of satisfaction with the management and the Management System degree of satisfaction with work conditions degree of satisfaction with the development of their careers and skills
	Other relevant Indicators	
process	input-output efficiency	manpower and human resources which the units should put into the set work manpower and human resources which the established service provided need
	the degree of implementation	the degree of implementation of the public goal Service commitments are clear and Implemented them completely emergency measure are clear to make goals and responsibilities clear and complete them
	Process improvement	To continuously Design, manage and improve the process To develop the products and services which are based on Customers or the public To make Customers or the public participate in the Process improvement
Learning and innovation	Human Resources	To Manage and improve human resources on the basis of Organizational strategic planning To Improve and develop the abilities of employees and to encourage them to participate
	Innovation Resources	Monitoring knowledge and resources which can be used by the organization To ensure that the employees have enough innovational knowledge and resources
	Innovation impetus	Monitoring the signal of internal changes Monitoring the signal of external technological changes

3.3 The Methods of Scoring in Public Service Units Performance Appraisal

Referring to the European Union's universal Appraisal framework, we build following scoring methods -- Stimulation-factors appraisal scoring method and Results-factors Appraisal method. In Appraising Result factors to which Organization activities lead, the actual achievement firstly must be considered, meaning while comparison with the same business and the progress also have to be taken into account. In evaluating the actual achievement, horizontal comparison with Benchmark and the starting point and progress all need to be considered. Stimulation factors are Conducts or activities of organization. The degree of progress about conducts and activities (also called PDCA) should be appraised according to the rule of Total Quality Management (TQM) in evaluating Stimulation factors.

In our designed institution performance evaluation indictors system, Public benefits and the public are Result factors; learning and innovation are Stimulation factors. In the second grade indictors of the Process-factors system that is relatively special, input-output efficiency is Result factor; however the degree of implementation and Process improvement are Stimulation factors. Specific details are given in the following table 2 and table 3. (3- means pass, 4-good, 5-excellent)

Table 2. Stimulation-factors appraisal marks

marks	expressional behaviors
1	planning correlation method (plan)
2	planning and doing correlation method (do)
3	planning doing and checking correlation method (check)
4	planning and doing correlation method, while evaluating and adjusting accordingly them on basis of the exemplar dates (act)
5	planning and doing correlation method, evaluating and adjusting accordingly them on basis of the dates ,also integrating The whole process above into the whole system
The reasons why the evaluation result was given	

Table 3. Results-factors Appraisal marks

marks	expression
1	The results have showed the poor and stable tendency
2	The results have showed moderate progress
3	The results have showed Substantial progress
4	Having gained outstanding results which are fine, compared with their own goal.
5	Having gained outstanding results which are fine, compared with their own goal and other related units ✓ exemplar
The reasons why the evaluation result was given	

3.4 The Model of Public Service Units Performance Appraisal

3.4.1 Weight-Variable Theory

One of The performance evaluation methods that are generally used nowadays in china is weighted sum (called Invariable Weight Appraisal (IWA) in following) in which

Weights are set on terms of Relative Importance of all kinds of Indicators that are invariable .however Mr li hongxing etc appointed that if only the importance of weights was considered and the predilection for the equilibrium level of appraisal value was neglected , sometimes the way of making Performance Appraisal like that would result in getting an unscientific decision[1]. The IWA′s bluntness of the advantaged and disadvantaged items that are used in Performance Evaluation makes the evaluation results centralizing and averaging . Most of Public Service Units Performance Appraisal results are pass level (Generally there are four levels – fail.pass.good and excellent),which lead to the difficulty with distinguishing the results good and bad. The phenomenon above can be explained by Invariance of weights or IWA, from the point of view of its computing Technique. Total scores can exceed mean level under the conditions of giving up IWA and changing invariable weights according to the deviate degree of extreme value from the average.

As the one-sidedness of Invariable Weight Appraisal always appears in practical decision, the Weight-Variable has begun to be taken into account seriously. Many papers about the research into eight-Variable Model appear at home and abroad, in papers[1-3]," weight-variable thought" were putted forward; in papers[4-8] the methods of building up balance function --including three types, Penaltical Variable Weight .inspirational variable weight and Mixed variable weight; in papers The comprehensive evaluation model with variable weight was set up. In a word, abundant literatures provide theoretical basis for our Establishing variable weights model.

3.4.2 Establishing Model

In this paper, a fuzzy AHP analysis model that is based on Variable weights will be established On the basis of the properties and characteristics of Public Service Units Performance Appraisal in practice. Analytic Hierarchy Process.fuzzy evaluation method and Weight-Variable Theory are Theoretical Basis of the model. As well as fuzzy evaluation method and Analytic Hierarchy Process that are mixed congenitally, variable weights are put into the model, which lead to a new Public Service Units Performance Appraisal model.

1. Determining variable weights

Variable Weights which are based on invariable weights change and move according to imbalance of indicator values. Referring to the paper[11] ,we have built up the following Calculation Formula of weights.

$$w_i(x_1, \ldots x_m) = w_i^{(0)} x_i^{a-1} / \sum_{k=1}^{m} w_k^{(0)} x_k^{a-1} \tag{1}$$

In the Formula, wi is the variable weight of factor I, xi is indicator value of appraisal factor I.

Generally, when only a little imbalance of indicators is taken into account, $\alpha > 1/2$ is put to use; when serious defects of some factors could not be accepted, $\alpha < 1/2$ is applied; when a=1, the model is equal to Weight-Variable type. The unreasonable phenomenons in which too many appraisal factors and too low appraisal value result in evaluation are effectively resolved because of the use of variable weights. Considering the severity of performance appraisal, and in order to simplifying calculation, we assign all of ‘α’s as 0.2. Indictor characteristic values need to be standardized before

Calculating Variable Weight, when their dimension is not standardized and value range is not same.

In determining weight values, indicator′ invariable weight is assigned as 0.3, indicator value as 40, the weights of other factors in the same system has been given, a=0.2. According to Formula (2), its weight value, w2, is presented in the following.

$$w2\ (x1,\ldots,x5) = 0.3 \times 0.40.2 - 1/\ \sum_{k=1}^{5} w_k^{(0)} x_k^{0.2-1} = 0.39 \qquad (2)$$

It is presented that Variable weights are used because of the lower this factor value, which makes the weight value rise.

2. The process of Multi-level variable weight assessment methods (Analytic Hierarchy Process)

Referring to the paper[5] and Public Service Units Performance Appraisal system we have built , we present the process of Weight-Variable appraisal—two-hierarchy multi-dimensional appraisal as example.

(1)To establish two-hierarchy Multi-objective system which separate General objective performance into four goal sets{ V1, V2,..., Vn } $Vi \cap Vj = \Phi, i \neq j$. Vi is further separated into { V1i, V2i,..., Vkii}, so obviously there are ki factors in Vi.

Solving:

The set of General objective performance Appraisal V (Matrix) is presented by Table 4. bij is factors of The set of Appraisal V which express the membership degree of original scores.

Table 4. The set of General objective performance Appraisal V (Matrix)

0.1	0.5	0.2	0.2	0
0.1	0.3	0.5	0.1	0
0.2	0.5	0.3	0	0
0	0.3	0.1	0.6	0
0	0.1	0.6	0.3	0
0.2	0.2	0.3	0.3	0
0.1	0.6	0.3	0	0
0.1	0	0.3	0.6	0
0	0.3	0.7	0	0
0.2	0.6	0.2	0	0
0.4	0.5	0.1	0	0
0	0.3	0.6	0.1	0

(2) To make the second grade Weight-Variable appraisal of kis in Vi on m levels (5,4,3,2,1)of appraisal results

First, Weight-constant Vector W =(W1,W2,...,Wn),0< Wj<1, $\sum_{i=1}^{n} W = 1$ (n= 1.2,...,n), Weight-Variable vector $\tilde{W}$ changes around Weight-constant.

Solution: Known first constant weight W1 =(0.3,0.35,0.15,0.2)and second weight W2=(0.4,0.35,0.25,0.4,0.3,0.3,0.3,0.5,0.3,0.3,0.2,0.5)

Conclusion of Variable Weight, unification of evaluation and specification of weight value, according to formula (5-1),secondary Variable Weight is

W2b=(0.41, 0.36 ,0.23, 0.42 ,0.31, 0.27 ,0.23 ,0.52 ,0.26 ,0.28 ,0.17 ,0.55)
W2b=(0.41, 0.36 ,0.23, 0.42 ,0.31, 0.27 ,0.23 ,0.52 ,0.26 ,0.28 ,0.17 ,0.55)
Secondly, establishment of secondary Variable Weight evaluation Matrix R2b.

$$R_{2b} = \begin{bmatrix} \tilde{W}_{1i} B_{1i}^{T} \\ \tilde{W}_{2i} B_{2i}^{T} \\ \cdots\cdots \\ \tilde{W}_{mi} B_{mi}^{T} \end{bmatrix} = (r_{1i}, r_{2i} \ldots\ldots r_{mi})^{T} \quad (3)$$

(i =1,2,...,n)

Among them, $\tilde{W}_{ki} = (\tilde{W}^{1}{}_{ki} \tilde{W}^{2}{}_{ki} \ldots \tilde{W}^{n}{}_{ki})$, row vector of Bi is B1i,B2i,...,Bmi, composed by bij

Solution: Know from calculation, secondary constant Weight evaluation matrix R2 and secondary Variable Weight evaluation matrix R2b.

The calculation of secondary constant Weight evaluation matrix ,is change Variable Weight (3-3)into constant Weight, the result is in table 5

Table 5. Secondary constant Weight evaluation matrix R2

0.4	1	0.875	0.4	0
0.666667	0.666667	0.75	1	0
0.166667	0.6	0.6	1	0
0.4	0.6	1	0.166667	0

The calculation of secondary Variable Weight matrix is based on (3-3),the result is in Table 6

Table 6. Secondary Variable Weight evaluation matrix R2b

0.4	1	0.895528	0.4	0
0.638765	0.638765	0.728494	1	0
0.166667	0.442897	0.5	1	0
0.333333	0.501907	1	0.166667	0

(3) n subgoal in V on m evaluation grade ,Weight-Variable Comprehensive again, obtained, get first decision matrix R1b.

$$R1b = WR2b \quad (4)$$

Among them W is first Weight-Variable, we get it from (3-1)

Solution:

From (3-1), first evaluation Weight-Variable W1b=(0.315, 0.34, 0.129, 0.212).

Finally we obtained first evaluation Weight-Constant matrix R1 and first evaluation Weight-Variable matrix R1b.

R1=(0.667 , 0.857 , 0.857 , 1.000 , 0.000),Finally we obtain the value of Weight-Constant evaluation is 3.35 through reducing calculation

R1b=(0.639,0.857,0.857,1.000,0.000), Finally we obtain the value of Weight- Variable evaluation is 3.338 through reducing calculation.

Visible, finally the value of evaluation is decreased, Disadvantage exist in the nuit.

(4)Disadvantage judgment.

The establishment of Weight-Variable Model means that it not only give out the total score which is different from constant weight; but also can find out the reason that causes changes, The factors which can improve the total score is called "advantage", The factors which can reduce the total score is called "advantage". We should find out "disadvantage" in order to improve organizational performance of Public Service Units.

The judgment of disadvantage is as follows: General evaluation relative to constant weight, we can believe that it is the factor which weight is bigger that changed the general evaluation. If the objective evaluation value of the factor is low, while the general evaluation come from variable weight is smaller than that from constant weight, then we can judge the factor is "disadvantage".

In this example, the result calculated before, secondary constant weight W2 =(0.4,0.35,0.25,0.4,0.3,0.3,0.3,0.5,0.3,0.3,0.2,0.5), and secondary Variable Weight W2b=(0.41, 0.36 ,0.23, 0.42 ,0.31, 0.27 ,0.23 ,0.52 ,0.26 ,0.28 ,0.17 ,0.55).From the contrast we know that the fourth and the last item of Variable Weight is bigger than constant weight, but both of them are in the lower level of evaluation, so we can call it "disadvantage".

4 Conclusion

We get the following conclusion through the model:

1)Introduction of Variable Weight, the result from it is different from constant weight. The evaluation results from Variable Weight affected by equilibrium of the date, that is to say items of high score or low score have greater impact on evaluation results. Based on the principle, we find that the advantage or disadvantage item can avoid the centralized or average evaluation results; what's more this method can find out the key of next work.

2) Fuzzy evaluation method which apply to Public Service Units performance appraisal provides the possibility of Objective Evaluation, it has solved the problem of qualitative project evaluation and comprehensive evaluation affected by many factors.

3) Analytic Hierarchy Process(AHP) mix Fuzzy Evaluation and Weight-Variable together, so we can build new model by the framework of AHP. The reasonable evaluation logic and mature Processing Method of AHP provide us with solid platform.

References

1. Li, H.: Factor space theory and Mathematical Frame expressed by knowledge (VIII). Fuzzy System and Mathematics 9(3) (1995)
2. Wang, P.: Fuzzy set and random sets falling shadow. Publishing house of Beijing Normal University, Beijing (1985)
3. Wang, P.Z.: A Factor Space Approach to Knowledge Representation. Fuzzy Sets and System 36, 113–124 (1990)

4. Liu, W.: Balance function and it apply to Variable Weight Synthesizing. Theory and Practice in System Engineering 17(4), 58–64 (1997)
5. Li, D., Li, H.: Analysis of variable weights effect and selection of appropriate state variable weights vector in decision making. Control and Decision (11), 1241–1245 (2004)
6. Li, D., Li, H.: Properties and structure of state variable weights vector. Journal of Beijing Normal University 38(4), 455–461 (2002)
7. Liu, W.: The punishment of variable weight synthesizing—Incentive utility. Theory and Practice in System Engineering (4), 41–47 (1998)
8. Liu, W.: Incentive strategy and solution of Variable Weight Synthesizing. Theory and Practice in System Engineering (12), 40–43 (1998)
9. Zhang, J., He, X.: Dicision method and application of Variable Weight. Journal of Yunnan University
10. Li, H.-X., Li, L.-X., Wang, J.-Y., Mo, Z.-W., Li, Y.-D.: Fuzzy Decision Making Based on Variable Weights. Mathematical and Computer Modelling 39, 2–3 (2004)
11. Miao, Z., Xu, J., et al.: Research on evaluation model of part Variable Weight. Prediction (4), 51–52 (1997)

The Study on University Library Consortium Based on Cloud Computing

Hai-bo Yang[1] and Ning Yu[2]

[1] Library of Hebei United University, 063000 Tangshan, Hebei, China
[2] Qian'an College of Hebei United University, 063000 Tangshan, Hebei, China
yhb523@126.com

Abstract. Cloud computing is a new computing model, which focuses on user and affords powerful data storage and web service functions. Cloud computing will have influence on many aspects of library. This paper describes the meaning of the cloud computing and library consortium, the application about cloud computing in the university library consortium, and made some comments and suggestions on cloud computing applications in the university library consortium.

Keywords: Cloud Computing, Library Consortium, University Library.

1 The Meaning of Cloud Computing

Cloud computing is a super-computing model based on Internet, which is stored in personal computers, mobile phones and other devices on the wealth of information and processing resources together and work together [1]. The basic principle of cloud computing is distributed computing by making a large number of distributed computers, rather than the local computer or remote server, the library data center operations will be more similar to the Internet. This makes the library will be able to switch to the needs of resources application, according to the demand access to computer storage system.

From a technical perspective, the realization of the basic functions of the cloud computing depends on two key factors, one is the data storage capacity, the other is the distributed computing power. Therefore, the cloud computing "cloud" can be further subdivided into "storage cloud" and " calculation cloud ", that "cloud computing is storage cloud and calculation cloud." "Storage cloud" is a large-scale distributed storage systems. "Storage cloud" is for the third-party users to open memory interface, the user can purchase according to their needs to the appropriate capacity and bandwidth. "calculation cloud" includes parallel computing and resource virtualization. The main significance of virtualization is with less resources to do more. Introduced in the calculation of cloud virtualization technology, is to seek to run on fewer servers for more parallel computing, the cloud computing resources which are applied to the fast and optimized configuration.

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 259–262.
springerlink.com

2 Library Consortium

Library Consortium refers to the number of libraries in order to achieve resource sharing, the purpose of mutual interest, combined with relevant information resources system, according to a mutually agreed protocols and contracts, in accordance with uniform technical standards and procedures, implement a or more cooperative function of the joint [2]. Library consortium is an important modern library organization form, it is mostly the same type of cooperation between libraries, it can be combined between different types of libraries. The purpose is to solve the contradictions of individual libraries can not solve the shortage of resources and conflicts between the needs of user information problems in the modern environment, which is in the mutual interests of all parties within the Union Library, in the enjoyment of rights, but also undertake corresponding obligations.

Early in the 20th century from the late 50's to share documents, interlibrary loan books and periodicals had just introduced a simple model species, more recently, the rise of library consortia model has shrunk, the United States and other Western countries have formed regional Library Alliance. Regional Library is a regional library for the joint coordination center of the organization. Construction began in China's Library Consortium late 1990s, the provincial, city, area-based unit, and then extend to larger areas, mainly in the China Academic Library System, the National Science and Technology Library, literature resources in Shanghai Collaboration, Tianjin, Jiangsu and other places on behalf of the alliance of the University Library. After ten years of development, coalition-building has made great progress, but still in its early stages, the number is also one of the few in the international consortium of library consortia, only CALIS is a member.

3 The Application of Cloud Computing in the University Library Consortium

3.1 Cloud Computing for the University Library Union Brought More Convenient, Secure Service

Cloud computing provide the most reliable, most secure data storage center for the university library consortium. We do not have to worry about data loss, virus attacks and other problems. Many people feel that their data is only stored in the visible, tangible computer was the most secure, is not the case, your computer may be because they do not accidentally damaged or virus attacks, resulting in data on the hard disk can not be restored , and then have the opportunity to make contact with your computer may use a variety of opportunities for criminals to steal your data. However, cloud computing strict rights management strategy can help you to safely share with the people you specify the data. In this way, you do not spend money to enjoy the best, most secure service, even more than saving money in the bank is also convenient.

3.2 Cloud Computing Can Easily Realize the Data between Different Devices and Application Sharing, Can Easily Realize the Museum and the Museum of Data Sharing

The University Library now has a diverse and complex equipment characteristics, cloud computing equipment, taking into account the different methods of a wide range of data synchronization, complex operation, in which many different devices to a preservation and maintenance of the latest contact information, the library needs to pay countless time and effort. At this time, we will use cloud computing to make everything easier, network applications in the cloud computing model, data is only one, save in the other side of the "cloud", you only need all the electronic equipment connected to the Internet, can simultaneously access and use the same data. So you can achieve a deeper level of the library of data sharing.

3.3 We Can Achieve Union between the Joint Purchasing of Digital Resources in the Cloud Computing Environment

On the common database, you can use Group buyout approach, the Alliance to bear the full cost, for members of the museum free of charge; for more integrated database, you can use Group buy-out, cost sharing by the Centre and members of the union; database for thematic , can be negotiated by the group, members of the union need to buy. To establish a performance evaluation system of digital resources, the cost of the database, content, utilization, and other services to assess the situation. Resources and build on the purchase of resources for integration of resources, to establish a unified search platform Museum.

3.4 Library Consortium Share a Hardware and Software Resources among Cloud Computing Services

University Library can be carried out cloud services, purchase of automated integrated systems, database development software, hardware deployed in data centers or libraries with better conditions, a rental service for the members of the museum, take advantage of hardware and software resources within the Union, small library with a smaller cost can be automated or integrated systems to manage the database to build their own characteristics.

3.5 The Union as the Center Set Up Several Regional Consortium

University Library Interlibrary can grant general read cards, the library can be carried out conditional interlibrary booking service. In the Union's development, based on the geographical union can be established near the library, library open to readers of the members of the literature review, lending, electronic reading, printing and photocopying services, logistics and transport through the books. The union can carry out the original foreign language periodicals and electronic resources coordinated establishment, the combined reference and consultation work and joint science and technology novelty search, according to area and subject complementary advantage, implementation area of university document resources and knowledge sharing.[3]

The emergence of cloud computing model, not only to the university library alliance brought opportunities, but also give us the future of cloud computing applications in the Library Association have endless source of fascination, the cloud computing applied to the university library, not only can improve resource utilization, conservation of resources of the member libraries can also be brought to a new cloud computing applications. At present, although cloud computing applications still in the exploratory stage, I believe that with the cloud computing applications in other fields mature, the library will also have a bright future.

References

1. Qin, W.: Exploration of Information Resource Sharing and the Mode of Regional Library Consortia. Modern Information 30(2), 35–38 (2010)
2. Xu, M.: A Comparative Study on Library Consortium at Home and Abroad. Library Tribune 29(1), 16–18 (2009)
3. Gao, X.: A probe into the application of cloud computing in the library. Journal of Academic Library and Information Science 28(2), 61–62 (2010)

Research on County Tourism Development in Hainan

Xiang-rui Chen[1,a], Fan-hua Meng[1,b], Lan Li[1,c], and Shi-chen Fan[1,d,*]

Tourism College of Hainan University, Haikou, 570228, China
[a] 251016412@qq.com, [b] gyhmfh1986@163.com,
[c] lean_1349@yahoo.com.cn, [d] fansc1230@163.com

Abstract. It was found out that Hainan's county tourism development level was low with large internal diversity. County tourism development level was closely related with its geographical site, economic development level, tourism resource quality and policy support. It was suggested that Hainan county should take the significant opportunity of International Tourism Island construction, and pay much attention to transportation, marketing and tourism service improvement in order to achieve better achievement.

Keywords: County tourism, Hainan, Development level, Spatial structure.

1 Related Concept

1.1 County

County is a specific form of a region, it is an administration concept including administrative units of town, township and village. County is the closest geographic space combining city and countryside, it is basic administrative region which is relatively of independence and integrity. It is an important component of China's administrative unit. The total land area of county occupies 94% of all, total amount of population takes up 85.3% of whole, and its GDP was 63.4% of whole. County studied in this article is regional space on the basis of administrative unit at the county level, which would focus more on its social comprehensive properties rather than administrative one.

1.2 County Tourism

County tourism was relative to regional tourism, it is a level of regional tourism. Based on definition of county, county tourism could be defined as tourism activity carried out in the administrative level of county. Because of the specialty and importance of county in Chinese administrative structure, in addition, tourism industry is am important component of county economy, the development of county tourism plays an important role in improving city's image, demonstrate attractive investment environment, promote regional specialties and dealing with unemployment. To enhancing tourism industry within county can drive local economy development[1], hence, county tourism was gradually thought highly of.

* Corresponding author.

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 263–269.
springerlink.com

2 County Tourism Development Situation

County studied in this article contains sixteen county-level administrative area. The total area of county is 31129.58 km2 which takes up 88% of whole land area of Hainan. Its population is 6.3155 million which was 73% of all[1] Development of county economy will impose vital influence on whole economy of Hainan. From the establishment of special economic zone, Hainan has been practising small government, large society, a special administrative system which means provincial government directly in charge of county government. Influence of county's developmentt on whole economic entity becomes larger and larger, showing that its development is of significance to that of Hainan.

2.1 Overall Situation of County Tourism Development in Hainan

Achievements were achieved on the aspect of county tourism from the time special economic zone was established. According to Statistics yearbook 2010, total amount of county scenic spots tourists was 7240329 in 2009, taking up 42.3% of that of all in Hainan. There is 25 high level (A level and natinal park) scenic spots in Hainan's counties, occupying 65.79% of all. It shows that tourism construction in Hainan was incrasing rapidly. However, it can also be seen from other indexes that profitability was low, and ratio of tourist population and ratio of tourism resource number was not match, showing that county tourism development level is still low and there is much more potential waiting for grubbing.

Table 1. Font sizes of headings. Table captions should always be positioned above the tables.

Region	Tourist population of Scenic spots (person trip)	No.of Highly Qualified Scenic Spots	Touris t hotel	Operation Revenue of Scenic spots (million)	Tourist population of tourist hotel (person trip)
Provinc	17108844	38	440	865.12	16410952
County	7240329	25	174	201.40	5875981
Ratio (%)	42.32	65.79	39.55	23.28	35.82

2.2 Analysis on County Tourism Development Level in Hainan

Tourism development characteristics of each counties can be found out through analysis on relative indexes. Data in table below are come from Statistics Yearbook 2010, GDP per capita, number of tourist hotel(TH), tourist population of tourist hotel(TPTH), tourist population of tourism scenic spots(TPTSS), operation revenue of tourism scenic spots(ORTSS), and number of highly qualified tourism scenic spots (HQTSS) were selected as original data. Because dimension of these indexes was different, in order to eliminate shielding effect between data, it is necessary to have normalization on them. Range normalization transformation was used, its formula is as shown below:

$$B_i=(I_i-I_{min})/(I_{max}-I_{min}). \quad (1)$$

In this formula, Bi is normalized value, Ii is original value, and Imax , Imin is the maximum and minimum original value in its line[2].

Obvious Individual Tourism Development Differences within County. From the view of tourism scenic spots operation revenue, difference of index value between Lingshui and Wuzhishan, which ranks No. 1 and No.6 respectively, is 0.97. From the view of tourist population of tourist hotel, difference of index value between rank first county Wanning and the second county Qionghai is approximately 0.7. These kind of thing happens here and there, showing that there is large differentiation of tourism development level within counties in Hainan.

Low Overall Tourism Development Level within County. It is shown in Table 1 that overall tourism development in Hainan's county is low, and it is proved by data in Table 2. It can be found that index value of most counties is lower than 0.2 by comparison of tourist population of tourist hotel and tourist population of tourism scenic spots, showing a low level of development.

Table 2. Statistics of county tourism development index(DI)

Area	GDP per capita	No. of TH	TPTH	TPTSS	ORTSS	No. of HQTSS	DI
Wuzhishan	0.0711	0.1087	0.0602	0.0079	0.0298	0.1667	***5***
Wenchang	0.7262	0.2174	0.106	0.0047	0	0.1667	***5***
Qionghai	0.8303	0.5435	0.3049	0.6314	0.8164	1	***6***
Wanning	0.3957	1	1	1	0.9018	0.8333	***6***
Ding'an	0.08	0.087	0.0813	0.0967	0.5199	0.3333	***4***
Tunchang	0.0803	0.087	0.0157	0	0.0151	0	***1***
Chengmai	0.6059	0.1087	0.0442	0			***2***
Lingao	0.4645	0.0652	0.0118	0			***1***
Danzhou	0.2865	0.3261	0.1664	0.0572	0.0749	0.6667	***5***
Dongfang	0.8887	0.087	0.0414	0		0	***1***
Ledong	0.0061	0.1087	0.0098	0.0041	0.0013	0.1667	***1***
Qiongzhong	0	0	0	0.0011	0.0003	0.1667	***1***
Baoting	0.0579	0.087	0.0519	0		0.1667	***1***
Lingshui	0.1264	0.2609	0.2719	0.342	1	0.3333	***5***
Baisha	0.1547			0		0	***0***
Changjiang	1	0.0435	0.03	0		0.1667	***2***

2.3 Spatial Structure Analysis of County Tourism Development Level in Hainan

Tourism development index I was chosen as indicator evaluating county tourism comprehensive development level. Tourism development is sum of number of index ranks in top seven. Mapinfo was used to form schematic for county tourism development index differentiation in Hainan (Figure 1).

It can be clearly seen in Figure 1 that counties whose tourism development index is relatively high are centralized in Qionghai, Wanning, Wenchang, Lingshui, Danzhou, Wuzhishan, and Ding'an. Seeing from the point view of spatial distribution, number of those counties located in eastern Hainan is five, Danzhou locates in western

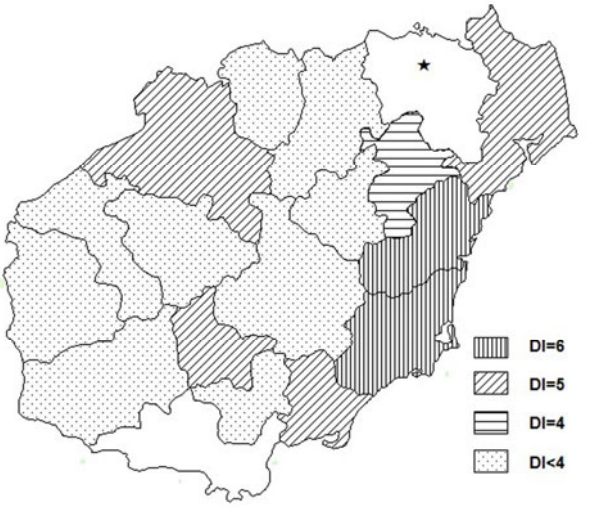


Fig. 1. Schematic for County Tourism Development Index Differentiation in Hainan

Hainan, and Wuzhishan locates in the middle of Hainan. It demonstrated that important tourism area centralized in eastern region, besides, counties in middle and weatern region represented by Wuzhishan and Danzhou also developed well.

The other nine counties, whose tourism development index are fall in between 0 and 2, are considered as low development level area. It was shown that tourism development level among different counties was uneven, and it was also proven that spatial differentiation was existed

3 Analysis on Effecting Factor to County Tourism in Hainan

3.1 Geographic Location

There is much relationship between county's geographic location and its tourism development level. The uneven tourism development of eastern, middle and western Hainan is one of basic characteristics of county tourism in Hainan. Tourism development bipolarity in Hainan is apparent, tourist access to Hainan through two landing cities, Haikou and Sanya, while most of county tourists also enter from Haikou and Sanya. What we can find in the figure is that cities whose development index was high are centralized in eastern part of Hainan, because transportation of eastern expressway is well and there were many tourist routes between Haikou and Sanya, in addition, island tour was massively practised. It helped a lot to the accessibility of Wuzhishan to build highway. It can be seen from Figure 1 that area which is close to Haikou or Sanya is developed better. So, distance between county tourism resources and central cities is an important affecting factor to county tourism development.

3.2 County Tourism Development Level

County economic development level is positive correlation to its tourism development level. Cities and counties which have higher number of tourism development index are concentrated in eastern coastal developed area, the other area, Danzhou and Wuzhishan, are located in central zone of western and middle region with higher development level than other counties in its region, respectively.

3.3 Number of Tourism Resource

It is shown in Table 1 that Qionghai, Wanning, Lingshui, Ding'an, and Danzhou are counties owns many high level tourism scenic spots, in contrast with Figure 1, it was found that counties owning higher tourism development index and counties boosting of many tourism scenic spots are almost the same. Tourism attractions is key part in tourism. Hence, it becomes important for counties owning much tourism resources to developing these qualified tourism resources to be qualified tourism resorts.

3.4 Policy Support

All five county-level cities is area where county tourism developed well. In the system of county-level city, many policies can be practised easier. At the same time, counties which have better urbanization have better supporting facility. And it was found that there are tourism related website in some eastern cities' governmental website, such as Qionghai, Wenchang, Wanning, it pushed development of local tourism.

4 County Tourism Development Strategy

Under background of establishing International Tourism Island, a big change would be promising in Hainan. Tourism industry in Hainan can also take opportunity and develop much better and faster. There are some factors influencing Hainan's county tourism development, hence, attempt can be made focusing these aspects.

4.1 Transportation Should Be Improved

Construction of high-speed rail around Hainan is a hallmark events in transportation construction. Operation of east part of high-speed rail brought opportunity to counties along the line and gave them better location, and hence, it could weaken shielding effect of Haikou and Sanya and drove the whole eastern part a common prosperity. What's more, counties in middle of Hainan, such as Ding'an and Baoting[3], were associated developed. Another function of eastern part of rail is public transportation among cities, which would bring counties closer and bring in more frequently flow of human and material. Fact that tourist can make a return journey within half a day can push county tourism a better development for it brings in big amount of local tourists, and expand tourist routes deep in counties. Haitun express way, two bridge and one road in Wenchang further tightened relation between counties and cities while its transportation within counties are still waiting for improvement[4]. It was found when

undertaking a research in countryside of Hainan that countryside high way net was basically finished. However, most of the road's width was only 3 to 4 metres and road hardening was not completed yet, in addition, lacking of money and management, roads connecting county and tourism scenic spots exists potential safety hazard. In the future, it should be a focus to deal with problem of transportation within counties and form a high quality and multi-level transportation net.

4.2 Marketing Should Be Intensified

Time of Good wine needs no bush has past. Nowadays, to get higher market share, what we need is reasonable publicity. Counties in Hainan boosts of abundant of tourism resources, each counties has their own specialties. So what should be done is develop and demonstrate its own specialties specific to tourism market and carry out depth marketing. Influences should be impose and broaden by the use of festival advertising, Internet, and media. Integrate and popularize tourism resources within counties after well planning. Internet would be main marketing approach, so it would be a important choice for counties to take advantage of tourism resources and promote it in Internet.

4.3 Service Should Be Improved

First of all, a series of high quality tourism resort and scenic spots should be build up. After beginning of construction of International Tourism Island, there begins lots of large project, the biggest problem was no longer capital shortage but how to position, find out own specialty and build creative scenic spots. Then, because county's economic development level was not high with tourism reception facility need to be completed, besides, characteristic of county tourism was different from coastal tourism with different tourist demand, which is focusing more on nature and different life in village, facilities built should be paid more attention to express idea of nature, ecology, and health. Lastly, attention should be paid to training of tourism talents. Talent absence exists in family hotel, rural inn, and tourism scenic spots. Advanced operation idea is hard to practise in family operation mode. Specialized talents would be a driving force to Hainan's county tourism development.

5 Epilogue

It is an important opportunity for tourism development in Hainan to construct International Tourism Island, especially under background of pour tourism development foundation and different development level. Improving transportation, practising depth marketing, and improving tourism service are approaches to push county tourism development and bring county economy a lasting and healthy development as well as make contribution to International Tourism Island construction.

Acknowledgement. Financial supports from: the National Natural Science Foundation Project of China in 2009 (NO.40961005), the National Soft Science Item of China in 2010(NO.2010GXS5D252), the National Education Ministry Planning

Project(NO.09YJC840008), University Research Project of Hainan Ministry of Education in 2010(NO.Hjsk2011-28), Hainan University Education Project Fund in 2010(NO.HDJY1008).

References

1. Stiatistical Bureau of Hainan province, Survery Office of National Bureau of statistics in Hainan Hainan Statistical Yearbook, p. 72. China Stiatistics Press (2010) (in Chinese)
2. Fan, S.-C., Song, T.: Appreciation and Comment on Regional Differentiation Characteristics of Intracounty Capacity for Sustainable Development in Hainan Special Economic Zone-From Visual Ang le of Processes Coupling. Journal of Henan University (Natural Science) (in Chinese)
3. Zhou, X.: Diagnosis on the problems of Hubei province's county area tourism development and strategic idea. Journal of Wuhan Polytechnic University (12), 113–116 (2007) (in Chinese)
4. Zhang, P.: On the Sustainable Development of County-region tourism in Sichuan province. Sichuan Normal University, 22–24 (in Chinese)

The Research of Human Resources Flexibility

Changna Qian[1], Wen Zhou[1], Liang Chen[2], and Wenbing Zhang[2]

[1] Guilin University of Electronic Technology, 541004 Guilin, P.R. China
[2] College of Information Science and Technology, Donghua University, 201620 Shanghai, P.R. China
yumiguilin@163.com

Abstract. Human resource (HR) flexibility is vital in entrepreneurial ventures that need to respond to the changing challenges of growing the new business. This paper discusses the human resources flexibility concept, classification, and flexibility firm model, the goal is to provide a theoretical reference for human resource management. Future research directions based on this model and practical applications are discussed.

Keywords: Human Resources, Flexibility, Classification, Flexible Firm Model.

1 Introduction

With the further development of market economy and with the rapid development of information technology and knowledge economy era, companies are facing various changes, such as the emergence of virtual organizations, environmental uncertainty and increasing complexity, etc., all of these changes gave an unprecedented impact on businesses, and requiring enterprises to speed up decision and response, to response the changing environment flexibly, and the rigidity of traditional human resource management (such as fixed working hours, task fixation) has been unable to adapt the business development needs. Therefore, companies want to have a competitive advantage, the key is to motivate and manage employees effectively. In this context, a new management model emerged - people-oriented flexible management of human resources. Of course, human resource flexibility management model is not deny the traditional rigidity human resource management, but further improve the framework of management, and removing the unsuitable management ideas and methods, and enrich the traditional rigidity human resource management.

2 The Definition of Human Resources Flexibility

Flexibility is from the Latin, and the original meaning is flexible, steer able, or adjustable, variable [1], it can be translated flexible. That is generated for change adapts ability, pliability, and responsiveness [2]. For organizations or enterprises, the scholars are given the different views about the definition of flexible. Eppink thought that the flexibility has a characteristics that enables organizations to reduce their injury and look for a favorable position and successful to adapt the external

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 271–275.
springerlink.com

environment. One way to reduce injuries including change the direction of business operations (business strategy, product type, products quality, workflow, human structure etc.) [3], the definition of flexibility is different due to the different industries, but it is consistent to effective integration and flexible use of various resources to achieve organizational goals [4]. In summary, flexibility is a Characteristic that organizations or enterprises applied a series of efficiency and effectiveness measures to adopted Internal and external environmental pressures.

Human resources are an important resource in the business or organization, and human resources flexible are not only one of the management objectives, but an approach to promote to human resource management to people-centered. To the concept of human resource flexibility strategy, there has not a single and clear defined, but in general, human resource flexibility strategy is a human management methods, and it is people-centered, and on the basic of researching human psychology and behavior, and using flexible and non-mandatory measures to make employees from an inner loyalty to the organization, so that the organization will become a conscious action of the Compass staff. Specific methods include adjusting staff functions (such as job rotation, rich staff content, etc.), numerical (such as hiring temporary workers, part-time, etc.), wage of financial (such as combined with the performance or rewards etc.), working hours (such as work done under the premise of employee discretionary time) and so on. The most important feature is the 'people-oriented' that companies respect their employees' wishes, and applying democratic management, rather than one of the superiors make decision, and stimulating their creativity and a sense of belonging from the heart to adapt to the changing external environment.

3 The Dimensionality of Human Resource Flexibility

In 1984, Atkinson proposed human resources flexibility for the first time in the "Personal Management", he divided it into task or functional flexibility, numerical flexibility, financial flexibility. The following years, more and more scholars research the dimensionality of human resource flexibility, for example, EAPM (1977) redefined the human resources flexibility into five dimensionality, there are working-time flexibility, working-time career flexibility, contractual flexibility, workplace flexibility, elasticity flexibility and so on. Vosko, Leah F(1998) remand four dimension that it is reward flexibility, functional flexibility , numerical flexibility and distance flexibility . Now many scholars has been in using four categories proposed by Blyton, that is task or functional flexibility, numerical flexibility, temporal or working-time flexibility, wage of financial flexibility.

3.1 Task or Functional Flexibility

Task or functional flexibility has the characteristic of mobility, adaptability or versatility when the staffs execute their task or work in certain extent; they can make more rapidly response for the work supply change and technology development [6]. The certain extent includes horizontal or vertical mobility, horizontal or vertical mobility, thronging the job rotation or alternative work and so on to improve staff

diversity. That is, companies can take advantage of education or training to let employees work versatility.

Task or functional flexibility adopt job duty enlargement, work autonomy and team work etc to deal with the vagaries of the market environment. Some research also found that working with diversity, integrity, significance, autonomy, and the feedback can produce an effective incentive effect on staff, and make the work motivation and job satisfaction higher [5]. Therefore, in order to achieve these tasks or functional flexibility need, the organization required a series of measures including the goal which provide initiative attractive conditions for employees to learn to enhance skills and create learning opportunities and provide a complete career planning and so on.

3.2 Numerical Flexibility

Numerical flexibility refers to that the company be ale to adjust or manage the labor supply quickly when the labor supplies change. Companies exploit external labor market (such as part time and temporary contracts and fixed-term etc) and outsourcing and labor assignment to change the number of required labor in time. In 1991, Sloane and Gasteen divided numerical flexibility into primary flexibility and secondary flexibility. According to the British Human survey found that the companies prefer to hiring temporary labor or part-time staff when the labor supply fluctuates can be expected. And the companies prefer to work overtime when the company labor supply fluctuates can be unexpected. Therefore, it is called primary flexibility that hiring temporary or part-time employee to meet expected labor supply fluctuates. And it is secondary flexibility hiring regular staff .work overtime to deal with unexpected labor supply fluctuates.

3.3 Temporal or Working-Time Flexibility

Temporal or working-time flexibility means that company manage the time of staff flexibility under the staff voluntary or companies needed. Flexible management including working hour flexibly and work overtime and compensatory time-off and so on [5].Temporal or working-time flexibility rise in the 1970s, and has two ways, one way is to provide the number of hours per day, that is, as long as the employees complete working hours regulated per day, staff can control other times freely according to individual circumstances. This form of employment is popular in the current and attracts a lot of staff, especially women that they are required to take care of the family [6]. Another way is to regulating working hours per day. That is, the premise of a fixed weekly working hours, you can increase or reduce per day working time, in this way, not only improving employee productivity, but strengthen the capacity of the company deal with unexpected business growing.

3.4 Wage of Financial Flexibility

Wage of financial flexibility means that the salary of employees is changing, and it is a salary structure based on performance and combined with profit sharing mechanisms. It has changed the traditional single salary structure and pay bonuses and welfare and salary and benefits based on individual performance, which people can produce more equitable treatment, increased job involvement and willingness,

and improved job satisfaction, and also enhanced recognition of the organization and further promoted the competitiveness of the organization. Such as the daily wage and piece rate policy are the typical examples of the wage of financial flexibility. Company would also consider the differences and relativity between employees when they hire staff. Wage of financial flexibility can stimulate their wishes of active learning and to foster the function of it.

4 The Utilization of Human Resource Flexibility

With the continuous development of information technology and global economy, the science and technology content of products is one of measurement criteria for enterprise. The essence of the measurement criteria is that the competition of human resource flexibility. Human resource flexibility can make optimization of human resource allocation. Furthermore, human resource flexibility not only motivates the potential of employee effectively, but also increases their sense of belonging to organization. A large number of studies show that human resource flexibility have closely relationship to job performance and job satisfaction.

4.1 The Relationship between Human Resource Flexibility and Job Performance

Towards the research of relationship between human resource flexibility and job performance, different researchers hold different point of view. Through the empirical study, most of scholars suppose that human resource flexibility and job performance has significant correlation, that is, the higher degree of human resource flexibility, the higher performance of employee. Poole and Warner [5] consider that to improve the work efficiency of staff can through some approaches including education or work shift. Besides, flexible salary can improve job performance effectively. On the contrary, some researchers think that different human resource flexibility can lead variously job performance. For instance, the utilization of task flexibility and quantity flexibility may influence the quality of work and living, even increase the turnover rate.

4.2 The Relationship between Human Resource Flexibility and Job Satisfaction

The research of relationship between human resource flexibility and job satisfaction has received relative little attention. Some researchers suppose that human resource flexibility is significance to human resource management activities. Human resource flexibility not only can inspire the work interest of employee, but also improve their job satisfaction. Hackman and Oldham [6] found that characteristics of work can influence the job satisfaction. The experiments show that the job motivation and satisfaction can be improved by change the work characteristics including work scope, work skills, wage and benefits etc. The higher the task flexibility is, the more the satisfaction employees have. The diversity of work can stimulate the working enthusiasm of the staff, and improve the job satisfaction.

5 Conclusions

It is can be seen from the above, since the 20th century, 90 years, many scholars study the theory of human resources flexibility, and it has always been the core of the field of human resources, and its success was due to providing a reliable theoretical basis for "correct" management of human resources. Of course, human resource flexibility theory has its flaws, for example, the first, most of the human resource flexibility theory only concerned about the research quantitative research, but lack of quantitative research, in short, human resource flexibility theory remains for further study, for example, on researching the relationship between human resources flexibility and work performance, not only to enhance the impact on the various dimensions of human resource flexibility to job performance, but different dimensions of human resource flexibility focus on how it affects the vary dimensions of job performance, and it is impact on job performance directly or thronging an intermediary variables to affect job performance.

References

1. Atkinson, J.: Manpower Strategies for Flexible Organizations. The Personnel Management 147, 165–167 (1984)
2. Zolin, R.: Human Resource Flexibility and Strong Ties in Entrepreneurial Teams. Journal of Business Research (2010)
3. Martinez, A.: The Dynamics of Labour Flexibility: Relationships between Employment Type and Innovativeness. Journal of Management Studies 48(4), 715–736 (2011)
4. Poole, M., Warner, M.: Handbook of Human Resource Management. International Thomson Business Press, London (1998)
5. Li, Z., Qiu, Q.: Review of Human Resource Flexibility. Value Engineering 8, 44–52 (2008)
6. Zhu, X., Ling, W.: Human Resource Management pliability. Business Research 3, 66–74 (2005)
7. Hackman, J.R., Oldham, G.R.: Motivation Through the Design of Work: Test of a Theory. Organizational Behavior and Human Performance 16, 250–279 (1976)
8. Cappeilli, P., Rogobsky, N.: New Work systems and skill requirements. International Labor Review 133(2), 204–208 (1994)

A Developed System for Business Management

Sharly Joana Halder and Wooju Kim

Yonsei University, Seoul 120-749, Republic of Korea
{sjhalder,wkim}@yonsei.ac.kr

Abstract. The readymade garment industry of Bangladesh has expanded dramatically over the last three decades. Traditionally, the jute industry dominated the industrial sector of the country until the 1970s. Since the early 1980s, the readymade garment industry has emerged as an important player in the economy of the country and has gradually replaced the jute industry. Although Bangladesh is a garments based country, in most of the garments, employees work manually though there are some error in the software and also those software do not work properly. We developed a system which highlighted the style sheet information, export L/C, back to back L/C, U/D preparation, input commercial invoice, export invoice and so on, which save time and money, increased data communication and transmission and faster data retrieval.

Keywords: Export, import, readymade, garments.

1 Introduction

The ready-made garment (RMG) industry of Bangladesh started in the late 1970s and became a prominent player in the economy within a short period of time. The industry has contributed to export earnings, foreign exchange earnings, employment creation, poverty alleviation and the empowerment of women. Shirts, T-shirts, trousers, sweaters and jackets are the main products manufactured and exported by the industry.

Bangladesh exports its RMG products mainly to the United States of America and the European Union. These two destinations account for more than a 90% share of the country's total earnings from garment exports. The country has achieved some product diversification in both the United States and the European Union. Recently, the country has achieved some level of product upgrading in the European Union, but not to a significant extent in the United States. Bangladesh is less competitive compared with China or India in the United States and it is somewhat competitive in the European Union. The "export-quota system" in trading garment products played a significant role in the success of the industry. This quota system came to an end in 2004. Therefore, the competitiveness issue needs to be addressed, with special attention given to the long-term sustainability of the industry [1].

The manufacturing industry in Bangladesh is hampered by inadequate infrastructure facilities that are essential to support the production of textiles and readymade garments, which include; power supply, port congestion, inadequate air transport facilities, communications, roads and telecommunications facilities [2].

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 277–284.
springerlink.com

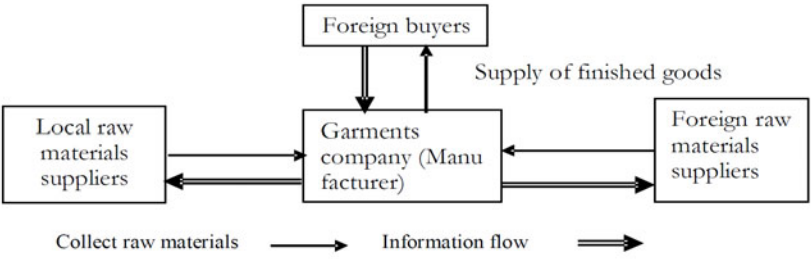


Fig. 1. Business structure of Bangladeshi RMG industries [6]

Readymade garment is a labor intensive industry and relatively simple technology compared to other high-tech industries. The RMG manufacturing units are like tailor's shop; getting order from the foreign buyers and then import raw materials specially fabrics from the foreign suppliers or sometimes buy from the local market as per order, then manufacture garments and supply those to the buyers.

Although Bangladesh is a garment's based country, we found that in most of the garments, employees work manually though there are some errors in the software and also that software do not work appropriately. It is very difficult and time consuming to do these works in manual fashion. We proposed a developed system where all these manual works can be done automatically.

The work processes of a garment company are, contact with buyers, buyers bring the contracts from the customers, then company get style sheet information from that. In that sheet all the descriptions about the contract is given so corresponding to that sheet- style sheet information, export L/C, back to back L/C, U/D preparation, input commercial invoice, export invoice happened step by step which we briefly described bellow:

Style sheet information: Buyers first collect contracts from customers then make the contract with Garments Company through contract number.
Export L/C: A L/C (L/C is an assurance paper about financial matters) is opened under the contract with a L/C number, which is called export L/C.
Back to back L/C: Under the export L/C a back to back L/C opened to import goods. Back to back L/C could be open either under the existing Export L/C or a new back to back L/C.
U/D Declaration: The next step is U/D Declaration. In U/D Declaration things are checked by BGMEA. BGMEA check all the transparency, e.g. how much goods are importing.
Input commercial invoice: The input commercial invoice is opened to check that how much goods are coming from foreign as because sometimes it is seen that goods and imported goods are not same by quality.
Export invoice: An export invoice is opened to export the order. Invoice for the goods are created for the amount of goods to be exported.

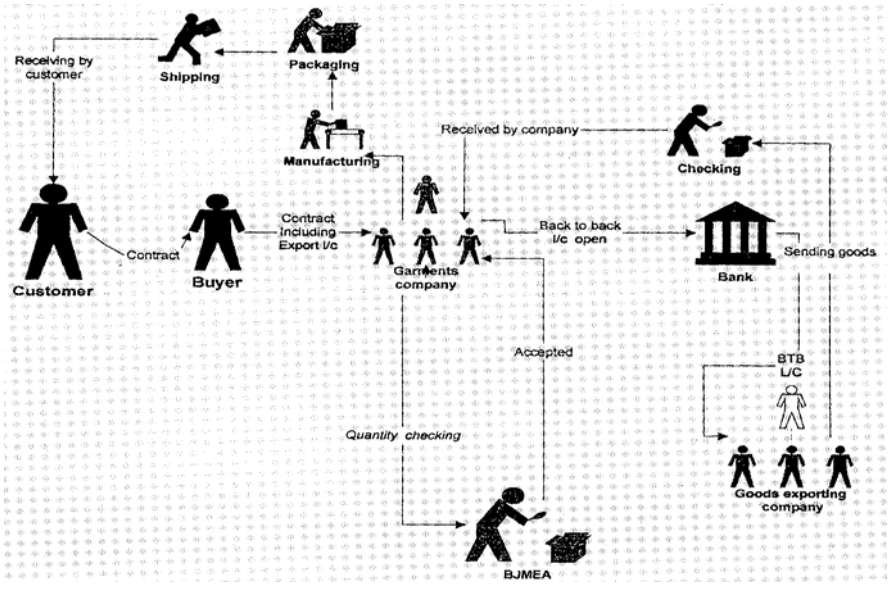


Fig. 2. Work flow diagram

Our system highlighted the style information sheet, export L/C, back to back L/C, U/D preparation, input commercial invoice, export invoice and so on, and is helps to save time and money, increased data communication and transmission and faster data retrieval.

2 Materials and Methods

We applied our system on a garments factory named "Right Choice Knit Fashion" established on 2006. The main office of "Right Choice Knit Fashion" is situated in Ahmed food Bhaban, Mirpur, Dhaka, Bangladesh. There were 120 employees in this company. Each employee had their different tasks to perform. Receiving the orders from buyers, processing and delivering the order each and every task is separated for the each employee. To develop the new system we required Oracle database [4], which is a relational database management system, presents information as rows contains collection of tables and each table preserving a set of one or more columns. We also used Visual basic [5] and TOAD (Third Party tools for Oracle) which is a set of development tools used for application development, database development, or business intelligence and assist developers in deploying Oracle-based applications and Web services on the Windows platform. To develop the proposed system the package developer was selected as a front-end oracle as a back end.

We also prepared a questionnaire and took interviews of employees. The objectives of our developed system are to make a system which is easy to handle, highly efficient, flexible, secure, could cut down cost over runs and delays, to provide proper control of data, avoidance of error and reduce the workload. By using our system

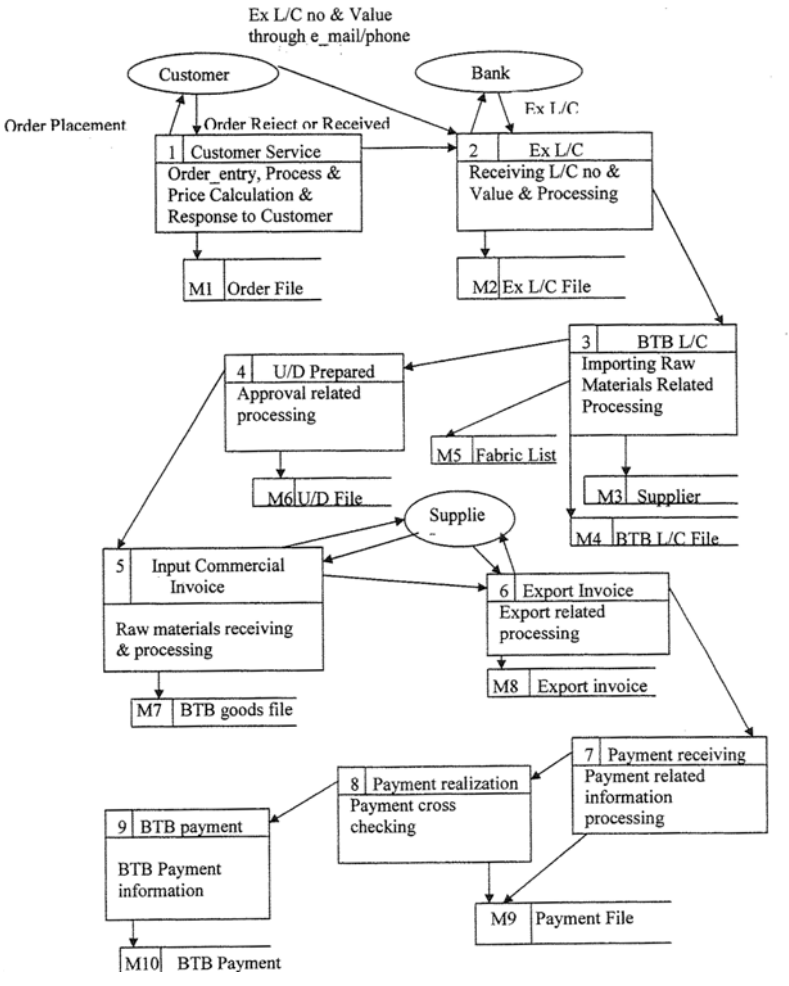


Fig. 3. 1st level DFD

company will get facility to monitor the different type of operation, causes of operations and the result of calculation process.

We also analyze the cost and benefit for four years by using this equation

$$NPV= 1/(1+i)^n$$

Where, n= number of years and I= Interest rate

The proposed system will be feasible to implement for the company. We find out the specific information needs of the company and also find out the basic resource requirements, costs and benefits. There are five phases i.e. economical, technical, behavioral, legal and operational for feasibility study.

Economical Feasibility: Implementation of our proposed system will ensure the best possible performance from the part of resources, including human recourses.

Technical Feasibility: Compatibility and upgrade ability of the system are the main factors that determine the technical feasibility. Here for our system we are using oracle 8.0.4 and developer 6.0 as GUI and database server respectively as these will be very feasible for any organization because they are able to run those two software with minimum requirements.

Behavioral Feasibility: In our system the data entry forms are very user friendly because oracle is very powerful as GUI and data entry has been kept to minimum.

Legal Feasibility: This factor mandates that no conflicts exists between the system under consideration and the ability of company to discharge its legal obligation.

Operational Feasibility: The proposed system has designed based on the environment of the company, existing procedures and personal. The employees will get the limit to adjust them to the new system gradually rather than jumping into it on out right.

To make an entity grid we followed the following steps

Step 1: Find out or select initial entities
Step 2: Placing the entities into a grid
Step 3: Converting the grid into grid into initial entity relation diagram
Step 4: Determine the degree of relationship
Step 5: Identify and add additional characteristics

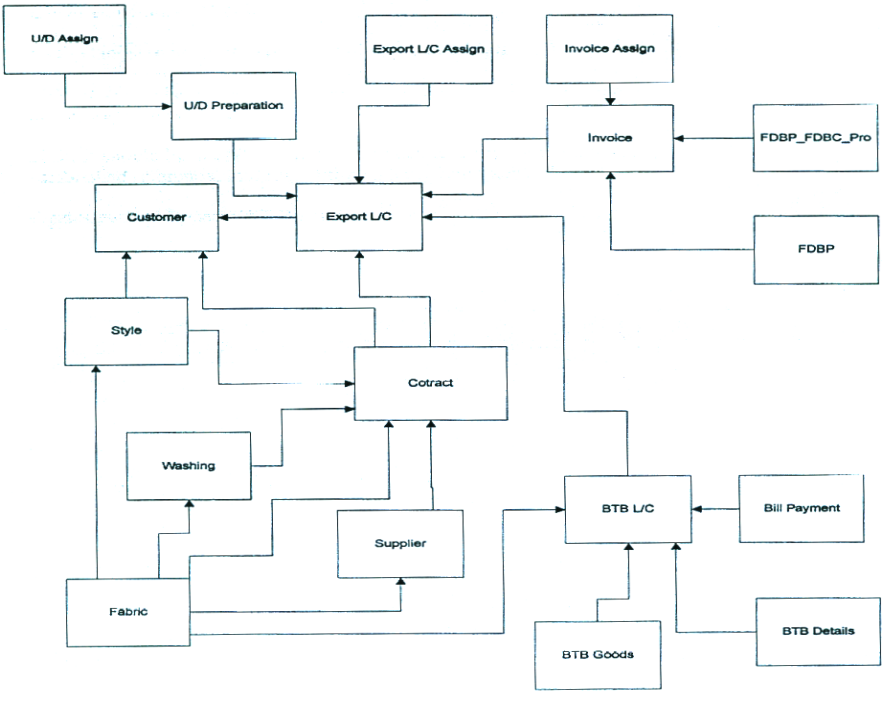


Fig. 4. Entity relation diagram

To normalize the process we followed the following rules

UNF: Normalized form
1NF: Choosing key and remove repeating group of key
2NF: It should be in 1NF and remove the partial dependency on key attribute or increase full dependency on key field
3NF: It should be in 2NF and remove inter dependency among nonkey attribute and established referential integrity (Primary and Foreign key)
Optimization: Repeat 3NF process

UNF	1NF	2NF	3NF	ENTITY
BTB_l/c_no BTB_l/c_Dt E_l/c_no E_l/c_val E_l/c_Dt E_Reff Bill_no Bill_val Pay_Amt Pay_Dt	**BTB_l/c_no** BTB_l/c_Dt E_l/c_no E_l/c_val E_l/c_Dt E_Reff **BTB_l/c_no** **Bill_no** Bill_val Pay_Amt Pay_Dt	**BTB_l/c_no** BTB_l/c_Dt E_l/c_no E_l/c_val E_l/c_Dt E_Reff **BTB_l/c_no** **Bill_no** Pay_Amt Pay_Dt Bill_val **Bill_no** Bill_val	**BTB_l/c_no** BTB_l/c_Dt *E_l/c_no **E_l/c_no** E_l/c_val E_l/c_Dt E_Reff **BTB_l/c_no** **Bill_no** Pay_Amt Pay_Dt Bill_val **Bill_no** Bill_val	**BTB L/C** **Export L/C** **BillPayment** **BillVal**

Fig. 5. Normalization of BTB payment

3 Results

User Interface Design:

1. **Style Sheet Information:** We used 8 buttons for Style Sheet Information. Those were Add new, Update, Delete, Cancel Update, Move first, Move last, Move next, Move back.
2. **Back to Back L/C:** We used 6 buttons- Add New, Update, Delete, Add Row, Save and Delete. Last 3 buttons are under the grid.
3. **Export L/C:** We used 6 buttons for Export L/C- Add New, Update, Delete, Add Row, Save and Delete. Again last 3 buttons are under the grid.
4. **U/D Declaration:** For U/D Declaration we used 5 buttons- Add New, Update, Delete, Add Row and Save. Last 2 buttons are under the grid.
5. **Input Commercial Invoice:** We used 2 buttons- Update and save.
6. **Export Invoice:** We used 4 buttons- Add New, Update, Delete, and Save where save button is under the grid.
7. **FDBP/FDBC:** For FDBP/FDBC we used 7 buttons- Add New, Update, Delete, Cancel update, Move first, Move last, Move next and Move back.
8. **FDBP/FDBC Realized:** Add New, Update, Delete, Cancel update, Move first, Move last, Move next and Move back are the buttons for FDBP/FDBC Realized.
9. **Back to Back Payment:** This phase is related to Back to Back L/C.

10. **System:** This is the form where all information is stored initially. There are some buttons which have some specific works. These buttons are customer, supplier, style, fabric and washing.

Back to Back Payment:

1). Input data and test output data (including all respective important fields)

Name of the forms	Input fields	Output fields	Important fields
Back to back payment	BTB lc no	All other fields.	BTB lc date, Elc no, Elc reff

2) Individual test for each form and functionality:

Name of the forms	Input fields	Output fields	Buttons
Back to back payment	This field BTB LC no is an input field here. We have to a number input in this field. Other input will give an error to this form	All other fields show data regarding to this	**Function of the buttons under the grid** **Function of Add Row** This button is used to add all the formations which are needed to enter into the grid. **Function of Save** This button is used to save all in formations which are entered into the grid.

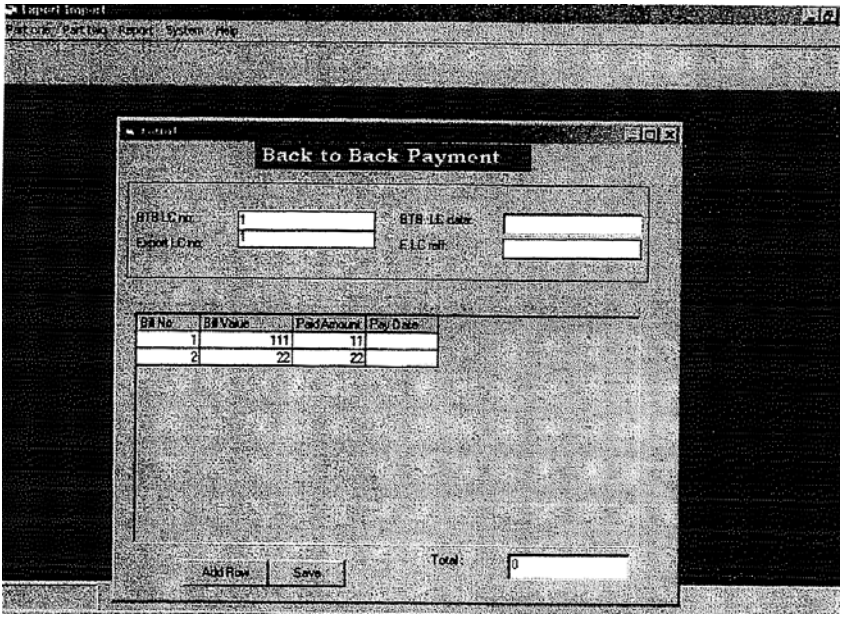


Fig. 6. Screen shoot of Back to Back Payment

FDBP/FDBC Realized:

1). Input data and test output data (including all respective important fields)

Name of the forms	Input fields	Output fields	Important fields
Fdbp_fdbc realized	FDBP/FDBC no	All other fields.	Foreign commission, Realization amount

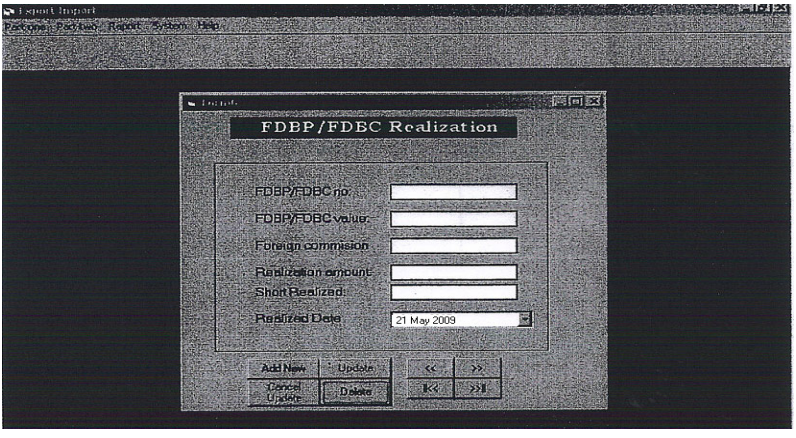


Fig. 7. Screen shoot of Export Invoice FDBP/FDBC Realized

4 Conclusions

Our proposed system improved the stability, security administration and performance of the previous system. It also enhances the support for database replication. This system has the capability to handle the larger number of concurrent users and also could provide a high database environment to support.

References

1. Bangladesh Development Series: Key Issues and Strategic Options for Bangladesh Readymade Garment Industry. The World Bank Office, Dhaka (2005)
2. Munir, Q., Salim, R.: Entrepreneurs and Economic development- The Remarkable Story of Garment Exports from Bangladesh, pp. 23–80. The University Press Limited, Dhaka (2000)
3. Nuruzzaman, M.: Developing Export of RMG products in Bangladesh: Analysing the lead time. Management Trends 4(1), 1 (2007)
4. ORACLE 8.0.4, `http://www.oracle.com/index.html`
5. Visual Basic 6.0, `http://msdn.microsoft.com/en-us/vbasic/`
6. Nuruzzaman, Ahasanul, H.: Lead Time Management in the Garment Sector of Bangladesh: An Avenues for Survival and Growth. European J. of Sci. Res. 33(4), 617–629 (2009)

Reliability Model Research on E-Business System Composited with Web Services

Yugen Du

Software Engineering Institute, East China Normal University, 3663 Zhongshan North Road, Shanghai, China
ygdu@sei.ecnu.edu.cn

Abstract. With the development and application of Web services technology, E-business composite with Web services represents new paradigm in modern E-business development. The reliability research of this E-business system is different from tradition one: the update of web services elements and dynamic nature of their composition shall be considered as well as the open environment of Internet. This article analyzed basic structure of E-business system composited with Web services, made reliability research on service elements and their composition, introduced the concept of service update into reliability model, and provided reliability model for E-business system based on service update under open environment. Its application method and feasibility has been presented with a real case at the end.

Keywords: E-Business, Reliability Model, Web Services, Service Update.

1 Introduction

E-business generally refers to business activities and transactions by electronic means. Its application field can be divided into three parts: First is international electronic business and trades. Second is trade between enterprises, such as signing orders through internet. Third is trade between consumers and enterprises, such as internet shopping. Users with basic computer knowledge can carry on transactions on the web with convenience and low cost. E-business changes traditional way of transaction and has great influence on business activities of enterprises.

1.1 New Paradigm of EB System

But on E-business market, it is not realistic to request all participants to adopt the same model which is based on some language and platform. It becomes a trend to set up E-business system based on internet with application of services-oriented architecture (SOA) technology and composite with web services elements[1][2][3]. E-business composite with Web services represents new paradigm in modern E-business development [4]. It is made up of general software component and web services component. Since software is regarded as a service, such components can be generally called services. Web service is a loosely coupling software component, which can be described and published, discovered and deployed on the web. With combination of component oriented methods and web technology advantage, web

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 285–293.
springerlink.com

service has good adoptability and flexibility to communicate by standard internet protocol and XML data format [5]. Under the vast internet virtual computing environment, any system supporting these standards can be dynamically positioned and interchanged with other web services. Any customers can use any web services at any places, which has broken limits that traditional distributed computing model has on communication and field of application. Both enterprises and individuals can quickly set up and deploy their global applications at low cost.

This new paradigm E-business system is generally composed with many service components (including web services) and satisfied users' requirements on reliability. However, web services normally are provided by different providers, they have different physical distribution and running platforms. In reality, service options that provided for some functional requirement may be more than several, which makes web service reliability research become more complicated.

1.2 Basic Structure of New Paradigm EB System

Many service components need to be composed in new form EB system. Service composition refers to a new value-added service originated from a group of services with logic order and composed with some rules. The components of this "new service" can inter-correspond to each other. What's more, they are composed by certain logic.

The inner structure relationship of service composition is shown as below [6][7] (see Fig. 1):

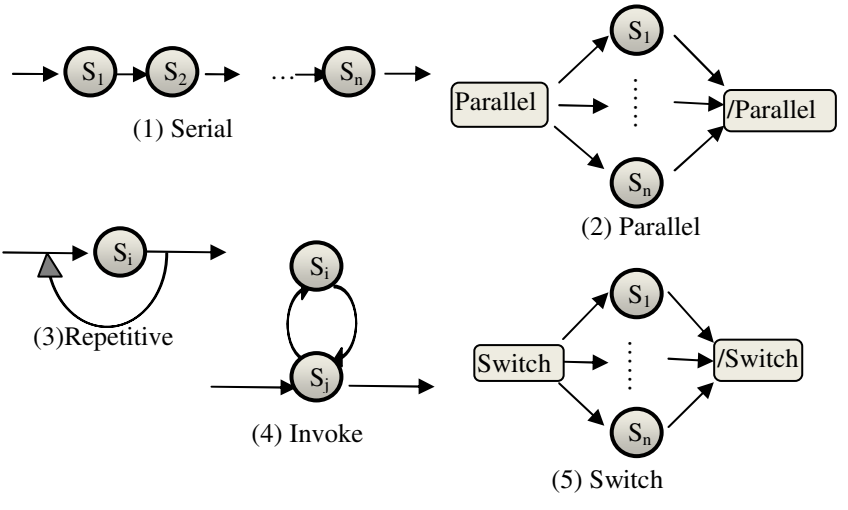


Fig. 1. This shows basic structure relationship between services. (1) Serial structure contains one or more services that need to be executed by turn. Arrows show sequential structure and execution priority of service. (2) Parallel structure provides concurrency and synchronization of services. (3) Repetitive structure requests repeat execution of assigned services until given conditions no longer fit. The back arrow shows repetitive service structure. (4) Service invoke structure corresponds a service execution needs to invoke outer service. The arrow between services S_i and S_j shows their invocation relationship. (5) Switch structure refers that a branch execution is chosen either by an inner condition or an outer request.

2 System Reliability Analysis

2.1 Basic Service Reliability Analysis

Reliability of service execution is mainly analyzed by the following data [8]:

1. Service invocation failure: refers to many errors or bugs in service program codes. When input makes program trigger the bug, it will cause unexpected return result.

2. Cumulative failure function *N(t)* refers to total failure count from the start of service test to present moment. It is a random process and its mathematical expectation function $\mu(t) = E[N(t)]$.

3. Failure intensity function $\lambda(t)$ is change rate of $\mu(t)$, $\lambda(t) = \frac{d}{dt}\mu(t)$.

4. Unreliability $F_S(t)$ a service *S* is the probability of failures happen in the range of [0,*t*] (*t* is time), its change rate is failure intensity function $f_S(t)$.

5. Reliability $R_i(t)$ of service S_i is the probability that service S_i still can work (*t* is time), $R_i(t) = 1 - F_{S_i}(t)$, $R_i(t) = \int_t^{\infty} f_{S_i}(x)dx$.

6. Failure rate, or risk rate of a service *S*, $r_S(t) = \lim_{\Delta t \to 0} \frac{P(t \le T \le t + \Delta t \mid T > t)}{\Delta t}$, *T* refers to the time when failure happens, then $r_S(t) = \frac{f_S(t)}{R_S(t)}$.

7. Service invocation failure rate P_{ij} refers to probability that failure happens when service S_i invokes S_j under open, dynamic and inconstant Internet environment.

8. Transfer probability Q_{ij} refers to probability that service S_i transfers to service S_j.

2.2 Static Service Composition Reliability Analysis

Reliability of services composition depends not only on each individual service, but also on relationship of connection structure of services composition. Services composition can be divided into two groups by different composition mechanism: static composition and dynamic composition. Static composition refers to static binding of E-business system and single service that fulfills the business function. This situation is suitable for interchanges within a united service organization or between fixed partners. Both parties have already reached some agreement at first to ensure a long time stable cooperation. Service provider guarantees services are accessible at any time and service interface will not be changed. To analyze reliability

of services composition, inner structure relationship (which is static composition relationship) must be obtained. There are following service connection relationship [9][10]:

1) Serial system. Among it, failure of any one service will cause failure of the whole system. System reliability is $R_{sys}(t) = \prod_{i=1}^{n} R_i(t)$.

2) Parallel system. Among it, if only one component can reach target function, the whole system will not fail. System reliability is $R_{sys}(t) = 1 - \prod_{i=1}^{n} (1 - R_i(t))$.

3) M/N system. There are total N components in the system, at least M components are needed to collaborate to complete target function. System reliability is $R_{sys}(t) = 1 - \sum_{i=0}^{m-1} \binom{n}{i} R(t)^i (1 - R(t))^{n-i}$.

Obviously when M=1, it degenerates to parallel system and when M=N, serial system.

4) Repetitive system. One service S_i passes through n times repetitive use. System reliability is $R_{sys}(t) = R_i(t)^n$.

5) Invoke system. Service S_i invokes S_j. System reliability is $R_{sys} = R_i P_{ij} R_j P_{ji}$

6) Switch composition. System reliability is $\sum_{i=1}^{n} Q_i R_i(t)$. Among it, Q_i is probability of transfer to service S_i and $\sum_{i=1}^{n} Q_i = 1$.

2.3 Dynamic Web Service Composition Reliability Analysis Based on Service Update

Under internet environment, static services composition normally is not enough. Enterprises always hope to find better partner to get better and more economic services. Therefore, dynamic composition is a requisite function when E-business system composes services (especially web services).

Dynamic composition means no fixed service is assigned in the process of setting up services composition. Detailed binding is delayed and dynamically finished until composite services are executed. Dynamic composition requires that at stage of designing composite services, only certain contents such as functions to be fulfilled, service type, address need to be searched and strategy of choosing multiple searched results are described. While detailed service provider is bound in the process of its execution. This guarantees composing services better fit very high dynamic application environment of internet [11[12].

In the composition process of web services, service update is usually encountered. What is web service update? Service update includes two levels of meanings: one level is updates that web service itself makes to modify defects or add new functions; the other level is updates of service binding decided by UDDI choosing strategy to realize dynamic web services composition or collaborative participants dynamically leave or join the system due to its outer environment dynamic evolution.

Service updates cause service failure. A service failure count is a random process. Its Cumulative failure function is determined by two other random processes [10].

1) Service failure caused by service component itself failure. Its cumulative failure function is *N(t)*;

2) Service failure caused by service updates. Its cumulative failure function is *M(t)*.

While cumulative failure count of EB system composite with services is also a random process. It is two dimensional joint distribution of *N(t)* and *M(t)*.

In order to research system reliability in the process of service updates, we introduced service update intensity into Musa-Okumoto model and provided an E-business system reliability model based on service update. Model hypotheses with reference to [13][14][15] are given as below:

(1) When time t is 0, failure count is 0 and web service update count is 0.

(2) Since web service failure caused by service entity's fault or bugs will reduce along with progress of testing, thus failure intensity $\lambda(t)$ decays exponentially:

$\lambda(t) = \lambda_0 e^{-\theta\mu_1(t)}$.

(λ_0 is initial failure intensity, θ is failure intensity decay parameter, and $\mu_1(t) = E[N(t)]$)

(3) Suppose failure increase caused by service changes is stable which is a homogeneous Poisson process and web service update intensity $\delta(t)$ is a constant $\tilde{\delta}$, and then there exist

$$\delta(t) = \frac{d\mu_2(t)}{dt} = \tilde{\delta}, \quad \mu_2(t) = E[M(t)]$$

(4) Suppose no new bugs occur after service entity's malfunction has been repaired; suppose if web service updates happen, service entity's original failure rate and failure decay rate after service updates equal to their values before service updates.

(5) Within a small interval(t, $t+\Delta t$), probability of one web service failure is $\lambda(t)\Delta t + o(\Delta t)$, probability of more than one failure is $o(\Delta t)$, $\lim_{\Delta t \to 0} \frac{o(\Delta t)}{\Delta t} = 0$.

From supposition (2), differential equation can be deducted as

$$\frac{d[e^{\theta\mu_1(t)}]}{dt} = \lambda_0\theta,$$

Its integral is $e^{\theta\mu_1(t)} = \lambda_0\theta t + C$, C is integral constant.

From supposition (1), $\mu_1(0)=0$,so C=1。 The mean function is

$$\mu_1(t)=\frac{\ln(\lambda_0\theta t+1)}{\theta} \quad . \tag{1}$$

From supposition (3) differential equation $\frac{d\mu_2(t)}{dt}=\tilde{\delta}$, integral is $\mu_2(t)=\tilde{\delta}t+C$, C is integral constant, and from supposition (1), C=0 and

$$\mu_2(t)=\tilde{\delta}t \quad . \tag{2}$$

Since web service update random process and Web service entity failure process are independent to each other, from supposition (4), the mean value function of web service failure can be calculated

$$w(t)=\mu_1(t)+\mu_2(t)=\frac{1}{\theta}\ln(\lambda_0\theta t+1)+\tilde{\delta}t \quad . \tag{3}$$

From supposition (1) and (5), web service failure is a Poisson process. Therefore, probability of $w(t)=m$ is $P_r\{W(t)=m\}=\frac{[w(t)]^m}{m!}e^{-w(t)}$, web service reliability distribution is

$$R_i(\Delta t\mid t_{i-1})=e^{-[w(\Delta t+t_{i-1})-w(t_{i-1})]} \quad . \tag{4}$$

By formula (3), (4)

$$\begin{aligned}R_i(\Delta t\mid t_{i-1}) &= \exp\{-[\frac{\ln\lambda_0\theta(\Delta t+t_{i-1})+1}{\theta} \\ &+\tilde{\delta}(\Delta t+t_{i-1})-\frac{\ln\lambda_0\theta(t_{i-1})+1}{\theta}-\tilde{\delta}t_{i-1}]\} \\ &=[\frac{\lambda_0\theta t_{i-1}+1}{\lambda_0\theta(\Delta t+t_{i-1})+1}]^{1/\theta}e^{-\tilde{\delta}\Delta t}\end{aligned}$$

3 Case Analysis

This section will apply the above model based on web services update to analyze service feasibility first, then analyze connection structure of web services composition, and provide a reliability measurement to the whole system.

Here is EB system of on-line markets. There are many stores in the market providing goods. Consumers can buy goods on line and choose multiple on-line payment ways. This is a typical EB system with both B2B (market to stores and banks) and B2C (market to consumers) features. Below shows basic flow chart (see Fig. 2.).

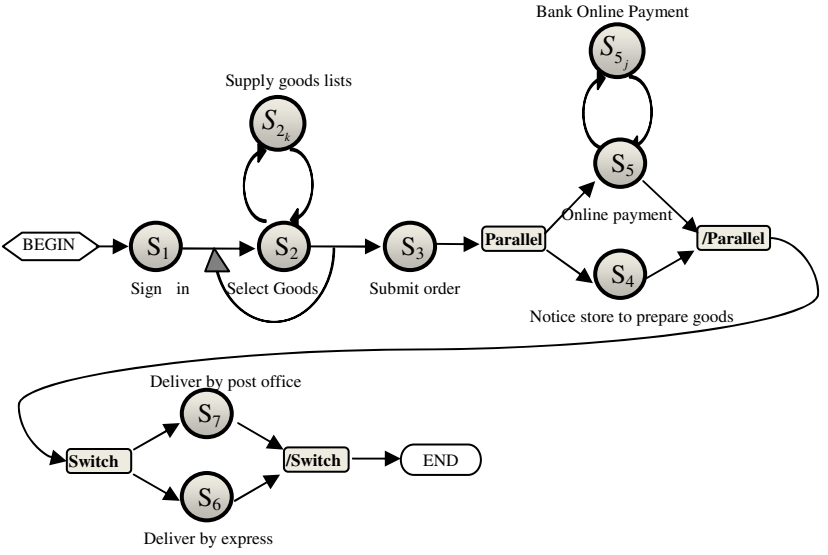


Fig. 2. This shows a services flow of EB system. A user logs on the goods transaction platform, then he searches the goods he needs. He can invoke some service S_{2_k} provided by a store *k*. After he chooses the goods, he can put it in a temporary storage zone and start to search for next one. After selecting all needed goods, he submits orders. Two processes happen concurrently: to inform the store to confirm and supply and to pay on line. On line payment service S_{5_j} is provided by bank *j*. Finally, goods are delivered by postal or courier services according to user's request.

The reliability of the whole system is analyzed as below:

First by reliability measurement model mentioned in the above section, reliability of web services S_1, S_2, S_3, S_4, S_5, S_6, S_7, S_{5_j} and S_{2_k} can be computed as R_1, R_2, R_3, R_4, R_5, R_6, R_7, R_{5_j} and R_{2_k}, R_{2_k} is reliability of web service S_{2_k} provided by store *k*, R_{5_j} is reliability of web service S_{5_j} provided by bank j. User chooses total n pieces of goods. Then below can be got by applying reliability of web service connection structure:

$$R_{SYS} = R_1 \prod_{k=1}^{n} (R_2 P_{2,2_k} R_{2_k} P_{2_k,2}) R_3 [1-(1-R_4)(1-R_5 P_{5,5_j} R_{5_j} P_{5_j,5})] \times (Q_6 R_6 + Q_7 R_7) \quad . \tag{6}$$

Suppose a single store *k*=9 providing goods; single bank *j*=8 providing on-line payment. We use a group of simulate data λ_0, θ and $\tilde{\delta}$ of services S_1, S_2, S_3, S_4, S_5, S_6, S_7, S_{5_8} and S_{2_9} to analyze reliability of dynamic web services based on update services. Suppose our observation hour (CPU hour) t_{i-1}=1024h, then within Δt=2h,

reliability R_1-R_9 of S_1, S_2, S_3, S_4, S_5, S_6, S_7, S_{5_8} and S_{2_9} can be computed as Table 1 shows.

Suppose invoke failure rate of S_2 to S_{2_9} : $P_{2,2_9}$=0.99, $P_{2_9,2}$=0.98; invoke failure rate of S_5 to S_{5_8} : $P_{5,5_8}$=0.98, $P_{5_8,5}$=0.99;probability of choosing post office or express services by user's outer request, and suppose transfer probability of S_6 to S_7 is Q_6=0.3, Q_7=0.7 respectively.

According to formula (6), EB system reliability of such on line shopping is R_{sys}=0.792876.

Table 1. Services reliability data

Service	λ_0	θ	δ	R_i
S_1	23.6	0.066	0.0031	0.964885888
S_2	25.5	0.049	0.0013	0.958496133
S_3	21.8	0.057	0.0027	0.961168952
S_4	44.2	0.069	0.0045	0.963416451
S_5	35.4	0.054	0.0067	0.951690622
S_6	36.8	0.078	0.0054	0.964826349
S_7	29.7	0.041	0.0005	0.952607196
S_{5_8}	19.7	0.068	0.0023	0.967273944
S_{2_9}	39.4	0.076	0.0049	0.965155948

4 Conclusion

The article discussed new paradigm of E-business system composite with web service, researched system reliability in the process of service updates, introduced service update intensity into Musa-Okumoto model and provided an E-business system reliability model based on service update, which is an extension of Musa-Okumoto model under composite web services. The model proposed a reliability measurement of E-business under web service composite, which provides effective support to E-business system reliability research under open environment. Its application method is demonstrated by a real case at the end.

References

1. Gold, N., Knight, C., Mohan, A., et al.: Understanding service—oriented software. IEEE Software 21(2), 71–77 (2004)
2. Piccinelli, G., Sallé, M., Zirpins, C.: Service-oriented Modelling for e-Business Applications Components. In: Proceedings of the 10th International Workshops on Enabling Technologies: Infrastructure for Collaborative Enterprises (WET ICE 2001), pp. 1080–1383. IEEE (2001)
3. Booth, D., Haas, H., Mccabe, F., et al.: Web Service Architecture [EB/OL] (2004), http://www.w3.org/TR/2004/NOTE2ws2arch220040211/

4. Marton, A., Piccinelli, G., Turfin, C.: Service provision and composition in virtual business communities. In: Proc. 18th IEEE - ISRDS, Int. Workshop on Electronic Commerce, Lausanne, Switzerland (1999)
5. W3C. Extensible markup language (xml) 1.0 (2nd edn.) (2000), http://www.w3.org/TR/2000/REC-xml-20001006
6. Hamadi, R., Benatallah: A Petri Net-based model for Web services composition. In: 14th Australasian Database Conference, Adelaide, Australia (2003)
7. Hamlet, D., Mason, D., Woit, D.: Theory of software reliability based on components. In: 3rd International Workshop on Component-Based Software Engineering, pp. 361–370. IEEE Computer Society, Toronto (2001)
8. Singpurwalla, N.D., Wiilson, S.P.: Statistical methods in soft ware engineering: reliability and rick. Springer-Verlag New York Inc. (1999)
9. Yang, W.L., Wu, Y., Chen, M.H.: An architecture-based software reliability model. In: Proceedings of Pacific Rim International Symposium on Dependable Computing. IEEE, Hong Kong (1999)
10. Du, Y., Li, Y.: Research on Internetware Reliability Measurement Model Based on Service Update. Journal of Frontiers of Computer Science and Technology 2(4), 431–438
11. Mennis, D., Pagurek, B.: An Architecture to Support Dynamic Composition of Service Composition. In: Proceedings of the 5th International Workshop Component-Oriented Programming, WCOP 2000 (2000)
12. Poladian, V., Sousa, J.P., Garlan, D., Shaw, M.: Dynamic configuration of resource-aware services. In: Proc. of the 26th Int'l Conf. on Software Engineering (ICSE), Edinburgh, pp. 604–613 (2004)
13. Lu, D.Q.: Stochastic processes with applications, p. 8. Tsinghua University press (1986) (in Chinese)
14. Musa, J.D., Iananino, A., Okumoto, K.: Software reliability: measurement, prediction, application. McGraw-Hill, New York (1987)
15. Musa, J.D.: Software reliability measurement. Journal of Systems and Software 1, 223–224 (1980)

Discuss in the Mission and the Curriculum Construction of Chinese College Physical Education in Transformation Period

Wu Jianjun and Zhang Anhong

Yunyang Teacher's College, Shiyan, China 442000

Abstract. The method of literature consultation was used to analyse the mission of college physics curriculum in social transform period .And trying to raise some feasible suggestion for the construction of college physical curriculum.

Keywords: Social transform, college physical education, curriculum.

1 Introduction

In the economic globalization and cultural pluralism reform tide,the world's whole pattern are peace and development,cooperation and competition at present. The Chinese nation are experiencing a great reformation in that history background. The general characteristic of social change is social transformation in China today. The international competition are mainly the competition of science and economic. And the essence of the competition is talents competition. Institutions of higher education shoulder heavy responsibilities for training higher talents.And physical education is an important part of higher education.The mission of training higher talents determines the college physical education must refraining so that it can meet the talent requirement in social transformation period.

1.1 The Influence of Social Transformation on the Structure of Social Physics

Social reformation which is a whole social revolution course is general future of social changes. It is also a course of social modernization.The social transformation constitutes of two profound changes. One is structure transition which from planned economy system to market economy system. The other is structure transition which change from agricultural and closed traditional society to industrial and open modern society. The change of the social structure is the main contents of the social transformation that involves the complete transformation in all thefields of the society. The economic development model of the state decides the development model of the physics .The change of economy and structure affect various fields,and it will affects the reform and development of social physics.

In the planned economy period, sports have some kind of welfare. In order to improve the health of workers and labor productivity, promote cultural development; enhance cohesion, agency employees carry out various sports activities in the unit of the agency, industry, or specific area. With the socialist market economic system

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 295–299.
springerlink.com

gradual establishment, sports activities changing from the "welfare" to "consumer" is a big trend. With the promoting of market oriented economic system reform,the ftransformation of government functions and the establishment of modern enterprise system, market economy will require for desocialization of units function.

Desocialization of units function is a process of separation and relocation of social functions,including the supply of sports welfare function . Under conditions of market economy, units especially as companies have not fully occupied working time, human ,material or financial resources of the units to organize a variety of sports activities. Social ports has been management in units, but it will be more and more restrictions. Service functions of the social sports, including social welfare services, gradually transferred to the society and the community. Sports is gradually becoming one of leisure activities of workers themselves . The profit-orientation of the companies and the amaturizition of sports activities reduce the possibility of organizing trade union activities only rely on administrative . People's consciousness of dominating their free time according to their own hobby is growing in transition period. That makes the original mode of operation mechanism been challenged.However, with globalization people are faced with increasing competition . In such context, more and more people, especially who are in the pillar of the status of society and family, their physical activity can not be guaranteed. When observed groups participating in physical exercise, we can easily found the people who were the most active the elderly. Women accounted for the vast majority of the elderly.The youngs were minorous. This transition can not but to admit it is the regret and lack of social sports.

Of course, with higher levels of economic development, a substantial, advanced material conditions, rich sports life means and methods are provided for the masses to participate in sports ; Secondly, it enable people enjoy the tremendous physical fun freely; thirdly, higher socio-economic security improve people's ability to pay for sports, so that more people are taking sports activities.

1.2 The Mission of College Physical Education in Social Transition Period

In 21st century, facing with the rapid development of mass sports world and our fitness plan, college physics can not settle for traditional teaching content and teaching methods only , and it should assume a new mission.

First, college physical education shoulder teaching motor skills, enhancomg physical fitness of students and providing scientific training methods for students. Because of academic pressures, students in primary and secondary school system are difficult to have time for physical education and exercise for a long time. A national survey results in 2005 showed that students gradually increased the level of morphological development of physical and nutritional status continued to improve.But the level of vital capacity continued to decline, speed, explosive strength endurance levels decline further. While the prevalence of obesity continues to rise, vision poor detection rate remains high. One of the reasons is institutional cause.But the main reasons are the students and the public understanding of sports is not enough and the students lacking scientific method of exercise. After entering university, relatively speaking, students have more free time, students taking more time to engage in physical exercise becomes possible. Therefore, college physical education should guide students

to scientific exercises and help them learning one or more scientific methods of exercise to lay a solid foundation for lifelong physical lay.

Secondly, college physical education also shoulder the tasks of teaching knowledge of health care and self-cultivation. Though after ten years of primary school to university physical education, we find that people' s common sense about sports fitness is very poor. We often came across people kneading immediately when ankled valgus or varus (commonly known as Wei feet). They do not realize the hurt can not be kneaded. The best way should be using handle ice or cold water to wash affected area immediately.And use recovery techniques 24 hours after . We also came cross many students stopped and rested or drinking lots of water immediately after long-distance running. Those wrong moduses show college physical education are also a long way to go, there are many things to be done.

Finally, colleges are the the last leg for students to entry the society. It shoulders the an important task that links the proceeding with the following.And it is also the last chance for education system. Therefore, how to make the results of school physical education become the beginning of sports, so that students develop a sense of good sportsmanship and benefit to society for life is also the mission of college physical education.

1.3 Curriculum Construction of College Physical Education in Transitition

In the past, the monotony of college physical education was a unified model because of the impact of competitive sports and the traditional teaching mode. Such single teaching mode has not adapted to China's economic construction of the personnel requirements. Only to build a variety of teaching in accordance with teaching objectives especially the quality target, so that different modes complement each other and play to their capabilities object, Physical Education Teaching can gradually build teaching models with Chinese characteristics and adapted to the characteristics of college teaching.And can make our Physical Education Teaching more perfect. College sports must constantly open up new areas of physical and mental function based education and awareness education according to the reality of today's society, civilization and ways of life. Modern sports has become a way of life. Especially with the the development of society and science,it shows its positive role. Modern sports not only enhance the function of human body, but also protect people's mental health and promote quality of life and lifestyle. Therefore, increasing the reformations of physical education models is new requirements for the quality of talent and size requirements with changes of age, which is the fundamental reform of physical education motivation.

Currently, the obvious contradictions between college physical education and the needs of the community is reflected in the teaching content, extent, methods, and social awareness of sports, fitness ideas and methods. In order to train students to a lifelong sports awareness people, college sports should combine personal and social development goals organically and give full play to the role of college sports. And brought college sports into the fitness track , into the orbit of lifetime sports. Only in this way, college sports and community sports can achieve seamless and better services for the masses and socialist market economy.

First,we must reform its curriculum, so that the course structure consistent to training objectives.The contents of current PE curriculum are above normal.Throughout the former

Ministry of Education promulgated the "Higher general physical education curriculum ", consists of eight categories, namely: basic theoretical knowledge of sports and health knowledge , track and field, gymnastics, basketball, volleyball, soccer, martial arts, swimming.Only from the contents of the individual track and field settings,we found it consisted both the basic theoretical knowledge and various forms of running, jumping, throwing.The long running consisted of running, short distance running, relay, hurdles.Jumping consisted of a high jump (straddle, Fosbury Flop) and long jump (squat-style, to come forward type).Throwing grenades consisted turning back or side pushshot put.Which we can easily find that more of these elements follow a set of sports content, as opposed to ordinary exercisers are not targeted.College sports shoulder the important historical mission of transition to social sports. Therefore, setting the contents of courses should be targeted to a wide audience loved.Popular sports that people love should arranged into the curriculum, some more competitive items should be strongclean up the classroom project instead putting on amateur training.For example, wide variety of ball games that optimistic about market prospects now, Aerobics, dancing, Tai Chi, Sanda, taekwondo ,swimming (if conditions can also be set up), outdoor sports and other special projects.In addition,we should build some characteristic curriculum relying on localregional characteristics.China have vast territory and various nations.So there are many national and local characteristics of sport has yet to be exploited.Such as the Dai peacock boxing, jump on thin Ah Yi, Bai rattle, Gaoshan basket ball, Naxi Dongba dance, bullfighting Hui, Korean swing, the Tibetans' Bixiu "Tu wheels fall, Li jump bamboo, the Uighur's "Dawa Zi," Zhuang altitude lion, the horse picked up silver Kazak, Mongolian camel racing and horse racing, the Diaoyang Tajik, Daur "Plendil wide"and so on.And There is much water in south China,the South have more water, swimming diving and other items can be arranged. And the north, you can arrange ice-skating according to weather and climatic conditions.Only reforming curriculum in accordance with social needs to break mainly competitive project-based curriculum and establish curriculum mainly cultured lifelong awareness.So that college physical education can make a real college sports for the community, and integrate with socialphysical.

Meanwhile, the establishment of complementary curricular and extra-curricular teaching system can create conditions for comprehensive development of for college sports. We know that the current sports system of our state is coexistence of public sports and athletic competition.

In addition to completing the task of teaching curricular college sports for the sports training fitness professionals, the college sports also bear the appropriate sports tasks.

Second, developing more stringent security system,increasing efforts to protect the university's physical education class to ensure completion of tasks of college sports. "National College Physical Education Teaching Guidelines" (2002) stipulates "The students in grade one and two must be opened of university physical education curriculum (four semesters totaling 144 hours). To obtain required credits one of necessary condition that students graduated with a degree. "Many colleges arranged a physical education class once a week (2 hours), 18 weeks per semester basis, 4 year term up to 2 144 hours, However, a variety of recreational activities and holidays together with the annual school track meet often delay a certain number of teaching hours in actual operation,hours are often not protected. A considerable part of teaching is unable to complete the task ,reach a good teaching or practical effect of physical exercise. And some colleges open physical education in the first grade only due to various reasons . Therefore, we need a strong

system to protect the teaching of physical education and need to bring physical teaching into training evaluation system, and achieve a veto. In addition, the annual track and field games remain competitive ideas which needed to be innovative. We can design a new multi-functional organization or individual competition combined teaching contents, such as the Three-Basketball, Futsal, or seven-a-soccer, martial arts competitions, aerobics competitions,etc. And link the game scores with PE scores so that training promotes competition and competition promote learning.

Third, increasing investment in high-tech and introducing multimedia sound and light to physical education. Students have not only know how to learning or only teachers from the "nice children" now.They are characteristic and access to knowledge by more broadly means . Therefore, we should use all available means and carrier, so that students obtain various sports knowledge and skills from different channels. For example, you can use multimedia remote video technology to enable students to follow the world-class coaches to learn technology.

Finally, the form should not be limited to class, but according to the phases of different levels of teaching techniques to meet the needs of students at different levels. Grouping by interests should be more placement, rather than prescriptive splitting classes. For example, a centralized placement in the lower grades for teaching basic techniques. Using club form of high school electives can increase student autonomy and learning initiative. Theoretical part of the lecture format is used, rather than the boring scripted to increase student interest.

In short, about the curriculum of college physical education, we must expand the college sports training objectives to meet the needs of the development of modern society.And grasping well the relationship between general and special courses.Training students in life-long physical concept to make the curriculum with epochal character, social values ,life-long training for personnel training and exploration of bold reforms.

Guided by life-long sports college sports reform, which is inevitable social trend of scientific and popularity of modern sports.It is also the popularity of mass sports and in-depth development. In the new historical period, our college sports journalists only bear in mind the mission and courage to explore, to build a college physical education model that meet the needs of social transition period.So that we can provide better protect system for students to develop the concept of lifetime sports and train qualified community expertise.

References

1. Luo, L.: College sports programs to build should reflect the "lifetime sports" the purpose. Qiannan Teachers College (Philosophy and Social Science) 4 (1997)
2. Lei, X.: China's social transition coaches gymnastics training and quality of construction. New West 16 (2007)
3. Ren, H.: On the new era of China's workers face the challenges and opportunities sports. Beijing University of Physical Education 1 (1994)
4. Wang, K., Ren, H.: China's social transformation and urban social change in sports management system. Beijing University of Physical Education 4 (2004)
5. Wang, Y.: On the quality of education and reform of college sports. China Adult Education 5 (2007)
6. Wang, H.: The social environment influence the development of college sports. Technology Education Research 8 (2005)

Accounts Payable Analytics - Indispensable for Optimizing Cash Flow

Partho Bandopadhyay, Jaywardhan Semwal, Kumar Suneet, and Tulika

Hewlett Packard, Embassy Prime, Bagmane Techpark, Bangalore-93, India
partho.bandopadhyay@hp.com, jaywardhan.semwal@hp.com,
kumar.suneet@hp.com, tulika@hp.com

Abstract. Operating Cash flow is the, or one of the, most important parameters in gauging the financial health of an organization. Some of the not so visible cost components arising out of accruals system of accounting; needs detailed evaluation in bringing out cost efficiencies. Significant amount of cash can be saved by streamlining payables; one of the key components of operating cash flows. Payables management becomes an insurmountable challenge for organizations that have to deal with inconsistent data coming from heterogeneous information management systems, which mostly is a result of ongoing inorganic growth adopted by organizations. At Hewlett Packard, we have developed an analytics framework that is core to an end to end effective payables management which enables management to take effective control of the vendor payment systems by enhancing visibility and control.

Keywords: Accounts payable, operating cash flow, vendor payment, analytics, data management, DPO, decision support, payables cycle, risk management.

1 Introduction

Most of the organizations with significant cross-border balance-sheet exposure grapple with the challenges of managing their payables effectively. These exposures, that are as wide as considerable geographical spreads, differentiated product lines; to multiple contract terms with numerous vendors, require a comprehensive payables management solution. This paper introduces an analytical framework that can be employed for effective payables management.

Hewlett Packard with its ever expanding reach has all the above mentioned exposures that demanded an effective payables management tool for improved short term liquidity and cash flow forecasting. In addition, inorganic growth strategy at HP resulted in a complex web of several heterogeneous systems with inconsistent data definitions posing an insoluble challenge for payables management. With the given background, our detailed investigations led us to identify fluctuating operating cash flow as one of the key challenges that HP faced.

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 301–307.
springerlink.com

We devised a multi-stage solution in order to optimize the cash flow and increase DPO the various stages are as follows: 1) Standard data definition to get reliable and consistent data, 2) Improved visibility of the payables & trend analysis for geography, suppliers and products, 3) Cluster analysis to classify countries based on performance, 4) Focused control on critical and non-performing accounts such as overdue and early payments accounts, and 5) Risk analysis using effectiveness index and derived risk matrix for the payables landscape.

The insights that management derived from proposed framework contributed in enhancing visibility and improving controls on payables management which contributed to optimized cash flow.

2 Problem Statement

Analysis of high fluctuations in cash flow at HP indicated the problem with vendor payables. Days payable outstanding (DPO) had reduced significantly affected overall cash flow as well as the cash cycle.

In addition, a drill down analysis of operating cash flows suggested that root cause analysis of the problems related to payables was not possible due to data taken from heterogeneous systems, which led to inconsistent data definitions and lack of procurement cycle data at one place. All these caused challenges for the management in identifying payables inefficiency areas and implementing effective vendor payment controls.

3 Analytical Approach for Effective Payables Management

Fig. 1 shows the framework of accounts payable analytics and its linkages to the overall payable system to enable an effective payables system.

3.1 One Size Fit All Database

HP is carrying a number of legacy systems operating in silos owing to various reasons such as systems inherited due to inorganic growth and country and regional complexities. Payables data extracted from the legacy systems had inconsistent data structures and data definitions from different business units. Data inputs were reconciled to produce standardized format, which helped in identifying relevant data and avoid redundancy and inconsistency. Further, all the structured data were consolidated into a standard database with standard querying tools and designs (using SQL, advanced macros and other technical querying methodologies).

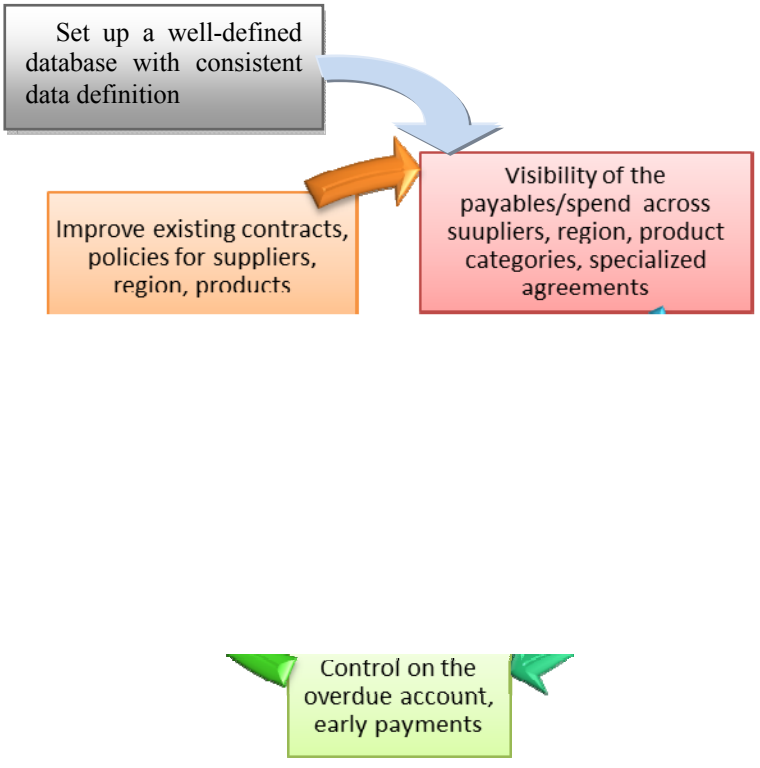


Fig. 1. Closed loop framework for effective accounts payable management

3.2 Enhanced Visibility

Spend value, contractual purchase order term (PO term) and actual payment days were important elements were identified as the critical business parameters for the payables managements. Distribution and trend analysis of inventory (expense used for product manufacturing) and non-inventory spend (royalties, supplies, salary and wages etc.) provided deep insights in operational effectiveness and measurable payment trends.

Custom SQLs were designed to get spend data to deduce valuable insights detailing product categories, regions, countries and suppliers. Also, cross tab analysis of spend (as shown in Fig. 2) across different spend category in different regions helped payment team monitor paid term & manage payables.

Pareto analysis on spend data revealed that about 85% of total spend was with ~25 suppliers. A reporting package was designed to monitor payments for these suppliers closely as deviations could adversely affect cash flow. Further, the Pareto analysis was extended from supplier to product & business category and countries to enhance visibility of inconsistent or problematic area which further helped in root cause analysis.

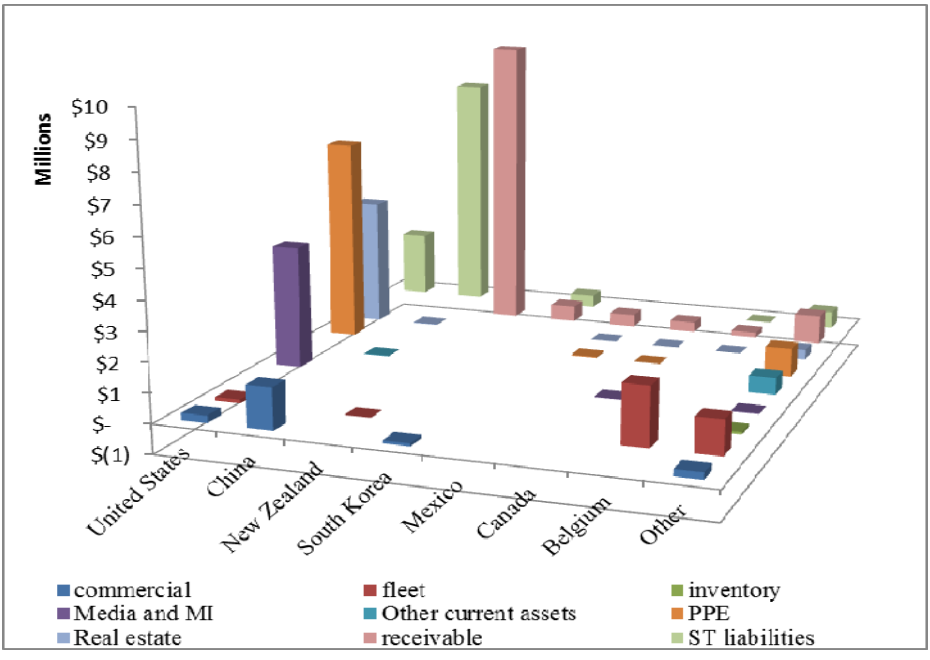


Fig. 2. Cross sectional early payment analytics for countries, spend category and early payment amount (numbers for illustrative purpose)

Purchase order term (PO) and *Actual days paid (DP)* are two important factors for deciding payment flow to supplier. Two indices were defined "weighted average purchase order *(WAPO)*" and "weighted average days paid *(WADP)"*as performance parameters.

WAPO for purchase order = invoice amount in PO * PO contract term / (sum of all invoice amounts). (1)

WADP for purchase order = invoice amount in PO * Actual days paid term/ (sum of all invoice amounts). (2)

These indices are easy to compute, use and compare the performance. Trend analysis of deviation in WADP from WAPO across product category, different countries and different suppliers help identifying the inefficient countries and suppliers which helped in renegotiate the terms.

3.3 Improved Control

While consistent data definitions increased visibility of payables management; there was a need for improved control for cash flow streamlining and in turn improve DPO. These objectives were achieved in twofold initiatives.

Work with Various Stakeholders in Payable Cycle
Actionable insights from Spend, purchase order term and actual days paid analysis were continuously shared with procurement, finance and different business units. These insights provided them visibility to focus on inefficient processes, suppliers, geographies and also renegotiate terms as and when required. Often suppliers demand payments to be made earlier than the PO term to maintain higher liquidity. Analytical insights empowered the payments team to enforce pre-payment discounts to HP as per the guidelines provided by treasury team. In order to deter these early payments and avoid early payments without discount; closed invoice for every quarter were analyzed for four ranges of actual paid days (0-10, 10-29, 30-45, >45 days). Also, it helped in modulating overdue payments as HP would have lost money in late payment penalty and lost trade credit. Further, it also provided the payment team with visibility of cash it should have to close all invoices. Trend analysis of special contract terms like buy-sell contracts (HP sells raw and finished products and procures products from a number of vendors), early payment discount were done to find out the payable performance of such terms.

DPO for different organizations in the industry, different types of procurement and best practices were assessed. These insights were shared with stakeholders for possible renegotiation of terms with the vendors wherever necessary.

Effective Decision Support System
In order to set-up an effective decision making, an effectiveness index was defined as following:

$$\{(WADP1 - WAPO1) * \text{spend sup1} + (WADP2 - WAPO2) * \text{spend sup2} + \ldots\} / (\text{Total spend})\} . \quad (3)$$

Note: "spend sup1, spend sup2.. .spend sup n." are the value of spend with suppliers. Supplier can be replaced with product category, country or any other parameter that one wants to measure for payment effectiveness.

Effectiveness index a financial metric which provides that information for the relative performance of a unit among a category (suppliers, product category, region, countries etc.) or against a benchmark value and thus suggest whether that particular unit is efficient or requires high level of monitoring or to be delved into reasons of non-performance. Ideally this index should be 0.

A positive number indicates that DPO would be healthy for the organization but one should be cautious with higher index due to two reasons: 1) Vendors can feel pressure on their cash flow & relationship with strategic vendors may be strained. 2) It can attract contractual penalties.

Using effectiveness index and annual spend value with supplier; "Risk matrix" can be plotted as shown in Fig 3. Four quadrants of risk matrix present different level of risk associated with the supplier and expected management action. For high risk and moderate risk areas, payables status is being sent to the stakeholders every quarter.

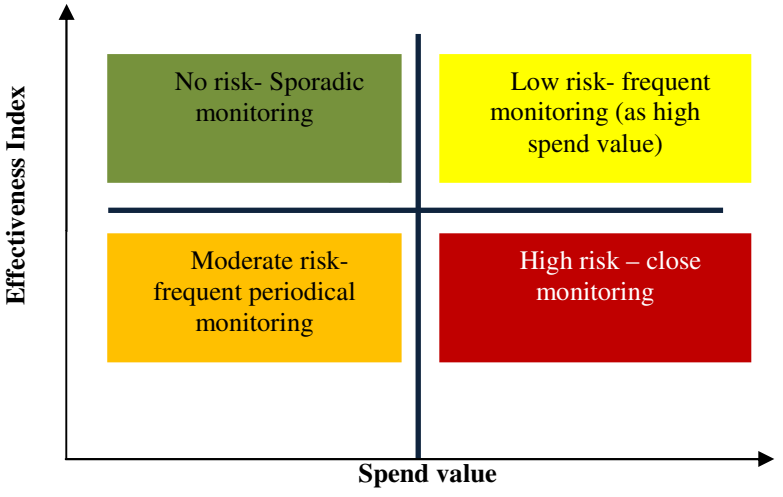


Fig. 3. Risk matrix to classify critical and non-performing countries with control approach

Cluster analysis on payment data for different countries was performed using country-specific parameters such as: API & APNI spend, (WAPD-WAPO), DPO, Account Payment Days; APD (0-10 days), APD (10 -29 days), APD (30-45 days) & APD (>45 days). We received three clusters which were distinct on early payment behaviors as following:

Cluster 1: Low DPO, negative (WAPD – WAPO), high percentage of APNI spend
Cluster 2: Moderate DPO (comparable to standard term), (WAPD – WAPO) positive
Cluster 3: High DPO, Low percentage of API & APNI spend.

It provided payments teams to concentrate on the countries in first cluster as this cluster has high spend value accompanied with low payment performance.

4 Conclusion

The presented case highlights the value generated by having an effective payable management systems coupled with strong analytical methodologies. These insights contributed immensely in enhancing visibility and improving controls on payables management. The adopted analytical methodology empowered decision makers in predicting cash flows with greater confidence and effectively managing payables. This enabled senior management to effectively control the payable systems and achieve an increment of DPO by 10%; thereby increasing shareholders' value. This methodology and the basic ideas can also be replicated to other working capital management components (viz. receivables & inventories) and HP is already exploring the same.

References

1. Hewlett Packard Company annual reports (2008–2010)
2. Filbeck, G., Krueger, T.M.: An Analysis of Working Capital Management Results Across Industries. Mid-American Journal of Business 20(2)
3. Turner-Samuels, M., Thorling, A., Berg, J.: Working Capital Management – Why it needs the treasuresr's touch. Treasury Management International (May 2001)
4. New DPO Metrics Help You Assess Creditworthiness. Institute of management and administration (November 2005), http://www.ioma.com

Prospect and Business of Leapfrog Development of Hainan Rural Tourism against the Background of International Tourism Island

Shi-chen Fan[1] and Tao Song[2,*]

[1] Applied Economics, Postdoctoral Research Station of Southwestern University of Finance and Economics, Chengdu, 611130, China
[2] Finance Department of Hainan University, Haikou, Hainan, 570228, China
{fansc1230,songtao7221}@163.com

Abstract. Based on parsing on the significance and Prospect of the Leapfrog Development of Hainan Rural Tourism against the Background of International Tourism Island, the paper made choices of the tourism business of Hainan rural tourism.

Keywords: International Tourism Island, Leapfrog Development, Prospect deconstruction, business choice.

1 Significance of the Leapfrog Development of Hainan Rural Tourism

At present, the development of Hainan Rural Tourism is in the initial stage, and the overall situation is still sporadic, small-scale, low level and introverted. As to the development of rural tourism, there is great gap between Hainan and the other world famous tourist destination islands, for example, the Bali island of Phuket. Apparently the current situation is not enough to take the mission of constructing international tourism island[1]. Therefore, Hainan's great-leap-forward development of rural tourism is an inevitable strategy and plays an important role and practical significance to promote international tourism island construction.

1.1 Fostering the Support of Being Formats for Hainan International Tourism Island Construction

The essence of Hainan island international tourism development strategy is to realize the economic and social development of the whole province based on regarding tourism construction as pilot and breakthrough. Obviously, whether tourism can take the lead in development or not is the major step that affects Hainan international tourism island construction. As the organic component of tourism, the leapfrog development of Hainan rural tourism is not only the key points of the whole province, but also the basic part of the construction of Hainan international tourism island. Therefore, there

* Corresponding author.

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 309–316.
springerlink.com

are some measures that can support formats for Hainan international tourism island construction. For example, active exploration, first trial, integrating production factors including rural land relying on hainan unique natural and human country landscape, and actively promoting supernormal leapfrog development of hainan rural tourism.[2]

1.2 Providing Industrial Security That the Whole People Share the Construction Achievements of International Tourism Island

The construction of international tourism island is not only the commanding heights or the breakthrough, but also a difficult and complex system engineering that relates to the welfare of the whole province 8.5 million people in need. The government has made it clear that the hainan international tourism island construction will be a great significant project to enrich common people. Leapfrog development in hainan rural tourism can promote the countryside industrial structure adjustment, increase employment by widely absorbing local farmers, increase the farmers' income, improve farmers' life and rural situation. Therefore, leapfrog development in rural tourism can bring real benefits to the farmers.

1.3 The Important Point of Balancing the Urban and Rural Development and Realizing the Integration of Urban and Rural Areas

Through the Leapfrog Development of Hainan Rural Tourism, it will effectively connect the demand and supply of urban and rural, realize the overall coordination of urban ,rural and regional development, make the farmers based on its own reality realize agricultural modernization and the characteristic, and enhance the sustainable development of rural economy ability. It will create a new path to make the farmers locally move towards modernization without leaving home and soiland and balance the urban and rural development of the whole province.

1.4 Important Ways of Adjustment and Optimization of Rural Industrial Structure

Rural tourism great-leap-forward development of hainan can effectively help developing tourism services elements, promote agricultural and sideline products, the quality of the structure adjustment and the agricultural and subsidiary products scale, standardization, industrialization and pollution-free development, drive the development of rural hotel management, kind of animal breeding, agricultural and sideline products processing, transportation, construction and the cultural industry to repair, etc, optimize the industrial structure of rural and realize the traditional agriculture and tourism organic fusion and mutual development[4].

1.5 Effective Absorption of Rural Surplus Labor Force, Promoting Them Getting Rid of Poverty and Becoming Better Off

Rural tourism great-leap-forward development of hainan can effectively use the rural human resources, attracting old, weak, and the women vulnerable groups, such as employment realize farmers local employment, to enhance the agricultural added value near to increase farmers' income.

1.6 Promoting Protection and Improvement of the Rural Ecological Environment

Rural tourism Leapfrog development of hainan not only can make broad peasant masses recognize that well ecological environment, beautiful natural scenery and neat fields is also valuable homes production material and rich source, but also can make the farmers consciously improve the ecological environment and optimize the living environment, which is good to improve the rural ecological environment.

1.7 Effective Protection and Inheritance of Local Traditional Culture

Through fully using and tapping the traditional cultural resources of local country, it will naturally make these cultural resources get a higher degree of attention and more reasonable protective arrangement. Through the cultural exchanges with tourists, it can make local residents more fully understand and learn about native traditional culture, which will further improve the traditional culture protection consciousness and more attention of mining and finishing the traditional culture connotation, and then the strength of cultural spread will be Increased.

1.8 Training and Improving Farmers' Civilization Consciousness and the Modern Quality

In the process of hainan rural tourism leapfrog development, tourists' demand , Management and administration of tourism enterprise not only improve the accommodation health conditions to a great extent ,but also bring the knowledge of science and technology and advanced management idea to local villagers.It improves the management efficiency,promotes the farmers and modern civilization consciousness and promotes the improvement of the quality of the spiritual civilization construction and development.

2 Picture Deconstruction of Hainan Rural Tourism Great-Leap-Forward Development

2.1 Strategic Positioning of Hainan Rural Tourism Great-Leap-Forward Development

Based on the necessity and feasibility of hainan rural tourism great-leap-forward development, according to the basic position of tourist industry as pilot industry international tourism island construction in 《Hainan island international tourism development plan for the construction》 , considering village tourism industry is the industry to belong. Organic component, Hainan great-leap-forward development of rural tourism strategic positioning are as follows.

Pilot Tourism Formats of International Tourism Island Construction. Actively exploring, first try first, relying on hainan unique natural and human civilization country landscape, integrating industry elements including rural land elements, and actively promoting hainan rural tourism leapfrog and hypernormal development,

making Hainan rural tourism brand new, and strong pilot tourism formats in Hainan international tourism island construction and effectively sharing the regional development function of pilot tourism breakthrough.

Reform and Innovation Experimental Plot of Our Country Rural Tourism. Giving full play to the advantages of hainan special economic zone, active exploration, first test, and Playing the basic role of market allocation of resources ,speeding up the system and mechanism innovation, promoting hainan rural tourism and related modern service industry in reform, opening up and the scientific development going in the front of the whole country, making Hainan rural tourism reform and innovation experimental plot of our country rural tourism, providing experience for reference and example reference for rural tourism transformation and the great-leap-forward development of our country[4].

Asian First-Class, the World's Leading Seaisland Type Rural Tourism Destination. Giving full play to its own location and resources.According to the prevailing international standard service to promote rural tourism, rural tourism elements, to further improve the transformation and upgrading of rural tourism infrastructure and services and facilities, developing the characteristic rural tourism products, to regulate the market order of rural tourism, comprehensively promote hainan rural tourism management and service level. Making Hainan rural regional Asian first-class, the world's leading seaisland type rural tourism destination, providing regional tourism function support for realizing the grand goal of Hainan island being the world first-class leisure vacation island tourist.

2.2 Strategic Target of Hainan Rural Tourism Great-Leap-Forward Development

Based on the necessity and feasibility of hainan rural tourism great-leap-forward development, according to the basic position of tourist industry as pilot industry international tourism island construction in《Hainan island international tourism development plan for the construction》, considering village tourism industry is the industry to belong. Organic component, Hainan great-leap-forward development of rural tourism strategic targey are as follows[5].

By 2012, it will consume 2 years of time to lay solid foundation and optimization of the rural environment ,connect the deployment of international tourism island construction, plan and construct a group of key rural tourism infrastructure and characteristic rural tourism projects, and realize rural tourism market order, rural tourism service quality, the international tourism attraction, and social influence obvious improvement. Rural tourism receipts in the domestic and foreign tourists and total tourist income reached 4.75 million people time respectively and 4.8 billion yuan, the proportion of up to 15%.

By 2015, rural tourism management, marketing, service and product development of international market were significantly improve rural tourism, the scale of the industry, the quality and the benefits, obviously improve the rural tourism in rural economic growth and social development of the driving role further enhanced. Rural tourism receipts in the domestic and foreign tourists and total tourist income reached 9.52 million people time respectively and 10.8 billion yuan, the proportion of up to 20%.

By 2020, rural tourism service facilities, management and service level and the international rural tourism service standards fully connected, hainan rural tourism international reputation, reputation greatly improve rural tourism, the scale of the industry, the quality and the benefits reach the international advanced level, initially build Asia first-class, the world's leading HaiDaoXing rural tourism destination. Rural tourism receipts in the domestic and foreign tourists and tourism revenue of 23 million people time respectively and 37.2 billion yuan, the proportion of up to 30%.

3 Choices of the Tourism Business of Leapfrog Development of Hainan Rural Tourism

Mirroring the domestic and foreign advanced experience, and combining the actual situation of Hainan, thinking around strategic picture of Hainan's great-leap-forward development of rural tourism, including resources environment, the advantage location and the convenient traffic, Hainan can take some measures as the following on the base of local conditions to provide effective support formats for great-leap-forward development of Hainan's rural tourism.

3.1 Relying on Ecological Leisure Resorts- to Build the Original Ecological Leisure Vacation Manor

As to the town with the ecological environment located in the beautiful countryside, making good use of convenient traffic advantage,taking rural ecological landscape, cultural landscape and farmers produce life as the foundation, government can build rural tourism formats in the form of rural tourism activities, such as to inspire the family in reception unit to provide leisure vacation rural tourism products. Some original ecological leisure vacation manors around Haikou city and Sanya city have made great progress in tourism ecological leisure business.In fact, these resorts are advantage to create more original ecological leisure vacation manors. The key point in the process of building the original ecological leisure vacation manors is to make scientific planning and rational layout, to mine culture connotation, to prevent simple copy, and to Implement quality standards, strengthen the industry training and standardizing services.

3.2 Reling on High-End Scenic Spot –to Build Functional High-End Scenic Manor

As to the towns located in the famous scenic spot in the villages near distance,making good use of convenient traffic advantage, it should take part service functions to reduce the burdon of scenic spots, in the meantime, farmers are expected to participate in the process to provide visitors with tourism commodities and agricultural and sideline products, which is also the mature and successful formats of Hainan's rural tourism. The key point in the process of building functional high-end scenic manors is to strengthen the supporting infrastructure and public service facilities, and to strengthen the organization, guidance and training of farmers [6].

3.3 Relying on Tourism Amorous Feelings Towns-to Build Small Special Tourism Amorous Feelings Town

As to towns with rich tourist resources, it should create the organic combination of tourism development and urban construction. Obviously, tourism development can promote the urban construction, and urban construction can impel tourism development in turn. The model that taking the urban construction to driving tourism development is also one main form of Hainan rural tourism, which helps to create the world famous special tourism amorous feelings. The key point in the process to build small special tourism amorous feelings town is to make serious tourism planning and land use planning, and to increase input in the construction of tourist facilities in multi-channels.

3.4 Relying on Coconut Island and Li & Miao Nationality Flavor - to Build Original Li & Miao Nationality Flavor Villages

It is to point to in the rural areas, ethnic minorities with original culture landscape and unique national amorous feelings as the foundation, reasonable protection and development of tourist project, scientific and reasonable planning to build infrastructure and tourism facilities, promote the development of rural tourism rural tourism form formats. For hainan Li &Miao Nationality flavor of the trait, this format form is the most potential formats of hainan. It will be conducive to creating some original Li &Miao Nationality flavor villages, the implementation of point is scientific and reasonable planning, protection and development, emphasizing the original pay equal attention to natural and humanistic environment protection, training and guidance minority farmers' participation of tourism and service activities.

3.5 The Characteristic Industry Promoting Type- to Build the Characteristic Industrial Tourism Experience Base

It refers to that in certain rural regional scope, based on region unique advantage industry (tropical agriculture, fisheries, aquaculture, NaYao, etc.) of professional production and operation ,using special industry production process of knowledge, interest, participation and entertaining, sightseeing, leisure vacation development, experience and feeling, rural tourism products to promote rural tourism development of rural tourism form formats. The model helps to make some special industry tourism experience base, implement the point is to scientific and reasonable planning, pay attention to tourism and characteristic industry organically combined and integration, and pay attention to the development of independent innovation and creative planning and market research and market promotion, strengthen the propaganda promotion efforts.

3.6 The Ecological Civilization Demonstration Type - to Build Civilized Ecological Construction Achievements Exhibit Village

It refers to that in some parts of the economy developed in rural areas, adjusting measures to local conditions and making the best use of the circumstances to attract tourists to visit, to show the new socialist rural construction achievements and new rural scene,

developing sightseeing and leisure resorts tourism products promoting the development of rural tourism rural tourism form formats as attraction and unique selling proposition of entironment. In view of the primacy of the natural ecological environment of hainan superiority eco-villages civilization, This format form now has become hainan civilization ecological construction and independent and continuous development of important industrial support. It can be called Civilization eco-villages and rural tourism development of the organic combination of model. Through the key development civilization ecology villages YiTuoXing rural tourism format, it helps building some civilization achievements exhibit village of ecological construction. The point is to implement scientific and reasonable planning, strengthen properly handle the rural tourism development and other industry development of the relationship, positive training and guide peasants to participate in tourism and service activities, pay attention to tour operator and ecological environment protection consciousness of the tourists ascension, strengthen the ecological construction and environmental protectionand prevent Ecological degradation caused by rural tourism development.

3.7 Relying on QiongYa Red Village-to Build QiongYa Red Revolutionary Education Manor

As to towns with "red tourism" resources, it can combine red tourism development and sightseeing tourism activities, tourists drive the development of rural tourism rural tourism form formats, which model helps to make some QiongYa education manors. The key point in the process to build QiongYa red revolutionary education manor is to outstand the characteristics of ”red” lines, to give full play of tourism revolutionary traditional education function.

To sum up, under the background of the building of international tourism island, the leapfrog development of Hainan rural is a key stone that we can not evade, and it neither can not develop without choices of the tourism business and cultivate new kinds of businesses. Therefore, it ‘s especially important to mirror the domestic and foreign advanced experience, to combine the actual situation of Hainan, to think around strategic picture of Hainan’s great-leap-forward development of rural tourism, in the end, it will provide related tourism format support for the construction of international tourism island.

Acknowledgement. Financial supports from: the National Natural Science Foundation Project of China in 2009 (NO.40961005), the National Soft Science Item of China in 2010(NO.2010GXS5D252), the National Education Ministry Planning Project(NO.09YJC840008), University Research Project of Hainan Ministry of Education in 2010(NO.Hjsk2011-28), Hainan University Education Project Fund in 2010(NO.HDJY1008).

References

1. Wang, Y.: Policy experience and reference of international rural tourism development. Travel Journal 17(4), 45–50 (2002)
2. Li, S., Fan, S.: The study of land security of Leapfrog Development of Hainan Rural Tourism under the background of the building of international tourism island. Reserving and Finishing Trade Center of Hainan Land (2), 1–45 (2011)

3. Fan, S., Meng, F.: Analysis on urbanization regional differentiation characteristics about the counties of Hainan. China Collective Economy (2), 42–43 (2011)
4. The state council. The state council's some opinions on promoting the construction and development of international tourism in Hainan island (DB/OL) (December 31, 2009), http://www.gov.cn/zwgk/2010-01/04/content_1502531.htm
5. NDRC. Outline of the construction and development of international tourism in Hainan island (DB/OL) (2010-2020), http://wenku.baidu.com/view/241e094d2e3f5727a5e9622c.html (June 20, 2010)
6. Zhou, Y.: Discussion of the development of leisure agriculture under the background of international tourism island. Tropical Agricultural Sciences 30(4), 67–70 (2011)

Spatial Economics Research on Formation Mechanism of Commercial Circle Based on Consumer Searching Cost*

Juan Feng** and Jian-Wei Wu***

School of Economics and Management, Tongji University, Shanghai, 200092, China
juan_von1982@foxmail.com, jwwutz@gmail.com

Abstract. Based on Nelson's Searching Cost Saving related theory, through introducing concepts of Shopping Cost, Taste Searching Cost Ⅰ, Taste Searching Cost Ⅱ, and Location Searching Cost, explaining consumers' behavior of searching and shopping in depth, this paper constructs the Spatial Economic Model on Consumers Searching Expect within Commercial Circle. The results of the study show that: first, the satisfaction provided by Commercial Circle goods increases with the improvement of consumer demand and the reduction of Searching Cost for a new store within Commercial Circle; second, the satisfaction provided by Isolated Enterprise goods increases with the improvement of consumer demand, the increase in the numbers of enterprises within Commercial Circle and the reduction of Searching Cost for a new store within Commercial Circle; third, the influence of Location Searching Cost of goods and distance between Commercial Circle and Isolated Enterprise is decided by satisfaction comparison of goods provided respectively by Commercial Circle and Isolated Enterprise.

Keywords: Commercial Circle, Formation Mechanism, Searching Cost, Shopping Cost, Spatial Economics, Isolated Enterprise.

1 Introduction

After the development lead by agricultural and industrial economics, the business economics is stepping in as the main mode. With the Economic factors including capital, work forces, technology highly aggregated as well as the industry division deepened, Commercial Circle is becoming the new grow point in different communities, cities, regions and even countries as the most common mode of Commercial Agglomeration. The Commercial Circle Refers to a certain spatial range formed by gathering network of Commercial enterprises, generating interaction with

* Fund: Chinese Ministry of Education, Humanities and social science research projects "Theoretical Research on Spatial Size and Efficiency of Network City"(11YJA630150).

** Juan Feng(1981 –), female, born on Mizhi County, Shanxi province, PhD student inTongji University, focus on City Industry Plan and Commercial Studies ; Jianwei.

*** Wu (1959 –), male, born on Yongkang County, Zhejiang province, Professor of Tongji University, Dr. of ReadingUniversity, focus on Urban Economics and Industry Economics.

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 317–326.
springerlink.com

surrounding enterprises and consumers in certain special regions in the cities [1]. Inside the Commercial Circle, high density and specialization of business operation are the key features, and consequently high efficiency and high added value support the relatively expensive rents and operational cost[2]. Many researchers have studies the key factors in the formation and how to build up the Commercial Circle.

Harold Hotelling [3] presented the principles of minimum differentiation. He supposed that under the following hypothesis: in the finite linear homogenous market; consumer needs is the same and non-elastic; consumer is fully price-oriented; homogenous products are sold in all shops; transport cost keeps constant, then shops will ultimately chose to operate their business aggregately in order to optimize profits. Michael Webber [4] introduces uncertainty and risk lowering behaviour to Hotelling model, supposing a permanent location decision strategy, and the conclusion is generated that Commercial aggregation is necessarily formed in the centre of market. Avijit Ghosh [5] presents a modified central area theory and states that, behavior of shopping with multiple objectives would reduce cost of consumer travelling and safekeeping, so consumers tend to shop with multiple objectives. In the two-dimension homogenous market with homogenous customers, consumers' multi-objectives shopping behavior stimulates lower-class shops centers to higher-class ones, making both luxury and basic products are sold in these areas. Phillip Nelson [6] attributes Commercial aggregation to consumers' tendency to reduce searching cost.

Chinese researcher Lijun Lu[7] conducted longitudinal research to enterprise clusters at YiWu in Zhejiang Province. He supposes division economy, network external effect and increasing returns to scale are the key reasons for YiWu Commercial aggregation. Jizeng Peng [8] studies the problem from the perspective of Cournot Competition Model and spatial economics, and analyzed location cost advantages and location economic combination advantages. He believes that essential driving forces come from benefits generated by relation of interdependence highly overweighs the lost generated by relation of competition. This paper is based on Nelson's searching cost abatement theory together with Jizeng Peng's simulated market assumptions and analytical framework, describes consumers' searching and shopping behaviors, and builds up Economic Model on Consumers Searching Expect within Commercial Circle to illustrate Formation mechanism of Commercial Circle.

2 Concept Explanation

2.1 Shopping Cost

During the process of shopping, consumer needs to pay for the price of product as well as additional necessary expense including time, energy, risk, etc. We called these expenses Shopping Cost, defining it as the additional cost expect for price which consumer pays for in order to obtain certain utility [9][10]. From the perspective of economics, shopping cost is one of transaction cost, it includes searching cost, purchase implementation cost and risk bearing cost.In these three costs, the first two are the main part of shopping cost, because risk bearing cost is high unpredictable and shows positive relevance with former two parts.

2.2 Searching Cost

As stated above, the Shopping Cost mainly comes from searching cost and purchasing cost, the latter one consists of transport cost for commuting between shops and home. In this paper we define the searching cost as two parts: buying options searching is defined as taste searching cost, and transport cost is defined as location searching cost.

Searching Cost = Taste Searching Cost + Location Searching Cost

In this research we further divide Taste Searching Cost into type I and type II. Type I stands for the lost effectiveness caused by buying substitute product because consumer cannot find ideal product in the market (including Commercial Circle and isolated enterprise); Type II stands for the total cost of continuous action to look for ideal product conducted by consumer in Commercial Circle.

So, Searching Cost = Taste Searching Cost Type I + Taste Searching Cost Type 2 + Location Searching Cost.

3 Basic Ideas and Assumptions on Model Construction

Commercial circle is formed by the commercial agglomeration. And reason for the formation is due to the improved efficiency caused by spatial concentration of the business activities, and the resulting transaction cost savings and increase in income and effectiveness. The fundamental reason of the benefits is due to the division of labor and specialization within the Commercial Circle interests, the interests of economies of scale, scope and external economic interests, economic interests [11].

Many researchers have explored the proposition that the different species dislocation management within the district will be gathering economic effects, while the same kind of concentration of the industry will lead to competitive damages, resulting in the loss of customers [12] [13]. In reality, however it has been found that stores within the Commercial Circle with the same industry types gather operations often results in an increase in the amount of customers, in other words, that the increase in customers gained from gathering operations is greater than the loss in customers caused by the competitions. This may be due to species of the same industry within the district shop gathering operations, will reduce the costs for consumer tastes and location searching.

In this paper, more in-depth analysis is performed based on the analog market assumptions and analytical framework proposed by Chinese scholar Jizeng Peng [14]. In the market simulation, enterprise sells similar but not identical goods and consumers are uniformly distributed in a particular interval, customers are aware the type of merchandise sold on the market, but not knowing exactly the location and price that the product is sold at.

For convenience, we describe consumer tastes and corporate style using geographical feature space for consumer and enterprises. In the framework of the simulated market, different types of the same product distribute uniformly on a unit length of circle C, r_i is the location of i corporate in the feature space C, location r represents the ideal product location for a consumer. Therefore, a consumer can be defined by two parameters: their location x in the geographical space [0,1], and the ideal product location r in the feature space C. Assuming these two distributions

are independent, the consumer distribute (x, r) uniformly on the cylinder with the bottom C and a height of 1.

Thus, the taste searching cost could be expressed as $s|r - r_i|$, indicating the utility loss when consumer can only find goods from i corporation but not the ideal product, where s is the marginal utility loss of the consumer not getting the ideal product, $|r - r_i|$ is the gap between r and r_i along the relatively shorter arch, presenting the taste utility difference between i enterprise and the ideal goods in the feature space C. Assuming the non-convexity in transportation, and every time consumer does a separate searching a fixed cost t_0 is involved (with certain corresponding terminal conditions, e.g., time for parking and waiting for buses), y_i is the endogenous location of i enterprise, and $t|x - y_i|$ represents the location searching cost of a single one-direction search. If a consumer needs to know which product is sold in certain shop, the utility function for the consumer to visit buy product r_i from i enterprise on the feature space C can be defined as:

$$V_i(r, x) = U - p - s|r - r_i| - 2(t_0 + t|x - y_i|)$$

Where U represents the utility from consumer buying products, p is the price of general merchandise of enterprise, as consumers need to return after shopping, location searching cost doubles compared to that of one way searching. Although consumers can see the location y_i of enterprise i, they can not see the products r_i sold by the company. In order to see the products they sell, the consumer must visit enterprise i with the corresponding searching costs.

In summary, simulations were made with the following assumptions:

- Assumption 1: Store operations gathering with same product type in the business circle, could reduce tastes and location searching costs for consumers.
- Assumption 2: The only way for consumers to find out which store sell a particular product is to visit the store with the corresponding searching costs.
- Assumption 3: Typical consumers only know the size of district, but not the composition of district. Once consumers are at the Commercial Circle, the location searching costs will be reduced substantially due to effect of business scale.
- Assumption 4: Prices for the similar products searched by the consumers are fixed and there is no difference among stores. The only determining factor that affects consumers to whether continue searching is whether the quality of goods they encounter is the same with the ideal products.

4 District Formation Mechanism Analysis — Consumers Shopping Search Expectations Economics Model

A total of M enterprises are in the Simulated market, in which $M - 1$ enterprises are concentrated in the geographical layout of y_C, remaining alone M in the location y_1, of

which $y_1 \neq y_C$ and $|y_1 - y_C| = b > 0$. At this point, if we can prove the layout of company M in the Commercial Circle location y_C to get greater benefits ,then it can explain the reasonability and necessity of formation of Commercial Circle.

For consumer's part, there may have two possible plans for searching desirable merchandise: First, start searching one-by-one from those $M-1$ enterprises in the commercial circle until you find the most desirable one, or continue to search to the M separated company; Second, on the contrary, start searching from the M separated company until you find the most desirable one, or continue to search to the last enterprise in the Commercial Circle.

In each searching plan, consumers follow certain rules, described as Assuming 4, no difference in price between the goods, whether to a new store to continue searching based only on whether the product quality is consistent with the ideal product. However, given the continuing searching cost a (Cost with browsing one of the new businesses within the commercial circle, which is independent of the taste searching cost I $s|r - r_i|$ and location searching cost $t|x - y_i|$), Therefore, consumers can not have a Unrestricted search until finding the best one, but only search to a store where the benefits received is equal to the cost .

Typically, when consumers go to commercial circle or isolated enterprise for shopping, as long as the searching cost difference between the two locations is not too large, consumers will go to the commercial circle first. Assuming that consumers have browsed a number of shops in the commercial circle, the gap between the most desirable merchandise can be found and the ideal one is D, defined as the Taste Retention Gap of a number of shops browsed, $D \in C_1 \in C$, C_1 represents the Feature Space of enterprise visited in the commercial circles in C .Therefore, consumer's utility by continuing to visit a new store within the Commercial Circle is as follows:

$$V(D) = \int_0^D [(U - p - s\delta) - (U - p - sD)]d\delta$$

$$= \int_0^D (sD - s\delta)d\delta \text{ , among them, } \delta = |r - r_i|,\ \delta \in C, \delta \notin C_1 \text{。}$$

Then Taste Retention Gap D_C of the whole Commercial Circle is gotten by the following formula:

$$V(D) = \int_0^D (sD - s\delta)d\delta = a \text{ , then } D_C = \sqrt{2a/s} \text{。}$$

According to Hypothesis 2, consumers can find the Taste Retention Gap D_C of the whole Commercial Circle, only when searching every store in the Commercial Circle. In this process, consumers' Expected Utility continuing to search in the Commercial Circle consists of two parts: the first part describes Expected Utility Value that the merchandise which has been found and bought is the most desirable one in the Commercial Circle, but not the ideal one; the second part describes the Taste Searching Cost II consumed by searching the desirable one individually, expressed as A. Therefore, the consumers' Expected Utility by searching one by one in the Commercial Circle is as follows:

$$V_C(r,x) = U - p - sD_C - A$$

When consumers search only one shop in the Commercial Circle to find the most desirable goods, Searching Probability $P_1 = C_{M-1}^1$, here,there is no need to search for new stores, so the Searching Cost paid is 0; When consumers find the most desirable goods after searching for twice, at this moment, the consumers searched one new store, and they need to pay Searching Cost a;When the third search to the most desirable products, probability and cost are respectively $P_3 = C_{M-2}^1 C_{M-3}^1 C_{M-3}^1$ and $2a$; Similarly, When the k time to search, the probability and cost are $P_k = C_{M-2}^1 C_{M-3}^1 C_{M-4}^1 \cdots C_{M-k}^1 C_{M-k}^1$ and $(k-1)a$ .So,

$$\begin{aligned} A = {} & C_{M-1}^1 * 0 + C_{M-2}^1 C_{M-2}^1 * a + \\ & C_{M-2}^1 C_{M-3}^1 C_{M-3}^1 * 2a + C_{M-2}^1 C_{M-3}^1 C_{M-4}^1 C_{M-4}^1 * 3a + \cdots + \\ & C_{M-2}^1 C_{M-3}^1 C_{M-4}^1 \cdots C_{M-k}^1 C_{M-k}^1 * (k-1)a + \cdots + \\ & C_{M-2}^1 C_{M-3}^1 \cdots C_{M-(M-1)}^1 C_{M-(M-1)}^1 * (M-2)a \\ = {} & \frac{(k-1)a}{\sum_{k=1}^{M-1} P_{M-2}^{k-1} * (M-k)} \quad \text{,in most cased, } M \geq 3 \end{aligned}$$

A is relative to M and a, and it decreases with the increase of M ,and increases with the increase of a. When M is very large, the first two items are often considered only, that $A = \frac{a}{(M-2)^2} + \frac{2a}{(M-2)(M-3)^2}$.

Similarly, the gap between the goods provided by isolated shop and consumers' ideal goods is D_1, and the consumers'Expected Utility when shopping in isolated shop is as follows:

$$V_1(r,x) = U - p - sD_1$$

As consumers must overcome Transport Cost between the two locations when shopping from Commercial Circle to isolated shop, therefore, if and only if consumers' Expected Utility shopping from the isolated shop exceeds that from the Commercial Circle, and the excess part is enough (i.e. not less than) to compensate Transport Cost between Commercial Circle and isolated shop, consumers will be shopping from the isolated shop. It can be expressed as follows:

$$U - p - sD_C - A \leq U - p - sD_1 - (t_0 + bt), \text{ here } b = |y_1 - y_C| > 0$$

$(t_0 + bt)$ is Transport Cost between Commercial Circle and isolated shop. Then $s(D_C - D_1) + A \geq t_0 + bt$ here, $D_C > D_1$, means consumers can't find the most desirable goods in Commercial Circle, and those in isolated shop are more desirable than anyone in Commercial Circle.

On the contrary, when consumers go to isolated shop first, then go to Commercial Circle for shopping, as follows:

$$U - p - sD_1 \le U - p - sD_C - A - (t_0 + bt)$$

Then $s(D_1 - D_C) - A \ge t_0 + bt$,here $D_1 > D_C$, means consumers can have a chance to find the most desirable one only when going to Commercial Circle for shopping.

Whether $D_C < D_1$ or $D_C > D_1$ happens, D_1 are available through any crucial situation of any formula. When $D_C < D_1$ and $D_C = \sqrt{2a/s}$, then

$$D_1 = \frac{\sqrt{2as} + t_0 + bt + A}{s}$$

Here $A = \frac{a}{(M-2)^2} + \frac{2a}{(M-2)(M-3)^2}$

When $D_C > D_1$, $D_1 = \frac{\sqrt{2as} - t_0 - bt + A}{s}$

As is analyzed above:

(1) D_C decreases with the increase of s, but increases with the increase of a. Therefore, the marginal utility loss increases when it is not the ideal commodity that has been consumed, which results in consumers' higher quality demands for the commodity. Therefore, consumers are spurred to continue to search for more desirable commodity in the Commercial Circle; and the increase of Searching Cost for new stores within the Commercial Circle will hinder consumers' search, which ultimately result in the fact that consumers can't find the most desirable commodity.

(2) D_1 decreases with the increase of s and M, but increases with the increase of a, which indicates that isolated enterprise will have more motivation to provide consumers with more desirable commodity when consumers have higher quality demands or the number of enterprises in the Commercial Circle increases; Simultaneously isolated enterprise will be hindered to offer desirable commodity when Searching Cost for new stores in Commercial Circle increases.

(3) The influence of Transport Cost (t_0 and t --Location Searching Cost) and the distance b between Commercial Circle and isolated enterprise depends on the cases. If $D_C < D_1$, D_1 increases with the increase of t_0, t and b. If $D_C > D_1$, D_1 decreases with the increase of t_0, t and b. In fact, the dynamic changes of D_1 and b play a decisive role in whether the M enterprise can get the largest profits at the location of y_C. When $D_C < D_1$, the most desirable commodity is sold in the Commercial Circle. The reduction of distance between the isolated enterprise and the Commercial Circle force the isolated one to compete with those in the Commercial Circle, which spurs the isolated enterprise to improve commodity quality to meet the demands of consumers so that consumers' satisfaction can be improved. when $D_C > D_1$, the most desirable commodity is sold in the isolated enterprise, whether traffic conditions worsen, or the distance between the isolated enterprise and the Commercial Circle increases, consumers' satisfaction with isolated enterprise will not be influenced, instead, they will provide special commodity to consumers alone.

(4) In more general case, when M or b is large enough, every consumer will find there is a commodity in the Commercial Circle whose Taste Retention Gap is less than the ideal commodity's, which means there is no consumer to visit the isolated store.

5 Case Discussion -----Based on the Analysis of Xujiahui

5.1 Profile of Xujiahui Commercial Circle

Xujiahui Commercial Circle is located in the southwest of Shanghai, and is the sub-center and city-level commercial center in Shanghai. Xujiahui Commercial Circle covers a total area of 1,200-1,250 thousand square meters, its core area of 240,000 square meters. The Commercial Circle Integrated shopping, entertainment, dining, travel, finance, information and other functions into one, with a total construction area of 2 million square meters.

5.2 Analysis of Formation Mechanism of Xujiahui Commercial Circle

Before the 1980s, Xujiahui Commercial Circle was only the commercial network with "Lin's Store" type distribution, which had only a few thousand square meters. Until the early 1990s, with opportunities of major public constructions such as subway Line 1, Caoxi Road overpass and comparative advantages of the unique geographic, location, transportation, a long history, and human resources, Xujiahui Mall Group was set up to form the existing Xujiahui Commercial Circle, which includes Shanghai Sixth Department Store, Pacific Department Store, Zhongxing Department Store, Daqian Food Forest, Joint Commercial Exchange Store, West Asia Hotel, Huijin Plaza, Metro City, Meiluo Building and Gateway Plaza, except for East Shopping Center. Now Xujiahui Commercial Circle becomes one of city sub-center, the commercial center, the southwest Gathering Place for people flow, logistics, capital flow and information flow in Shanghai.

According to the constructed Spatial Economics Model on Consumers Searching Expect within Commercial Circle, the formation of Xujiahui Commercial Circle is not only due to the advantages of necessary preconditions in location, transportation, history and culture, but also due to the function for saving consumer' searching cost in Commercial Circle which is the main strength to gather a variety of resources flows. In fact, the convenience of location and transportation can also save consumer' searching cost ultimately. The analysis is as follows:

According to conclusion (2) and (3), D_C decreases with the increase of s, and increases with the increase of a. D_1 decreases with the increase of s and M, and increases with the increase of a. The increase of s indicates that it is not easy to meet the consumers' requirement with their taste improvement. Facing with increasingly demanding consumers, the Commercial Circle can meet the rising requirement of consumers only by constantly improving the quality of goods. The increase of searching cost for new stores in Commercial Circle will hinder the continual searching of consumers and correspondingly reduce consumers' satisfaction ultimately. The increase of M indicates that the more is the number of commercial

enterprises, the more fierce is the competition in Commercial Circle . And the Commercial Circle also has an incentive to improve the quality of goods, which leads to the improvement of consumers' satisfaction.

According to the conclusion (3),for the isolated enterprise 1, which run independently out of the Commercial Circle, when $D_C < D_1$, the most desirable commodity is sold in the Commercial Circle, and the reduction of t_0 (Location Searching Cost), t and b indicates more convenient transportation or less distance between the enterprise and Commercial Circle. At this moment, the isolated enterprise needs to compete with other enterprises in the Commercial Circle, which motivates it improve the quality of goods to get customers more satisfied. In the most extreme case, if the isolated enterprise is located at y_C in the Commercial Circle (the distance is reduced to 0), it can provide the most desirable goods with customers and gain the largest profits.

If $D_C > D_1$, the isolated enterprise which sells the most desirable goods will have less effect on the customers satisfaction even if the transportation is getting worse or the distance is enlarging, on the contrary, it motivates it to offer more desirable goods. If the distance is enlarged to infinity or such a degree that the customers find no way to compare the products, the isolated enterprise can get the customers the most satisfied and gain the largest profits. As a matter of fact, the two can both meet the demands of customers respectively without any competition if the distance is enlarged to some degree. So, it becomes a matter of course that the isolated enterprise gained the largest profits at the condition of $D_C > D_1$, such as the cases of some remote special stores.

As is mentioned above, the dynamics of D_1 (customers' satisfaction of the Mth enterprise) and b determines the location of the enterprise. It is the very reason that Xujiahui Commercial Circle can develop rapidly with its convenient transportation and location advantages. Represented by Xujiahui Mall Group, the rise of sales, profits, customers, mall rents and the number of commercial enterprises of Xujiahui Commercial Circle give further evidence to the formation and development of Xujiahui Commercial Circle.

6 Conclusions

This paper makes a detailed analysis and verification on formation mechanism of the Commercial Circle through construction and analysis of the Spatial Economics Model on Consumers Searching Expect within Commercial Circle and discussion the case of formation of Xujiahui Commercial Circle.

It should be noted that the Commercial Circle formation is often promoted by various forces, including government planning, history and human factors, market factors and other forces, or it is formed with one factor playing the decisive role and others playing supportive role. Regardless of the process of formation, consumers search for less cost is always the most essential reason. Therefore, the conclusion of this paper provides the government planning with policy basis, that is to say, the

government should be committed to efforts to reduce consumers' searching cost, including improvement of transport convenience, the provision of public facilities and recreational facilities and encouraging commercial enterprises to improve their management, adjust format, operate hierarchically and so on.

References

1. Feng, J., Wu, J.: Modeling Research of the Expanding Scale on Urban Commercial Circle. In: International Conference on Computer and Network Technology (2011)
2. Wu, J.: Major Planning: Business, Commerce, Services and Housing, vol. (5), p. 161. Chinese Building Industry Press, Beijing (2011)
3. Hotelling, H.: Stability in Competition. The Economic Journal (3), 41–57 (1929)
4. Webber, M.J.: Impact of Uncertainty on Location, pp. 52–53. M.I.T. Press, Cambridge (1972)
5. McLafferty, S.L., Ghosh, A.: Multipurpose Shopping and the Location of Retail Firms. Geographical Analysis 18(3), 215–226 (1986)
6. Nelson, P.: Information and Consumer Behavior. Journal of Political Economy 78(2), 311–329 (1970)
7. Lu, L.: Share-based Market, Local Specialization and Regional Economic convergence——Evidence from "Yiwu Commercial Circle". Statistical Research (4), 90–95 (2008)
8. Peng, J.: Commercial Cluster Analysis Based on Industry Cluster Theory: A Location Model Construction of Commercial Enterprises. Finance and Economics (11), 53–55 (2006)
9. Zhou, A.: Shopping Cost and Commercial Circle. Economics and Management of East China 17(2), 130–132 (2003)
10. Feng, Y.: Shopping Cost——A New Perspective on Defining Commercial Circle. Journal of Guizhou Commercial College 20(3), 38–41 (2007)
11. Zha, D., He, H., Wu, J.: Evolution of Urban Centers and Formation Mechanism of CBD. Journal of Chongqing University 26(3), 110–113 (2003)
12. Luo, J., Liu, Z., Jiang, M.: Correlation Analysis on Scale and Performance of Commercial Circle. Provincial CPC Party School Journal of Sichuan (4), 27–29 (2007)
13. Stahl, K.: Differentiated Products, Consumer Search, and Locational Oligopoly. The Journal of Industrial Economics 31(1/2), 97–113 (1982)
14. Peng, J.: Running Mechanism on Commercial Cluster Based on Spatial Theory. Reform of Economic System (4), 145–148 (2008)

Study on the Construction of Hainan Theme Park under the Background of International Tourism Island

Fan-hua Meng[1,a], Xiang-rui Chen[1,b], Shi-chen Fan[1,c], and Tao Song[2,d,*]

[1] Tourism College of Hainan University, Haikou, 570228, China
[2] Financial Department of Hainan University, Haikou, 570228, China
[a]gyhmfh1986@163.com, [b]251016412@qq.com, [c]fansc1230@163.com,
[d]songtao7221@163.com

Abstract. The construction of international tourism island is in great need of influential, high quality tourism projects, including the theme park to promote the development of Hainan. This paper put forward several suggestions in the construction of theme parks on the basis of reviewing the development of theme park at home and abroad and analysing current situation of Hainan.

Keywords: Theme park, International tourism island, current situation.

1 Overview on the Theme Park at Home and Abroad

1.1 Overview on the Foreign Theme Park

Several decades ago, the first theme park was born in the United States, and spreaded quickly in the following days. With progress and changing of the society, theme park has made constantly updating in the content and form in the past 60 years, roughly including four stages :"street casino - city garden - mechanical amusement park - theme park". With decades' development, now theme park has widespread all over the world, including Europe, Asia, Oceania, etc. Analysis on the region, there is no doubt that North America owns numerous large-scale and cut-throat competition theme parks, such as Disney, Universal Studios, Six Flags. Those famous brands have already formed oligarchs situation, ranking among the best of the top 20. In recent years, the growth of theme park in Europe and the Asia-Pacific region is rapid, especially the Asia-Pacific region. The Asia-Pacific region theme park mainly concentrated in Japan, South Korea and Hong Kong, China. [1]

1.2 Overview on the Domestic Theme Park

Our country theme park is mainly affected by the construction and operation of foreign theme park (especially the United States Disneyland), and is also on the influence and inspired of the building of base for the film and TV shooting in early 1980s. The establishment of shenzhen "splendid China" made the beginning of the theme park in China, our country came into a vigorous development period in the

* Corresponding author.

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 327–333.
springerlink.com

following years. Though many theme parks do not make success in China, new personality theme park are still been building and into commencement to attract leisure travelers acceptance. Only in no more than thirty years short development, though the gap between Chinese theme park and foreign ones is still wide, but build quality and management standards has greatly improve. In the meantime, Chinese theme park has experienced great transition from completely imitate foreign style to local design, from dependence on foreign technology to independent research and development.

According to not complete count, by 2010, China's various types of theme park is about more than 2500 or so, among them on A level scenic spots is 191 in total. Two theme parks are on the level of 1A, 39 theme parks are on the level of 2A, 44 theme parks are on the level of 3A, 93 theme parks are on the level of 4A, 13 theme parks are on the level of 5A. 116 theme parks locate in the eastern region in central park, 33 in the midland, the rest 42 in the western. [1]

On the analysis of spatial agglomeration, there are three main cluster center, they are: pearl river delta region, the Yangtze river delta region and Beijing-Tianjin-Tangshan area around the Bohai sea. On the analysis of spatial distribution, Chinese theme parks are mainly in the eastern coastal region, and many are in large scale. Several individual provinces such as Shandong province and Yunnan Province, in outstanding number, yet their performances can't compare with pearl river delta, Yangtze river delta area. This fact matches with Chinese economic development, on the other hand, also highlights the theme park development in need of certain economic base.

The Chinese theme park development of 30 years is also our country villages and towns urbanization process of 30 years, we can say that Chinese theme park construction has been integrated into the many places of the urbanization of steps. On one hand, the building of theme park is in an area of large scale, lots of money, in close to the suburban, what directly promotes the suburb of urbanization modification; At the same time, the theme park surrounding infrastructure and other service facilities also promote the urbanization. In the 1990s, development of theme park promoted the development of real estate industry, increased the local real estate value, stimulated real estate developers's enthusiasm on villages and towns. The construction of theme parks effectively change the local investment environment, promote land value, improve the living conditions, theme park surrounding area is often one of the price highest buildings among the local buildings.

2 Current Situation of Hainan Theme Park

As domestic and international famous tourist destination, contrasting to the achievement of Hainan's tourism industry is outstanding, the development of theme park is relatively slow. At the end of the 1990s, Hainan's began to try kinds of construction models of theme parks, which accumulated a certain amount of experience. That Times is a chaosworld in the history of Chinese theme park, for lack of bright sight and profound realization, many theme park had closed down before open.(Table 1)

Table 1. Construction of Hainan theme parks

NAME	OPENING TIME	ADDRESS	THEME	SCALE OF INVESTMENT (¥)	STATE OF OPERATION
Lingshan amusement park	1991	Haikou	amusement park	300 000 000	Closed in 1994
Nanshan Cultural Tourism Zone	1995	Sanya City	Buddhist culture	1 000 000 000	On business
Xinglong Tropical Botanical Garden	——	Wanning	tropical plants garden		On business
Dongshanhu animal park	1995	Haikou	the natural zoo imitated the wild concept	120 000 000	On business
Hainan village of the Chinese national culture	1997	Wuzhishan City	National park	130 000 000	Closed after business 8 months
Nantai crocodile lake	1999	ChengMai old ciity	the natural zoo imitated the wild concept	180 000 000	Closed in 2004
Tropical Ocean Park	2000	West coast of Haikou	Sea theme park	800 000 000	Closed in 2004
Love in Azure Sea	2005	Shenzhou peninsula of Wanning	Sea theme park	150 000 000	Closed before open
Ah norda rain forest tourism	2008	Baoting	Tropical rainforest theme park	300 000 000	On business

With the advancement of international tourism construction, Hainan island theme park is starting a new round of prosperity. In the next 10 years, Hainan will build ten large-scale theme parks. Hainan theme park construction is expected to usher a new opportunity time.

In the next 10 years, 10 top theme parks will become an important part in Hainan tourism system. Specific goal goes as the following:

Short-term Objectice(2011-2013). On the basis of optimizing, improving Nanshan religious culture theme park and Haikou volcanic site theme park, the government will newly build Wenchang space theme park, Lingshui Ocean Park, Haikou dynamic amusement park.

Medium-term Objectice(2013-2015). This period will continue to adjust, optimize, and improve existing theme parks, and complete Wenchang space theme park, Lingshui Ocean Park, Haikou dynamic amusement park. At the same time, it starts to construct the Sanya anime culture theme park, Lingao navigation technology theme park, Wuzhi-shan folk culture and tropical rainforest theme park.

Long-term Objection(2015-2020). This period will complete the construction of main 9 top theme parks, including new-built Ledong future world theme park, and start the construction of another Ocean park named “Soul of the South China Sea”. [2]

Table 2. Planning of Hainan theme parks in the next 10 years

NO.	NAME	ADDRESS	THEME	AREA (hectare)	SCALE OF TOURISTS (10000/year)
1	Lingshui Ocean Park	Lingshui	Rely on the ocean, making a cultural experience area	90-130	600-800
2	folk culture and tropical rainforest theme park	Wuzhishan	making comprehensive tourism experience paradise	50-100	300-400
3	space theme park	Wenchang	Aerospace science and technology culture tourism center	120	300-400
4	dynamic amusement park	Haikou	Multi-culture,leisure and fashion, amusement places	50	300-400
5	navigation technology theme park	Lingao	Navigation, education, entertainment	50	300-400
6	religious culture theme park	Sanya	Entertainment, religious culture	90-130	400-500
7	volcanic site theme park	Haikou	Volcanic ecological culture experience park	50	200-300
8	anime culture theme park	Sanya	Fashion park, collecting research	50	200-300
9	future world theme park	Ledong	Based on primeval forest, show the future Hainan in all degrees	50-100	300-400
10	Soul of the South China Sea	Xishan islands	Explore the mysterious south China sea	100-200	400-500

3 Theme Park and Construction of International Tourism Island

3.1 Sufficient Conditions to Develop Theme Park

Policy. To construct Hainan International tourism island has been a national strategy, for this goal, Beijing has come on many preferential policies to guard the project's smoomh going, so it is the construction of theme park. Because of the powerful economic radiation of theme parks, Hainan provincial government has also set theme park construction as one of the important projects. "The planning of Hainan's theme parks" points out the bright direction for the construction of Hainan's theme park, and puts forward the standard development system.

Passenger Flow Direction. From a global perspective, developed regions founded theme parks almost at the time that national per capita GDP was about $10000. At present, China is beginning to accord with the condition step by step: China is no longer just a country with large population, but own numerous high income population, and the size is considerable. Such as Beijing, Shanghai, Guangdong province, per capita GDP have come close to or over $10000 [3]. Hainan has become a famous domestic and foreign tourist destination, with huge tourism passenger flow, that provides a certain market foundation for theme park. In 2010, Hainan's tourism revenue is RMB 2.5763 billion, an increase of 21.7% year-on-year. The overnight travel has over 25.8734 million, an increase of 15.0% year-on-year, thereinto, 25.2103 million is domestic tourists (increased 14.8%), 663100 is inbound tourists (increased 20.2%). What's more, growing large-scale tourists are swarming into Hainan international tourism island. The huge volume is becoming a great advantage resource for the development of Hainan's theme parks.

Traffic Facilities. In the construction of the international tourism island, traffic network may the top priority. High Speed Rail, in the east of Hainan island, has opened the whole line, which all further enhances the Hainan's accessibility. Now it only takes one and a half hours if you trip from Haikou to Sanya. In of Hainan, the building of high speed rail has been put on the agenda, with the foundation of that, Hainan will make a inter-city life circle, tourists can visit the whole island in a half day. At the same time, Hainan's route to the outside world is under the construction. Haikou and Sanya, two big airports' in and out of tourists per year are in pole position in China. By the construction of excellent communications, Hainan can improve the accessibility further, which provide the good foundation for construction of Hainan theme parks.

3.2 Theme Parks Promote the Construction of the International Tourism Island

Theme Parks Promote Upgrade the Region Tourism Industry. At present Hainan tourism way has presented multilayered transformation, from Moderate amount dimensions to quality and efficiency; tourism products transform from sightseeing tour to leisure vacation travel; travel market structure changes from low to high. Hainan tourism industry has come into a stage of speeding development. Obviously, theme parks will improve Hainan tourism's transformation and upgrading. Firstly, theme parks can adjust the tourism resources configuration, optimize regional industrial structure; secondly, it can rich tourism product structure, speed up the

upgrading of the industry; finally, it also can help to form new regional industry chain, promote the region core competitiveness, and promotion the regional socio-economic comprehensive development [4].

Theme Parks Make for Rational Allocation in Tourism Resources. At present, the Hainan tourism market is basically rely on outlanders, tourism products are not abundant, which cause a series of bad results, such as clear short tourist season, short average staying time, not rational consumption structure, and so on. Theme parks are to alleviate the influence factors, to rich tourist holiday life, to prolong the visitors staying time. We hope to to expand the extensive tourism income and to reasonably allocate resource through the travel amusement projects pluralism, amusement area enlargement, and the tourism industry enrichment.

Theme Parks Promote the Construction of Cultural Industry. Each theme park intend to attract visitors through outstanding theme and creative stuff. Hainan planning and construction of the 10 theme park to Marine culture, tropical ecological theme, fashion shopping, modern science and technology, health and fitness regimen, move feeling amusement and so on six big theme for the leadership, take and colorful island, sports, the volcano remains, regional customs, culture art, rural interest and so on six big theme is characteristic, build hainan theme park product system. These topics covers most of hainan cultural resources. If successfully formed the theme park comprehensive system will be able to showcase to the tourists hainan excellent culture, also be helpful for hainan's outstanding culture protection and inheritance. Promote cultural industry construction is hainan international tourism island construction is the essence of international tourism is the meanings of the construction of the island, the construction of theme parks for this contribution will be due power.

4 Suggestions

4.1 Close to Market, Reasonable Development

In the short pass twenty years, Chinese theme parks have develop to more than 2 500, but no more than 10% are profitable firm, for the phenomenon of blind construction and serious follow the trend. Theme park is a commercial project, it needs to follow business rules, operate in the canon of economic. Before the construction of theme parks, we must make full market research, carefully study the main market capacity, consumption level, the needs of the customers and the competitive environment.

4.2 Stressed Characteristics, Superior Quality

Characteristics is the core of theme park, namely a theme park need connect with the nantive culture, show its difference and have a unique style. With the same blue sky, the similar local customs, it is inevitable that most towns own similar points, but there are still each other’s unique. According to local features and project planning reasonable arrangement, each theme park should present its theme to the full extent. Government ought to make full use of the existing superiority that tourism industry has accumulated, to realize that theme park meets the traditional tourism with no boundaries. On one side, theme parks need to highlight characteristics of Hainan with

closely integrating with Hainan tourism image positioning; on the other side, theme parks need to take Hainan's beautiful landscapes as its logo, in order to realize the organic combination of them[4].

4.3 Emphasize Marketing, Brand First

To seek greater development, theme parks must recognize the importance of marketing, established a set of perfect marketing model. Although with congenital advantage of natural resources, Hainan province is till facing the predicament that it just started a few years ago with weak foundation, low popularity and poor attraction, what's worse, the island's tourist market is too narrow, etc. We can make up these development disadvantages through good marketing. Therefore, the construction of the theme parks in hainan must be comprehensive and methodical, and do effective marketing.

5 Conclusion

Through reviewing and analysis on the domestic and international theme parks' construction and Hainan province's actual situation, I believe that construction of Hainan's theme parks will make success with the following points: taking culture as the theme, taking characteristics as the foundation, taking humanism as the core, taking product as the carrier, taking experience as the essence, taking transcendence as the form. Successful theme parks will change Hainan's currunt situation, make this tropical island more attractive and more dynamic.

Acknowledgement. Financial supports from: the National Natural Science Foundation Project of China in 2009 (NO.40961005), the National Soft Science Item of China in 2010(NO.2010GXS5D252), the National Education Ministry Planning Project(NO.09YJC840008), University Research Project of Hainan Ministry of Education in 2010(NO.Hjsk2011-28), Hainan University Education Project Fund in 2010(NO.HDJY1008).

References

1. Xu, C.: Research on the current situation and influence factors of theme parks in China, pp. 26–28. East China Normal University (2010)
2. Hainan government nets: "Planning of Hainan's theme parks"-construct 10 top themes park in the next ten years (EB/OL) (December 23, 2010), http://www.hainan.gov.cn/data/news/2010/12/119663/
3. Wei, S.: Planning:4 advantage regions to build theme parks in Hainan. Hainan Daily (January 18, 2010)
4. Wei, S.: Hainan can build 8 kinds of theme parks. Hainan Daily (January 18, 2010)

Human Library: A New Way of Tacit Knowledge Sharing

Yong-hua Zhai[1], Ying Zhao[1], and Rui-mei Wang[2]

[1] Library of HeBei United University, 063000 Tangshan, Hebei, China
[2] HeBei United University Affiliated Hospital, 063000 Tangshan, Hebei, China
Caliszyh@126.com

Abstract. The problem of tacit knowledge capture and sharing is a central theme in the field of knowledge management. The difficulties of tacit knowledge transfer and sharing not only for its own characteristics, but for provider's intellectual monopoly and the recipient's resentment. The Human Library is an innovative method designed to promote dialogue, reduce prejudices and encourage understanding. The paper aims to introduce the means of Human Library and its origin and development, points out the Human Library is a new way of tacit knowledge transfer and elaborates the advantage and role of human library in tacit knowledge transfer and sharing.

Keywords: Human Library, Human Book, tacit knowledge, knowledge transfer, knowledge sharing.

1 Introduce

Knowledge is well recognized as one of the most important sources for an organization or personal. Knowledge may be either of explicit or tacit nature, while explicit knowledge is easy to express both verbally and in written form, possible to codify, and therefore also easy to transfer; tacit knowledge is complex and hard to codify ,express or share, and is transferred through personal interactions and rich social interplay[1].

According to Pan and Scarbrough (1999 p362) "Tacit knowledge is not available as a text. . . . It involves intangible factors embedded in personal beliefs, experiences, and values."[2] The tacit knowledge is subconsciously understood and applied, difficult to articulate, developed from direct experience and action. Usually, shared through highly interactive conversation, story-telling and sharing of experiences[3]. How can we know people in our organizations who have the specific and specialized knowledge that we need to tap into to do our jobs more effectively? How can we capture tacit knowledge? The difficulties of tacit knowledge transfer not only for its own characteristics, but for provider's intellectual monopoly and the recipient's resentment. The paper aims to introduce the Means of Human Library and its Origin and development, then elaborates the Human Library as a new way of tacit knowledge transfer.

2 Human Library

2.1 What Is the Human Library

The Human Library is an innovative method designed to promote dialogue, reduce prejudices and encourage understanding. The main characteristics of the project are to

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 335–338.
springerlink.com

be found in its simplicity and positive approach. The Human Library works exactly like your local Library - readers can borrow a "book". The difference is, books in the Human Library are human beings: "Human books", with whom the reader can have a conversation. In its initial form the Human Library is a mobile library set up as a space for dialogue and interaction. Visitors to a Human Library are given the opportunity to speak informally with "people on loan"; this latter group being extremely varied in age, sex and cultural background. The Human Library enables groups to break stereotypes by challenging the most common prejudices in a positive and humorous manner. It is a concrete, easily transferable and affordable way of promoting tolerance and understanding[4].

2.2 Origin and Development of Human Library

There was a young and idealistic youth organization called "Stop The Violence" which founded by Dany Abergel and his friends from Copenhagen. The aim of the organization was to raise awareness and use peer group education to mobilise danish youngsters against violence. In 2000 Stop The Violence was encouraged by then festival director, Mr. Leif Skov, to organize human library for Roskilde Festival Events that would put focus on anti-violence, encourage dialogue and build relations among the festival visitors. And the Human Library was born with a total of 75 books available[5].

The conclusion made was that with so many different people, put together in a rather small space for a long time, they are bound to start reading each other. From the moment they ask the other book what their title is. And that will be the opening question of all books on the first day. The services of the Human Library has always been free to its public. Today a majority is hosted within the public library sector. Others are located in educational institutions, festivals, books fairs and other relevant settings.

Human Library appeared for the first time since 2000 in Denmark , rapidly spread in Europe, then spread to America, Australia, etc. According to the list of past human libraries on its official web site, there were 47countries have hosted the 257 Human libraries[6].

3 Human Library Could Be a New Way of Tacit Knowledge Sharing

3.1 Human Library Provide a Comfortable Environment for Tacit Knowledge Tranfer

Human library could be as a tool in promoting diversity awareness in a fun and relaxed way but like the Gurteen knowledge cafe it could be used much more widely to get people talking. Human library can make the tacit knowledge transfer smoothly in good communication with both sides talking in the good atmosphere. Reading activities held in safe, comfortable environment, both sides could sit in a quiet reading room, or walk in the park or near the playground. In the process of the Reading, volunteers can use props, such as veterans' war medal, athletes' running shoes and the Rangers' suitcase, to increase the attraction of dialogue and enrich its content.

3.2 Human Library Foster Independent Learning Skills through Asking Questions and Sharing Information

A human Book is a person, that has chosen to be a public representative of a certain group, may be experts in their field, have a particular passion or hobby, or have significant experiences to share. For everyone, this is an opportunity to learn from and better understand one another.

The human library held in TonJi University in china, provided several human books of the theme for occupation plan, scientific research, foreign language study and study abroad[7]. The human book volunteers were experts in some fields, scholars, members of the community, students and teachers. Students could learn the method of foreign language study through talking with a Korean Students who master three foreign languages, or read the book titled how to have a good life on university campuses to obtain some experiences of university life.

The Human book speak about their lives and experiences to interested users. So, for example, on the human library of the University of British Columbia, a user could sign up to talk to an Afghani student about life in Afghanistan, his wartime experiences and struggles. Someone else may want to speak to an esteemed female doctor who has been recognized for her exemplary research and her success in a male-dominated field. And yet another participant may want to talk to a writer about the creative process, and about the ways his views on identity, gender and sexuality inform his writing[8].

3.3 Human Library Break the Barriers between the Population through Tacit Knowledge Communication

Human Book stories are very diverse ranging from people from varied religious, social or ethnic backgrounds to people who have had interesting life experiences. By borrowing a Human Book you can get to know more about the people that make up our community. You might not normally get to interact with everyone in your community and the Human Library is a great way to meet different people and exchange ideas in a safe, friendly environment.

Human Library enables readers to ask the questions they have always wanted to ask; while books are given the opportunity to share stories about their personal experiences in a safe and welcoming environment. This enables individuals to actively contribute to building inclusive communities of understanding. People come from different backgrounds would talk with each other, such as the policeman sitting there speaking with the youth activist, the politician in discussions with the graffiti writer and the football fan in deep chat with the feminist. Human Library would promote understanding between people of different social backgrounds in communication, breaking the barriers between the population and increasing the social cohesion.

The idea is to introduce people to a collection of individuals who come from different walks of life and realities compared to their own. It's an attempt to promote diversity and share a broad range of views and opinions all within a context of respect and intellectual curiosity.

References

1. Flåten, B.-T., Shneor, R.: The role of social interaction in cross border tacit knowledge transfer: An Integrative Model. In: 34th European International Business Academy Annual Conference (2008)
2. Pan, S.L., Scarbrough, H.: Knowledge M The role of social interaction in cross border tacit knowledge transfer: management in Practice: An Exploratory Case Study of Buckman Labs. Technology Analysis and Strategic Management 11(3), 359–374 (1999)
3. Bhirud, S., Rodrigues, L., Desai, P.: Knowledge Sharing Practices In Km: A Case Study In Indian Software Subsidiary. Journal of Knowledge Management Practice 6 (2005)
4. What is the Human Library, http://humanlibrary.org/what-is-the-living-library.html
5. The history of Human Library, http://humanlibrary.org/the-history.html
6. A list of past Human Library events, http://humanlibrary.org/list-of-past-human-libraries.html
7. Zhai, Y.H.: Living library and its development in China. Chinese Journal Medical Library and Information Science 20(7), 28–30, 55 (2011)
8. The living library, http://www.ikebarberlearningcentre.ubc.ca/living

Strategies for Increasing the Key Capabilities of Organization Agility (Case Study: Tehran Urban and Suburban Railway Co.)

Mohammad Mahmoudi Maymand[1], Mohammad Zare[2], and Sedigheh Asefi[3]

[1] Member Faculty of Business Administration Department, Tehran Payam-e Nour University, Minicity, Artesh Blvd, Tehran, Iran
Drmahmoudim@yahoo.com
[2] MA in Executive Management, 15 km to Karaj, Karaj Special Road, Saipa Company, Tehran, Iran
zare@saipacorp.com
[3] MBA, Urban Rail Tehran
Aasefi_s@yahoo.com

Abstract. Due to the increasing changes, the organizations need a new form naming organizational agility to exploit the agility capabilities as a competitive privilege besides matching the environmental changes so that they can use the potential opportunities in a turbulent environment and can gain the superior position. Based on a survey on Tehran Urban and suburban Railway Corporation, and using Vokurka & Fliedner's Sand Cone Model, this study tries to investigate the relations between the improving factors of key capabilities of organizational agility on one hand, and the key capabilities of or organization agility on the other hand. The results confirm the relations between reliability, agility, and key capabilities of the organization; moreover, the results determine the effective indices on improving the key capabilities of organization agility and show their priorities.

Keywords: Agile Company, Agility Key Capabilities, Competitive Privilege, Tehran Urban & Suburban Railway Co.

1 Introduction

The emergence of new commercial period with the change as one of its main characteristics has led to essential review in commercial priorities, strategic viewpoint, and methods used for organization's survival [14]. Today, they are more emphasis on consistency with the changes of business environment. One of the modern methods for facing the challenges of changes is the organization agility which moves the organization forward to reach a better performance and being successful in business [3]. So, the agile system is a set of abilities and competences which cause the survival and improvement of any agency in a competitive environment [4]. So this study is going to provide a means to measure the effectiveness and the importance of the factors that improve the key capabilities of organization agility.

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 339–346.
springerlink.com

2 Literature Review

The concept of "agility manufacturing" (IS) evolved after 3 periods of change in manufacturing industry. These changes came from manual manufacturing to mass production, then to pure manufacturing, and now in current period, to agility manufacturing. The first period can be characterized as the very low level of production, decentralized organization, and very high prices of the product. At the second period, due to the high prices of the products, the mass producer avoids the innovations and creativity as much as possible. The philosophy of the third period was the pure production or manufacturing at a global level [16, 17]. The aim of the pure production was to fill the gaps in the philosophy of mass production. Consequently, in the third period, the interests of the customers and respecting their rights were paid especial attention.

At the beginning years of 1990s, pure manufacture or World Class Manufacture (WCM) was a modern category in manufacturing management which had been proposed to answer the increasingly growth of competitive factors. These factors went beyond the price and quality because the factors like interest in new products, on-time delivery, and being flexible to compete on an international level got more important [5, 6, 7, 8]. The fourth period, i.e. agility manufacturing is the newest manufacturing method whose aim is to satisfy the customers, to increase the effects of information and individuals in organizations, to dominate over the changes and instabilities, and to increase the competitiveness of competitions via cooperation [9]. The new business environment has changed the competitive priorities and operational strategies of the organization. In the new environment, competitive prices and high qualities are necessary but these factors are not enough to ensure the business success; instead, the speed toward reaching the market and prompt and flexible reply to the customers are being regarded as a main principle [10]. Therefore, the organizations can reach the increasing and stable development in their strategic competitive capabilities by reinforcing and equipping their abilities [16].

2.1 The Key Abilities and Capabilities of Organization Agility

Agility capabilities are the abilities to be established in organization in a way that the organization find the needed power to be responsible against the changes. These capabilities which are regarded as the bases for maintaining and developing the agility are as follow [2]:

a) *Responsibility power*: the ability of realizing the changes and prompt reply to them.
b) *Sufficiency*: a set of wide range of abilities which support all the activities toward the organization's goals.
c) *Flexibility*: the ability of manufacturing and presenting different products and reaching to different goals with the same resources and equipments.
d) *Speed*: the ability of doing the operation at the shortest possible time.

2.2 Conceptual Model of Improving the Key Capabilities of Organization Agility (Sand Cone Model)

Sand cone model investigate several issues. Firstly it approaches to the changes in competitive environment and the strategic responses to these changes. Secondly it tests the suggested strategic priorities and presents the priorities of such capabilities. Thirdly it provides a format for accessing the priorities of competitive program in supply chain. Sand cone model presents a complete framework for improving the constant agile supply chain and process by which the strategic competitive privilege of the organization will be gained. The model suggests that the organizations have to focus on collective attempts to reach the quality, reliability, agility, and finally, effectiveness and economization of the costs. Sand cone model offers a theory which shows how the organizations make a goal for each of their priorities in their program [11].

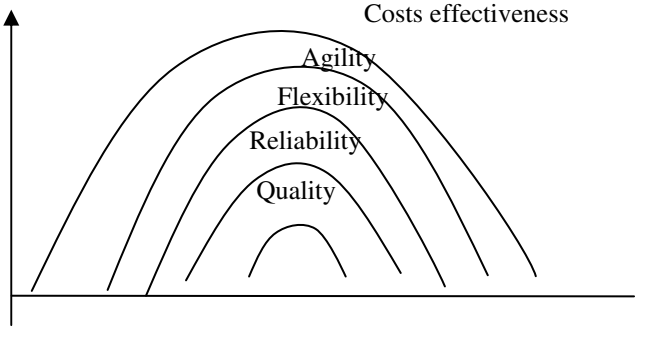


Fig. 1. Sand cone conceptual model

3 Research Method

This study has been done on the base of Sand Cone Model and has used the survey method to investigate the relations between 5 factors which improve the key capabilities of agility on the one hand, and the key capabilities of organization agility on the other hand. In this regard, the authors made a questionnaire containing the mentioned variables and used it in Tehran Urban and Suburban Railways Co. The samples were 187 employees of Tehran Urban and Suburban Railways Co. The samples were determined with Simple Random Sampling to a final sample of 123 employees. Moreover, the authors used AHP approach to determine the measure of importance of the factors which improve the key capabilities of organization agility.

4 Hypothesis of the Research

There is a significant relationship between improving the key capabilities of organization agility (i.e. quality, reliability, agility, and costs effectiveness) and improvement of key capabilities of organization agility.

5 Reliability of Measurement

The validity of measures of the questionnaire is confirmed by the applied definitions and the comments of the respected experts. The results of Cronbach alpha coefficient confirm the reliability of the questionnaire.

6 Results of the Research

In order to investigate the accuracy of the research's hypothesis, data obtained from organization's experts were collected for static analysis. Results of the student's t-test shows that the significant number in all hypothesis is nearly zero, which is lower than the level of significance (0.05); so the research's hypothesis have been proved by confidence of 95%. (Table 2.).

Table 2. Results of statistical analysis of the research

Variable name	Test statistic	Test result	Research result
Quality	25.228	Null hypothesis rejected	Hypotheses confirmed
Flexibility	3.915	Null hypothesis rejected	
Reliability	1.326	Null hypothesis rejected	
Agility	0.389	Null hypothesis rejected	
Costs effectiveness	25.228	Null hypothesis rejected	

Therefore it can be concluded that there exist a significant relation between five improving factors and key capabilities of Tehran Urban and Suburban Railway organization agility. Results of the Friedman variance analysis related to research variables are as follows.

Table 3. Ranking average and priorities of the research variables

Index name	Items	Average	Priority
Reliability	Offering especial services on holidays and especial days	3.67	First
	Congruence of the variety of provided services and customer's needs		
	Distributing metro stations per the different urban zones		
	Offering different new services by the organization to customers		
Agility	The ability of recognizing the changes in due time by the organization	3.27	Second
	The ability of respond quickly to the environmental changes		
	Cooperation network among the organization and other organizations		

Table 3. (*continued*)

	Using creative solutions for providing services		
Flexibility	Doing the services on promised time	2.84	Third
	Easy accessibility to the services		
	Offering compensating and alternative services when a problem emerges		
Quality	Making a plausive environment for the customers of the organization	2.66	Fourth
	Suitable dealing with the customers		
	Attempt to increase the customers' trust for their further usage of the services		
	Being obligated to the rules and principles in order to prohibit any discrimination among the customers.		
	Employees' responsibility for their probable mistakes		
	The ability of the employees to solving the problems of the customers and offering suitable solutions to them		
Cost effectiveness	The ability of performing the processes for executing the goals of the organization	2.57	Fifth
	The capacity of maximizing the services		
	Optimal synthesis of the factors of service providing with regard to their prices		
	The ability of reaching the specified goals with lesser costs		

Comparing the obtained mean in table shows that the reliability and agility components in rail-way organization can be assumed as the most effective factors on key capabilities organization agility. Moreover cost's effectiveness factor is the least related to the increase of key capability organization agility.

On the other hand, couple comparing matrix table (AHP) according to some experts were applied, in order to determine the importance rate of any five improving key capabilities agility variable in their relation to key capability organization agility. Results are shown in Fig.2.

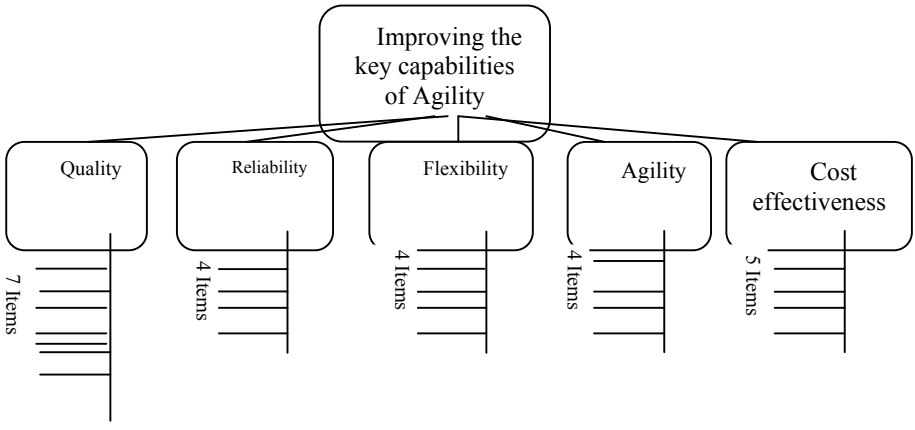


Fig. 2. AHP tree diagram of improving the key capabilities of agility

As table 4 shows, the importance of five improving key capabilities of Tehran Urban & Suburban Railway Co. agility is ordering as: reliability, agility, quality, flexibility and cost's effectiveness.

Table 4. The priorities of the independent variables according to their weighting importance

#	Variable name
1	Reliability
2	Agility
3	Quality
4	Flexibility
5	Cost effectiveness

Then the score of every variable in relation to all characteristics of improving key capabilities agility has been obtained afterwards. The mean of the obtained matrices in experts' poll were calculated in order to obtain the score of any variables characteristics:

$$C_i = [W_V \times A]$$

Where:

C_i : The weighted score of the dependent variables

W_V : Weights of independent variables

A : weighted mean of variables

Weighted score of any dependent variable related to independent variables are calculated as follows and obtained weighted score have been multiplied by 100 for similarity.

Table 5. Weighted score of each independent variables in relation to dependent variables

#	Variable name	Weighted score of each independent variables in relation to dependent variables	Weighting score *100
1	Quality	0.0298	2.98
		0.0266	2.66
		0.0294	2.94
		0.0288	2.88
		0.0292	2.92
		0.0276	2.76
		0.0286	2.86
2	Flexibility	0.03567	3.5
		0.03554	3.5
		0.03596	3.5
		0.03683	3.6
3	Reliability	0.0744	7.4
		0.0738	7.3
		0.0727	7.2
		0.0721	7.2
4	Agility	0.057	5.7
		0.0577	5.7
		0.058	5.8
		0.0575	5.7
5	Cost effectiveness	0.0264	2.6
		0.0267	2.6
		0.0277	2.7
		0.026	2.6
		0.0252	2.5

Since the obtained weighted scores were calculated according to the weighted importance of the factors effective in increasing the key capabilities organization agility and also according to the mentioned approach (AHP); therefore all results are determined and ordered based on them which have been investigated and analyzed as follows.

7 Conclusions and Suggestions

Reliability factor is introduced as the most relevant factor to the improvement of key capabilities of organization agility. Among the studied items, "Doing the services on promised time" with weighted score of 7.4 and "suitable time of customer's wait to receive the services" (7.3) have been the most relevant factors to the improvement of key capabilities of organization agility of Tehran Urban & Suburban Railway Co. Out of the reliability, the agility factor is the most important and most relevant factor to the improvement of key capabilities of organization agility of Tehran Urban & Suburban Railway Co. The results show that the item of "Cooperation network among

the organization and other organizations" (weighted score 5.8) is the most relevant factor to the agility among all the studied factors.

The results of the present research show the strong relation between the factors of improving the key capabilities of the organization agility on one hand, and the key capabilities of organization agility of Tehran Urban & Suburban Railway Co. on the other hand. This issue has been especially proved for the reliability and agility factors. Moreover, the AHP approach of the research also confirms the results. Thus the organization can improve its key capabilities of agility by reinforcing and promoting these relevant factors.

References

1. Jafarnejad, A., Shahaee, B.: An introduction to Organizational Agility and Agile Production. Mehraban Publish Institute (2007)
2. Rajabzadeh, A., Shahaee, B.: Presenting an Evaluative Model for Agility in Governmental Organization; A Glance to the Effects of IT on Organizational Agility. In: Second International Conference on Management of ICT (2005)
3. Khoshsima, G.H., Jafarnejad, A.: Agile Production System; Frameworks and Abilities. In: Second International Conference on Industrial Engineering (2001)
4. Khoshsima, G.H.: Presenting a Model for Measuring the Agility of Productive Organizations in Iranian Electronics Industry Based on Fuzzy Logics. Tehran University, Management Department (2001)
5. Hormozi, A.M.: Agile manufacturing: The next logical step. BIJ 8(2), 132–143 (2001)
6. Kidd, P.T.: " Agile enterprise strategy" A NGM concept. IEEE (1997)
7. Duguacy, C.R., Landy, S., Pasin, F.: From mass production to flexible/agile production. IJOPM 17(12), 1183–1195 (1997)
8. Gunasekaran, A.: Agile manufacturing: Enablers and an implementation framework. IJOPR 36(5), 1223–1247 (1998)
9. Kidd, P.T.: Agile Manufacturing: Forging New Frontiers. Addison-Wesley, Reading (1994)
10. Greis, N.P., Kasarda, J.D.: Enterprise Logistics in the Information Era. California Management Review 39(3), 55–78 (1997)
11. Vokurka, R., Fliedner, G.: The journey toward agility. Industrial Management & Data Systems 98(4), 165–171 (1998)
12. Ferdows, K., DeMeyer, A.: Lasting improvements in manufacturing performance: in search of a new theory. Journal of Operations Management 9(2), 168–183 (1990)
13. APICS Dictionary 36 (1998)
14. High, R., Mills, S., Jensen, C.T.: Achieving business agility with BPM and SOA together: Smart work in the smart enterprise (2009)
15. Pisapia, J.: The strategic leader. New tactics for a globalizing world. Information Age Publishers, Charlotte (2009)
16. Doz, Y., Kosonen, M.: Fast Strategy: How Strategic Agility Will Help You Stay Ahead of the Game. Wharton School Publishing, Philadelphia (2008)
17. Varghese, K.C.: Do multinationals really understand globalization? Bloomsburg Business Week (retrieved August 2, 2010), http://app.businessweek.com/UserComments/combo_review?Action=all&style=wide&productId=108280&productCode=spec

Using Fuzzy Delphi Method and Fuzzy AHP for Evaluation Structure of the Appeal of Taiwan's Coastal Wetlands Ecotourism

Chien-Chi Lin[1] and Laurence Zsu-Hsin Chuang[2]

[1] Ph.D. Candidate, Institute of Ocean Technology and Marine Affairs, National Cheng Kung University. No.1, University Road, Tainan City 701, Taiwan
[2] Asistant Professor, Institute of Ocean Technology and Marine Affairs, National Cheng Kung University. No.1, University Road, Tainan City 701, Taiwan

Abstract. Coastal wetlands are important resources. Many developed countries have actively promoted coastal wetlands ecotourism in order for the public to understand the importance of coastal wetlands. The key to the successful promotion of coastal wetlands ecotourism lies in finding out their appeal. In order to achieve this research purpose, this paper is divided into several stages. First, through literature review and expert interviews, coastal wetlands ecotourism related appeal candidate factors were collected. Then, through the Fuzzy Delphi Method, the important factors were filtered. Finally, through the Fuzzy Analytic Hierarchy Process, the relative weights of the related factors were estimated, and a hierarchical structure also can be established to evaluate appeal on coastal wetlands ecotourism.

Keywords: Coastal wetlands, ecotourism, Fuzzy Delphi Method, Fuzzy Analytic Hierarchy Process.

1 Introduction

The Ramsar Convention is an intergovernmental treaty adopted in 1971. Under the terms of the Convention, wetlands are defined as "areas of marsh, fen, peatland or water, whether natural or artificial, permanent or temporary, with water that is static or flowing, fresh, brackish or salt, including areas of marine water the depth of which at low tide does not exceed six meters". According to the definition of wetlands from the Convention, wetlands comprise inland and coastal/ marine wetlands, and human-made wetlands [1].

In the past, coastal wetlands had often been regarded as useless mudflats by many countries, thus the large-scale development of the agricultural, industrial, or urban lands [2]. According to statistics, over the past few decades, about 30%~50% of the coastal wetlands around the world have degraded, human destruction being the main reason other than climate change [3]. In fact, coastal wetlands are important natural resources, some of the many functions include: groundwater recharge, water purification, climate regulation, breeding of fish and shellfish, the development of ecotourism, and so on [4].

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 347–358.
springerlink.com

As the functions of these coastal wetlands are gradually being taken seriously, in addition to legislation banning against coastal wetlands conversions, many developed countries such as Britain, the Netherlands, and the United States have actively promoted coastal wetlands ecotourism, hoping to preserve the coastal wetlands through ecotourism development that have the least impact on the local residents and environment while enhancing the general public's awareness of the importance of wetlands and promoting the local economy.

Taiwan is surrounded by water. The formation of many coastal wetlands is due to the natural features, but the general public in Taiwan has little knowledge of coastal wetlands, and there are few ecotourism attractions developed from the wetlands. As appeal is the main driving force of ecotourism promotions, and in consideration that coastal wetlands ecotourism that still requires more active efforts in Taiwan, in addition to considering the ecological/environmental impact of its development on the local residents, the coastal wetlands ecotourism appeal factors are worthy of further study. The purpose of this study is to find out the coastal wetlands ecotourism appeal factors and their respective weights, and to establish a hierarchical structure for evaluating appeal on coastal wetlands ecotourism.

2 Literature Review

2.1 Sustainable Development and Ecotourism

The concept of "Sustainable development" began to sprout from the 1960s, but was not noted by many people in the early phase. In 1962, the book "Silent Spring" indicated that water, soil, plants, and animals are closely related to humans, so pursuing only economic development will endanger the environment of human life. In 1972, the Club of Rome that specialized in the study of global issues published a book "The Limits to Growth" that used the System Dynamics model to simulate the interaction between human and the global environment. It found that the human behavior to pursue economic growth damaged the balance of nature, and called human beings to face the subject of the limits to growth with the attitude of sustainable development. In 1980, the International Union for Conservation of Nature (IUCN), United Nations Environment Programme (UNEP), and World Wildlife Foundation jointly published the "World Conservation Strategy," pointing out that "Man and the biosphere are closely interlinked, and only by conservation of resources can humans achieve sustainable development." In 1987, the "United Nations World Commission on Environment and Development (WCED)" published a report "Our Common Future," pointing out that "Sustainable development meets the contemporary needs without damaging future generations' ability to pursue the development of their own needs."

Jacobs and Sadler (1992) [5] pointed out that economic objectives, social objectives, and environmental objectives of the three pillars of sustainable development. In terms of economic objectives: sustainable development encourages economic growth; in social objectives, sustainable development improves and enhances the quality of life, and solves the social problem of poverty; in environmental objectives, sustainable development is based on environmental protection, to make human development not exceed the carrying capacity of the Earth's environment.

Ecotourism is a management tool that seeks sustainable development of natural resources in the economy, society, and environment. In the early days, the natural resources was managed by establishing reserves, vacating residents in the reserves, and prohibiting local residents and general public from entering the reserves. This type of natural resource management obviously deprived the local residents of their rights, making it difficult for the public to understand the importance of the natural resources. After years of consideration, conservationists proposed the concept of ecotourism, thereby enabling the public to understand the importance of the environment and bring economic benefits to local residents.

The Mexican ecologist Hetzer (1965) introduced the term "ecotourism" and identified four normative principles. According to Hetzer ecotourism should have (1) minimum environmental impact, (2)minimum impact on and maximum respect for host cultures, (3)maximum economic benefits to the host country's grassroots, and (4) maximum recreational satisfaction to participating tourism [6]. Ceballos-Lascurain (1996) defined it as 'tourism that consists in travelling to relatively undisturbed or uncontaminated natural areas with the specific objective of studying, admiring, and enjoying the scenery and its wild plants and animals, as well as any existing cultural manifestations (both past and present) found in these areas [7]' The Taiwan Sustainable Development Land group ecotourism White Paper defined ecotourism as a form of tourism in a natural region, emphasizing the concept of ecological conservation, and sustainable development as the ultimate goal. Ecotourism to meet this definition must through the explanation lead visitors to understand and appreciate the local special natural and cultural environment, provide environmental education to enhance the environmental awareness of tourists that leads to responsible environmental action, and feedback economic interests to the visited places [8].

Taking all these scholars' and organization's definitions of eco-tourism, we can summarize the eco-tourism to include:

(1) Tourism to enjoy the natural resources;
(2) By the process of activities to inspire visitors attention to environmental protection;
(3) To maintain the welfare of local residents and promote the development of local communities; and
(4) Minimize the impact on the environment.

Therefore, ecotourism is not simply recreational activities to enjoy wildlife, it also reaches the functions of promotion of local economic development, protection of local residents, conservation of natural resources, and so on, namely to keep economic, social and environmental sustainable development.

2.2 Coastal Wetland Ecotourism Appeal Factors

Although there are few literatures in the past on the "coastal wetlands" type of ecotourism appeal, there are numerous literatures on the analysis of the appeal of generalized types of other specific types of ecotourism, as summarized below:

Deng (2002) constructed the assessment framework of the appeal of natural reserve ecotourism in Australia, which consists of several levels. The first level includes five factors: tourism resources, tourist facilities, accessibility, local communities, and appeal of the surroundings. "Tourism resources" includes two factors: natural resources and cultural resources; the "natural resources" includes two factors: "geography" and "environment"; "geography" includes five factors: climate, water, mountains, animals, and plants; "animals" and "plants" include four items: scientific value, aesthetics, diversity, and scarcity. In addition, the tourist facilities include three sub-factors: educational facilities, recreational facilities, and infrastructure; the local communities include three factors: social impact, economic impact, and cultural impact; the appeal of the surroundings include two sub-factors: well-known and quantity [9].

The International Union for Conservation of Nature (IUCN), World Tourism Organization (WTO), and UN Environment Programme (UNEP) set up a set of guidelines for the development of nature reserves in 1992. The guidelines include a list of nature reserve site selections, including: accessibility, star species, beautiful landscapes, special wildlife observation methods, the probability of successfully spotting wildlife, food and accommodation services in the surroundings, other cultural heritage, and other tourist activities [10].

Chen & Chen(2006) studied the factors affecting the appeal of the Taiwanese black-faced spoonbill reserve for tourists, including: noise, waste, water pollution, air pollution, the crowds, tourists' waiting time, the number of parking lots, the placard setting, facilities and environmental harmony, the ability of the guides, and the number of guides [11]. Liu (2007) assessed the sustainable development of Taiwan's mountain ecotourism and pointed out the factors that attract tourists, including: low proportions of artificial facilities, media commentary, guides, and other nature related experiences [12].

Reinius and Fredman (2007) studied Swiss nature reserves and found that the reasons that attract tourists to visit the nature reserves include: landscaping, less waste, serenity, not too crowded, convenient transportation, and special artifacts [13]. Ethos Consulting (1991) studied the ecotourism of British Columbia, Canada and found that the appeal factors include: accessibility, topography, climate, plants, trees, minerals, fish, wildlife, landscape, and cultural heritage [14]. Arabatzis and Grigoroudis (2010) studied Dadia National Park in Greece and found that the important appeal or the factors affecting tourists satisfaction include: good service and attitude of the management, unique flora and fauna, unique landscape, pavement, accommodation facilities, tourist centers, bird-watching pavilions, viewing platforms, toilets, map indicators, and so on [15].

Yeh (2001) studied Taiwan's Penghu Gibe sea sports tourism and its appeal to tourists and found that clean environment, clear water, and beautiful, natural environment are the main appeal factors [16]. Hu (2008) studied the appeal of Taiwan's tourist attraction the Bali River and found that beautiful natural landscapes, ecological diversity, historic monuments, temples, squares, streets, fishing, and so on are the important indicators for appeal [17]. Wang (2009) studied the boat activities held in the Four Grass Wetlands in Taiwan and found that the important appeal factors include: natural landscape, the quality of guides, boat safety, transportation convenience, placard settings, and so on [18].

Table 1. The ecotourism appeal factors compiled from relevant literatures

Literatures / Appeal factors	A	B	C	D	E	F	G	H	I	J	K	L	M
1. With star species (such as endangered species)		✓					✓					✓	✓
2. Easily able to observe diversified and rich organisms	✓	✓				✓	✓		✓				
3. With well-known historic sites	✓	✓			✓	✓			✓			✓	✓
4. Low proportion of artificial facilities		✓		✓									
5. Beautiful or unique natural landscape	✓				✓	✓	✓	✓	✓	✓			
6. Environmental cleanliness (Including little garbage and good water and air quality)	✓		✓		✓			✓			✓		
7. Adequate infrastructure capacity	✓						✓						
8. Adequate safety facilities	✓									✓			
9. Convenient transportation	✓		✓		✓	✓				✓	✓		
10. Adequate parking lots			✓							✓			
11. Adequate public toilets			✓				✓						
12. Low degree of crowdedness perceived by tourists			✓		✓								
13. Short equipment (service) waiting time for tourists					✓								
14. Excellent management and service capabilities of management units							✓						
15. The area planned as reserves or national parks											✓		
16. Low impact of tourist activities on community	✓			✓									
17. Low impact of tourist activities on animals and plants	✓			✓									
18. With interpretation personnel			✓	✓						✓			
19. With appropriate interpretation placard settings (Refers to appropriate placard quantity and size; clear contents)			✓	✓						✓			
20. Experiential activities (Such as boating)		✓		✓							✓		
21. With bird-watching facilities			✓				✓						
22. With viewing platforms or scenic trails							✓						
23. Restaurants are available locally or in the vicinity		✓											
24. Other recreational sites are available locally or in the vicinity.	✓	✓							✓				
25. Accommodation is available locally or in the vicinity.		✓					✓						
26. High degree of support from local residents	✓										✓		

A: Deng, King, and Bauer [9]
B: McNeely, Thorsell, and Ceballos-Lascuráin [10]
C: Chen & Chen [11]
D: Liu [12]
E: Reinius and Fredman [13]
F: Ethos Consulting [14]
G: Arabatzis and Grigoroudis [15]
H: Yeh [16]
I: Hu [17]
J: Wang [18]
K: Garrod and Fyall [19]
L: Valentin [20]
M: Jafari and Wall [21]

Garrod and Fyall (2000) studied the tourist attraction of British sites and pointed out that the degree of protection, transportation accessibility, education, entertainment, supportive local community, and environmental quality are important appeal factors [19]. Jafari and Wall (1994) and Valentine (1993) both pointed out that rare or endangered species and local cultural features are the important ecotourism attractions [20, 21].

In view of the above literatures, the abovementioned factors affecting the ecotourism appeal were compiled into 26 items, as shown in Table 1.

3 Methodology

This study is divided into three stages. In the first stage, the document analysis and expert interview methods were adopted to collect the candidate factors contributing to the appeal

of Taiwan's coastal wetlands ecotourism; in the second stage, the Delphi Method was adopted to select the important factors from the abovementioned candidate factors; in the third stage, the Fuzzy Analytic Hierarchy Process (Fuzzy AHP) was adopted to estimate the relative weights of the respective important factors. The Fuzzy Delphi Method and Fuzzy AHP applied in the second and the third stage are described below.

3.1 The Fuzzy Delphi Method

The Fuzzy Theory was first proposed by Zadeh (1965). It is use to convert unclear and ambiguous phenomena in the environment in mathematical terms. Because the human perception and feelings are vague, subjective, and unascertained, the Fuzzy Theory is applied to more precisely describe the feelings of Man [22].

The Fuzzy Delphi Method is the combination of the Fuzzy Theory and the traditional Delphi Method. The Fuzzy Delphi Method takes into consideration the fuzzy concepts involved to aid the group opinions in reaching a consensus, thus ensuring more reasonable analysis results. The method has been applied by many scholars in factor screening related researches, such as Cheng, Chen and Lee [23]. The steps of the Fuzzy Delphi Method are as follows:

1. The experts give interview values corresponding to each linguistic variable.
2. The linguistic variables ticked under each factor by the experts are to be converted into corresponding interval values.
3. To establish the triangular fuzzy number of the most optimistic cognition in each factor $T=(L,M,U)$ and the triangular fuzzy number of the most conservative cognition $t=(l,m,u)$, of which L, M, and U are the minimum value, geometric mean, and maximum value of the upper bound of interval values for the factor; l, m, and u are the minimum value, geometric mean, and maximum value of the lower bound of interval values of the factor deemed by the expert.
4. As shown in Fig. 1, if $u>L$ and the gray interval G $(G=u-L)$ are greater than $C(C=M-m)$, it means the experts have reached a consensus on the perceived factors, or they have not reached a consensus and another round of the questionnaire needs to be conducted. If a consensus has been reached, the intermediate values within the gray interval $(1/2G)$ shall be used to represent the expert group's evaluation values of the factors.
5. An appropriate threshold is to be set. If a factor's $1/2G$ is greater than the threshold value, the factor will be retained; if not, it will be deleted.

3.2 Fuzzy AHP

The AHP was proposed by Saaty [24]. Through the creation of pairwise comparison matrices, the relative weights among the criteria were obtained. In the AHP, the evaluators used the linguistic variables to express the relative importance of

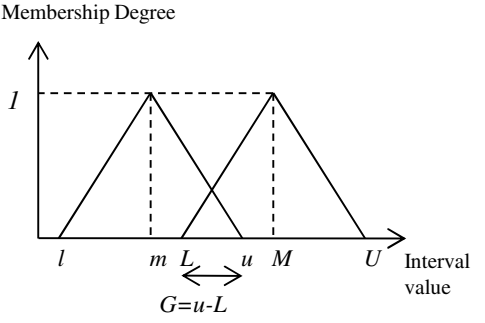


Fig. 1. The gray interval of the Fuzzy Delphi Method

the criteria, but the linguistic variables displayed fuzzy features. Therefore, some scholars such as Csutora and Buckley [25], Liu and Wang [26] combined the Fuzzy Theory and AHP to ensure more reasonable analysis results. The steps of the Fuzzy AHP are as follows:

1. To establish hierarchies. The assessment criteria of each level should be independent of one another.
2. To test the consistency of the pairwise comparison matrices.
3. To establish positive reciprocal matrices: Convert the relative importance between two criteria ticked by the evaluators into triangular fuzzy numbers and establish the fuzzy matrices.
4. Fuzzy weight calculation: In this study, the Lambda-Max proposed by Csutora and Buckley [26] was adopted to calculate the fuzzy weights. The steps are as follows:

a. Set $\alpha = 1$ to obtain the corresponding matrix T_m^k when the membership function is at the maximum; set $\alpha = 0$ to obtain the lower matrix T_l^k and upper matrix T_u^k, and then obtain the corresponding eigenvectors W_l, W_m, and W_u, as defined below:

$$W_l = \left\{w_{1l}^k, w_{2l}^k, ..., w_{nl}^k\right\}^T,\ W_m = \left\{w_{1m}^k, w_{2m}^k, ..., w_{nm}^k\right\}^T,\ W_u = \left\{w_{1u}^k, w_{2u}^k, ..., w_{nu}^k\right\}^T \quad (1)$$

b. To calculate the weight adjustment coefficients Q_l^k and Q_u^k as:

$$Q_l^k = \min\left\{\frac{w_{im}^k}{w_{il}^k} \middle| 1 \le i \le n\right\},\ Q_u^k = \min\left\{\frac{w_{im}^k}{w_{iu}^k} \middle| 1 \le i \le n\right\} \quad (2)$$

After the adjustment, the upper bound and lower bound of the weights are R_l^k and R_u^k:

$$R_l^k = \left\{r_{1l}^k, r_{2l}^k, ..., r_{nl}^k\right\}^T = Q_l^k W_u^k,\ \ R_u^k = \left\{r_{1l}^k, r_{2l}^k, ..., r_{nl}^k\right\}^T = Q_u^k W_u^k \quad (3)$$

By combining R_l^k, W_m^k, and R_u^k, the fuzzy weights obtained are:

$$fW_i^k = \left\{r_{il}^k, w_{im}^k, r_{iu}^k\right\}^T, \ i = 1,2,3,...n \tag{4}$$

5. Aggregate fuzzy weights: Through the arithmetic mean, the fuzzy weights of *n* number of evaluators were integrated.

6. Defuzzification: Defuzzification refers to the conversion of fuzzy weights into definite values in order to determine the relative weights of the criteria. In this study, the Center of Area Method was adopted to engage in defuzzification.

4 Results

4.1 First Stage: Collection of Candidate Factors

The 26 coastal wetlands ecotourism appeal factors listed in Table 1 were obtained based on the summarized literature reviews. Although these literatures are related to the ecotourism related researches, not all focused on Taiwan's coastal wetlands. In order to collect more complete the factors related to the appeal of coastal wetlands ecotourism in Taiwan, the 26 items listed in Table 1 shall serve as the basis while five senior guides, two scholars, two ecology illustration volunteers, and three Tourism Bureau officials, a total of 12 experts, were interviewed to explore whether or not appeal factors deemed important by the experts were omitted in Table 1. According to the survey results, seven factors were added, and some factors were combined to finally sort out 32 candidate factors, as shown in Table 2.

4.2 Second Stage: The Fuzzy Delphi Method Conduction

As this study is intended to analyze the appeal of coastal wetlands ecotourism in Taiwan for tourists, the most senior guides who had the best knowledge of the tourists' feelings were adopted as the expert group in this part of the questionnaire; however, government officials and academic experts were excluded from the expert group. In addition, tourists were not the interviewed subjects in this part of the questionnaire mainly because the development of coastal wetlands ecotourism in Taiwan has just sprouted. If tourists had been taken as subjects for interview, the analysis results would not have been applicable for other coastal wetlands, or deviated analysis results might have taken place because the majority of tourists had limited travel experiences, have only visited a certain part of the coastal wetlands or have never been to one. In this stage, 13 senior guides were adopted as the interviewed experts, all had over five years of tour guiding experience, over six coastal ecotourism activity guide experiences in the recent two years, and have tour guide licenses.

Table 2. Candidate factors and results of screening through the Fuzzy Delphi Method

Candidate Appeal Factors	*t=(l,m,u)*	*G*	*T=(L,M,U)*	*1/2 G*	$> \alpha$
1. With star species (such as endangered species)	(73,82.42,93)	(93, 80)	(80,93.14,100)	86.5	✓
2. Easily able to observe diversified and rich organisms	(73,83.41,90)	(90, 80)	(80,92.12,100)	85	✓
3. Low proportion of artificial facilities	(67,72.46,78)	(78,73)	(73,80.46,85)	75.5	
4. Beautiful or unique natural landscape	(70,79.83,93)	(93,80)	(80,90.33,100)	86.5	✓
5. Environmental cleanliness	(70,78.96,85)	(85,80)	(80,86.35,93)	82.5	✓
6. Safe	(65,73.72,81)	(81,80)	(80,83.65,88)	80.5	✓
7. Public transportation convenience	(73,77.52,85)	(85,80)	(80,83.12,93)	82.5	✓
8. Convenience of Driving Cars	(75,78.48,85)	(85,80)	(80,81.89,90)	82.5	✓
9. Adequate public toilets	(50,68.02,75)	(75,67)	(67,77.2,90)	71	
10. Low degree of crowdedness perceived by tourists	(70,80.69,90)	(90,80)	(80,91.23,100)	85	✓
11. Short equipment (service) waiting time for tourists	(65,71.15,77)	(77,75)	(75,81.15,85)	76	
12. Excellent management and service capabilities of management units	(70,79.04,85)	(85,80)	(80,85.93,93)	82.5	✓
13. The area planned as reserves or national parks	(55,68.02,73)	(73,70)	(70,78.32,85)	71.5	
14. Low impact of tourist activities on community	(55,69.43,75)	(75,70)	(70,78.32,85)	72.5	
15. Low impact of tourist activities on animals and plants	(55,67.32,77)	(77,75)	(75,78.23,90)	76	
16. With interpretation personnel	(70,75.69,83)	(83,80)	(80,85.93,93)	81.5	✓
17. With appropriate interpretation placard settings	(73,77.95,85)	(85,80)	(80,83.05,93)	82.5	✓
18. Experiential activities	(68,76.35,90)	(90,80)	(80,87.03,100)	85	✓
19. With bird-watching facilities	(50,68.95,83)	(83,70)	(70,79.32,93)	76.5	
20. With viewing platforms or scenic trails	(70,75.67,83)	(83,80)	(80,82.1,90)	81.5	✓
21. Food is available locally or in the vicinity	(73,79.8,85)	(85,83)	(83,89.6,95)	84	✓
22. Stores selling local specialties are available locally or in the vicinity	(73,79.54,83)	(83,80)	(80,88.32,95)	82.5	✓
23. Recreational places or historic sites are available locally or in the vicinity(excluding items 21and 22)	(75,78.6,83)	(83,80)	(80,85.63,93)	82.5	✓
24. Accommodation is available locally or in the vicinity	(50,67.32,78)	(78,68)	(68,77.32,90)	73	
25. High degree of support from local residents	(68,73.6,77)	(77,75)	(75,86.3,90)	76	
26. In just a short distance from the city	(68,75.78,85)	(85,75)	(75,77.68,85)	80	✓
27. There are bicycle lanes along the connecting roads	(73,76.31,83)	(83,80)	(80,85.28,90)	81.5	✓
28. Overall community building	(70,79.54,83)	(83,80)	(80,87.32,93)	81.5	✓
29.With mangroves	(55,68.26,78)	(78,65)	(65,75.43,81)	71.5	
30.Buildings for environmental education	(55,66.31,83)	(83,65)	(65,75.02,90)	74	
31. Coordinated buildings and natural landscapes in the area or vicinity	(72,82.76,90)	(90,80)	(80,92.43,100)	85	✓
32.With green lables	(70,72.6,75)	(75,79)	(79,83.45,90)	77	

Based on the data collected from the expert questionnaire, using the Fuzzy Delphi Method, and in reference to suggestions of Cheng and Chen and Lee [23], those with intermediate values exceeding 80 in the gray interval were deemed important. The threshold value α was set as 80. After screening, there were 20 factors that exceeded the threshold value α, which were listed as important factors, as shown in Table 2.

4.3 Third Stage: The Fuzzy AHP Conduction

The 20 important factors selected in the second stage were compiled into two hierarchies, through which the Fuzzy AHP questionnaire was designed to calculate the weights of the dimensions and factors, as shown in Table 3 and Table 4.

Table 3. Hierarchy Structure for Evaluation factors of the appeal of Taiwan's coastal wetlands ecotourism

Dimension	Factors	
C1 Natural resources	C11	With star species
	C12	Easily able to observe diversified and rich organisms
	C13	Beautiful or unique natural landscape
	C14	Coordinated buildings and natural landscapes in the area or peripheral
C2 Transportation	C21	Public transportation convenience
	C22	Convenience of Driving Cars
	C23	In just a short distance from the city
	C24	There are bicycle lanes along the connecting roads.
C3 Environment	C31	Clean and tidy
	C32	Low degree of perceived crowdedness
	C33	Safe
	C34	Excellent management and service capabilities of management units
C4 Recreation and education	C41	With guiding personnel
	C42	With appropriate interpretation placard settings
	C43	Experiential activities
	C44	With viewing platforms or scenic trails
C5 Local and vicinity support	C51	Food is available locally or in the vicinity
	C52	Stores selling local specialties are available locally or in the vicinity
	C53	Recreational places or historic sites are available locally or in the vicinity
	C54	Overall community building

Table 4. Deffuzzied weight of dimensions and factors

Dimension	Defuzzied weight	Factors	Defuzzied weight	Global defuzzied weight	Ranking
C1	0.416	C11	0.254	0.1057	3
		C12	0.284	0.1181	2
		C13	0.359	0.1493	1
		C14	0.103	0.0428	10
C2	0.128	C21	0.285	0.0365	12
		C22	0.614	0.0786	5
		C23	0.052	0.0067	19
		C24	0.049	0.0063	20
C3	0.154	C31	0.48	0.0739	6
		C32	0.261	0.0402	11
		C33	0.164	0.0253	15
		C34	0.095	0.0146	16
C4	0.153	C41	0.095	0.0146	17
		C42	0.192	0.0296	13
		C43	0.544	0.0838	4
		C44	0.169	0.026	14
C5	0.149	C51	0.338	0.0504	7
		C52	0.302	0.045	8
		C53	0.289	0.0431	9
		C54	0.071	0.0106	18

5 Conclusion and Recommendations

1. Among the five dimensions, "natural resources" has the highest weight (0.416), followed by "environment" (0.154), "recreation and education" (0.153), "local and vicinity support" (0.149), and "transportation" (0.128), of which "natural resources" has a significantly higher weight compared to the other dimensions. The remaining four weights show only minor differences. In addition, in view of the factor ranking, three out of the four factors of "natural resources" are ranked top three and the other one ranked tenth, indicating the high importance of the four factors in the dimension.

2. The factor ranked eighth "C52 with stores selling local specialties are available locally or in the vicinity" rarely appeared in past literatures on ecotourism, but played a rather important role in the factors of the appeal of Taiwan's coastal wetlands ecotourism. In consideration to "C52", "C43 experiential activities" ranked fourth, "C51 food is available locally or in the vicinity" ranked seventh, as well as "C53 other recreational places or historic sites are available locally or in the vicinity" ranked ninth, it can be seen that unlike other ecotourism types, Taiwan's coastal wetlands ecotourism, experiential activities locally or in the vicinity, local snacks, local specialties, and other activities are also important appeal factors. However, the expert group gave a poor rating for C43 in terms of its performance in the wetlands. It is therefore suggested that more activities be included under the premise of the least ecological disturbance.

3. "C31 cleanliness" ranked sixth and "C32 low level of perceived crowdedness" ranked eleventh, indicating low pollution and low crowdedness are also important influential factors when tourists engage in the nature related travel. From a comprehensive view of C43, C51~C53, C31~C32 factors, the tourism industry locally and in the peripheral flourished. With the experiential activities or specialty foods, more tourists were drawn to the place, but when there were too many tourists, it not only devastated ecology and disturbed residents, but also led to increased pollution and crowdedness, thus reduced appeal.

4. Although this study included the appeal of ecotourism as the study scope in order to understand the feasibility of promoting coastal wetlands ecotourism and ways to enhance the appeal. However, the competent authorities must comply with the principle of sustainable development in planning coastal wetlands ecotourism in the future. In other words, the development must take into consideration the environmental, social, and economic goals in order to achieve the purposes of natural resource preservation, environmental education, local community welfare protection, and local economy promotion.

Acknowledgment. This research was supported by the National Science Council grant NSC 100-2221-E-006-017 in Taiwan. The authors would like to offer their great thanks to the Agency.

References

1. Ramsar Convention Secretariat: 2006 The Ramsar Convention Manual: a guide to the Convention on Wetlands (Ramsar, Iran, 1971). Ramsar Convention Secretariat, Gland, Switzerland (2006)

2. Kim, S.G.: The evolution of coastal wetland policy in developed countries and Korea. Ocean & Coastal Management 53, 562–569 (2010)
3. Valiela, I., Fox, S.E.: Managing coastal wetlands. Science Magazine, 321–323 (2008)
4. Yang, W., Chang, J., Xu, B., Peng, C., Ge, Y.: Ecosystem service value assessment for constructed wetlands: A case study in Hangzhou. Ecological Economics 68, 116–125 (2008)
5. Jacobes, P., Sadler, B.: A key to tomorrow: On the relationship of environmental assessment and sustainable development. In: Jacobes, P., Sadler, B. (eds.) Development and Environmental Assessment: Perspectives on Planning for a Common Future, pp. 149–165. Canadian Environmental Assessment Research Council, Hull (1992)
6. Fennell, D.A.: Ecotourism in Canada. Annals of Tourism Research 25, 231–235 (1998)
7. Ceballos-Lascuráin, H.: Tourism, ecotourism, and protected areas. The World Conservation Union, Gland (1996)
8. Taiwan Sustainable Development Land group ecotourism White Paper, http://www.cpami.gov.tw/chinese/index.php?option=com_content&view=article&id=10052&Itemid=53
9. Deng, J., King, B., Bauer, T.: Evaluating natural attractions for tourism. Annals of Tourism Research 29, 422–438 (2002)
10. McNeely, J.A., Thorsell, J.W., Ceballos-Lascuráin, H.: Guidelines: Development of National Parks and Protected Areas for Tourism. WTO/UNEP/IUCN, Madrid, Spain (1992)
11. Chen, G.Y., Chen, Y.Y.: Black-Faced Spoonbill Conservation Area, Recreation Experience, Recreation Impacts Perception. Journal of Tourism and Travel Research, 1–18 (2006)
12. Liu, C.R.: A study on sustainable development evaluation indictors in alishan's dabang ecotourism destination. Journal of Tourism Studies 13, 235–264 (2007)
13. Reinius, S.W., Fredman, P.: Protected areas as attractions. Annals of Tourism Research 34, 839–854 (2007)
14. Ethos Consulting: Natural Resource Based Tourism in Northwestern British Columbia. Ministry of Development, Trade and Tourism, Vancouver, Canada (1991)
15. Arabatzis, G., Grigoroudis, E.: Visitors' satisfaction, perceptions and gap analysis: The case of Dadia-Lefkimi-Souflion National Park. Forest Policy and Economics 12, 163–172 (2010)
16. Yeh, M.S.: A study on tourist satisfaction of coastal recreational activities in the Jyibey island of Penghu. Master thesis, Department of Tourism, Shih Hsin University, Taipei (2001)
17. Hu, H.H.: A Study on the attractiveness of the tourism and recreation resources in Bali waterfront. Master thesis, Graduate Institute of Building and Planning, National Taiwan University, Taipei (2008)
18. Wang, R.M.: The tourist importance-performance analysis of ecotourism: A aase study on ecotourism of fishing raft in Sihcao, Tainan. Master thesis, Department of Fisheries Production and Management, National Kaohsiung Marine University, Kaohsiung (2009)
19. Garrod, B., Fyall, A.: Managing heritage tourism. Annals of Tourism Research 27, 682–708 (2000)
20. Jafari, J., Wall, G.: Sustainable tourism. Annals of Tourism Research 21, 667–669 (1994)
21. Valentine, P.S.: Ecotourism and nature conservation: A definition with some recent developments in Micronesia. Tourism Management 14, 107–115 (1993)
22. Zadeh, L.A.: Fuzzy sets. Information and Control 8, 338–353 (1965)
23. Saaty, T.L.: The Analytic Hierarchy Process. MaGraw-Hill, New York (1980)
24. Csutora, R., Buckley, J.J.: Fuzzy hierarchical analysis: the lambda-max method. Fuzzy Sets and Systems 120, 181–195 (2001)
25. Liu, H.T., Wang, C.H.: An advanced quality function deployment model using fuzzy analytic network process. Applied Mathematical Modelling 34, 3333–3351 (2010)

Application of Support Vector Machine Regression in Stock Price Forecasting

Zhongxin Ding

Mathematics Department, Imperial College London
London, The United Kingdom
zhongxin.ding09@imperial.ac.uk

Abstract. An investigation into how support vector machine can be used in the regression process of financial forecasting. A novel stock pricing model has been proposed based on the well-developed fundamental factors model and a combination of factors used in the model have been carefully selected to predict the common stock price. Several classical regression techniques therefore are applied separately in the predicting process and comparison has been made on the correctness of the predicted result. Support Vector Machine Regression has shown very strong competitivity throughout the test.

Keywords: Support Vector Machine, stock pricing model, regression technique.

1 Introduction

The financial market is a complex, evolutionary, and non-linear dynamical system [1]. The field of financial forecasting is characterized by data intensity, noise, non-stationary, unstructured nature, high degree of uncertainty, and hidden relationships. Therefore it's very difficult to predict various financial terms such as stock price, different index price etc. Based on the development of different factor models and empirical analysis, certain factors have stand out and shown strong relationship to a company's financial movements. A novel approach to relate a company's stock price to some factors is proposed in this paper and tested later.

In the stock price forecasting model we proposed and other vast majority of financial models, regression of a small sample of high-dimensional, non-linear data is always required. Therefore the accuracy of regression results hugely affects the predicting accuracy of the model. Support Vector Machines (SVM) has made huge impact to different area of studies such as bioinformatics and pattern reorganization due to its ability in dealing with non-linear and high-dimensional data set [2]. Therefore it's desirable for us to use it in our model and compare it with other classical regression techniques to see whether or not SVM regression is a suitable regression technique in financial analysis.

2 Stock Pricing Model Construction

2.1 Framework of the Model

The central idea of modern financial economics is that the average return of a stock is the payoff to the shareholder for taking on risk. Factor models express this risk-reward

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 359–365.
springerlink.com

relationship. Factors are explanatory variables that represent different types of risk. A factor model shows that the average stock return is proportional to the stock's exposure to the risk that the factor represents and to the payoff for each unit of exposure to the risk. One of the most commonly used factor models is fundamental factor model which uses stock characteristics such as the P/E ratio and market capitalization as factor to measure the exposure to risk for certain stock.

In this paper, we are only interested in testing the ability of different regression techniques when dealing with regression in modern financial analysis; therefore we simplified the fundamental factor model [3] to only consider predicting stock prices using factors that measure its exposure to risk which also known as factor exposure. That is the stock price is proportional to the factor exposure.

Hence by denoting the factor exposure of K factors as $f_1,\ldots,f_K$, stock price r for a certain stock can be estimated by the following equation:

$$r = C_1 f_1 + C_2 f_2 + \ldots + C_K f_K + b \quad (1)$$

Where $C_1,\ldots,C_K$ are weights for each corresponding factor exposure, b is the bias which reflects the random nature of stock price. The equation is typically estimated by a time-series regression using observations made at various time periods. Therefore we can take factor exposures, which are the variables of interest that affect stock price, and regresses stock price at every time interval in the data period on the corresponding weights. For example, the factor exposure f_i might be the market capitalization. The weight C_i found by estimating the regression shows how sensitive stock prices are to market capitalization. Different regression techniques can be used in this process, the suitability of Support Vector Machine Regression is looked closely in this paper and comparisons to other classical regression methods have been made.

2.2 Factor Choice of the Model

A factor is any variable that may represent a firm's financial condition and therefore have influence on the firm's stock performance. There are all kinds of factors, some of which describe stock characteristics and some of which describe conditions in the overall market. These fundamental factors are observable characteristics of the stock itself, and they usually can be read or calculated from financial statements.

Eight factors have been chosen carefully based on empirical analysis. Specific descriptions and reasons to use them in the model are given blew:

- Price-to-earnings ratio (P/E): P/E ratio is a valuation factor which attempt to measure whether stocks are relatively cheap or expensive. P/E ratio is defined as price of a stock divided by the earning per share. It has been found that low P/E ratio stocks outperform stock with higher P/E ratios over the long run.
- Price-to-cash-flow ratio (P/CF): P/CF ratio is also a valuation factor. It's defined as price of a stock divided by its cash flow per share. Lower P/CF ratio companies have been found to outperform high P/CF ratio firms. P/CF have been chosen along with P/E ratio is because cash flow is less vulnerable to accounting manipulation.
- Dividend yield (D/P): D/P is a valuation factor. It's defined as dividend per share of a company divided by its stock price. D/P serves as an invaluable aid in

determining a company's maturity and growth prospects. Stable, mature businesses tend to generate sufficient cash flow and offer relatively high dividend yields. High dividend yielding stocks shown to outperform over the long-run.

- Size: Size is a very classical valuation factor. It's defined as the market capitalization of a company. Smaller size stocks have been found to outperform larger size stocks over the long run.
- Current Ratio (CUR): CUR is a solvency factor which attempt to measure a company's ability to meet future short-term obligations. CUR has been defined as current assets of the firm divided by current liabilities. It measures the ability of a company to cover liabilities with varying degrees of liquid cash. Firms with high CUR have less likelihood of going into bankruptcy and thus can withstand paying their current liabilities.
- Net Profit Margin (NPM): NPM is an operating profitability factor, which tells us how well management is running the company. NPM is defined as net income divided by net sales. This represents the true income of the firm. A very important measure of profitability for companies.
- Debt-to-equity ratio (D/E): D/E ratio is a financial risk factor, which measure financial health of a company. It's defined as the long-term debt of a firm divided by its common equity. A very common measure of the financial risk of a firm.
- Momentum (MM): Momentum is a technical factor which is factor we use to express trend of stock movement we noticed from past data. Momentum is a very popular indicator of future performance. Various academic studies have found positive autocorrelation in stock returns, meaning that positive returns in one period tend to lead to positive returns in the next period [3]. In this paper the momentum has been defined as $\mathbb{MM}_{t,t+1} = \left(\frac{p_{t+1}}{p_t}\right) - 1$, where P_i is stock price at time period i.

3 Mathematical Formulation of Support Vector Regression

Complete formulation and deviation of Support Vector Regression can be found at [2, 4].

Briefly we are trying to solve the following problem:

$$\begin{aligned} &\text{minimise: } \tfrac{1}{2}\|\mathbf{w}\|^2 + C\textstyle\sum_{i=1}^{n}(\xi_i + \xi_i^*) \\ &\text{subject to: } y_i - (<\mathbf{w}, x_i> + b) \le \varepsilon + \xi_i \\ &\qquad (<\mathbf{w}, x_i> + b) - y_i \le \varepsilon + \xi_i^*, \forall i = 1, \dots, n \\ &\qquad \xi_i, \xi_i^* \ge 0; C > 0 \end{aligned} \tag{2}$$

Where **w** is the weight, C is a control factor we need to define ourselves to optimize the regression problem and ε is defined as our tolerance to error for each data points.

Then we can introduce Lagrange Multipliers to transform above to dual form:

$$\begin{aligned} &\text{minimise: } \tfrac{1}{2}\textstyle\sum_{i,j=1}^{n}(\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*)\,K(x_i, x_j) - \\ &\qquad \varepsilon\textstyle\sum_{i=1}^{n}(\alpha_i + \alpha_i^*) + \sum_{i=1}^{n} y_i(\alpha_i - \alpha_i^*) \\ &\text{subject to: } 0 \le \alpha_i;\ \alpha_i^* \le C; \\ &\qquad \textstyle\sum_i \alpha_i - \sum_i \alpha_i^* = 0;\ i = 1, \dots, n \end{aligned} \tag{3}$$

The Kernel function K(…) allows us to map the original non-linear input space to a higher-dimensional feature space inside which the original data are linear and thereby solve the optimization problem.

Finally the regression hyperplane can be expressed as $f(x) = \sum_{i=1}^{n}(\alpha_i - \alpha_i^*)K(x_i, x) + b$.

The enormous success SVM has achieved last decade is mainly attributed to four factors: superior generalization capacity, a globally optimal solution from a convex optimization problem, the ability to handle non-linear problems using the "kernel trick", and the sparseness of the solution which makes it possible to have specialized fast algorithms. These properties have made SVM very robust in dealing with small training set, non-linear and high-dimensional problems, and these are exactly the type of problem we are facing both in our model constructed above and in most of financial regression cases. Most companies only publish their financial report semiannually and some companies even publish them annually. This has hugely limited our resource of historical factor data since most of them can only be obtained twice a year which lead to very small historical data or training points available to us. Also the complexity of stock has made it very hard to determine which factors affect it the most, so it's very common to involve a lot of different parameters in the model to comprehensively describe the stock. This has made most of the model we are interested in high-dimensional. Therefore we expect SVM regression to have superior performance in financial analysis. Experiments are carried out to test which regression technique gives best result for our stock price prediction model.

Before carrying out the SVM regression, few control variables need to be determined.

We have set C = Infinity andε=0 throughout the experiment. This is mainly because C measures the tradeoff between complexity and loss of outliers. In our model, the training set is very small therefore each training point carries very precious information and hence they all should be regard as 'unmissable'. Therefore we set C = Infinity and ε=0.

Another thing is that we need to choose a proper Kernel function. In this paper RBF Kernel function has been chosen due to its reputation in previous studies.

RBF Kernel is given by:

$$K(x_i, x_j) = \exp\left(-\|x_i - x_j\|^2 / 2\sigma^2\right) \tag{4}$$

Where σ is a parameter that needs to be optimized. The larger the value of σ, the larger the area that the support vector has an influence on. Hence σ controls the smoothness of the decision surface. To find the optimal value of C, a very classical method called Cross-Validation has been used. This involved selecting a certain proportion of the data points to be our training set on which we ran SVM regression and then testing the model on the remaining data, called the holdout set. Mean square error is calculated for each choice of σ, and the one with smallest MSE is chosen as our final choice of σ.

4 Experiments and Results

4.1 Experiment Construction

In this section we present the experiments that were performed to validate the ability of SVM regression. Briefly descriptions of competing methods are given below:

- Support Vector Regression with RBF Kernel (SVRR): Method we described in previous chapter. Control variable σ is determined by cross-validation.
- Support Vector Regression with Linear Kernel (SVRL): Classical Support Vector Regression with Kernel function chosen as inner product of x_i, x_j. To test the importance of a suitable Kernel function in dealing with complicated non-linear data set.
- Random Walk (RW): This method is used as a base line. The method implies in each period of time, the price of certain stock can move up or down randomly by as much as ￥100 independent of all factors. The boundary is set according to past stock price data and in our method predicted stock price has been calculated 10 times then the average is taken as the final price.
- Ordinary Least Squares Regression (OLSR): One of the most classical linear regression methods. OLSR minimizes the sum of squared vertical distances between the observed responses in the dataset, and the responses predicted by the linear approximation.
- Backpropagation (BP) [5]: Backpropagation is a classical method of Neutral Network Training. They are nonlinear sophisticated modeling techniques that are able to model complex functions. Numbers of neurons we use are also determined using cross-validation.
- Radial Basis Function Networks (RBFN) [5,6]: A radial basis function (RBF) is a function which has built into it a distance criterion with respect to a center. Such functions can be used very efficiently for interpolation and for smoothing of data. RBF networks are universal approximators on a compact subset of $\mathbb{R}^n$. This means that a RBF network with enough hidden neurons can approximate any continuous function with arbitrary precision.

Three companies are selected as our sample for the experiment. They're CITIC Securities Co.,Ltd.; Ping An Insurance Group.; ICBC Limited. They are leading companies in their industry with biggest weights in HuShen 300 index, and therefore can express the trend of the stock price for their industry very well. Stock price and factor data are gathered from each company's financial reports.

Training data are sampled from December 2007 to December 2010, and test data is June 2011. Error is defined as absolute value between the predicted stock price of June 2011 and the real price. Small error indicates better regression method. Percentage error is calculated last to show the cross-sectional performance of each regression technique.

4.2 Experiment Results

Table 1. Data set for ICBC Limited

Date	Price	P/E	P/CF	D/P(%)	CUR	NPM(%)	D/E	size	MM
2011.06	4.46	8.83	2.79	2.06	1.07	43.6	14.89	13458.62	0.04
2010.12	4.24	16.24	2.01	2.17	1.06	46.9	14.89	12960.38	-0.25
2010.06	4.06	14.32	4.42	2.09	1.06	41.8	14.03	11785.05	0.01
2009.12	5.44	27.1	4.63	1.56	1.06	45.05	13.8	11435.09	0.53
2009.06	5.42	10.73	2	1.52	1.07	36.2	11.99	9757.65	-0.29
2008.12	3.54	26.11	5.39	2.33	1.06	41.86	11.37	9399.96	-0.39
2008.06	4.96	33.88	9.91	1.34	1.07	32.05	9.75	8684.29	0.62
2007.12	8.13	41.75	5.22	0.82	1.06	35.19	9.51	8303.99	-0.19

Since we have assumed equal weight of each factor, we need to standardize each column before carrying out regression.

Table 2. Standardized data set for ICBC Limited

Date	Price	P/E	P/CF	D/P	CUR	NPM	D/E	size	MM
2011.06	4.46	-1.15	-0.69	0.63	1.20	0.62	1.08	1.40	0.08
2010.12	4.24	-0.52	-0.99	0.84	-0.73	1.25	1.08	1.14	-0.69
2010.06	4.06	-0.69	-0.05	0.69	-0.73	0.28	0.69	0.54	0.00
2009.12	5.44	0.40	0.03	-0.34	-0.73	0.90	0.58	0.36	1.38
2009.06	5.42	-0.99	-0.99	-0.42	1.20	-0.79	-0.25	-0.49	-0.79
2008.12	3.54	0.32	0.33	1.16	-0.73	0.29	-0.53	-0.68	-1.06
2008.06	4.96	0.98	2.09	-0.77	1.20	-1.58	-1.27	-1.04	1.61
2007.12	8.13	1.65	0.26	-1.78	-0.73	-0.98	-1.38	-1.24	-0.53

Table 3. Standardized data set for CITIC securities co.Ltd

Date	Price	P/E	P/CF	D/P	CUR	NPM	D/E	size	MM
2011.06	13.08	-2.12	0.30	0.91	1.45	-0.45	-1.00	-0.71	-0.04
2010.12	12.59	0.64	-0.92	1.03	0.11	-2.10	-0.64	0.16	-1.19
2010.06	11.70	0.42	0.16	1.26	-0.22	0.26	-0.28	0.92	0.04
2009.12	31.77	0.97	0.73	-0.77	0.74	1.04	-0.49	-0.95	0.78
2009.06	28.26	0.85	-0.12	-0.64	0.86	0.14	-0.51	-1.20	-0.57
2008.12	17.97	-0.42	-1.51	0.13	-0.89	0.76	1.32	1.73	-1.35
2008.06	23.92	-0.32	1.74	-0.38	-0.47	-0.38	1.80	0.40	0.97
2007.12	89.27	-0.03	-0.38	-1.54	-1.59	0.72	-0.19	-0.36	1.37

Table 4. Standardized data set for Ping an insurance Group

Date	Price	P/E	P/CF	D/P	CUR	NPM	D/E	size	MM
2011.06	48.27	-0.43	0.17	0.44	-0.93	-0.13	1.18	1.68	0.07
2010.12	56.16	-0.38	0.64	-0.52	-0.64	0.08	1.41	0.96	-0.59
2010.06	46.81	-0.41	-0.31	0.90	-0.93	-0.06	-0.39	0.46	-0.10
2009.12	55.09	-0.22	-0.91	-0.52	-0.36	-0.67	0.87	0.20	1.32
2009.06	49.46	2.47	-1.28	-1.56	-0.36	-1.98	-0.55	-0.48	-1.17
2008.12	26.59	-0.37	0.17	1.32	0.21	1.07	-0.52	-0.82	-1.32
2008.06	49.26	-0.36	1.95	0.78	1.65	0.81	-0.75	-0.81	0.60
2007.12	106.10	-0.31	-0.42	-0.83	1.36	0.89	-1.25	-1.19	1.19

To fix the best choice of σ, Cross-Validation is used and specific result for ICBC Limited. are given below, σ choice for the other two companies are fixed in similar manner.

Table 5. Cross-Validation result for ICBC limited

σ	Predicted Price	Error
0.1	28.62	24.38
0.5	6.97	2.73
0.6	4.65	0.41
0.62	4.28	0.04

Hence σ values are fixed as 0.62, 0.61, 0.76 for ICBC, CITIC securities, PingAn Insurance respectively.

Final comparison results for all six regression methods are given below:

Table 6. Absolute error for different regression techniques

	SVMR	SVML	RW	OLSR	BP	RBFN
ICBC	0.21	1.09	30.48	38.74	3.65	1.13
CITIC securities	0.89	15.84	54.65	45.69	10.76	1.87
PingAn Insurance	2.97	20.33	34.26	184.12	8.47	3.23

Table 7. Persentage error for different regression techniques

	SVMR	SVML	RW	OLSR	BP	REFN
ICBC	4.71	24.51	683.41	868.61	81.84	25.34
CITIC securities	6.80	121.10	417.81	349.31	82.26	14.30
PingAn Insurance	6.15	42.12	70.98	381.44	17.55	6.69

We can clearly see that SVMR outperforms other common regression techniques by a huge amount. Especially techniques like OLSR which performs even worth than the baseline RW. Small training set, need of high-dimensional model, non-linear data set etc. all these properties of financial market has limited the use of some very successful regression technique such as Ordinary Least Squares Regression and Backpropagation. This suggests that Support Vector Machine Regression has very huge potential application in the field of financial analysis.

6 Conclusion

In this paper, a stock pricing model has been constructed based on the fundamental factors model. A combination of factors have been selected and proved to be very successful in reflecting a company's stock price. Using desirable regression method, the predicted stock price is significantly close to the real price. This can be very useful to stock trader, company strategists or any financial analyst. Several common regression techniques are compared under close scrutiny. Support Vector Machine Regression outperforms other classical methods by a significant amount. This suggests a very demanding application of SVMR in the future study of financial analysis.

References

1. Abu-mostafa, Y.S., Atiya, A.F.: Introduction to financial forecasting. Applied Intelligence, 205–213 (1996)
2. Vapnik, V.: Statistical Learning Theory. Wiley (1998)
3. Chincarini, L.B., Kim, D.: Quantitative Equity Portfolio Management: An Active Approach to Portfolio Construction and Management. McGraw-Hill Library of Investment and Finance (1963/2006)
4. Smola, A.J., Scholkopf, B.: A Tutorial on Support Vector Regression. Kluwer Academic Publishers (2003)
5. White, H.: Economic prediction using neural networks: The case of IBM daily stock returns. Neural Networks in Finance and Investing, 315–328 (1993)
6. Haykin, S.: Neural Networks: A Comprehensive Foundation. Prentice Hall, Inc., New Jersey (1999)

Empirical Analysis of the Influence on the Sino-US Trade Balance by the RMB Real Exchange Rata

Bangyong Hu

School of Economics and Management, Chongqing Three Gorges University, Wanzhou, Chongqing, China
njhby@126.com

Abstract. By using co-integration analysis, vector error correction model and granger causality test, and based on the quarterly data between China and the United States, this paper give an empirical analysis on the relation between the real exchange rate and Sino-US trade. The conclusions are as follows: factors affecting Sino-US trade balance includes U.S. real GDP, China's real GDP and real exchange rate, the most important factor to impact on Sino-US trade is the U.S. real GDP, and there is a clear J-curve effect between Sino-US trade. Based on the research results, at last the author gives the proposition of the RMB exchange rate's reform.

Keywords: Trade balance, ADF test, RMB exchange rate, Vector error correction model.

1 Introduction

Since reform and opening up, China's economic development has made remarkable achievements. China's trade exports from 206 billion dollars in 1978 to 2.9728 trillion dollars in 2009. China has become the world's largest trade surplus country. With the expansion of Sino-US trade surplus, appreciation of the RMB exchange rate is facing tremendous pressure, RMB exchange rate has become an important factor which affect Sino-US trade relations. RMB exchange rate in the U.S. balance of trade is to play a role in? RMB exchange rate to what extent play a role? The existence of China-US trade J-curve effect? This series of studies has a strong theoretical and practical significance for our country.

On the relationship between exchange rate and trade balance, foreign scholars have extensively studied. With developed countries for the study sample, Kenen and Rodrik studied the impact of Short-term exchange rate volatility on imports and exports, the results showed that: Frequent fluctuations in the exchange rate has significantly inhibited trade between developed countries. Krugman and Moffett believed that U.S. real exchange rate depreciation will help to improve the U.S. trade imbalance. Marquez using 1973-1985 quarterly data found that currency devaluation in developing countries can improve their trade balance. Rose By using the United States and Britain 1974-1986 years of data for statistical test, results showed that there is no co-integration relationship between trade balance and real exchange rate in developed countries,

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 367–372.
springerlink.com

Marshall-Lerner condition is also not established. By comparing developed and less developed countries, Arize found that exchange rate fluctuations in the short term and long term import were strongly inhibited import and export both developed and developing countries. Marquez believed that exchange rate changes have little effect on China's trade balance; There are many studies about the relationship between exchange rate and trade balance in the domestic. By building a gravity model of international trade and using import and export data between China and major trading countries, Chen Ping and Xiong Yan found that Exchange rate volatility is not conducive to increased exports to China. Li Guang-zhong believe that the depreciation of the RMB exchange rate depreciation have a positive impact on China's imports and exports; Cao Yang and Li Jian-wu think that exchange rate depreciation is no significant effect on Import and export in the short term, but in the long term the devaluation of the RMB has a positive impact.

2 The Establishment of Model and Data Processing

According to traditional international trade theory, Real exchange rate appreciation and the increase in national gross domestic product will increase the country's imports, otherwise reduce the imports. The real exchange rate depreciation and the increase in foreign GDP will increase the country's exports, otherwise reduce the country's exports. Therefore, the key factors affecting the trade balance include its gross domestic product, foreign GDP and the real exchange rate RER. So import equation and export equation reflect the Sino-US trade as follows:

$$LnIM_t=\delta_1+\beta_1 LnRER_t+\beta_2 LnGDP_t+\varepsilon_t \quad (1)$$

$$LnEX_t=\delta_0+\alpha_1 LnRER_t+\alpha_2 LnGDP_t^*+\mu_t \quad (2)$$

$LnIM_t$ and $LnEX_t$ are the logarithm value of the import and export. $LnGDP_t$ and $LnGDP_t^*$ are the logarithm value of the China and U.S. real GDP. $LnRER_t$ is the logarithm value of the real exchange rate of RMB. In this paper, the key data include Sino-US trade data, China and the U.S. real GDP, real exchange rate of RMB, Consumer price index etc. Sino-US trade data is from the U.S. Federal Reserve System database, the quarterly GDP of China and the U.S, consumer price index, the nominal exchange rate of RMB and other data are from the International Monetary Fund's International Financial Statistics Yearbook over the years. In this article, in order to eliminate the heteroscedasticity of the model, we will make each variable logarithmic transformation.

In this paper, the RMB exchange rate data is the real exchange rate. The real exchange rate is calculated as RER=NE×P*/P, RER represents nominal exchange rate under the direct quotation. P and P* represent the price level in China and the United States. In this paper, with China and the U.S. consumer price index replace the two countries' price levels, then come the real exchange rate equation: RER=NE×CPI*/CPI. In this paper, the main research methods include co-integration analysis, vector error correction model, impulse response functions and Granger causality test. Co-integration analysis is mainly to study the long-term relationship among variables, vector error correction model is mainly to analyze the short-term dynamics between the various variables affecting the mechanism, impulse response function is in order to

analyze the impact of real exchange rate on short-term dynamics of import and export. Granger causality test is to examine the causal relationship between variables.

3 Model Estimation and Interpretation of Results

3.1 ADF Test

In the last century 80's, Engle and Granger proposed the concept of co-integration. Co-integration means that maybe there is a stable linear relationship among the variables. The stable linear reflect the long-run equilibrium relationship among the time series. Before co-integration analysis, we should make unit root test for time series in order to avoid spurious regression phenomenon. The test principle is:

$$\Delta x_t=\delta+\alpha T+\lambda x_{t-1}+\Sigma\eta_i\Delta x_{t-i}+\varepsilon_t \quad (3)$$

By ADF test, we can see that time series LnGDP, LnGDP*, LnRER, LnEX, LnIM are non-stationary series, but ΔLnGDP, ΔLnGDP*, ΔLnRER, ΔLnEX, ΔLnIM are stationary series. Maybe there is a co-integration relationship among the time series.

3.2 Co-integration Test

According to ADF test results, all variables are first-order lag variable, which meet the prerequisite for co-integration equation. There are may have long-term stable equilibrium relationship among the time series. Co-integration test are usually two ways: Engel-Granger two-step method and Johansen maximum likelihood method. EG two-step method first is to use least squares regression analysis of each variable, and then the regression residuals for unit root test. EG two-step method co-integration parameters have a strong coherence and effectiveness, But its defect is: (1) Under the conditions of a limited sample, the resulting estimator is not unbiased, and the smaller sample size, the greater the deviation of its estimator; (2) When more than one co-integration vector, EG two-step method seems to be quite helpless. In order to overcome the small sample size of defects under the EG two-step method, in this paper we use a multivariate Johansen co-integration test methods to verify whether there is a co-integration relationship among the economic variables. Johansen method is based on vector auto regression model test, the key issue of VAR model is to determine the lag order, lag order are generally based on the principle of AIC and SC. Johansen co-integration test results for export equation in Table 1.

Table 1. Co-integration test results of export equation

Null hypothesis	Eigen value	Trace statistic	5% threshold value	Maximum Eigen value statistic	5% threshold value
rk=0	0.3828	41.3357	32.0819	22.4213	15.3315
rk≤1	0.1561	21.0735	25.0097	15.1018	19.3873
rk≤2	0.0918	5.0817	12.5179	6.8771	11.0916

From the test results we have seen, trace statistic and maximum eigenvalue statistics reject the null hypothesis in the 5% significance level, it means that there is a co-integration relationship between China's exports to the United States, the U.S. real GDP and RMB real exchange rate. According to test results, we can draw export co-integration equation by Eviews 5.0 as follows:

$$LnEX=-28.11819+4.17183LnGDP^{*}+1.52131LnRER \quad (4)$$
$$R2=0.872 \quad DW=1.82 \quad F=51.231$$

From model (4) we can seen that regression results is significant and the DW value is moderate, regression model results are obvious. Model (4) reflects the following information: First, U.S. real GDP growth and RMB real exchange rate depreciation will help increase China's exports; Second, the long-term elasticity of U.S. real GDP on China's exports is 4.17182, which means that U.S. real GDP increased by 1% makes China's exports increased by 4.17182%;Third, the long-term elasticity of RMB real exchange rate on China's exports is 1.52131, in other words, 1% depreciation of the real exchange rate of RMB will increase China's exports 1.21782%. Finally, from the T-statistic results, U.S. real GDP and real exchange rate have a significant impact on Chinese exports, but from the coefficient of elasticity can be seen that U.S. real GDP impact on china's exports is much higher than the RMB exchange rate depreciation. In the same way, we can get the import co-integration equation is:

$$LnIM=-10.88166+2.220458LnGDP-1.347978LnRER \quad (5)$$
$$R2=0.891 \quad DW=1.98 \quad F=31.217$$

We can see from the above equation, during the first quarter of 1990 to the fourth quarter of 2009, the elasticity of China's real GDP on imports is 2.220458, in other words, China's real GDP increased by 1% makes China's imports increased by 2.220458%. The elasticity of RMB real exchange rate on China's imports is 1.347978, which means that 1% depreciation of the real exchange rate of RMB will reduce China's exports 1.347978%.The two elasticity can pass the T-test, it indicates that the RMB real exchange rate and China's real GDP impact on imports are significant. Compare the two elasticity absolute value, the former is bigger than the latter, it means that china's real GDP shows a greater impact on China's imports than RMB real exchange rate.

Compare equation (4) and equation (5) can be found, the same degree depreciation of RMB exchange rate, reduction in U.S. imports from China is bigger than increase in exports to the U.S. The main reason may be related to the structure with the China Import and Export. China's exports products are low-tech and labor-intensive goods, while the imported products are technology-intensive or capital-intensive goods, and the labor-intensive products is strong sensitivity to the real exchange rate changes relative to capital-intensive goods.

3.3 Vector Error Correction Model

Johansen test can only show the existence of a stable equilibrium relationship among various economic variables. But in the short term, the unbalanced relationship may exist among the variables. In this paper, we use the vector error correction model to

study the dynamic relationship among the various time series in the short-term. We get the equation error correction term shown in Table 2 by using Eviews5.0.

Table 2. Import and export equation error correction term

Error correction term	ΔLnEX,	ΔLnIM
Coefficient	-0.04187	-0.31895
Standard deviation	0.02198	0.08902
t-Statistics	-3.12157	-4.52762

Error correction term coefficient indicate the speed of adjustment to the long-run equilibrium each variable. If the variable coefficient is statistically significant, it means that what proportion of dependent variable non-equilibrium can be corrected in the next years. We clearly see from Table 2: error correction term coefficient of the export and import equations are negative and the correction coefficient all can be through T-test. View from the export equation, about 4.19% of the gap between long-run equilibrium value and actual value of LnEX has been corrected, it shows adjustment speed is slow to long-run equilibrium when the variable LnEX is by short-term interference. From the export equation, we learn that about 31.90% of the gap between long-run equilibrium value and actual value of LnEX has been corrected, it shows adjustment speed is fast to long-run equilibrium when the variable LnIM is by short-term interference.

3.4 Granger Causality Test

Co-integration and vector error correction model can explain long-term equilibrium relationship and short-term dynamics among the economic variables, indicate a causal relationship among variables, but can not indicate which variable is the reason and which variable is the result, therefore require further validation. Granger causality test results as shown in Table 3.

Table 3. Granger Test Results

Observed variable	Null Hypothesis	Lag order	F-statistic	P-value	Conclusion
LnEX	LnEX does not Granger Cause LnRER	2	4.8912	0.1018	accept
	LnRER does not Granger Cause LnEX	1	7.1019	0.0016	reject
LnIM	LnIM does not Granger Cause LnRER	1	0.9017	0.2118	accept
	LnER does not Granger Cause LnIM	2	4.9836	0.0017	reject

From the Granger causality test results, in 5% significance level, RMB real exchange rate does Granger Cause China's import and export, however, China's import and export does not Granger Cause RMB real exchange rate. It indicates that the RMB real exchange rate and imports & export are one-way causal relationship.

4 Conclusion

Based on the traditional classical theory of international trade balance, this article study the relationship between the real exchange rate and China-US trade by using co-integration analysis, impulse response function analysis and Granger causality test. By the co-integration test, we can learn: there is existence of a stable equilibrium relationship among various economic variables in the long term. View from the export equation, U.S. real GDP and real exchange rate have a significant impact on Chinese exports, but the influence of the former is far greater than the latter. From the point of view the import equation, the main factor affecting imports of China's are real GDP and real exchange rate of RMB, but the former's influence is still dominant. By Granger causality test found that: in the case of the optimal lag, there is a causal relationship between the RMB real exchange rate and trade balance. RMB real exchange rate depreciation improves the trade balance.

Based on the above findings, suggestions are given as follows: first problem is to expand domestic demand. From the preceding analysis, real GDP for the impact of China's trade surplus is much higher than the real exchange rate effects. China should adjust the economic structure, improve the social security system, actively expand employment and promote urban-rural integration, by expanding domestic demand to reduce China's huge trade surplus. The second is to adjust the export structure, analysis shows that the labor-intensive export products affected by exchange rate fluctuations is greater. China should strengthen IPR protection, enhance the core competitiveness of foreign trade enterprises, promoting the upgrading of product structure and industrial innovation, improving the international competitiveness of export products. Third, reform the exchange rate mechanism. Government should relax controls over the RMB exchange rate, expand the exchange rate range and exchange rate is more flexible, implement comprehensive managed floating exchange rate system.

References

1. Arize, A.C., Osang, T., Slottje, D.J.: Exchange-Rate Volatility and Foreign Trade: Evidence from Thirteen LDC's. Journal of Business & Economic Statistics 18(1), 10–17 (2000)
2. Marquez, J.: Bilateral Trade Elasticities. Review of Economics and Statistics 72(1), 70–77 (1990)
3. Narayan, P.K.: Examining the Relationship between Trade Balance and Exchange Rate: the Case of China's Trade with the USA. Applied Economics Letters 13(8), 507–510 (2006)
4. Krugman, P.R., Baldwin, R.E.: The Persistence of the U.S. Trade Deficit. Brookings Papers on Economic Activity (1), 1–55 (1987)
5. Marquez, A., Schindler, J.: Exchange-rate Effects on China's Trade. Review of International Economics 15(5), 837–853 (2007)
6. Kenen, P.B., Rodrik, D.: Measuring and Analyzing the Effects of Short-Term Volatility in Real Exchange Rates. Review of Economics and Statistics 68(2), 311–315 (1986)
7. Bahmani-Oskooee, M.: Devaluation and the J-curve: Some Evidence from LDCs. Review of Economics and Statistics 67(3), 500–504 (1985)

Evaluation of the Influence of Selected Production Factors on Makespan for Different Routes in FJSP Problem

Tadeusz Witkowski, Arkadiusz Antczak, and Paweł Antczak

Warsaw University of Technology,
Narbutta 85 Street, Warsaw, Poland
tawit@poczta.onet.pl, {arkadij,nauka16}@op.pl

Abstract. In the paper flexible job shop scheduling problem FJSP (an extension of the classical job shop scheduling) is analyzed. We discuss the influence of different production factors on makespan. For the analysis of the production process the greedy randomized adaptive search procedure (GRASP) heuristic was used. Experiments with different levels of factors have been considered and compared. The GRASP algorithm has been tested and illustrated with results for the serial route and the parallel one.

Keywords: Serial and parallel route, simulation modeling, factors of production, flexible job shop, makespan.

1 Introduction

Simulation modeling is a common paradigm for analyzing complex systems. The greatest overall benefit of using simulation in a production environment is that it allows a manager or engineer to obtain a system wide view of the effect of "local" changes to the manufacturing system.

Simulation is well known as a powerful tool supporting the design, layout or redesign of production systems. Recently, many successful applications proved that it can also support the operation of manufacturing systems, especially in the area of scheduling and control.

There are a number of other well-known, general-purpose simulation packages [1-4]. The broad goal of production operation management, such as a resource constrained scheduling problem, is to achieve a coordinated efficient behavior of manufacturing in servicing production demands, while responding to changes on shop-floors rapidly and in a cost effective manner. Shop floor scheduling, such as resource constrained scheduling problems in general, is complex, NP-hard problem, thus is unfeasible to be solved computationally by the sole use of conventional operations research approaches. The development of decision-making methodologies is currently headed in the direction of simulation and search algorithm integration. This leads to a new approach, which successfully joins simulation and optimization. Hybrid techniques using domain specific heuristics are necessary to guide the search and to provide satisfactory solutions in due time.

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 373–382.
springerlink.com

2 Formulation of the Problem

In job shop scheduling problem (JSP), there are n jobs and m machines, each job is to be processed on a group of machines satisfying precedence constraints. Each operation of job is to be processed only on one predetermined machine. Though the JSP has been well studied, its application to real-word scenarios is often undermined by the constraint of the one-to-one mapping of operations to machines. Hence, FJSP problem [6] extends the JSP problem by allowing each operation to be processed on more than one machine. With this extension, we are now confronted with two subtasks: assignment of each operation to an appropriate machine and sequencing operations on each machine.

The FJSP problem is formulated as fellows. There is a set of jobs $Z = \{Z_i\}$, $i \in I$, where $I = \{1, 2, ..., n\}$ is an admissible set of parts, $U = \{u_k\}$, $k \in 1, m$, is a set of machines. Each job Z_i is a group of parts Π_i of equal partial task p_i of a certain range of production. Operations of technological processing of the i-th part are denoted by $\{O_{ij}\}_{j=\xi}^{H_i}$. Then for Z_i , we can write $Z_i = (\Pi_i \{O_{ij}\}_{j=\xi}^{H_i})$, where $O_{ij} = (G_{ij}, t_{ij\,(N)})$ is the j-th operation of processing the i-th group of parts; ξ_i is the number of operation of the production process at which one should start the processing the i- th group of parts; H_i is the number of the last operation for a given group; G_{ij} is a group of interchangeable machines that is assigned to the operation O_{ij}; $t_{ij\,(N)}$ is an elementary duration of the operation O_{ij} with one part d_i that depends on the number of machine N in the group (on the specified operations); t'_{ij} is the duration of set up before the operation O_{ij}. The most widely used objective is to find feasible schedules that minimize the completion time of the total production program, normally referred to as makespan (*Cmax*).

The GRASP algorithm [5] has been tested and illustrated with results for the serial route and the parallel one in FJSP problem.

3 GRASP Metaheuristic

Metaheuristic algorithm include - but are not restricted to - Ant Colony Optimization (ACO), Evolutionary Computation (EC) including Genetic Algorithms (GA), Iterated Local Search (ILS), Simulated Annealing (SA), Greedy Randomized Adaptive Search Procedure (GRASP) and Taboo Search (TS). In this paper GRASP metaheuristics have been used for the optimization of production flow.

The GRASP procedure consists of two main phases: the phase of the constriction of the initial solution and the phase of the local search. The feasibly solution for the task FJSP is generated at the stage of construction of the initial solve and its neighborhood is analysed at the stage of the local search phase. These phases are repeated till the moment that the stop criterion has been satisfied. The best of founded solutions among all iterations is returned as the result of application of the procedure. The construction phase builds a feasibly solution, one element at a time. The being of a candidate for a set up" is included into the construction of such solution [5].

4 The Computer Experiments

In this experiment there are $k = 6$ factors and we want to get an initial estimate of how each factor affects the response. We might also want to determine if the factors interact with one another, i.e., whether the effect of one factor on the response depends on the levels of the others. On way to measure the effect of a particular factor would be to fix the levels of the other k-1 factors at some set of values and make simulation runs at each of two lewels of the factor of interest to see how the response reacts to changes in this single factor. A much more ecomical strategy for determining the effects of factors on the response with which we can also measure interactions, called a 2^k factorial design, requires that we choose just two levels for each factors and then calls for simulation runs at each of the 2^k possible factor-level combinations, which are sometimes called design points. Because we are using only two levels for each factor, we assume that that the response is approximately linear over the range range of the factor.

In this work we examined the serial route and the serial-parallel route of the production flow. In a serial route (PS) an entire batch of parts is processed on one machine and only when all of the products in the batch have been processed are they sent to the next machine. In a serial - parallel route (PSR) individual items of the batch are sent to the next machines as soon as they have been processed on the previous machine.

For the solution of the FJSP problem special software to realize the heuristic was created. Experiments were carried out for data presented in [5], for the number of operations - 160, and the number of machines - 26. For the solution of the FJSP problem special software to realize the heuristic was created. In both cases the experiment series were carried out with 5 replications.

Below we considered how C_{max} - the total time in which all production tasks are completed - depends on different of production factors. The research was based on a six-factor two-level simulation experiment. In the analysis we took into account such factors as C1 - the times of the technological operations T_{ot} (the basic "+" level and the "-" level, whose values were twice as low as those of the basic one), C2 - the times of machine set up T_{pm} (the basic "+" level and the "-" level, whose values were twice as low as those of the basic one), C3 - transport operation times T_{otr} (the "-" level equals 2 min, the "+" level is 4 min), C4 - the number of transport devices L_{ut} (the "-" level is 2, the "+" level is 4 items), C5 - the number of tasks L_z, (the basic "-" level and the "+" level, whose values were twice as high as those of the basic one), C6 - the size of the production batch W_p (the basic "+" level and the "-" level, whose values were twice as low as those of the basic one). Further on during the analysis we also use TO, TP, TT, LU, LZ, WP respectively to stand for the 6 factors.

The loading/unloading times for the transport operations were left the same and equal to 2 min. The number of each combination of factor levels (CFL), the C_{max} values from the repeated experiments and average C_{max} values are evaluated for the serial route (tab. 1), and for the serial-parallel route (where the factor leves "-" and "+" were marked as "m" and "d" respectively).

The average C_{max} values for the factor level combinations in the serial route are shown in Table 1.

Table 1. The C_{max} values for various combinations of factor levels (serial route)

Combi-nation number	Factor level combination	Values of C_{max} and CPU time			Average C_{max} and CPU time
1	mmmmmm	15198,8	15240,8	15198,8	15212,8
		13,668	12,991	13,107	13,3
2	dmmmmm	28443,1	28443,1	28443,1	28443,1
		13,132	12,84	12,861	12,9
3	mdmmmm	15590,35	15516	15451,4	15519,3
		13,447	13,012	12,83	13,1
4	ddmmmm	28721,8	28647,4	28651,7	28673,6
		13,426	13,23	13,447	13,4
5	mmdmmm	19770,65	19570,4	19642,45	19661,2
		14,187	13,907	14,493	14,2
6	dmdmmm	31923,1	32078,1	32224,5	32075,2
		13,961	13,649	13,611	13,7
7	mddmmm	19832,65	19798,85	19932,55	19854,7
		14,4	14,144	14,305	14,3
8	dddmmm	32535,4	32113,6	32486,3	32378,4
		12,519	12,832	13,056	12,8
9	mmmdmm	15198,8	15347,2	15288	15278,0
		11,065	11,029	10,564	10,9
10	dmmdmm	28480,7	28443,1	28443,1	28455,6
		10,701	10,628	10,66	10,7
11	mdmdmm	15389,3	15389,3	15465,3	15414,6
		10,777	10,78	10,936	10,8
12	ddmdmm	28924,7	28789,8	28633,6	28782,7
		10,821	10,483	10,287	10,5
13	mmddmm	18769,95	18806,95	18820,6	18799,2
		10,686	10,421	10,583	10,6
14	dmddmm	31923,1	31923,1	31966,4	31937,5
		10,563	10,712	10,624	10,6
15	mdddmm	18869,3	18899,55	18889,55	18886,1
		10,702	10,662	10,561	10,6
16	ddddmm	32113,6	32321,4	32236,2	32223,7
		10,608	10,421	11,051	10,7
17	mmmmdm	19210,2	19349,1	19113,75	19224,4
		83,198	82,589	82,294	82,7
18	dmmmdm	37166,4	35786,3	37172,1	36708,3
		82,571	83,407	83,283	83,1
19	mdmmdm	19443,55	19962,2	19274,6	19560,1
		83,476	84,42	82,955	83,6
20	ddmmdm	36862,9	37586	37431,6	37293,5
		81,426	82,103	83,267	82,3
21	mmdmdm	26874,4	26560,65	27452,25	26962,4
		81,451	81,747	82,746	82,0
22	dmdmdm	41222,1	41642,5	40575	41146,5
		98,744	99,854	98,888	99,2
23	mddmdm	26503,5	26237,1	27031,15	26590,6
		83,088	85,382	85,139	84,5
24	dddmdm	41596,6	40658,8	39807,5	40687,6
		96,48	96,567	97,176	96,7
25	mmmddm	18263,85	18827,4	18285,3	18458,9
		32,683	31,918	31,249	32,0
26	dmmddm	35965,1	36179,3	35661,9	35935,4
		34,524	34,556	34,836	34,6
		18751,5	18512,7	18994,75	18753,0

Table 1. (*continued*)

27	mdmddm	32,106	32,083	31,703	32,0
		36041,4	35979,9	35496,9	35839,4
28	ddmddm	34,15	35,257	34,766	34,7
		22622,45	22164,2	22004,2	22263,6
29	mmdddm	38,225	38,213	38,109	38,2
		39087,5	38427	39012,4	38842,3
30	dmdddm	36,911	36,708	36,106	36,6
		22994,3	22801,55	22664,4	22820,1
31	mddddm	37,817	37,552	37,994	37,8
		39089,6	39243,5	39199,2	39177,4
32	dddddm	36,367	37,222	36,881	36,8
		30205,05	30198,35	30298,05	30233,8
33	mmmmmd	13,378	13,345	13,216	13,3
		56671,7	56671,7	56924,9	56756,1
34	dmmmmd	12,997	12,963	13,189	13,0
		30373,6	30553,5	30717	30548,0
35	mdmmmd	13,322	12,745	13,111	13,1
		57097,2	57760,5	56862,2	57240,0
36	ddmmmd	13,058	13,213	13,432	13,2
		38601,35	38484,35	38492,65	38526,1
37	mmdmmd	14,602	14,727	14,624	14,7
		64305,4	63885,1	64547,6	64246,0
38	dmdmmd	13,556	13,33	13,087	13,3
		38726,8	38972,6	38926,55	38875,3
39	mddmmd	14,605	14,637	14,337	14,5
		64229,5	63944,2	63822,2	63998,6
40	dddmmd	13,537	13,791	13,465	13,6
		30299,55	30210,75	30183,1	30231,1
41	mmmdmd	10,656	10,701	10,776	10,7
		56824,9	56671,7	56931	56809,2
42	dmmdmd	10,764	11,061	10,656	10,8
		30459,25	30373,6	30394,6	30409,2
43	mdmdmd	11,062	10,733	10,852	10,9
		56862,2	56862,2	56862,2	56862,2
44	ddmdmd	10,655	10,687	10,771	10,7
		37289,05	37284,7	37143,1	37239,0
45	mmddmd	10,562	10,811	10,639	10,7
		63808,3	63728,4	63854,2	63797,0
46	dmddmd	10,756	10,561	10,593	10,6
		37524,55	37491,05	37333,6	37449,7
47	mdddmd	10,677	10,671	10,703	10,7
		63822,2	63822,2	64046,2	63896,9
48	ddddmd	10,81	10,671	11,044	10,8
		38282,4	38244,9	39107,95	38545,1
49	mmmmdd	84,732	85,102	84,914	84,9
		73209,5	74788,4	73800	73932,6
50	dmmmdd	84,462	85,513	84,071	84,7
		39390,65	38575,4	38391,7	38785,9
51	mdmmdd	85,81	84,655	86,687	85,7
		74822,7	73718	74179,4	74240,0
52	ddmmdd	83,424	84,047	83,392	83,6
		53849,4	52221,45	53904,65	53325,2
53	mmdmdd	84,23	83,656	82,065	83,3
		84277,2	82582,8	84675	83845,0
54	dmdmdd	97,777	101,204	99,132	99,4
		54078,15	53085,3	54018,75	53727,4
55	mddmdd	85,223	84,219	85,272	84,9
		82946	82796,8	79891,9	81878,2

Table 1. (*continued*)

56	dddmdd	98,762	98,313	98,463	98,5
		36848,65	36850,35	37063,25	36920,8
57	mmmddd	34,11	33,463	33,064	33,5
		71858,5	72176,6	70657,6	71564,2
58	dmmddd	35,191	34,942	36,13	35,4
		36291,05	36581,5	37412,5	36761,7
59	mdmddd	33,07	33,292	33,385	33,2
		71489,9	71089,1	73336	71971,7
60	ddmddd	35,598	35,978	35,093	35,6
		44657,35	44656,45	45595,55	44969,8
61	mmdddd	39,228	38,736	39,408	39,1
		77516,2	75630,7	78523,4	77223,4
62	dmdddd	36,147	36,599	37,83	36,9
		45573,45	42868,2	45224,35	44555,3
63	mddddd	38,447	37,021	37,74	37,7
		78402,6	78843,7	78207,1	78484,5
64	dddddd	37,249	36,735	35,788	36,6

Let us consider some of the factor level combinations, marking with Cx the x number of the C factor and with Kx - the number x of the CFL combination. When analysing the results of the experiment the C_{max} value [min] was described with the following linguistic variables: C_{max} below 15.000 - especially small; $15.000 < C_{max} \leq 30.000$ - very small; $30.000 < C_{max} \leq 45.000$ - small; $45.000 < C_{max} \leq 60.000$ - average; $60.000 < C_{max} \leq 75.000$ - big; $75.000 < C_{max} \leq 90.000$ - very big.

Based on the results of the experiment we can assign to individual CFL (where * stands for any level of the factor) the C_{max} values in the following way: d***dd - big; d*m*md - average; ****dd - average; d*d**d or d*d*dd - big; d***mm or d***dm or m***md or m***dd- low, and m***mm or d***mm- very small.

The analysis of the result also shows that for a two-fold increase of the transport appliance number (factor C4) from 2 to 4 illustrated in K56 and K64 respectively in the PS route the C_{max} value drops relatively little, i.e. by (81878-78484)/81878 = 4 %. When C3 changes from the "m" level to "d" (combinations K60 and K64) the C_{max} value increases by (78207-71971)/71971 = 8 %, which results in a greater change of C_{max} than when C4 changes. When doubling the times of technological operations (factor C1), i.e. when comparing combinations K63 and K64 C_{max} increases by (28443-15198)/15198 = 87 %. On the other hand the influence of machine set up times (factor C2) on C_{max} is insignificant and, for example, when comparing K1 and K3 it increases by (15519-15212)/15212 = 2 %. The increase in the number of transport devices (factor C4) does not improve the C_{max} value (combinations K1 and K9, K2 and K10 as well as K17 and K25). Whereas a change in the transport operation times (factor C3) makes the C_{max} grow by (19661-15212)/15212 = 10 % when comparing K1 and K5.

Analyzing the obtained values for serial-parallel route we can see that for almost any combination m****m the C_{max} value <= especially low (comparison of combinations K1 and K31); while for almost any combination d****m is contained in the range very small < $C_{max} \leq$ small. In the PSR route almost all values of C_{max} do not exceed the value of C_{max}=60 000 min, and thus are below the linguistic value "very big".

We can see that for the K1 combination the C_{max} value in the PSR route is twice as low as in the PS route (15198,8/7298). Although for the K8 combination the C_{max} values are twice as low in the PSR route as in the PS one, when the L_{ut} number of transport appliances increases (factor C4) for K16 the C_{max} value decreases by (16074,9-13795,0)/16077,9 = 14 %, while in the PS route it remains on a similar level. Similarly, the increase of the L_{ut} number (factor C4) by 2 makes almost no change in the C_{max} value in the PS route, while in the PSR route C_{max} goes down by (7298-6839)/7298 = 6 %.

The increase of machine set up time (factor C2) for the examined data structure has a relatively small influence on the change of C_{max} (e.g. combinations K1 and K3, for which we have the values of 7298 and 7517 in the PSR route, and 15212 and 15519 in the PS route). The influence of factors C2 and C3 on C_{max} is stronger (combination K7) in the PSR route. In spite of the increase of transport operation times (factor C3) - and there are more transport operations in the PSR route - the C_{max} value decreases when the the L_{ut} number grows to 4 (combination K15) by (12112-8108)/12112= 33 %. Increasing both factors C3 and C4 at the same time (combinations K1 and K13) results in a relatively low C_{max} in the PSR route, i.e. by (7943-7298)/7298 = 8 % and a higher C_{max} in the PS route, i.e. by (18799-15212)/15212= 23 %.

Increasing the number of tasks L_z (combinations K1 and K17) results in a higher increase of the C_{max} value in case of the PSR route [(13218-7298)/7298 = 80%] than in case of the PS route [(19224-15212)/ 15212 = 26%]. For K21 and K29 the increase of factor C4 causes a relatively high decrease of C_{max}, i.e. by (23164-15516)/23164=33 %; for combinations K1 and K17 the increase of factor C5 causes an increase of C_{max} by (13218-7317)/7317 = 81 %. When comparing combinations K2 and K17 we can see that the increase of factor C1 results in a higher increase of C_{max} than the increase of factor C5 does for the PSR route, while the increase of factor C5 results in a significant decrease of C_{max} in the PS route.

In the analysis we can see that for combinations K52 and K64 we obtain a similar value of C_{max} at the "m" level of factor C3 and factor C4 (57153 min) and at the "d" level of factor C3 and factor C4 (56159 min). For combinations K56 and K64 changing factor C4 to level "d" causes a significant reduction of C_{max}, i.e. by (66225-56159)/66225 = 15%. When comparing combinations K48 and K64 we can see that a twofold increase of factor C5 results in exactly a twofold increase of C_{max} (56159/28104). A twofold increase of factor C6 causes C_{max} to grow by (56159-27464)/27464= 132573)/32573= 72%.

When analyzing the 2^6 factor experiment according to the values of the coefficient

$$R_{TPP} = C_{max}\,(\mathrm{PS}) / C_{max}\,(\mathrm{PSR})$$

we can see that for all combinations of factor levels of the type ****m* we obtain high values of R_{TPP} and inversely for the combinations of the ****d* low values. Factor C5 has the greatest influence on the C_{max} value.

Its increase causes a greater "parallelness" of production flow (e.g. when considering K48 and K64 we have, that factor C5 at the "d" level results in the R_{TPP} changing from 2.31 to 1.40). The other factors influence the size of the R_{TPP} relation significantly less.

Thus for the analyzed data structure the type of production flow can influence the C_{max} value changing the range of the R_{TPP} coefficient from 1.11 (combination K23 "mddmdm") to 2.36 (combination K14 "dmddmm"). Therefore there are combinations for which the C_{max} value is only slightly higher in the PS route than in the PSR route, but there are also those which result in a twice as high C_{max} value.

Fig. 1 shows the values of the R_{TPP} coefficient for the 64 combinations of factor levels (the 6-factor 2-level experiment).

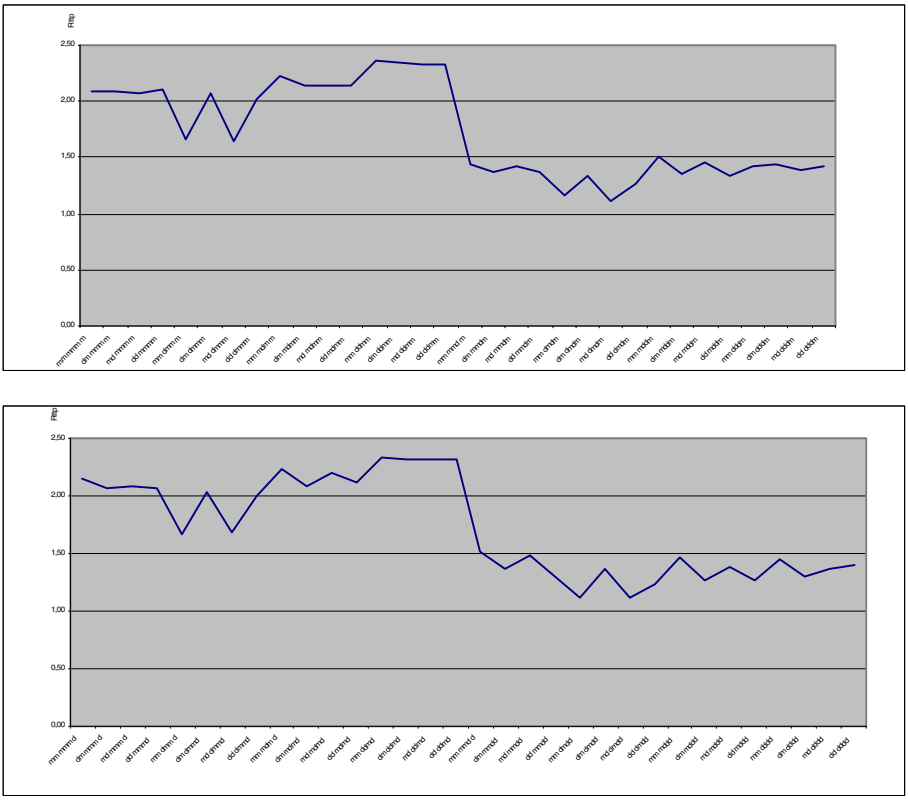


Fig. 1. The values coefficient R_{TPP} for the 64 combinations of factor levels (the 6- factor 2-level experiment)

5 Learning from Examples

The results obtained from the experiment analysis are the basis for formulating practical recommendations to be implemented in production processes. They can also serve as foundation for creating data bases and knowledge bases. For the evaluation of the process flow with regard to the C_{max} criterion we assume one of the following linguistic variables: ES - especially small, VS – very small; S – small; A – average; B – big; VB – very big.

Below we present some data representing the membership grade in fuzzy sets of especially small, very small; small, average, big, very big (quality of results) corresponding to some combinations of factors. For evaluation we use triangular membership functions [6]. The fuzzy sets unions can be considered as representing rules such as:
IF *E* THEN *H*
where *E* is the analyzed combination of factors and *H* is the fuzzy set union. E.g.

IF COMBINATION
TO(m)-TP(d)-TT(d)LU(m)-LZ(d)-WP(d)

THEN RESULT (C_{max})
(0.0 /ES + 0.8/VS + 0.2/S + 0.0/A + 0.0/B + 0.0/VB)
for serial route, and serial-parallel route.

IF COMBINATION
TO(d)-TP(m)-TT(d)LU(d)-LZ(m)-WP(m)

THEN RESULT(C_{max})
(0.0 /ES + 0.9/VS + 0.1/S +0.0/A + 0.0/B + 0.0/VB)
for serial route, AND
(1.0 /ES + 0.0/VS + 0.0/S + 0.0 /A + 0.0/B + 0.0/VB)
for serial-parallel route,

where the expression in parentheses is the fuzzy set union quality of the result.

There is a wide range of methods to be used for machine learning, e.g. artificial neural networks, Bayes classifier, *k*-nearest neighbors classifier, genetic algorithms, inductive learning, etc. Machine learning can be broadly classified into three categories: supervised learning, unsupervised learning and reinforcement learning.

In the paper [7] we present a few approaches to knowledge acquisition in solving problems of this kind, where inductive learning is applied.

Inductive learning methods are considered attractive for many real-life applications. Inductive learning constructs a description of a function from a set/output examples. An example is a pair (x, $f(x)$), where x is the input and $f(x)$ is the output of the function applied to x. The task of induction is to return the hypothesis h that approximates from a given set of examples. Then, the generates hypothesis is applied to the new examples to predict their class membership [5]. The selection of a suitable inductive learning method is a crucial phase in any system design methodology.

Nowadays we can choose from many inductive learning methods, ranging from linear discriminant and probabilistic approaches, to artificial neural networks and support vector machines, or even from a number of integrated design methodologies, where several identification methods are elicited to lead from a set of raw data to a full industry strong solution. Each approach behaves more or less well when applied to different problems in different domains. Machine learning techniques have been successively applied to the production problems.

6 Conclusion

In the paper the influence of the selected production factors for different routes in production flow discussed. Experiments with six factors have been tested and two levels of factors have been considered and compared. The algorithm have been implemented, tested, and illustrated with examples for the serial route and the parallel one. The results obtained from the experiment analysis make the basis for formulating practical recommendations for modeling and implementing production processes. They can also serve as a foundations for creating data bases and knowledge bases. From those bases, using data exploration methods (e.g. inductive learning) one can acquire knowledge for designing decision-making systems in production planning and management.

As a future perspective, it will be interesting to compare the proposed approach with the other methods and also to study the problem on a large set of benchmark data and real problems.

References

1. Law, A.M.: Simulation Modeling and Analysis. McGraw-Hill, New York (2007)
2. Banks, J., Carson, J.S., Nelson, B.L., Nicol, D.M.: Discrete-Event Systems Simulation. Pearson/Prentice Hall, New Jersey (2005)
3. Zeigler, B.P., Praehofer, H., Kim, T.G.: Theory of Modeling and Simulation. Academic Press, N.Y. (2000)
4. Witkowski, T., Antczak, P., Antczak, A.: Simulation Modeling and Analysis in Production and Logistics. International Journal of Factory Automation, Robotics and Soft Computing, 26–34 (2009)
5. Witkowski, T., Antczak, P., Antczak, A.: Application of GRASP procedure for production scheduling and its comparision with other methods. Journal of Automation and Information Sciences 37, 6(40), 35–40 (2005)
6. Witkowski, T., Antczak, A., Antczak, P.: Planning and Analysis of Production Systems Using Fuzzy Logic. International Journal of Factory Automation, Robotics, And Soft Computing (2), 88–95 (2008)
7. Witkowski, T., Antczak, P., Antczak, A.: Machine Learning – Based Classification in Manufacturing System. In: Proceedings IEEE International Workshop on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications, IDAACS 2011, Prage, September 15-17, vol. 2, pp. 580–585 (2011)

Foreign Direct Investment and Its Impacts on the Build-Up of Sustainable Development Ability of Regional Economy

Junhua Chen

School of geography science, Southwest University, Beibei, Chongqing, PR China
chenjhx@163.com

Abstract. Research on the impacts of foreign direct investment has been the theoretical frontier with the increasing impacts of foreign direct investment on globe economy. Based on the analysis of the effects on the development of regional economy and the research conditions of foreign direct investment, the article expounds the impacts of foreign direct investment on the buildup of regional economic sustainable development from six aspects, that is, structure, scale and technique of foreign direct investment. Based on the foregoing works, we bring forward regulation tactics and measures on foreign direct investment to promote sustainable development of regional economy.

Keywords: Foreign direct investment, regional economy, sustainable development, sustainability buildup.

1 Sustainable Development of Regional Economy: Ability and Build-Up

The ability of sustainable development of regional economy, is refers to the capacity which can promote economic growth and development of regional economic system in accordance with the objectives and requirements of sustainable development. It's the basic condition and reliable guarantee for realizing the economical sustainable development. In summary, sustainable development ability of regional economy has five basic contents, which are resources carrying capacity, regional productivity, environmental buffering capacity, advancement stabilizing power, and management & regulation capacity. Among them, the regional productivity and management & regulation capacity is its core.

The build-up of economical sustainable development ability is refers to the ability training and establishment of national and regional governments to formulate correct sustainable development goal as well as implement strategies for sustainable development. It's also refers to the capacity training and establishment of organize and optimize the economic elements of sustainable development resources to enhance development efficiency and effectiveness. The building process of economical sustainable development ability is not only a productive process of continuous development and improvement, but also a process of relation renewal and reconstruction between population, resources and environment, society and economy.

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 383–388.
springerlink.com

The sustainable development ability construction emphasizes on productive forces development, and attention to the improvement of resources carrying capacity, environmental capacity increases, and the balance of enhanced self-regulation is also focused on.

2 Foreign Direct Investment and the Build-Up of Sustainable Development Ability

The ability of regional sustainable development and its construction is one of the important content for regional sustainable development strategy. Although there are lots of research literatures of sustainable development, but very little is currently seen which are researches from the perspective of individual elements (such as foreign direct investment) to study its contribution to the ability of regional economic sustainable development. The capacity build-up for sustainable development has two ways, namely, the self-formation and external inpouring. The self-formation is a more ideal way, but it's not easy to obtain because of needing a long-term accumulation. Especially in the case of low economic level, poverty or poor ecological environment, the regional economy system is very difficult to form the corresponding capacity to promote the build-up of sustainable development. Therefore, the external elements inpouring through outside the system has become a key factor of starting a virtuous circle of capacity-building.

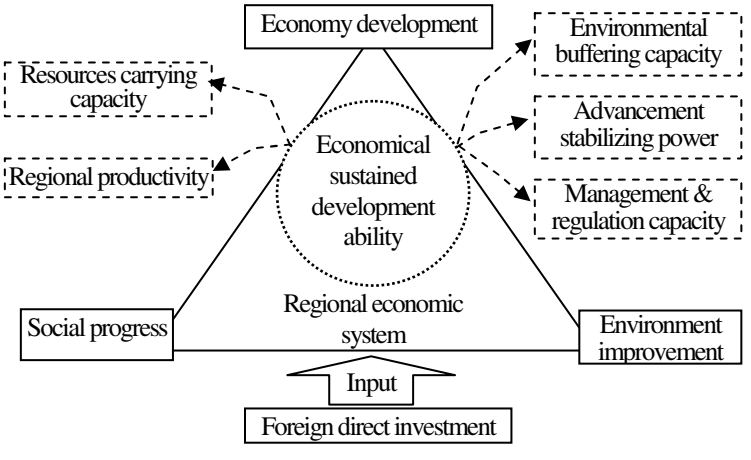


Fig. 1. Foreign investment and sustainable development ability build-up

As the main driving force of economic development for most of the world's countries, foreign direct investment is bound to have an enormous impact on the host countries capacity for sustainable development. We believe that foreign direct investment is just the external key factors to start the capacity build-up of regional economic sustainable development to form. Figure 1 has displayed the relations between foreign direct investment and ability construction of regional sustainable development.

We believed that the foreign direct investment is one kind of external force internalizes the influence to the sustainable development ability construction. Foreign direct investment needs to rely on the co-ordination of interior production element of regional economic to be effective. And this effect has both a positive side and negative side.

3 Effects of Foreign Direct Investment on the Build-Up of Sustainable Development Ability

3.1 Impacts of Foreign Direct Investment on the Carrying Capacity of Resources

Foreign direct investment scale's expansion causes the expansion of local production, also leads to the increase of resource consumption, thereby it increasing the pressure on the resource carrying capacity. In natural resource's development, Multinational corporations tend to develop the high-grade ore first, but give up the lean ore, which lead to a huge waste of resources. Along with the massively transfer of resource-intensive industries to the host country, the industry relies on resources and raw material much more. It inevitably accelerated rate of depletion of host country resources.

3.2 Influences of Foreign Direct Investment on Regional Productivity

Looking from the whole world foreign direct investment's situation, foreign direct investment flowed to developing host countries usually concentrates much in manufacturing industry and so on. Advanced production methods and technologies of developed countries promote host countries' regional production capabilities directly. Taking Asia "four dragons countries" as an example, its industrial structure's development have experienced labour-intensive industries primary in the 1960s, capital intensity industry vigorous in the 1970s, technology-intensive industry mainly in the 1980s. It is just consistent with the structure changes of direct investment and export of U.S. to Southeast Asia. It is also foreign direct investment flowed into China's manufacturing industry, which makes rapid development of China's manufacturing industry, thus China has become the world's manufacturing centre.

3.3 Effects of Foreign Direct Investment on Environment Buffering Capacity

Through Multinational corporations, developed countries massively transferred the resources-intensive, seriously polluting industries to the developing countries. And it has brought serious environment pressure to the developing country. At present, more than 90 percent polluting industries of developed countries have been shift to developing countries. These polluting industries bring massive industry waste material and trash, which has caused serious air pollution, water pollution to the host country, and aggravated the environment government difficulty.

3.4 Influence of Foreign Direct Investment on Advancement Stabilizing Power Construction

It is the focus threat of national economic security that host countries' industries may be controlled by Multinational Corporation, and the main control mode is the stockholder rights control. There are also technology, management, and human control. If there are no restrictions and interventions of foreign direct investment, for weak competitiveness of host countries, it is easy to form a multinational monopoly. This will lead to suppression of local industries, dependence on other countries' technology, lead to harm of the host country's macro-economic and even national sovereignty.

3.5 Impact of Foreign Direct Investment on Management and Regulation Capacity

Multinational Companies bring not only capital, but also to domestic enterprises with advanced management experience and international practice. It is certainly to improve local management level. Through the introduction of a series of management experience of Multinational Corporation such as advanced production management, quality management, sales and after-sales service management, personnel management, financial management, domestic enterprises enhance its management and labour productivity, and also its international competitiveness. The entry of Multinational Companies urge the host country to set up new law in many domains or to revise the old law such as company law, bankruptcy law, tax law, accounting and income export laws and regulations, etc. This is advantageous to the adjustment and reform of local economic management system.

4 Regulation Direction of Foreign Direct Investment on the Purpose of Economical Sustainable Development

Foreign direct investment for host country is a double-edged sword. It not only may be a great impetus on the local economic development, but also may have negative effects in a particular way. We may make the adjustment and control to the foreign direct investment, and cause it to contribute truly for the regional economies sustainable development.

4.1 To Make Sure Sustainable Development Is the Ultimate Objective

The introduction foreign direct investment is not the host country's final goal, short-term effects such as economic growth, trade expansion, increased employment and so on brought by foreign direct investment should not become the ultimate objective of host country also. Through the use of foreign direct investment to promote host country's sustainable economic development should be the fundamental purpose. Regional economies' sustained development emphasizes that foreign direct investment should meet the short-term effect with regional long-term development basic demand and benefit inter-coordination. It should achieve the long-term effects of sustainable economic development such as increased efficiency of resource allocation and utilization, accumulation and technological progress, and industrial structure optimization.

4.2 To Ensure Industrial Structure Optimization as the Core

The core of Sustainable economic development is just industrial structure adjustment, optimization and innovation. The emphasis of using foreign direct investment to promote sustainable economic development is bound to adjust and optimize industrial structure.

From the perspective of sustainable development, it must be clear, while foreign direct investment has improved the productivity levels of society, whether it leads to imbalance in industrial structure and regional distribution, whether hindered the normal upgrade process of their own national industrial structure, whether has achieved the comprehensive promotion of science and technology level to anticipated target.

4.3 To Assure Progress of Science and Technology Is the Key

Science and technology are primary productive forces, are the most powerful support force a country's economy to achieve sustainable development.

Through the introduction, digestion and absorption of advanced technologies of Multinational Companies, to promote progress of science and technology, it could be the key factor of using foreign investment to facilitate sustainable economic development. In the introduction of Multinational Corporation's advanced technology, we should choose those which have applicability and ease of digestion and absorption, especially those industrial technology of high efficiency, low energy consumption, light pollution. Thus, it can promote economic unceasingly change from the extensive production mode of high energy consumption, low output towards sustainable economic development mode.

4.4 To Be Sure Regional Coordination Is the Way

Sustainable economic development not only manifests in the inter-generational equity, but also required to achieve inter-regional equity. Obviously, the coordination of regional differences is not a foreign investor's objectives and tasks, but in the process of utilization of foreign investment, we can not ignore the issue of regional coordination.

It is the reality that due to the location preferences and the accumulative effect of foreign direct investment，it has aggravated regional disparities and imbalances. Therefore, to achieve the aim of using foreign direct investment to promote regional sustainable development, we must to give scientific guidance and regulation to foreign direct investment, making it an important way to promote regional coordination.

References

1. Yinxing, H.: Sustainable Development Economics. Commercial Press, Beijing (2002)
2. Yang, J., Zhang, Y.: China's foreign direct investment and industrial structure adjustment. Macroeconomic Management (July 2004)

3. Wang, H., Xie, X.: On the Connection of FDI, Regime and Economic Growth in Developing Host Countries. Journal of Northwest A&F University (Social Science Edition), 57–67 (2010)
4. Song, D.: A Study on the Effect of the Foreign Direct Investment on the Upgrading of the Regional Innovation Capabilities: Mechanism and Evidence. Soft Science, 62–65 (2009)
5. Wu, F., Yu, X.: FDI Utilizing Efficiency and Regional Economic Development Differences: from Perspective of Technological Progress. International Business, 70–77 (2010)
6. Manzocchi, S.: Foreign capital in developing countries: perspective from the theory of economic growth. Macmillan Press Ltd. (1999)
7. Zhang, H.-Y.: The Impact of FDI on the Independent Innovation Capacity of Chinese Home Manufacturing. Journal of International Trade, 72–81 (2008)
8. Li, N.: The Sustainable Development Stage Government Instructs Foreign Direct Investment Function Brief Analysis. Journal of Anhui Administration Institute, 44–48 (2010)
9. Moran, T.H.: Foreign Direct Investment and Development: the new policy agenda for developing countries and economics in transition. Institution for the International Economics, Washington DC (2000)
10. Blomstrom, M., Kokko, A.: Foreign Direct Investment: Firm and Host Country Strategies. Macllillan Press, Ltd. (2000)

Research on the Optimization Strategy of Normal Students' Teaching Behavior in Classroom*

Junhua Chen[1] and Yang Du[2,**]

[1] School of geography science, Southwest University, Beibei, Chongqing, PR China
[2] School of Chemistry and Chemical Engineering, Southwest University, Beibei, Chongqing, PR China
duy360@swu.edu.cn

Abstract. Classroom teaching is the main form of teaching activities and the basic component of primary and secondary education. Therefore, it is the main way to evaluate teaching quality and to reach the aims of education. The concept of the new curriculum calls for teachers to change classroom behavior. This article focuses on the normal students' classroom teaching behavior of geography teaching practice. By observing and analyzing their main teaching behavior in the geography class of high school, it will bring up corresponding strategy to normalize and optimize the main teaching behavior of geography "prospective teachers" for their problems of classroom teaching.

Keywords: Normal students, classroom, main teaching behavior.

Instruction

Classroom teaching practice is the main platform for development and training, regulating, optimizing geography teaching skills of prospective teachers' classroom behavior. Geography classroom teaching practice shows that during the practice of classroom teaching intern's behavior exist the following questions: ① Teaching language is not accurate and not enough scientific; ② The use of multimedia-assisted teaching, showing knowledge is too large, too much information redundancy; ③ The classroom blackboard writing capriciousness is arbitrary; ④ In the teaching process the teachers and students interact insufficiently, the classroom is quite boring ,etc. Its result is the teaching goal achieves is not high.

This research in the geography classroom instruction behavior observation's foundation, carries on the analysis to the intern geography classroom main teaching behavior, inquires about the reason, and take educates the theory as the basis, proposed that the standard, optimizes the intern geography classroom instruction behavior the raise strategy.

* This study was supported by the Southwest University Teacher Education Innovation Platform projects.
** Corresponding author.

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 389–394.
springerlink.com

1 The Observation Analysis on the Classroom Instruction Behavior of Geography Normal Students

1.1 The Type of Geography Teaching Behavior

Generally, the teaching behavior is refers to the teacher's all behaviors which to cause, to maintain and promote all students' learning. Professor Liangfang Shi considering that teachers' behavior in the classroom is divided into: Main teaching behavior, auxiliary teaching line and classroom administrative action. The main teaching behavior is refers to take the classroom instruction goal and the content as the orientation teaching behavior, it needs the teacher to have the essential specialized knowledge and the skill. The auxiliary teaching behavior needs the teacher to have certain classroom experience and the individuality accomplishment. The classroom administrative action mainly creates the condition for the smooth implementation teaching activity, also needs teacher's experience and certain skill. This article mainly uses Professor Liangfang Shi' theory about teacher classroom instruction behavior classification, carries on the observation analysis to intern's geography classroom main teaching behavior. The main teaching behavior may divide into the presentation behavior, the dialog behavior and the instruction behavior (Table 1).

Table 1. Classroom main teaching behavior classification and concrete way

Classroom main teaching behavior category	Teaching behavior concrete way	
the presentation behavior	Language presentation	Writing presentation
	Acoustic image presentation	Acts the presentation
the dialog behavior	Classroom inquiry	Classroom discussion
the instruction behavior	Reading instruction	Practice instruction
	Active instruction	

1.2 Geography Classroom Instruction Behavior Observation Content Determination

According to main teaching behavior type concrete way, we can definite behavior viewpoint (Table 2). Formulates the corresponding observation meter, unifies intern's geography classroom main teaching behavior to carry on the observation and the analysis.

1.3 Analysis of the Normal Students' Teaching Behavior in Classroom

The observation indicated that between the normal students has the big difference in the main teaching behavior way and the teaching request, may divide into outstanding, general and the bad three types roughly. The analysis discovered that the normal students' main teaching behavior in the geography classroom three aspect performance has the difference, forms this difference the substantial clause to lie in pedagogical lives own accomplishment, the different normal students the main teaching behavior which the foundation different displays to differ from.

Table 2. The main teaching behavior of normal students geography classroom observation table

Classroom main teaching behavior category	Classroom main teaching behavior observation content
the presentation behavior	How to explain? Explanation is valid or not? How does the blackboard writing present? Whether to have provided the help for the student? How does the media present? Whether suitable? Whether effective? How does the acts (Experiment, Demonstration, Manufacture) present? Whether standard? Whether effective?
the dialog behavior	How does the inquiry object, the type, the structure of questions, the cognition difficulty? Whether effective? Which auxiliary ways does the teacher have? Whether effective? Which topics have? How about the relationship of the topics and learning objectives?
the instruction behavior	How to guide students in independent study? Whether effective? How to guide students in cooperative study? Whether effective? How to guide students in inquiry study? Whether effective?

Below we further analyze from the classroom main teaching behavior's three aspects:

①**The Presentation Behavior.** The presentation behavior is mainly used in the teacher presenting the course content, the transmission teaching information behavior to the student. In the narration behavior aspect, the normal students basically can achieve the language to be smooth, but in the word usage expression's accuracy, the scientific nature is insufficient, the intonation lacks melodious, the rhythmical image is not strong, in the certain extent affects the teaching efficiency and the effect.

②**The Dialog Behavior.** The classroom instruction dialog behavior is the teachers and students take the language as intermediary, launches the mutual contact behavior regarding the classroom instruction content. Mainly includes the question and answer and the discussion behavior. What the normal students in the geography classroom instruction to appear are many are the question and answer behavior, but discusses the behavior nearly not to have. Some questions and answers are to realize the classroom teachers and students to interact to appear, has not undergone the careful design to ask the question the level to be not too high.

③**The Instruction Behavior.** Classroom instruction's instruction behavior is refers to the classroom instruction process the teacher to study the practice independently to the student the instruction. It includes instruction behaviors and classroom reading, work and activity so on .Very obvious may observe this in the pedagogical fresh classroom instruction, on the one hand the behavior are really few, some are only some to examine the student to study the knowledge whether to grasp feedback classroom practice. Teacher's instruction behaviors are not too many in the relation process, have not involved instruction classroom instruction behaviors nearly and cooperation study, inquiry study so on.

2 The Optimization Strategy of Normal Students' Teaching Behavior in Classroom under the New Curriculum Idea

2.1 The Optimized Basis of Normal Students' Teaching Behavior in Classroom

The classroom instruction behavior is and the related teaching idea relates. Certain teaching idea guiding corresponding classroom instruction behavior appears. The effective teaching is the goal which the teacher pursues, is also appraises the teacher teaching the basis. Teachers teaching students to solve major problems in behavior, reducing its outstanding and experienced the gap between geography teachers must help them in their teaching internship during the standard behavior. Requirements of the new high school geography curriculum "learning to live a useful geography," "necessary to develop a modern geographic literacy of citizens," "to meet the learning needs of students with different geographical," "emphasis on geographical issues to explore" other new curriculum ideas. Actually implement these ideas, the ultimate goal is to change behavior in the classroom teaching, which has been the lack of teaching experience of teacher classroom behavior of students put forward higher requirements.

2.2 The Geography Normal Students' Teaching Behavior Optimization in Classroom

2.2.1 The Optimization of the Teaching Presentation Behavior

①The Accurate Science's Narration Presents the Behavior

First, we should pay attention to science and geographical terms, science and strong. Students in the geography classroom teacher described the geographical terminology to describe geographic phenomena, to explain the law of geography, doubts about the geography and other things as well as geographical areas, the language must be accurate.

②The Discipline Characteristic Blackboard Writing, the Board Chart and the Board Picture Presents the Behavior

They may assist narrate the behavior, causes the content information which presents to have the coherence, the fresh perspicuity, moreover can also cause its key to play the emphasis prominently the role. Blackboard writing' writing is striking, is neat, the typeface and the size are suitable, uses the colored chalk appropriately, enables the blackboard writing to have the color change and the level contrast. The board chart and the board picture image concrete, is rich in the performance pleasant to the eye, artistic, the full reasonable use colored chalk becomes the chart, pays great attention the blackboard whole esthetic effect, enjoys, the relieve weary for the student by the artistic esthetic sense, to raise the efficiency.

③Choice at the Right Moment Audio and Video Presents the Behavior

The geography audio and video presents the behavior is refers to the teacher to utilize listens to, the visual media to study for the student provides the perceptual knowledge the behavior. We can selects the landscape picture, the television, media and so on self-made flash animation simulation demonstration, causes the student to observe the dynamic process which and evolution each stage the geography phenomenon evolves, can promote the student to understand effectively.

④**Chooses the Effective Body Potential to Present the Behavior**
The geography body potential presents the behavior is refers to the geography teacher presents the geography course content through own body position and the movement demonstration the behavior.

2.2.2 The Optimization of the t the Dialog Behavior

The classroom instruction dialog behavior is the teachers and students take the language as intermediary, launches the mutual contact behavior regarding the classroom instruction content, the question and answer behavior and the discussion behavior optimize importantly.

①**Attention Regulation Geography Question and Answer Behavior**
First, the inquiry must face all students. Second, reasonable choice and use inquiry way. Third, correct assurance question difficulty. In the question and answer behavior's question difficulty cannot have one's wish, strives for the difficulty control in the student "the recent developing area" on.

②**Face All's Geography Discussion Behavior**
Discussion's classroom production is very strong, how using the discussion behavior, to enable its to develop the positive function for the student speaking of the normal school which just mounted the platform to live variously the difficulty to be quite big. Therefore, in the geography classroom instruction should pay attention: First, prepares for the discussion. The teacher while prepares for aspect and so on skill which the discussion goal and the plan, the organization discuss to request to prepare for in discussion's role. Second, pays attention to all student, the encourage student participation positively. Must pay attention to student's individual difference, pays attention to beginning which the student studies, and carries on the communication from student's study angle embarking.

2.2.3 The Optimization of the Instruction Behavior

①**Read Guidance Specifically Requested**
The teacher when carries on the reading instruction needs to pay attention: First, proposed that is clear about the concrete reading request. Second, is the student provides the corresponding reading background knowledge.

②**Designed to Guide the Practice of Migration**
The geography teacher to student's classwork practice's instruction, needs to take steps from the below aspect: First, designs the concrete guide aim. Second, plans the practice duty and the step request. For example, after having studied affects urban the position factor, may design such question: "the utilization studies the knowledge, explained that Chongqing takes the super city to form the reason."

③**Building Cooperation Atmosphere Cooperation Instruction**
First, teacher should prepare for full. Second, teacher should build atmosphere. Third, teacher should regulate the process suitably. Patrols through the process, listens attentively and observes, for them suggested that resolves the contradiction and the question plan and the mentality; Encourages through at the right moment process, to urge the student to cooperate the active duty maintains with diligently; Through

measures to the group result looks up, supervision panel member's cooperation and diligently; Through to the group achievement's reward, drives the member and group's cooperation behavior. In the student groups the operation in the process, the teacher carries on the attention to student's operation activity, when the student meets the question or slightly has the achievement gives the corresponding direction or the appraisal, and encourages each group of schoolmates to unite the cooperation, forms between the group and group's competition and exchange consciousness, builds a good cooperation study atmosphere.

3 Conclusion

Classroom main teaching behavior regulated immediate influence geography classroom teaching effect. The normal students the geography classroom main teaching behavior the optimized research, to enhances the geography classroom instruction the validity and the raise fast growth geography normal students has the significant significance. Affects normal students the geography classroom instruction behavior standards to have many factors, therefore, to normal students the geography classroom main teaching behavior regulated research is also the process which has heavy responsibilities. This is also the researcher unceasingly direction diligently.

References

1. Zuo, X.: How to pay attention to the education theory in the teaching behavior the utilization. University Research (11), 182 (2010)
2. Gao, L., Zheng, W.: Discussion which transforms about the elementary education teaching behavior. Chinese Electric Power Education (35), 81–82 (2010)
3. Li, H.: In new curriculum teaching behavior transformation, vol. (21), pp. 56–58 (2010)

IT Outsourcing Business Process Model Innovation Based on the Theory of Change

Jing Mu and Yilei Pei

Beijing Union University

Abstract. This paper depicted the development of IT outsourcing at present, which proposes to utilize four different kind of change mode to IT outsourcing enterprise process including product developing process, service deliver process and so on. By analysis the case of IT outsourcing enterprise process, four key factors are found which affect on type of change mode during the different period of IT outsourcing enterprise developing. IT outsourcing enterprise process will increase changes, which requests IT outsourcing enterprise continuously to adjust and adopt new methods to manage business process.

Keywords: IT outsourcing enterprise, Process, Mode of change.

1 Introduction

With the increase development of Information Technology, demands on IT functions have grown in speed and flexibility of delivery, which request IT managers to do more with less combining with commercial benefit. In terms of this, gap occurs between IT provisioned locally and business needed, which represent a shortage in effort, time, skills or any of the micro components of the IT business process.

Business models are transforming in relation to request from the market; it is not only a question of wanting more. Provision of Internet business systems is an example here, as is the move towards solutions implementation rather than the more straightforward product delivery. It is critical to recognize both the gap between capability and business objective at the corporate level, and how this is tackled: “the sourcing strategy should build a bridge between the enterprise’s current status and capabilities and the desired future ability to fulfill businesses objectives” (Da Rold (2001)).

Reasons for outsourcing Analysts suggested three primary reasons for outsourcing:

1) IT improvement Quite clearly any idea of making tangible gains in this area must equate to an increase in systems performance.

2) Improved business performance and/or impact. If IT improvement is about local parochial gains, then to a degree outsourcing motives around business performances or impact could be seen as being independent of the IT function itself. Before moving on, perhaps I should clarify the definition of outsourcing is “the sub-contracting of IT provisioning responsibility for a specific and pre-defined technological deliverable, solution or product ”; this would encompass, for example, provision of help desk

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 395–402.
springerlink.com

function, off-size support for payroll application, or the delivery of an Internet presence. In terms of where outsourcing might be adopted with a business (and not IT) focus, we might consider examples such as collaborative e-business solutions where a third party provides an extranet-type facility for a cartel of related businesses, and software on behalf of the members of that community. Although this is an example of a business-driven outsourcing arrangement, it is still relevant to the IT manager because a) there may well be interfaces to and from local core systems, and b) even if there are not, the business will probably expect IT to manage the outsourcing relationships.

3) Commercial gain. The notion of outsourcing for commercial gain can, for some businesses, become incredibly broad. Rather than talking about outsourcing for particular projects or products, some companies have outsourced their entire IT capability, retaining few in-house IT staff-and these primarily managers to look after the third-party suppliers. This kind of arrangement is most often pursued as an extreme "do more with less" philosophy, with senior executives being lured by promises of multi-million-dollar savings against their traditional IT budgets,

Research on global computing services shows that there has been a massive growth in the number of large outsourcing deals signed in continental Europe, The Global Computing Services' contracts database tracks all outsourcing, integration, and consulting deals with greater values. These outsourcing contracts are not just exercises in cost reduction necessitated by economic problems, but they are also intended to advance the clients toward becoming real-time enterprises through the benefit of emerging real-time infrastructure and real-time delivery services.

Enterprise restructuring is expected to provide fertile ground for outsourcing in the business process outsourcing market segment. Some companies turn to outsourcing to foster change management during consolidation and integration. Consolidation, mergers and acquisitions result in integration needs for back-office processes, which are often met by outsourcing, divested companies need to grow entire back-office functions from scratch and look to external services providers to provider this process-management capability. Business process outsourcing includes enterprise services (human resources, finance and accounting, payment services, and administration), supply management (buying processes, storing processes, and moving processes) ,demanding management processes (customer selection, customer acquisition, customer retention, and customer extension), and operations. A typical business process outsourcing contract includes discrete project-based IT services, ongoing IT management services, and general process management (Gartner, 2004b).

Business processes within a company can be broken down into three categories: core; business critical noncore; and finally noncore, noncritical. Core processes are seldom outsourced, because they are the very essence of the business and the area that requires the most investment, Critical and noncritical noncore processes may be suited for outsourcing to a third-party supplier who will invest in them on the company's behalf. Process management has the highest expected growth rate in outsourcing. Business processing outsourcing is typically the outsourcing of a company's noncore or back-office business processes. Usually those processes are (or should be) IT enabled and hence can be transformed by the use a new or improved technology platform. The appeal of BPO is that it attempts to involve a new support

services model involving cost-effective, scalable, efficient services. The growth in demand for process outsourcing has also seen an expansion in the range of services being provided by suppliers, Processes typically outsourced include finance and accounting, procurement, human resources, and real estate (Honess, 2003).

2 Mode of Change

Challenges for the customer's IT organization, Outsourcing is just a different group of people doing something that, given the appropriate skills and resource, Managing an outsourcing suppliers is a different problem altogether, If asking internal teams to do something extra or make that additional effort, they are likely to comply because they are part of the organization, have company loyalties, know their efforts will be recognized at pay review time, and so on.

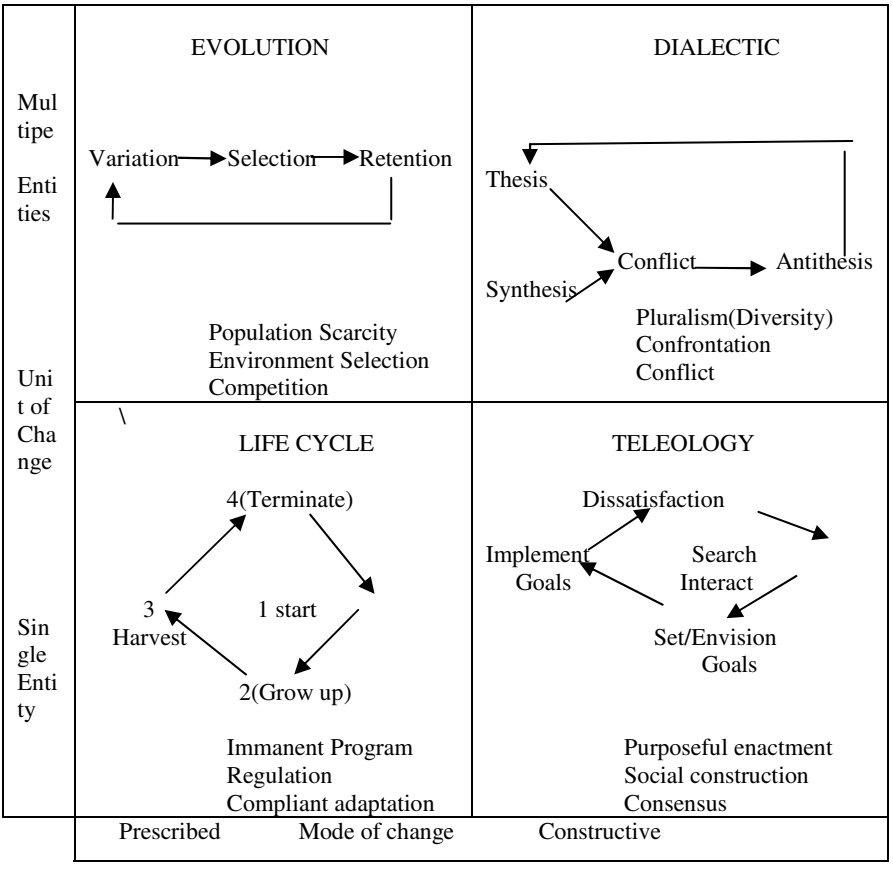


Fig. 1. Mode of change

Figure1 illustrated the four ideal type theories defined by Van de Ven and Poole(1995). As the cells of this figure illustrate, each theory views the process of development as unfolding in a fundamentally different progression of change events, and as governed by a different generative mechanism or motor, As table outlines, the theories can also be distinguished in terms of:

1) whether the end state of the process can be predicted from the outlet,
2) if the path of development is predetermined,
3) if the process is convergent or divergent,
4) if time is based on events or cycles, Variations of each of the basic theories will also be discussed.

A life-cycle model depicts the process of change in an entity as progressing through a necessary sequence of stages or phases. The specific content of these stages or phases is prescribed and regulated by an institutional, natural, or logical program prefigured at the beginning of the cycle. Change in a life-cycle theory tends to be morphogenesis, involving the progression from one stage to the next as the unit develops, While there may be continuous development within stages in a life-cycle theory, transitions from stage to stage involve a qualitative change in the unit and sometimes in the nature of the developmental process itself. As the name implies, time for a life-cycle theory is cycle: life-cycle models are comprised of repeating milestones that take the unit from inception to demise or fulfillment. Once the end of the cycle has been attained, the process is set to commence anew, with either the same or the different entity. There are three varieties of life-cycle theory, depending on whether the generative mechanism is regulated by natural, logical, institutional requirements, Sequences driven by natural or logical requirements exert a stronger determinism on the developing unit than do those driven by institutional norms and rules. They are adapted to different phenomena. Logical life cycles are founded in processes governing intangibles such as the development of ideas, decisions, and cultures or changes in language and symbols in an organization. Natural life cycles are more pertinent for tangible phenomena that are grounded in space and time, such as the life-cycle of organization or the growth of infrastructure.

A teleological process views development as a cycle of goal formulation, implementation, evaluation, and modification of actions or goals based on what was leaned or intended by the entity. This sequence emerges through the purposeful enactment or social construction of an envisioned end state among individuals within the entity. The theory can operate in a single unit or among a group of cooperating units who are sufficiently like-minded to act as a single collective unit. Because teleological processes are goal driven, the developmental path followed by the unit is not predetermined, but is generated by activities necessary to achieve the goal.

In dialectical models of development, conflicts emerge between entities espousing an opposing thesis and antithesis that collide to produce a synthesis, which in time becomes the thesis for the next cycle of a dialectical progression. Confrontation and conflict between opposing entities generate this dialectical cycle. Unlike life-cycle theories, the goal or endpoint of a change process is not clear at the beginning but emerges from the dialectical process. In some cases changes is driven by the conflict and contradiction itself. The developmental path of dialectically driven change is not predetermined. Units react to and cope with conflicts, contradiction, and tensions in

many different ways, and the resulting path will vary greatly from case to case. While basic moments of the dialectical process can be distinguished at a conceptual level.

An evolutional model of development consists of a repetitive sequence of variation, selection, and retention events among entities in a designated population. This evolutionary cycle is generated by competition for scarce environmental resources between entities inhabiting a population. The evolutionary motor drives change through the core process of variation-selection-retention (VSR). In the familiar explanation, variations in existing unit characteristics occur, and those that enable the unit to compete for scarce resources in the environment are selected for survival. Surviving units spawn others like them and retain the "blueprint" for competitive survival in the population. The VSR explanation operates at the level of the individual unit or organism and is the micro-level process by which populations of a species evolve and eventually or are extinguished.

Innovation diffusion is defined as the process by which an innovation is communicated through certain channels over time among the members of a social system. Four key elements determine the characteristics of the diffusion process of an innovation: innovation, time, social system, and communication channels. An innovation is any idea, object, or practice that is perceived as new by the members of a social system. Time relates to the rate at which the innovation is diffused or the relative speed with which it is adopted by members of the social system. The social system consists of individuals, organizations or agencies that share a common culture and are potential adopters of the innovation. Communication channels are the means by which information is transmitted to or within the social system (Ho.Ang,&Straub,2003)

3 Methodology

This paper will focus on responses from one case study of offshore service provider firm in China. The participant represented the global service offshore delivery firm in China ITO industry. The research team also decided to choose participant who has obtained capability maturity model and integration (CMMI) level 3 certification companies. In addition, the respondent is a Certified Outsourcing Professional in outsourcing field. Such informant definitely has a lot of experience and much more information to share for this in-depth insight. Initially, the researcher conducted in-depth interviews with five participants and decided to pick three of them to participate in the case studies to represent software development and service delivery outsourcing field.

Sourcing can be classified into the following four sourcing categories (Lacity et al,1996)

1) Total outsourcing: the decision to transfer IT assets, leases, staffs, and management responsibility for delivery of IT services from an internal IT function to an external IT provider that represents more than 80% of the IT budget.
2) Total in sourcing: the decision to retain the management and provision of more than 80% of the IT budget internally after evaluating the IT services market.
3) Selective outsourcing: the decision to source selected IT functions from external providers while still providing between 20% and 80% of the IT budget internally.

4) De facto in sourcing: de facto decision to use internals IT departments to provide products and services that arise from historical precedent, rather than from a reasoned evaluation of the IT services market.

This paper only one case study will be presented focusing on global offshore service delivery model of ITO. The researcher will not reveal the real name of the participant, as it is part of the ethics of doing research where the background of participant should be kept confidential. Therefore, SP1 (service provider 1) will represent global offshore ITO service delivery model for this case study. The case study was conducted over a few meetings and exchange of calls and emails for a few months with the subjects. The purpose of the case study is to gain as much information as possible about the participants' experiences in ITO process model. The case study was used as part of the data collection method for the actual research project. However, this article will only focus on the results related to process model innovation of SP1 respondent. The results were obtained through one case study which emphasizes on global offshore ITO service delivery model.

There are six phases in the IT outsourcing business process innovation, each phase is important:

Phase1: Vision

In an evolutional model, customers' needs changed, which creates a strategic vision of IT sourcing. The two main activities in this phase are identifying core IT capabilities and identifying IT activities for potential outsourcing. Typically, the customer senior business managers and customer senior IT managers are the primary stakeholders involved during this initial phase, which is first phase in the innovation process.

Phase2: Evaluation

The customer goal in this phase is to identify the best source for IT activities, The major activities during this phase include measuring baseline services and costs, creating a request for proposal, developing evaluation criteria, and inviting internal and external bids.

With dialectical models of development, We can observe that the customer IT users are primarily concerned with service excellence during the entire outsourcing evaluation. IT users sometimes question confidentiality and privacy of data with IT outsourcing, But in general, IT users typically support outsourcing because they perceive that suppliers, with their IT expertise, will increase service and provide new IT to the user community.

Phase3: Negotiation

In this phase, the intent is to negotiate a contact to ensure sourcing expectations that are realized. According to life-cycle model, the following activities may be included in this phase: conduct due diligence to verify claims in the request for proposals, negotiate service-level agreements for all IT services, create customer-supplier responsibility matrixes for all defined responsibilities, price all defined units of work, negotiate terms for transfer of employees, and agree on mechanisms for contractual change, including benchmarking, open-book accounting, no exclusivity clauses, and pricing schedules.

Phase4: Transition

The intent is to establish precedents for operational performance. For large contracts, transition activities may last from 18 months to more than two years. According to

life-cycle mode, there are eight activities in this phrase. They are distributing the contract; interpreting the contract; establishing post contract management infrastructure and process; implementing consolidation; validating baseline service scope, costs, levels, and responsibilities; managing additional service requests beyond baseline; fostering realistic expectations of supplier performance; and publicly promoting the IT contract.

Phase5: Improvement

The major activities in this phase include benchmarking performance, realigning the contract, and involving the supplier in value-added area.

Phase6: Mature

It provides an opportunity to learn from past experiences as well as to explore creative opinions when constructing a new deal.

The IT outsourcing process model that the firm operates is known as offshore delivery centre whereby the firm deploys standard global delivery model that most large companies deployed. Customers are in one country but delivery is from offshore, using telecommunication network offshore services. It is a very comprehensive delivery model. As explained by the informant, "It is based on four different components and essentially, the delivery model is based on four pillars, namely leadership, best practice, experience and human capital development."

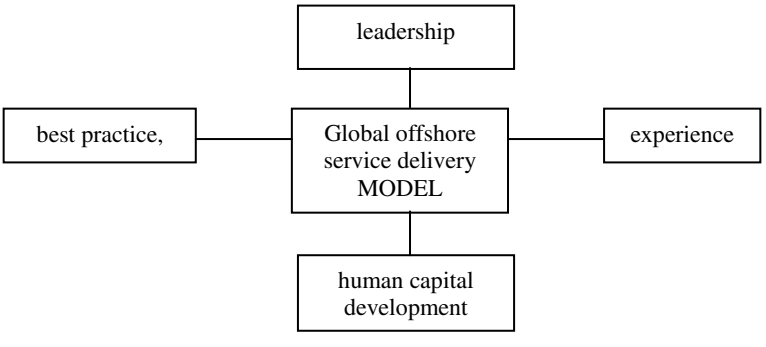


Fig. 2. Global Offshore Service Delivery Model Overview

As figure2 shows:

1) Experienced Leadership

Consultants are essentially people who have worked in multiple jobs industry such as they have been in the buyer side; supplier-side and some of them have been in the consultant-side. This type of multilevel experience in IT will actually bring comprehensive type of leadership at all levels of perspectives

2) Best Practice

To ensure global standards of delivery, benchmarking to the international standards are necessary. The firm establishes its delivery model and discipline which is complemented by certification such as capability maturity model and integration and ISO. With these `certifications, the firm could be compared equally with some of the large companies like IBM and others which have proper capabilities and competencies.

3) Experience
IT Outsourcing knowledge is a key to service delivery. This means that for instance, a service company, either IT company or other services, as long as the firm understands that the customer industry is important, and always keep abreast to the issues and development of the software then the firm is intact. In ITO, a firm is not offering a technology service but a customer service. Therefore, the firm must understand the country it is serving The firm needs to understand not only about outsourcing but also what goes on in that enterprise.

4) Human Capital development
The last component of the "pillars" is to have the right people and to grow them continuously. To develop an expert takes time. In fast-moving IT field, the environment changes rapidly and hence human capital also need to continuously upgrade themselves along with the changes.

4 Conclusion

The paper discusses soundly on the global service delivery model, many world-class ITO and BPO companies have established their bases or subsidiaries in China due to several advantages offered by unique Chinese positioning. Combined four kind of mode of change during the different period, IT outsourcing enterprise can adjust its process from product developing process to deliver service process and so on.

References

1. Aron, R., Singh, J.: Getting Offshoring Right. Harvard Business Review 83(12), 135–143 (2005)
2. Brandel, M.: Offshoring Grows Up. Computerworld (March 12, 2007), http://www.cio.com.au/index.php/id;1177022986;fp;4;fpid;51240
3. Van de Ven, A.H., Pools, M.S.: Explaining development and change in organizations. Academy of Management Review 20, 510–540 (1995)
4. Pools, M.S., Van de Ven, A.H., Dooley, K., Holmes, M.E.: Organizational change and innovation processes: Theory and methods for research. Oxford University Press, New York (2000)

American Brand Development of Citrus Industry and the Enlightenment to China Three Gorges Reservoir Area Developments Based on SRCT

Jian-hua Zhao[1,2], Yu-heng Zhao[3], and Si-ying Li[4]

[1] Research Center for Reservoir Resettlement, China Three Gorges University
butterfly_zjh@yahoo.com.cn
[2] College of Economics & Management, China Three Gorges University
[3] Yuanpei College, Peking University
[4] Yunnan Ankang Registry Safety Engineer Firm

Abstract. Florida ranks as the 4th largest state economy in the United States relying mainly on tourism, agriculture and transport instead of the heavy industries. Citrus are the predominant crops in Florida, accounting for about 70 percent of U.S. production. Based on the SRC Theory, we deduce the general principal of the citrus industry through the examination of citrus brand development in Florida and California, USA. After analyzing the SRC situation, we provide the suggestions to promote the citrus brand development for a better developing model in Three Gorges Reservoir Area.

Keywords: China, Citrus industry, Brand, Three Gorges Reservoir Area, SRCT.

1 Agricultural Brand Developing Characteristic Based on System Resource Constraint Theory (SRCT)

There are four characteristics of Famous brands, including High market share, High Reputation, Stable quality and supply chain as well as Sustainable innovation (2H3S). Since reformation and opening up in 1978, Chinese agriculture has maintained the characteristics of less input and return, smaller risk and scale, as well as lower technological added value. Therefore there are less China top brands in agricultural products. Development of thinking in the traditional agriculture influenced the support strength from government policy and transferring of industrial capital to agriculture.

We could deduce that the modern agricultural industry may be characterized by High investment, High tech, High risk, High income (4H) which need a large-scale economy to develop after investigating the brand development process of American citrus industry which also become the characteristics of Chinese modern agriculture system.

According to the theory of System Resources Constraint, the development of agricultural brand is constrained by system resources, including the limited planting area, capacity of transportation and quality assurance, seasonal changes of supply, market volume, demand concentration and rigidity and so on. It want of a lot of investment for Citrus industrial quality improvement, technological innovation,

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 403–408.
springerlink.com

industrial integration and brand promotion to get rid of the resource constraints, and investors must get the return from the market resource, otherwise obstacles will occur for continuous innovations.

2 American Citrus Brand Development

During the100 years of citrus brand development, there are at least five major technological and system innovations in the United States to breakthrough market and resources constraints for sustainable development.

Firstly, Emerging of industrialization and urbanization preliminarily impels brand development since 1850s; it boost the consumption of citrus fruit and push the large-scale producing leading to the foundation of quantity for brand development.

Secondly, railway system has connected the American domestic market. During the 1870s-80s, three railway lines were built which link California citrus with such eastern markets and unite American domestic market resources. American citrus commercial production started around in 1890, California Sunkist joint company arises at the historic moment. It made the citrus brand develop in the national market. Sunkist Company with consortium production integrates the scattered agricultural resources, which could clear the way for standardization and industrialization.

The thirdly, the market resources development promotes the citrus technological innovation to optimize the seed. America introduced orange seed from Brazil in 1870 as the fresh-eating fruit of California. Being big, seedless and tasty, it became very popular in the market.

The fourthly, system innovation integrates system resource. California joint company, Sunkist, was founded in 1893, which gradually built its famous brand by integrating produce-supply-sell chain, specializing production, intensively operating, managing by contract, diversifying competitions, and reversely integrated industrial chain and the agricultural resources with the help of market information, technology strength, and investment ability.

The fifthly, science and technology innovation further enlarge the global market. In the mid of 1940s, freezing enrichment orange juice industry first appeared in America, leading to a greater consumption market share which largely promoted the Florida citrus production. Florida Seagram, pure fruit juice producer, developed into the world's largest company.

According to the survey, America's Coca Cola and Pepsi food companies also extended their industry into the citrus juice industry. Relying on the original brand influence, the companies agglomerate global top team to make American citrus industry develop further and integrate international agricultural resources. It is the most important strength to develop the global resources, and also a threat to the industrial safety of developing country.

3 The Setup of Chinese Orange Brand Industry

China is the largest citrus-producing country in the world, with the annual output of 25 million tons and largest planting area, including Hunan, Hubei, Chongqing, Sichuan,

Guangdong, Jiangxi, Zhejiang, Fujian, Guangxi and other provinces. There are some regional brands of fresh fruits including GuangDongShaTangJu, JiangXiNanfengMiJu, GanNanQiCheng and FuJianMiJu. The citrus brands of China Three Gorges region competes in the marketplace with limited system resource. Processing and exporting of Citrus products focus on raw fruits, canned orange, jelly jam, non-frozen orange juice and so on, it is difficult to get the economical scale. Orange's target market of Three Gorges (Yichang) is in the North china, avoiding the competition with South Orange, and forming two-way logistics for northern Apple and Three Gorge Orange. Fresh fruit price is even higher than which most enterprises offer. Most factories had to use defective fruit for processing to reduce the cost. It is difficult to create a top brand for the citrus' industry, and size of processing industry has been gradually shrinking. When there are snow and other natural disasters, logistics of citrus get worse and processing capacity cannot match the entire citrus product. The vicious cycle prevent industrial capital from flowing to citrus industry. Building Three Gorges Citrus Logistics Center can only partially solve the problem.

According to the investigation, triple balance of powers exists in china to share the fruit juice market. The first include some enterprises which possess the background of Taiwan, such as Unity and Masterkong which exist in the market by the innovation of package and the excellent taste. The second are the famous enterprises within mainland including Huiyuan, Wahaha, and so on, which influence the customer's choice with brand effect. Huiyuan was once reported to be combined by Coca-Cola which shocked Chinese. The third are the multinational corporations such as Coca-Cola, Pepsi and so on, which occupy the major share by the new product. Now, many first-line beverage brands, such as Wahaha, Huiyuan, Nongfuguoyuan, TongYiXianChengDuo, are coexisting in the orange juice market.

The focus of competition in the industry of Chinese fruit production is transferring from downstream production chain to upstream raw material chain, and the ability to control the source is becoming a key element to win the game. And with the ever increasing transparency in the price of juice drinks, the profit of manufacturers and distributors is getting lower; seeking for new profit growth point becomes the urgent for the enterprises.

4 Comparison of American and Chinese Citrus Brand Developing Model and the Strategies of Development for China's

The most productions of orange in America come from California and Florida, the Sunkist of California corners the fresh fruit market, while the Seagram of Florida the juice producing field. The Coca-Cola and Pepsi-cola are trying to combine the traditional process and fruit juice producing together. Through the historical choice over the century, the huge companies gradually reversely conforms the chain of production. As a consequence, the manufactures have high abilities to invest, technological innovate, and marketing infiltrate. The loyalty and the credit of brand are of high level, which are very resources that Chinese agricultural industry want of, especially in orange production.

Therefore, we come up with the strategies of developing the brands as follows:

4.1 Unify the Mind, Raise Awareness to Strengthen the Security of Citrus Industry

The average juice consumption of Chinese will be 2.89L in 2015, close level of Japanese and Korean, if 20% of which is pure juice, that will be a great market resources, and many international magnates will infiltrate into Chinese market, in the situation that Chinese brands are not conformed they will take the chance to beat us one by one. To merger (to eliminate) Chinese brands like Huiyuan. So, we must unify our minds; raise the awareness to focus on the issue from a height of agricultural industry's security. Once foreign companies control Chinese brands, Chinese market is controlled. On the upper reach, it can control agricultural resources, so that many special policies for developing agriculture that government gives become the fund of profits for foreign groups. On the lower reach, it can deprive consumer's right of choosing, which may cause the crisis of orange industry even the whole agriculture. In the process of discussing with the American experts, they think that American brand come into China is a natural process of competition, but according to our research, the fight for the control of global agricultural industry has influence on the leadership in the negotiations of farm produce. The Central government should pay great attention on it.

4.2 Issue Policies to Improve the Environment for the Sustainable Develop of Agricultural Group

The ever worsening environment for investment makes many enterprises to give up the food industry and other manufacturing industry, while the brand like Huiyan may be swallowed by the international strategic investors. Commercial office should not only publish the antitrust laws to supervise the illegal merger, but also improves the operating environment of fruit industries. Supporting the Chinese brand to buffer the impact of international brand. Taking advantages of the strong abilities to invest, scientific innovate, and the rich human resources and market resources of famous producing group to conform the resources in reverse, china agriculture may develop the modern system.

4.3 Innovating System and Combining the Resources to Break the Constraint of System's Resources

By perfecting the Rural Economical Cooperation Organization, combining and adjusting the field resources, transferring status of fruit farmers, companies can extend the area of orchard and the portion of the market which they rely on, and develop in the process of breaking the bound of resources.

4.4 Standardizing the Evaluation of Top Brand to Enlarge the Effect of Policies

In the original evaluating system, the famous brand is for a certain product. Otherwise, in America, the award of national quality administration will be given to companies that have a great business achievement, so the brand of those companies can naturally get high loyalty. This system benefits the building of agricultural brand. There are rare agriculture brands in the china top brand during 2001 to 2007; the

system of Chinese famous brand is paused, but the Chinese Famous Agricultural Brand is still in the shadow of government. Besides the top ten of the famous fruit in China, and other reputations honored by all kinds of apartments make the customers puzzled. We advise that China should promote the national standardization to evaluate business achievement, and to promote agricultural brands.

4.5 Building the Modern Citrus Producing System with the Brand Strategy

Relying on the agricultural exploring group, Government may take the agricultural brand strategy, program and adjust the citrus brand across the country, form the modern industrial system marked with "the high-invest high-science high-hazard high-profit" thus build the new countryside.

4.6 Build the Sanxia's Orange Brand and Promote the Employment of Immigrants

Three Gorges area is a famous major place for orange's producing. Now the Coca-Cola, Pepsi-cola and the Huiyuan juice producer come into Zhongxian, Wanzhou, Yichang and so on. The Bronfmen citrus Chongqing Co., Ltd invests to develop the orange industry of Three Gorges area in the long run. China should attract companies to exploring orange industry so as to employ more immigrants of CTGP, and pay attention to the industrial security at the same time. We should put the factory of process into Yichang's industry developing area instead of in the Reservoir area following to the new concept of "separating industry and agriculture areas in suitable place". Taking advantages of the immigrant's preferential policies and using the Post Three Gorges' training project to extend the job training of immigrants to promote the brands development. While solving the problem of unemployment for the immigrants, it is possible to take the advantages of planting oranges to protect Three Gorges's environment, and to explore a new way of agricultural economic developing to build Three Gorges' eco-defense system.

4.7 Innovation the Three Gorges' Immigrant Policy

Three Gorges' immigrant later developing and training policy should pay attention to supporting the orange industry's building, using the brands to strengthen the achievements of industrialization. Taking advantages of industrialization to promote the build of standardize station .Thus making the immigrant's training regular .Finally, combining the immigrants' income, the Three Gorges' eco-defense's building, and the return of enterprises in the building of famous brands.

Acknowledgements. The research(2008KYM18) is Supported by Research Center for Reservoir Resettlement, Key Research Center for Humanities and Social Sciences in Hubei Province (China Three Gorges University).

The research is also based on the research project: System Resource Constraint Theory and Central China Regional Development (2009y023) supported by Hubei Education Bureau.

Reference

1. China top brand strategy promotion Committee: china top brand product management method (2001)
2. Zhao, J., Guo, Q.: System resource constraint theory and the practice. China Contemporary Economics press (2007)
3. Zhao, J.: Combining the agriculture brand resource to promoting modern agriculture development. CPPCC of Yichang (2008)
4. HeJin, Q.: Compare and Reference of China-foreign citrus industry Development. Economic Review 2 (2009)
5. Zhao, J.-H., Li, Y., Ye, J., Liu, Y.: China top brand strategy and innovation. Chinese Business Review (May 2007)

Author in Brief:
ZHAO Jianhua, Professor, PHD,mainly engaged in resource economics and management, regional economy. E-mail: butterfly_zjh@yahoo.com.cn

A Secure Machine to Machine-Based Framework for Service Provisioning in Cloud Computing Infrastructures

Giovanni Cagalaban[1], Jae Young Ahn[2], and Seoksoo Kim[1,*]

[1] Department of Multimedia Engineering, Hannam University, Daejeon-city, Korea
[2] ETRI Standards Research Center, Daejeon-city, Korea
gcagalaban@yahoo.com, ahnjy@etri.re.kr, sskim0123@naver.com

Abstract. M2M pose unique security challenges as the automation and lack of human intervention makes the network vulnerable to a variety of potential attacks and inadaptable to current situation where conventional security solutions are infeasible. Our framework comprises of security components which achieve certain level of security that enables the M2M systems to interface with intelligent devices that are deployed in customer facilities to sense real-world conditions and control physical devices, relay data from the devices in the customer premises to a centralized data center, and analyze the data and trigger business events. We also present performance results of our framework and show that our framework is capable of scalably and reliably handling concurrent events generated by different types of M2M devices, such as RFID tags, Zigbee sensors, and location tracking tags to achieve a high degree of security.

Keywords: M2M, Service Provisioning, Cloud Computing, Security.

1 Introduction

Cloud computing improves service delivery by applying engineering discipline and economies of scale through Internet architecture where users access the computational resources using a computer, notebook, tablet, smart phone, or other network devices. This computing paradigm shifts the location of the computing infrastructure to the network to reduce the costs associated with the management of hardware and software resources [1]. Cloud applications are provided and managed by the cloud server and data is also stored remotely in the cloud configuration and on-line services may be offered from a cloud provider or by a private organization. Cloud computing is capable of providing ease of use, on-demand self service, location independent resource pooling, dynamic provisioning capabilities, scalability, performance, reliability, virtualization, and pay per use services. In addition, service management capabilities are critical for dynamic provisioning and managing resources within the cloud.

* Corresponding author.

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 409–416.
springerlink.com

The recent convergence of Internet and wireless communications for cloud computing have become a premise for the machine to machine (M2M) technology. M2M communications involves the automated transfer of information and commands between two machines without human intervention at either end of the system [2]. M2M involves low-cost, scalable and reliable inter-machine interaction via wireless communication standards like GSM, GPRS, WLAN, Bluetooth, and Zigbee technology [3][4]. With M2M technology still at its early stage, wireless service is one of the many important links in a machine to machine deployment chain.

M2M technology is capable of building wireless M2M ecosystems covering a wide range of applications for a professional and personal everyday life. With increased processing power, it would enable to jointly deliver federated cloud services to users that fully leverage the power of cloud. With its capability of capturing and analyzing the massive amount of data available in all kinds of smart devices, M2M is a business concept used for automatic transmission of data from remote sources by wired, wireless, radio, and other transmission technologies.

Despite the technology existing in its various forms, M2M comprises of a number of separate technologies that needs to be integrated with cloud computing in the appropriate manner to enable a broad market deployment. While providing reliability, value chain, and cost optimization, deployment of M2M technologies by M2M innovators is hounded by challenging privacy and security issues. For instance, without the use of passwords or PIN codes, it poses security challenges when deploying M2M technologies as it has less possibility to control changes compared to a more traditional Human to Machine (H2M) technologies. Also, if the data traffic over GPRS from M2M devices becomes too large there is a risk of blocking out other data services and even normal phone calls as it can be susceptible to attacks by malicious hackers. In the provisioning of M2M services, privacy and security is a requirement for a reliable and secure M2M communication.

In this paper, we demonstrate the integration of M2m technologies with cloud computing to address security issues. Specifically, we define a secure M2M technology for service provisioning, data acquisition, and transmission for cloud computing. This secure service provisioning aids in providing massive amounts of M2M services through cloud computing environment.

2 Related Works

The high technological advancement in the field of communication and computation fueled by wireless mobile networks and sensor networks have allowed the development of a new technology called machine-to-machine communications [5][6] which has recently received considerable attention. Although still in its early phase, M2M is a promising technology. M2M networks make use of multi-hop routing in order to route data in a wireless network. Some works which treat the problem of energy by reserving the network connectivity are done by [7][8]. The growing interest in sensor applications has created a need for protocols and algorithms for large-scale self-organizing ad hoc networks, consisting of hundreds or thousands of nodes. Although M2M networks do not only consist of sensors, wireless sensor networks (WSNs) are key components of M2M communication that sometimes sensor networks are referred to as M2Mnetworks [9].

WSNs have been the topic of considerable research effort due to their potential for civilian and military applications and their ability of being incorporated in M2M networks.

There have been previous surveys on the characteristics, applications, and communication protocols in WSNs [10][11] which addressed several design issues and techniques for WSNs describing the physical constraints on sensor nodes, applications, architectural characteristics, and the protocols proposed in all layers of the network stack. Despite the keen interest in M2M and great value in building such a system, M2M is still relatively new and the technology faces several significant challenges. The challenges are wide ranging. In order to minimize the risks, many modern M2M applications pose complex design and software challenges and therefore demand a pre-integrated and well-tested software solution [12]. However, not much research has been done to address security challenges in M2M systems.

3 Machine to Machine Technology

M2M systems consist of wireless sensor networks connected to the outside world through the Internet. Devices are equipped with heterogeneous wireless sensors that can monitor behavior, conditions and can interface with virtually any type of mechanical, electrical or electronic system for an unlimited number of specific applications, which include access control and security, vehicle tracking systems, home automation systems, automotive systems, robotics, and medical systems. A wireless sensor network is composed of the sensors and their local interconnections, the gateway to the external world, a transport network and a service platform that handles the data and supports applications and users [13]. M2M technology is primarily a combination of various technologies such as wireless sensors, Internet, personal computers, and software technologies. In M2M, a field node or a group of field nodes gather data and send it wirelessly through a network where it is routed, often over the Internet, to a server or cloud of servers. A typical M2M technology for cloud computing which comprises the following basic components [14][15] is shown in Figure 1.

Fig. 1. M2M technology

M2M system for data acquisition and control belongs to the class of distributed, heterogeneous, network systems for data collecting, data processing and process control. It provides following features for remote data monitoring and control of subsystems, communication and control interfaces to industrial microcontrollers integrated into machine network system, building data acquisition tracking systems, cooperative task processing and evaluations between subsystems, building of heterogeneous networks based on wireless and wired communication technologies, and service of different type of embedded devices using wireless and wired interfaces.

Providing reliable services is complicated by the fact that different parts of the network are provided by different entities in the cloud. A service with multiple entities means that there must be several service level agreements (SLAs) between the different operators, the service providers and the customers [16]. This implicates that a service provider's ability to fulfill an SLA is dependent on several sub-providers ability to fulfill their SLAs, and they will probably be dependent on further SLAs.

3.1 M2M Service Provisioning

The service provisioning in M2M systems are implemented by several design strategies which include the push strategy, pull strategy and hybrid strategy. In this paper, we will focus on the implementation of the push strategy of the M2M infrastructure. To implement the push strategy for M2M communication, intelligent devices initiate the communications and sends data through an M2M domain over the cloud to a remote user. In this operation, the device recognizes predefined conditions and triggers itself to send alarms, alerts, e-mails, data and commands to an M2M domain. Subsequently, the gateway connects over the network through the Internet to send data to the remote users as shown in Figure 2. For instance, if the patient's medical information being monitored inside the house such as body temperature suddenly rises above normal level, an intelligent sensor device sends a message via the M2M domain to physician's cell phone. In this scenario, the gateway connects over the Internet only when the body temperature level is not normal.

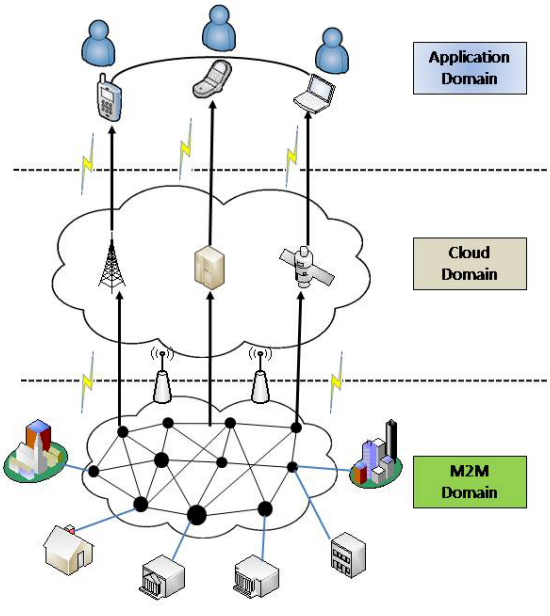


Fig. 2. M2M push strategy

3.2 M2M Service Provisioning Challenges

Generally, M2M service provisioning challenges are in the areas of security, privacy, reliability, robustness, latency, cost-effectiveness, software development and standardization. Security is one of the most important considerations while designing an M2M system, as the users do not want the hackers to break into M2M applications designed to control, for example, building security, environmental monitoring, vehicle tracking, etc. In order to prevent possible security violations, the most appropriate communication techniques must be used, because different types of communication techniques present different encryption and security features. For instance, Ethernet technology does not provide encryption and, it can provide only limited security with the use of a firewall. Thus, shown in Table 1 are the possible attacks and vulnerabilities in M2M service provisioning using Ethernet technology. Cellular operators, on the other hand, provide encryption and access authorization to data sent over the network. Reliability is another important issue. The intelligent devices used in an M2M network should be reliable by means of availability. Energy-efficient sensors and techniques must be developed to allow these devices to communicate over short distances using less power, or over long distances using line-powered bridges so that the battery energy is utilized more efficiently, for longer continuous operation without maintenance. Latency is also a concern in many M2M applications. For example, in the case of intruder detection, an alarm that cannot be sent on time can be useless since the intruder can leave the vicinity of event by the time that the alarm is received.

Table 1. M2M service provisioning attacks and vulnerabilities

M2M attacks	Secrecy	Privacy	Service Provision	Encryption Policies
Man-in-the-middle attack	No	No	Unreliable service provisioning	M2M service without encryption
Cipher-text attack	Yes	No	Invalid ciphertext causes incomplete service provisioning	M2M service encryption using private keys
Chosen plaintext attack	Yes	No	Service provisioning cannot be completed without encryption keys	M2M service encryption using permanent keys
Known plaintext attack	Yes	No	Service provisioning cannot be completed without encryption keys	M2M service encryption using temporary keys

4 Threat Analysis

In a typical fashion, system failures are classified according to the threat they pose to the M2M system. The threat T of a system failure is defined as the consequence C such that a system failure will be multiplied by the probability of occurrence P of the system failure for a specific component *i*.

$$T_i(x) = C_i \times P_i \quad (1)$$

Here the probability P of the entire system failing with the probability p_i for a component i failure. Assessing the security of a system is defined by the security of the individual components, so increasing security can often be done by introducing redundancy into the system. It is, however, not as trivial as it seems to utilize redundancy in order to improve the security of a system. Redundant components are an extra cost that needs to be justified in some way. There may not be sufficient will to invest in redundant resources, if the expense incurred by a failure is less than the cost of preventing it. A thing to note is that the failure rate of the individual components should be considered when introducing redundancy in a system.

The level of dependability of a system is in the end a trade-off between cost and how reliable the service needs to be. In defining the security features of the M2M system, reliability is a requirement. The reliability of a system has to with the quality of its measurement in providing uninterrupted service [17]. The reliability function R(t) of a system is defined as:

$$R_i(x) = P(T_i > t) = 1 - F(t) \tag{2}$$

Scalability is also another requirement as it provides desirable attribute of the M2M system. The concept connotes the ability of a system to accommodate an increasing number of elements or objects, to process growing volumes of work gracefully, and/or to be susceptible to enlargement [18].

In this security framework, we consider the hardware devices as wireless sensors, mobile devices, access network, and service platform for the reliability. Other components are assumed to be fault-free. The wireless sensors have a failure rate of *fws*, the access network has a failure rate of fan and the service platform have a failure rate of *fsp*. The security function can be defined in the following:

$$S_i(x) = R_i(x) - e^{-(fws + fan + fsp)} \tag{3}$$

The security function for the M2M system is defined by the reliability and the failure of network devices. Important to notice is that this implies that a component with a very large failure rate will be dominating the security of a system. If we assume that the failure rates of the sensors are several orders of magnitude larger than those of the access network, it will in essence be useless to introduce a redundant solution for the access network if the devices continuously fail.

5 Implementation and Evaluation

To illustrate the scenario of M2M communication, we implement the use of using wireless sensors and mobile devices to provide IP connection to the M2M smart application where the patient's prescription bottle can send updates on when the medicine has been consumed and also provides alerts to the patient and the monitoring physician. The monitoring of patients will request very high reliability from any service. This means that the service availability have a high value. Table shows details of some aspects of patient monitoring application. It is noteworthy that the worst-case availability offered by the M2M platform per device is actually lower than the stated requirement for the monitoring of patients. In addition, we also need to take into consideration network elements that may fail and consequently lower the availability of services. 95% might be achievable service

availability on average, but not per network devices. If this system is to use the public GPRS network in competition with mobile phone users and other M2M services the problems might be even worse. One way of solving this is by using for instance a fixed connection from the house. This will however lead to problems when the patient wants to get out of the house. These are aspects that need to be addressed if such a service is to be launched. Some way of enhancing the reliability for such services is needed to provide reliability as well as security.

Table 2. Patient Monitoring Details

Service Reliability	98.5%
Service Availability	95%
Response Time Average	4 seconds
Number of patients	100
Message Interval	2 minutes
Transmission Capacity	2 MB

6 Conclusion

Machine-to-machine communication is associated with the automated connectivity of remote machines through the Internet. Successful adoption of M2M technology requires a strategic approach to ensure that the technical solution is balanced with the business case to demonstrate an early return on investment. Simple solutions work best where the technology can be proven quickly and the benefits easily understood. There are many challenges in the successful adoption of M2M technology. In this paper, we provided an understanding of M2M, discussed the security issues, and propose a secure framework to a successful deployment of a resilient M2M solution.

The suggested system architecture gives possibilities for building more intelligent and autonomous wireless M2M system. It resolves the security issues in communication and control problems between different in technical characteristics machines that make them part of global Internet network. The secure software framework allows systems to function in different application domain. Providing reliable services is complicated by the fact that different parts of the network are provided by different entities in the cloud. In the future studies, we aim to extend our security framework to our scheme to adopt block-cipher encryption algorithms to provide higher security strength for service provisioning.

Acknowledgements. This research was supported by the ICT Standardization program of KCC(Korea Communications Commission).

References

1. Hayes, B.: Cloud computing. Communications of the ACM 51, 9–11 (2008)
2. Brown, A., Moroney, J.: A Brave New World in Mobile Machine to Machine (M2M) Communications. Strategy Analytics, Forecast and Outlook Snapshot (2008)
3. ZigBee Alliance, ZigBee Overview, `http://www.zigbee.org/documents/ZigBeeOverview4.pdf`

4. Milanov, S.: Bluetooth Interface, http://www.comexgroup.com/communications/bluetooth.htm
5. Lawton, G.: Machine-to-machine technology gears up for growth. Computer 37, 12–15 (2004)
6. Brazell, J., Donoho, L., Dexheimer, J., Hanneman, R., Langdon, G.: M2M: the wireless revolution. Texas State Technical College Publishing (2005)
7. Ye, M., Li, C., Chen, G., Wu, J.: An energy efficient clustering scheme in wireless sensor networks. Ad Hoc & Sensor Wireless Networks 3, 99–119 (2006)
8. Pandana, C., Liu, K.: Robust connectivity-aware energy-efficient routing for wireless sensor networks. IEEE Transactions on Wireless Communications 7, 3904–3916 (2008)
9. Kwok, R.: Use of sensors and radio frequency ID (RFID) for the national children's study, Technical Report, RTI International (2004)
10. Akyildiz, I., Su, W., Sankarasubramaniam, Y., Cayirci, E.: Wireless sensor networks: a survey. Computer Networks 38, 393–422 (2002)
11. Tilak, S., Abu-ghazaleh, N., Heinzelman, W.: A taxonomy of wireless micro-sensor network models. ACM Mobile Computing and Communications Review 6, 28–36 (2002)
12. Gilani, S.: The promise of M2M: How pervasive connected machines are fueling the next wireless revolution. White Paper, Mentor Graphics (2009)
13. Ferreira, J., Roque, R., Roadknight, C., Foley, J., Ytterstad, P., Thorstensen, B.: Sensor Telcos - new business opportunities. Technical report, Eurescom (2006)
14. Huff, M.: M2M device networking: components & strategies. Technical Report, MSI Tec., Inc. (2007)
15. Boswarthick, D., Mulligan, U.: M2M activities in ETSI. In: SCS Conference - Sophia (2009)
16. Ouelhadj, D., Garibaldi, J.M., MacLaren, J., Sakellariou, R., Krishnakumar, K.T.: A Multi-agent Infrastructure and a Service Level Agreement Negotiation Protocol for Robust Scheduling in Grid Computing. In: Sloot, P.M.A., Hoekstra, A.G., Priol, T., Reinefeld, A., Bubak, M. (eds.) EGC 2005. LNCS, vol. 3470, pp. 651–660. Springer, Heidelberg (2005)
17. Wei, G.: Reliability Evaluation Method of Tactical Communication Network. Acta Electronica Sinica 1 (2000)
18. Bondi, A.: Characteristics of scalability and their impact on performance. In: Proceedings of the 2nd International Workshop on Software and Performance, pp. 195–203 (2000)

Study on 4-Step BASE Mechanism Vulnerability Analysis and Design of Evasion Attack Method for Smart Network

Sungmo Jung[1], Jong Hyun Kim[2], and Seoksoo Kim[1,*]

[1] Department of Multimedia Engineering, Hannam University, Daejeon, Korea
[2] Electronics and Telecommunications Research Institute, Daejeon, Korea
sungmoj@gmail.com, jhk@etri.re.kr, sskim0123@naver.com

Abstract. Although cyber attacks are becoming smarter and use more advanced technologies, existing security mechanisms cannot promptly detect or respond to such attacks. For example, IP spoofing is effectively used to attack a security model using authentication process based on IP address or host name and, there is a large number of measures to detect such an attack. However, a malicious user can make a bypass attack by analyzing vulnerabilities. This study, therefore, has analyzed the vulnerability of 4-step BASE mechanism exploited by a bypass attack.

Keywords: Cyber Attack Detection, IP Spoofing, Evasion Attack.

1 Introduction

Recent cyber attacks on the Internet are becoming smarter and more intelligent using advanced technologies, increasing a threat to information networks of growing national importance.

Some of the current trends in cyber attacks are DDoS (Distributed Denial of Service), which paralyzes services of a specific commercial server employing a large number of attacking agents, and use of domestic networks by overseas hackers as a roundabout route. In this way, cyber attacks are increasingly used as powerful means of crimes.

DDoS attack is based on IP spoofing. Here, IP spoofing refers to a technique of creating a fake source IP address when a packet is sent. Using the method, attackers can hide their information and avoid detection, making IP traceback difficult.

Although we can prevent fake packets through Ingress Filter[1] and RPF(Reverse Path Forwarding)[2], such a method should be applied to a large number of routers. Though other methods have been developed over the past a few years, they are not fully employed due to complicated mechanism and degraded network functions. On the contrary, BASE mechanism [3] is relatively simple and do not slow down network functions.

In this study, we suggest a bypass attack of the BASE(BGP Anti-Spoofing Extension) mechanism and the vulnerabilities.

* Corresponding author.

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 417–422.
springerlink.com

2 Related Study : IP Spoofing

The defect of IP protocol was mentioned by Robert Morris [4] in 1985, and Kevin Mitnick [5] further developed the theory in 1995, which allowed him to attempt hacking attack and even arrested. Since then, the technique he used called IP spoofing, and the vulnerabilities of IPv4 have been exploited by hackers.

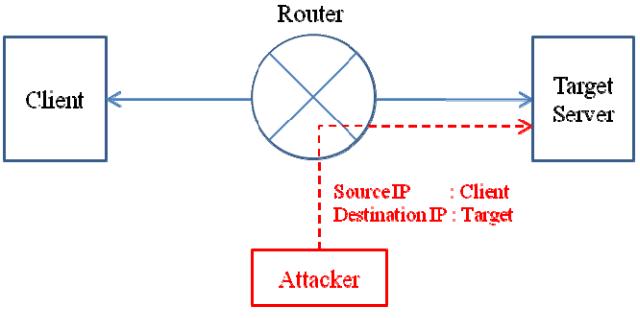


Fig. 1. Packet Transform using IP Spoofing Method

An attacker attempting IP spoofing uses the vulnerability that a user connected to the Internet can send a fake IP packet. A receiving host identifies a sending host only with the IP address contained in a packet so that the receiving host cannot know exactly from where the packet has been sent. In general, the structure of IP packet, sent/received between two computers, is as follows.

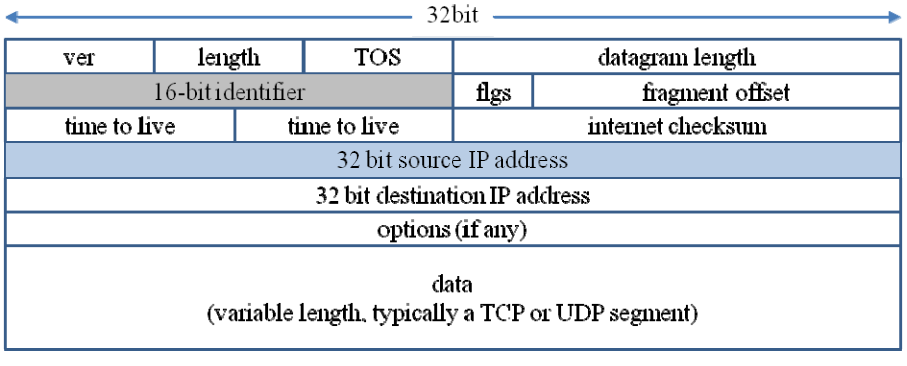


Fig. 2. IPv4 Packet Header Structure

As the figure shows, IP packet contains its source IP address and destination IP address. Attackers may change their 32-bit source IP address into that of a reliable host or use an IP address that may not exist in order to hide his location.

Although there are several methods of detecting/preventing IP spoofing such as Ingress Filter, RPF, BASE Mechanism, and so on, this study analyzes BASE

Mechanism, a method considered to be most effective, and suggests a method of a bypass attack.

3 4-Step BASE Mechanism Analysis

A protocol has the following three characteristics and can be practically applied to real life.

① Initial Benefit
② Incremental Benefit
③ Partial Deployment

However, existing methods of spoofing prevention fail to satisfy the three requirements and cannot be widely applied to real life.

On the contrary, BASE Mechanism has the three characteristics of a practical protocol. In case an attacker sends a large amount of spoofed packets to a destination at once, overload may occur in the network of the destination and, thus, BASE prevents such packets before they arrive at the destination. That is, if spoofed packets meet a router using BASE mechanism while they are being sent to the destination, the router checks if the packets have a correct marking value. If not, the packets are dropped by the router and, if so, the packets are sent to the next router.

A marking value refers to the value of the route that the packet has taken, and the value is marked in the 16-bit identifier of the packet header. When the marking value is distributed to the BASE filter in advance, BGP[6], which is Inter-AS routing protocol (considered standard today), is used for the distribution. This BASE mechanism has the following four step.

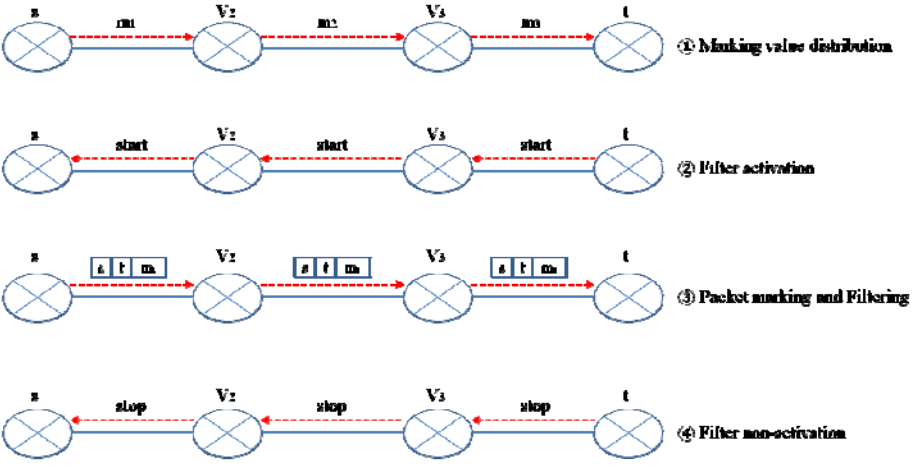


Fig. 3. 4-Step BASE Mechanism

BASE mechanism detects and drops a spoofed packet in the third process.

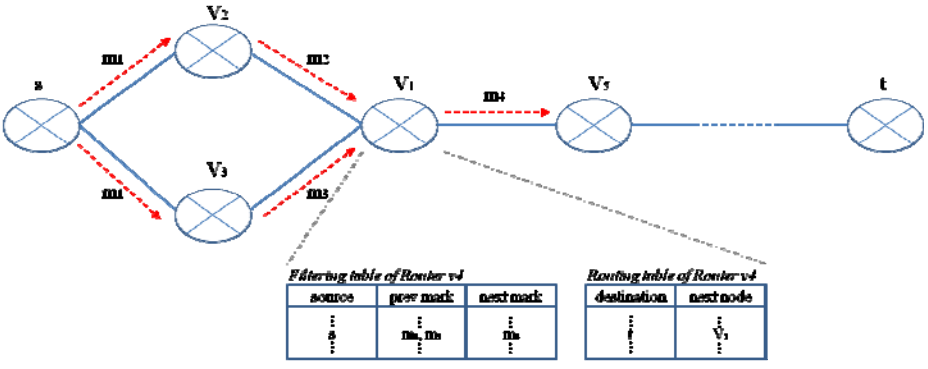


Fig. 4. Filtering Table and Routing Table in Router v4

As depicted in the figure above, each BASE filter contains a filtering table and a routing table through which a marking value and routing route are identified.

However, the marking value is included in a short 16-bit identifier so that the value may be easily revealed by a specific attack.

In this study, therefore, we have designed an attack method of obtaining a 16-bit marking value through Brute Force[7] and attempting packet spoofing with an idle PC.

4 Design of 4-Step BASE Mechanism Evasion Attack Method

The following figure shows how the BASE filtering works in general.

① An attacker attempts IP spoofing towards his target.

② The router operated by the BASE filtering compares marking values of incoming packets.

③ If judged as a spoofed packet, it is dropped by the router.

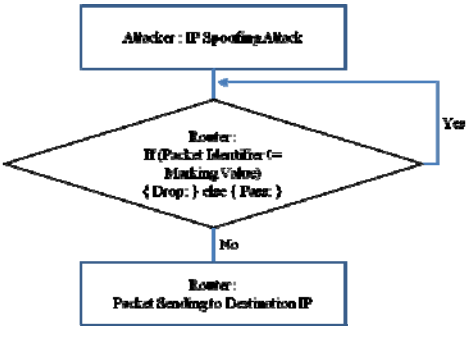


Fig. 5. General Operating of BASE Filtering

Since BASE mechanism compares marking values through the four steps, it can effectively filter out spoofed packets.

Nevertheless, this method can detect only simple IP spoofing attacks and, thus, can be vulnerable to a bypass attack.

This study has designed the following attack method.

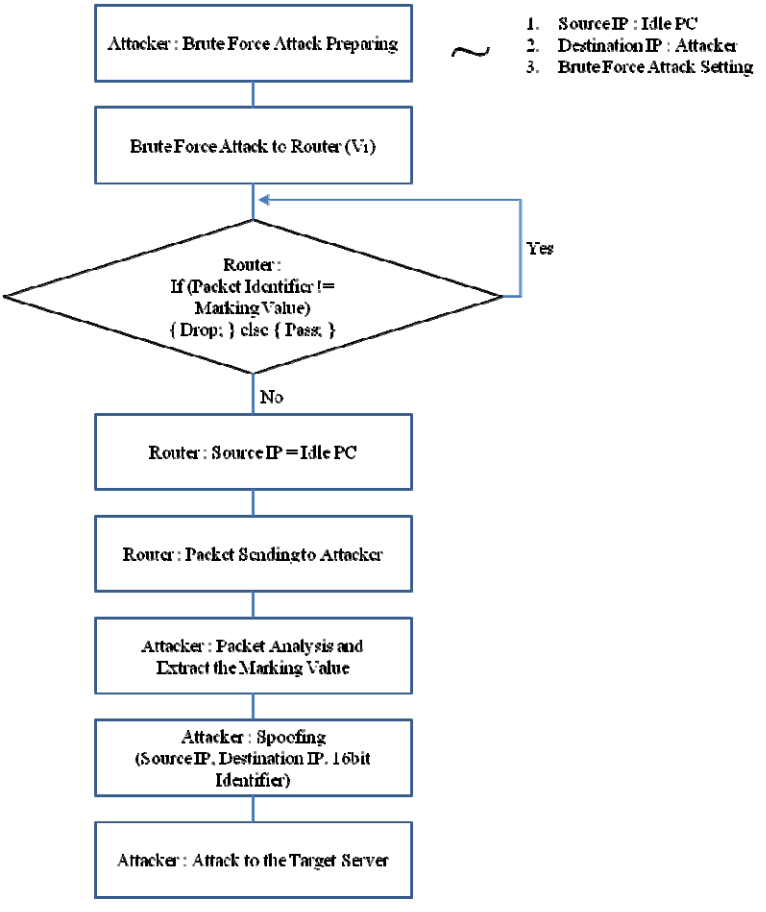


Fig. 6. 4-Step BASE Mechanism Evasion Attack Algorithm

① An attacker disguises a source IP as an idle PC and a destination IP as himself.

② The attacker makes a series of attacks on the router by Brute Force in order to obtain a marking value contained in a 16-bit identifier of a packet header.

③ If the router judges the packet as a fake, it will drop all the packets but, if not (if the Brute Force attack is successful), the router will consider the idle PC (source IP), which is recorded in the fake packet, to be the sender.

④ The router sends a packet to the attacker (recorded in the destination IP).

⑤ The attacker then confirms that the Brute Force attack is successful and obtains the marking value contained in Router V1 through packet analysis.

⑥ The attacker can modify a 16-bit identifier of a packet header using the marking value and makes DoS attack on the target server through IP spoofing.

In this way, although BASE mechanism can improve network functions by filtering out spoofed packets before they arrive at the target server through a router, an attacker may obtain a making value and causes the router to recognize spoofed packets as regular packets.

5 Conclusion

IP spoofing exploits vulnerability in the IPv4 design and implementation and, therefore, effective prevention is not possible unless we use a new protocol.

BASE mechanism, analyzed by this study, can also improve network functions by filtering out spoofed packets before they arrive at the target server through a router, an attacker may obtain a making value in an unexpected manner, making useless security measures.

Although we have analyzed the vulnerability of BASE mechanism and designed an algorithm for a bypass attack, the research on the solution is still underway.

Hence, the follow-up study is expected to suggest a hybrid security measure, which is technically improved, by further analyzing the vulnerability discussed in this study.

Acknowledgements. This paper has been supported by the Software R&D program of KCC. [10914-06002, Development of a Global Collaborative Integrated Security Control System].

References

1. Ferguson, P.: Network ingress filtering: Defeating denial of service attacks which employ IP source address spoofing, ferguson 2000 network (2000)
2. Dalal, Y.K., Metcalfe, R.M.: Reverse path forwarding of broadcast packets. Communications of the ACM 21(12), 1040–1048 (1978)
3. Lee, H., Kwon, M., Hasker, G., Perrig, A.: BASE: An incrementally deployable mechanism for viable IP spoofing prevention. In: Proceedings of the 2nd ACM Symposium on Information, Computer and Communications Security, pp. 20–31 (2007)
4. Gil, T.M., Poletto, M.: Proceedings of the 10th conference on USENIX Security Symposium, pp. 3–8 (2001)
5. Tanase, M.: IP spoofing: an introduction. Security Focus 11 (2003)
6. Murphy, S.: BGP security vulnerabilities analysis, murphy2006bgp (2006)
7. Kedem, G., Ishihara, Y.: Brute force attack on UNIX passwords with SIMD computer, pp. 8–12. USENIX Association (1999)

Dose Debt Financing Has Positive Governance Effect?* — An Empirical Study Based on Private Benefit of Control

Chaonan Lin**, Li Liu, and Huayu Shen

Department of Accounting, School of Management,
Xiamen University, Fujian, China, 361005
linchaonan@xmu.edu.cn, shy1130@126.com

Abstract. Based on the relationship between debt financing and corporate governance, the paper makes an empirical study on the relationship between debt financing and private benefit of control (PBC) of listed company in China. The findings are as follows: Firstly, considering the special ownership structure in China, debt financing governance effect is dispositive, the effect of debt financing on corporate governance and agency cost of substantial shareholder is insignificant; Secondly, special ownership structure in listed company has insignificant adjustment on the relationship between debt financing and PBC; Finally, the first and largest shareholder equity ratio is essential to PBC, the influence of other big shareholders' restriction in reducing the effects of PBC is not significant.

Keywords: Debt financing, PBC, debt financing, governance effect, ownership structure.

1 Introduction

Capital funding is always the focus of studies on the financial affairs in the company. The early study of debt financing mainly lays the emphasis on its financial effects, such as the investigation on tax shield and bankruptcy risks. Besides it, in recent years, western scholars believe that debt financing should have an effect of debt governance. The research was first conducted by Jensen in 1986. He thinks the prohibition of debt interests' payments restricts the ineffective behavior of the governors to a certain extent on one hand, for example, non-pecuniary compensation, unfair related party transaction, and overinvestment. On the other hand, the agency costs of the enterprises can be decreased. Meanwhile, it reduces the so-called free money and abates the controllers' investment ability. Recently, the capital structure and funding ways of the listed companies become the highlight among the native scholars, while its empirical test of the effect of debt governance is still weaker than shareholding funding (favorites, share allotment and effects of newly-issued shares). At present, Chinese

* This paper is Funded by MOE (Ministry of Education in China) Project of Humanities and Social Sciences (Project No. 10YJC630137).
** Corresponding author.

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 423–430.
springerlink.com

studies mostly belong to the verification of the relationship between debt funding and company's profitability. In other words, debt funding has an influence on the scheme of company's governance and finally affects its profitability. However, in Chinese capital market, the phenomena of earnings management are so serious that company's credibility standards have been widely questioned (Zingales, 1994; Nenova, 2003; Doidge, 2004). Private Benefits of Control, apart from the traditional standards, as another key one to estimate company's governance effect, have been gradually attached importance to by western scholars and diffusely applied into the empirical research papers. This conception was come up by Crossman and Hart in 1988. It refers to the benefit owned by controlling shareholders and the medium-sized and small ones have no access, which reflects the aggression from controlling shareholders upon those medium-sized and small ones. That is, PBC is the reflection of the agency costs caused by controlling shareholders' agent problems. Consequently, more and more foreign scholars treat it as another important index to cover the shortage of the traditional one in the company's profitability evaluation. Based on the above reasons, this thesis will examine the governance effect of debt funding through the inspection of the relationship between debt funding and PBC in the light of the new angle PBC.

2 Theoretical Analysis and Hypothesis

From Jensen's 1986 viewpoint, guarantee mechanism decreases Free Cash Flow allotted by managing people, and plays an important role in the reduction of agency costs. It means debt funding has an effect of governance and can be taken as an effective controlling mechanism to improve company's profitability and oversee the governors. William points out that in 1988, debt is not only a funding tool, but should also be considered as a management structure. In 1989, Jensen further sums up its several advantages. Firstly, it confines enterprises' free cash flow and managers' overinvestment. Secondly, the debt bondage has an ability to cause company's bankruptcy, through which investors' benefits are well protected. Thirdly, it is easy for debtee to observe the record of company's paying back and supervise the governors. Hart in 1995 even holds a belief that it's not so much crucial to provide business operators with controlling rights, but a reasonable funding structure ought to be regarded as a decisive question. Had a company's profitability been low, debt funding should be increased in order to restrain the governor and raise its efficiency. PBC comes from the ineffective behaviors inside the company such as Tunneling, so according to the above theories, if debt funding is really positive to company management, it will be seen to reduce the agency costs and company's PBC level, that is, there is a negative relationship between debt funding level and PBC. However, the deduction is true on condition that debt funding forms an effective bond on the managing class in a corporation, which depends on two preconditions. One is the company should pay back debt in accordance to the debt contract. The other is its bankruptcy mechanism functions normally. The company with financial situation in a bad shape will be bankrupt by law and its controller will be punished. If these two conditions disappear, no matter how high the debt level is, the controller will not be restricted. In the present China, on one hand, most listed companies are transformed from state-owned enterprises, so their serious debt situation in a large part owes to

historical reasons, not the enterprises' self choices. The phenomenon of state-owned non-tradable share dominating exclusively is widely existed in the listed companies. As a result, their debt financing behaviors show the quality of administration and irrationality (Zhizhong Huang and Yunxia Bai, 2002). Under such circumstance, it's hard to believe the debt producing in this way can play a positive part in governing the companies. On the other hand, the nature of state-owned non-tradable share dominating exclusively decides that the listed companies have a common benefit entity (country) with the banks which are not independent one. Companies' debt burdens are relaxed by the government. Banks borrowing money don't bring financial risks to companies, but become the concrete embodiment of listed companies' real power. The relationship between two sides gets more closed, a company will be more difficult to go bankrupt or de-list. Consequently, taking Chinese special share structure into consideration, it is concluded that debt financing's governance effect is weaken, and its negative relationship with PBC is likely not transparent.

On the basis of the above analysis, this thesis puts forward two hypotheses. One is without allowing for share structure, debt funding level is higher, and company's PBC is lower. The other is concerning the special share structure, the negative relationship between Chinese listed companies' debt funding and PBC is not evident.

3 Research Design

3.1 Sample Choosing Standard

The research object is the companies which have transferred state-owned non-tradable shares in the A-share market of Shanghai and Shenzhen from 2003 to 2004. The research principle is based on the ones obeyed by Dyck and Zingales (2004) to measure PBC. In addition, the thesis will combine with the realistic transfer of non-tradable shares in the national market. To sum up, the sample need satisfy the following conditions. First, after the transaction, the buyer will be the biggest holder. Second, it should be marked as non-related one and make the price known to the public. Third, both of them is willing to conduct the behavior, not including state-owned shares' free allocation in state-owned organizations, debt payment or court public auction. Fourthly, the transfer of state-owned shares in the sample has been permitted and defined by the relative governing departments. Fifthly, financial companies will be eliminated because their transfers' procedure for approval is different and the financial data is less feasible than other industries. At last, the effective samples in all are 97, distributed in 10 professions. In the paper, the transfer record of the state-owned shares and financial data origin from CSMAR' s China Listed Firm's Corporate Governance Research Database, China Listed Firm's Share Holders Database, China Listed Firm's Annual Financial Database and China Listed Firm's Financial Index Database.

3.2 Methodology

PBC, reflecting the big share holder's agency costs, is a vital index to judge corporate governing efficiency. Therefore, if debt funding has an effect of governance, it can reduce the agency costs. There is a negative relationship between debt funding and PBC.

3.2.1 Explained Variable

In the same way as that adopted by Guadalupe (2005), this thesis chooses PBC as the explained variable to evaluate debt financing's governing effect. The accurate measurement of PBC is essential in the research which will reference Chaonan Lin's model in 2006. Because it takes a full account of the specialty of the transfer pricing of corporate control under Chinese present stock market share structure, and divides control of corporate into pubic benefits and private ones. The specific models and detailed measuring are as follows.

$$CP = (TP - NA)/NA - EP = (TP - NA)/NA - \sum_{i=1}^{3} R_i * ROE_i/(R_1 + R_2 + R_3) \quad (1)$$

$$ROE_t = a_0 + a_1 ROE_{t-i}, \quad \text{i=3,2,1} \quad \textbf{(2)}$$

In the first model, CP is PBC; TP is each share's price which is among the block share transfer. NA is the latest seasonal report or each share's net asset in the annual one. EP is the buyer's reasonable expectation towards the target enterprise, which is pictured by weighted average net asset yield of the former 3 years. Its weight will be decided by the relevant index in the back result of the second model, among which t represents the year of transference.

3.2.2 Explanatory Variable

The studied explanatory variable here is corporate debt funding level. At the same time, in order to further inspect whether equity structure arrangement can affect the relationship between the debt financing level and PBC. We add in the model the equity structure arrangement variables and debt funding level and equity structure arrangement variable's interaction. The agent variable of debt financing level is represented by asset liability ratio (DEBT), that is, total liabilities/total assets. For the agent variable of equity structure arrangement, combining with the previous experience, this paper studies the two variables about proportion of the first largest shareholder and degree of ownership balance, among which the proportion of the first largest shareholder (TOP01) is denoted by (the first big shareholders holding/company total equity), and degree of ownership balance (Z2345) is rendered as (the sum of four-after big shareholders holding ratio / the proportion of the first largest shareholder), to reflect the balance power of the other big shareholders to the first largest one.

3.2.3 Control Variable

Share transfer ratio (SHARE), means the ratio of the transferred non-tradable shares in the total non-tradable shares.

Tradable share scale (STOCK), is represented by the tradable shares' natural logarithm in the end of the last year.

Enterprises' size (SIZE) is showed by the total corporate asset's natural logarithm in the end of the last year.

Proxy cost (PROXY) is displayed by the proportion of the sum of the cash and short time investment to the total asset scale.

Rate of return on common stockholders' equity (ROE) is the one in the end of the last year.

Industry dummy variable (IND) refers to industry virtual variable. According to the design in Listed Firms' Industrial Classification Guidance, when the firm belongs to the industry i,INDi=1 or INDi= 0.

3.2.4 Empirical Model

In order to examine Hypothesis 1 and 2, we found the following model:

$$PBC_{it} = \beta_0 + \alpha_1 DEBT_{it} + \alpha_2 TOP01_{it} + \alpha_3 Z2345_{it} + \alpha_4 TOP01_{it} * DEBT_{it} + \alpha_5 Z2345_{it} * DEBT_{it} + \gamma(Control\ \ Variable)_{it} + u_t + \varepsilon_{it} \quad (3)$$

In the above regression model, we in gradual add to debt funding level and equity arrangement variable's interaction (TOP01*DEBT) and (Z2345*DEBT) so as to further test whether equity structure arrangement will affect the relation between debt funding level and PBC one. β_0 is intercept; i is the company; t is the year. α_i (i=1, 2…, 5) and γ are regression index. u_t is year control variable. ε_{it} is random variable. Chaonan Lin's (2006) research shows that PBC has a clear industrial trait, so we add industrial virtual variable into the model.

4 Empirical Analysis

4.1 Variables' Descriptive Statistics

Table 1 offers the descriptive statistics result of each variable in the empirical model. It is reported that the average scale of the domestic listed firms' PBC is 26.9%, far more than that of the companies in the western developed countries, while close to the level of the counterparts in those countries with newly-emerged markets1. The result conforms to Donghui Shi's (2003). Besides, the biggest shareholders' controlling in the sample firms amounts to 40.16%. With the constant reform of the state-owned enterprises, the phenomenon of state-owned non-tradable share dominating exclusively has been gradually lightened to some extent, but not totally changed or disappeared, in virtue of its general high level.

Table 1. Descriptive statistics of variables

item	sample	mean	minimum	maximum	standard deviation
PBC	97	0.269	-0.437	3.117	0.695
DEBT	97	0.524	0.097	0.938	0.184
TOP01	97	0.4016	0.1153	0.7328	0.1506
Z2345	97	0.941	0.011	3.2251	0.664
SHARE	97	0.3156	0.0314	0.7328	0.1599
STOCK	97	17.772	16.245	18.972	0.597
SIZE	97	20.505	18.970	22.718	0.717
PROXY	97	0.135	0.004	0.448	0.104
ROE	97	0.034	-1.2347	0.236	0.179

4.2 Empirical Result

The data in the paper is collected from 2003 to 2004 and has cross-section traits. In order to overcome the impact of heteroscedasticity on the models, the thesis employs GLS to estimate return model and chooses variance by annual characteristics. We firstly conduct the regression analysis on the relationship between the sample corporations' PBC and debt funding level, and then adopt the examination steps of gradually bringing in other explanatory variables, and carry out the stability test on the relevance of debt funding and PBC.

Table 2. Results of Logistic Regression Modelling

variable	modelling1	modelling2	modelling3	modelling4	modelling5
β_0	0.489*** (0.000)	0.011 (0.959)	3.553*** (0.009)	3.284** (0.022)	3.452** (0.017)
DEBT	-0.379* (0.067)	-0.231 (0.283)	-0.359 (0.151)	-0.479 (0.713)	-0.548 (0.609)
TOP01		1.106*** (0.001)	1.357*** (0.001)	1.819** (0.043)	1.443* (0.076)
Z2345		-0.046 (0.448)	-0.076 (0.334)	-0.074 (0.358)	-0.258 (0.235)
DEBT*TOP 01				-1.020 (0.564)	-0.235 (0.905)
DEBT*Z234 5					0.361 (0.347)
SHARE			-0.546* (0.076)	-0.525* (0.092)	-0.547* (0.081)
STOCK			-0.042 (0.587)	-0.039 (0.615)	-0.021 (0.789)
SIZE			-0.125** (0.043)	-0.125** (0.048)	-0.132** (0.046)
PROXY			-0.548 (0.188)	-0.556 (0.185)	-0.561 (0.181)
ROE			0.012 (0.978)	0.009 (0.981)	0.021 (0.941)
IND_i	controlled	controlled	controlled	controlled	controlled
U_t	controlled	controlled	controlled	controlled	controlled
Adj-R^2	0.021	0.187	0.277	0.269	0.267
F	2.739* (0.089)	7.220*** (0.000)	2.822*** (0.001)	2.657*** (0.002)	2.555*** (0.003)

*, **and*** indicate statistical significance at the 10%, 5%, and 1% levels, respectively

In order to save the space, we only give relevant variables in the regression results, but omits the fixed effects of industry and year. The first model's simple regression results shows that without considering the share structure arrangement and other controlling variables, there is a significantly negative relationship between debt financing and PBC. Hypothesis 1 gets support.

But the regression model results in model 2 and 3 show that when introducing equity structure arrangement variables, the proportion of the first largest shareholder (TOP01) become important influencing factors of PBC, instead of debt levels (DEBT). Although between them, there is still a negative relationship, but is not obvious, and the regression results are not under the influence of other control variables. This suggests that in our special capital market, western widespread debt financing governance effect in the listed companies is weakened or does not yet exist in China. Hypothesis 2 gets support. In addition, the regression results show that, in general, other big shareholders' ownership restriction (Z2345) in reducing the effects of PBC is not transparent.

The existing documents show that corporate governance structure has an effect on the choice of capital structure. Take Zuoping Xiao (2005) for an instance. Through the factor analysis model, he displays that between the company's equity concentration and debt financing level, there is a significantly negative relationship. So in order to further understand the mechanism of equity structure arrangement and debt financing to together gain control over PBC, we add debt financing level and ownership structure arrangement variable to the model 4 and 5 and do further investigation. The regression results show that share structure arrangement and the debt levels' interaction (DEBT * TOP01) and (DEBT * Z2345) are not significant, and along with the interaction's introduction, model 4 and 5's adjustment R2 decline in different degrees, which suggests that under the circumstance of the debt controlling level and other controlling variable situations, equity structure arrangement's impact on the relationship between debt levels and PBC adjustment effect is not apparent. That is to say, the article has not found the direct evidence for corporate share arrangement's indirect influence on PBC level through affecting debt levels, and fully explains in the Chinese listed companies, the reason for debt financing governance effects' weakening is complex and diverse.

In the regression results of other factors influencing PBC, enterprise scale, the stock transfer ratio has a strong link with PBC, but other factors' functions are weak. It's similar to the domestic research conclusions of the same category.

5 Conclusions

Based on the theories about the relationship of debt financing and Corporate Governance, and combining with China's listed company ownership arrangement, this article takes the non-tradable shares transfer of state-owned listed company as a sample, in Shanghai and Shenzhen A stock market, from 2003 to 2004, and analyzes the relationship theoretically and empirically between our country debt financing of listed companies and private benefits of control (PBC), and has an empirical test on debt financing governance effect in the perspective of PBC. The results of the study show in the following aspects. First, the listed companies in China do exist in the high PBC and the average reaches to 26.9%, more than the level of developed countries, and close to the scale of some emerging market countries. Second, considering the ownership structure of the special arrangement system, in the Chinese listed companies, debt financing governance effect is weaken or do not exist on the whole. That means the function of debt financing is not significant to strengthen corporate management and to reduce the large shareholders' agency costs. Third, the listed company' special equity structure arrangement has no big influence on the adaptation

of the relationship between debt financing and PBC. The reasons for the weakening are complex and varied. Fourth, the proportion of the first largest shareholder's control is essential to PBC. The influence of other big shareholders' restriction in reducing the effects of PBC is not significant.

References

1. Dyck, A., Zingales, L.: Private Benefits of Control: An International Comparison. Journal of Finance 59, 537–600 (2004)
2. Johnson, S.R., Florencio, L.S., Shleifer, A.: Tunnelling. American Economic Review 90, 22–27 (2000)
3. Shleifer, A., Vishny, R.: A Survey of Corporate Governance. Journal of Finance 52, 737–783 (1997)
4. Tang, Z., Jiang, W.: A Study on the Expropriation Degree of Large Shareholders of China's Listed Companies. Economic Research 4, 44–50 (2002)
5. Ye, K.: A study on the return of corporate control-—exchange of non-circulation shares in China. Economic Science 5, 61–69 (2003)
6. Wang, H.: Debt financing, corporate governance, market value of listed companies. Economic Research 8, 28–36 (2003)
7. Huang, Z., Bai, Y.: The Listed Companies' Liabilities Behavior, Shareholders' Wealth and Stock Market Effect. Economic Research 7, 49–57 (2002)
8. Lin, C., Liu, X., Hao, Y.: Industrial Characteristics and private benefit of control: an empirical study from listed companies in China. Economic Science 3, 61–72 (2006)

A Study on the Recent Application of the Human Capital Theories in China

Liu Na* and Qin Ying**

Beijing University of Civil Engineering and Architecture
Beijing Normal University, PHD

Abstract. In China, the study on the application of the labor capital theories in recent years mainly focused on the following four aspects: 1. the perpetuation and expansion of the western labor capital theories; 2. the study closely related to the labor capital management of business; 3. the study on labor capital theories of rural areas; 4. the study on labor capital investment. The further application of the labor capital theories enriches its content and makes it possess certain values of academic research.

Keywords: labor capital theory of business, labor capital theory of rural areas, labor capital investment, applying study.

Once the labor capital theory was put forward in the 1960s and 1970s in western countries, it aroused strong interests among different countries and was accepted by the economists immediately, thus promoted the development of educational economics. Many economists began spread it and some of them even regarded as "a revolution in economics". As an economics theory, the labor capital has its merits and limitations as well. It has been doubted and challenged by some scholars who believe that the educational investment is not a kind of capital at all and cannot bring profits and values in a long run. Nowadays, however, it has become a well-accepted theory that education, training and health insurance can be regarded as a labor capital investment, thus promoting the changes from traditional personnel management to modern labor capital management. As a popular issue in economics, the application of labor capital theory has been studied intensively in China in recent years. These studies will be introduced, analyzed and commented in this paper in detail, hoping to provide some beneficial thinking and inspiration for the labor capital theory in China.

1 The Definition of Labor Capital

In 1960s, American economists Schultz、Becker and Mincer elaborated the contribution of labor capital investment, such as education, to economic growth and income

* Associate Professor.
[financing projects] The project of training men for professional outstanding genius of Beijing in 2009 (2009D005017000004); The project of Scientific Research Fund of Beijing University of Civil Engineering and Architecture(101004906).
** Professor.

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 431–435.
springerlink.com

distribution, and began to apply cost-benefit analysis to the behavior of labor capital investment of individuals and families. Different scholars defined labor capital in different ways. The representative of the first view is Schultz, who hols the idea that the labor capital is the sum total of the laborer's knowledge, skills and health condition. This definition considers the content of labor capital as the essential aspect and emphasizes that "the contribution of the increasing of the labor capital as laborer's knowledge, ability and health to the economic growth is far more important than the increasing of properties and the number of laborers." The second definition is put forward by Becker, who thinks that " The human capital is the investment on education 、job training、 health and emigration ". Li Zhongmin, as the representative of the third definition, concludes that the labor capital is the value that exists inside a person and can be materialized into commodities or services, increase their effectiveness and share the benefits. This definition puts emphasis on the abstract aspect of labor capital and believes that "the labor capital belongs to the value category and it is the product of conscious investment, which can promote the advancement of the social productivity and wealth increase, and share the profits as well as other kinds of capitals. The last view, concluded by Mr. Zhang Fenglin, focuses on the presentations of the labor capital. It believes that "the so called labor capital refers to the sum total of a person's various capacities being developed and formed by the labor capital investment. These capacities exist in one's body and can be divided into three types: 1. the physical power, including the physical strength, endurance, etc.; second, 2. intelligence or level of science and culture, such as knowledge, skills, cognitive and creative abilities, etc.; 3. the morality, including the moral standard, will power and attitude towards life." The definitions above all illustrate the essential features of labor capital: the labor capital is the power of a person within one's body, that is, the physical power, the capacities, knowledge and skills, among which the intelligence (one of the capacities) is the core element and can be regarded as the capital of profits. Therefore, we conclude that the labor capital refers to "the sum total of the knowledge, skills and capacities that exist within laborers." It includes the capital of physical power, intelligence, knowledge, creation and skills.

2 The Current Situation of the Studies on the Application of Labor Capital Theories in China

The studies of labor capital theories in China mainly focus on the following aspects:

2.1 The Perpetuation and Expansion of the Western Theories

In his book The theories and Application of Labor Capital (2007), Mr. Zhang Fenglin systematically illustrated the basic theories of labor capital investment. He explained in detail such problems as the form and features of the labor capital investment, the labor capital theories of family activities, the school education and the labor capital investment, medical care and labor capital investment, labor capital and income distribution, and labor capital and economic growth, etc. As a valuable expansion and complement of labor capital theories, Mr. Zhang also discussed some problems that have not been specifically tackled by other scholars, for example, the analyses of the demand and supply of labor capital investment, the external effects and public

strategies of investment, and the labor capital of the economic activities of business. In the analysis of the economic activities of business, Mr. Zhang explained the special feature of enterprise, proposing that it is a factory of talented persons, and produces material products and labor capital as well. The discussion makes people attach more and more importance on business training and labor capital development, and provides forceful theoretical support for the management and development of modern labor resources.

2.2 Studies Closely Related to the Labor Capital Management of Business

Liu Yeyun has explicitly extended the field of labor capital theory from education to business management and industry management in his book A Study on Labor Capital Management (2006). A prominent point of the book is the accurate distinguishment between labor resources and labor capital: "labor resources roughly refer to 'the total number of people who have labor ability and normal intelligence.'; labor capital, on the other hand, is 'the sum total of the knowledge, skills and capacities of the laborer.' Labor resource is a quantitative concept, while labor capital is a qualitative one; the former cannot reflect the differences in quality, while the latter is the result of the development of labor resource." What's more significant is that Liu clearly define the relationship between the two: "labor resource is the natural power of human being, and it is the prerequisite and basis of labor capital; labor capital is the promotion and distillation of labor resources." Although other scholars in China, Wang Xiaobo (2002), Jiao Binlong(1999), Xu Xiangyun and Liao Xiaoguan (2004), for example, also distinguished labor capital from labor resource, they didn't give distinctive explanation on their relationship, and some other scholars even held mixed ideas about it. As a result, as far as this issue is concerned, Liu Yeyun undoubtedly did great contribution to the study of the labor capital theory. What's more, Ye Zhengmao and Ye Zhengxin (2007) had an exploration from enterprise theory to the theory of organizational labor capital, which is an extension and application of the labor capital theory.

Moreover, Liu also made a breakthrough in management theory. Given the central position of labor capital in enterprise, Liu proposed some relevant management concepts, for example: the management of labor capital information, cost, manipulation, investment, transfer and strategy. These studies are the model examples of the combination of economics and management, not only expanding the research domain, but also providing important theoretical basis for further study.

Different from Liu's point of view, Li Baoyuan had a practical study about labor capital theory in his two books (2005): 1. the exploration on the 折股实践 of labor capital; 2. the stock corporation: enterprise mode advocating the labor capital rights and benefits; 3. the measurement of the investment of labor capital in enterprise; 4. the basic requirement and framework of the analysis of labor capital investment profits; 5. the practice of analyzing labor capital investment profits.

Both Li Yeyun and Li Baoyuan applied labor capital to business management and industry management, but they had different standing point. Liu focused on the management of labor capital; while Li put emphasis on the management of human resources.

2.3 The Study on the Labor Capital Theory about Rural Areas

Based on the definition of labor capital, Some scholars in China put forward the definition of rural labor capital. Bai Juhong and Yuan Fei (2003) suggested that the rural labor capital refers to "the capital attached to the labor power of countryside by means of education, training health investment and labor power transfer." Wang Fangfang (2005) concluded that the rural labor capital is "the physical power, knowledge, skills and other capacities to promote the rural productivity attached to the labor power of countryside, with the form of the reserve quantity of labor capital." Xu Lijie (2005) held the idea that the rural labor capital is "the non-material capital that is attached to the labor power in countryside by education and training, health investment, transfer investment, etc. and presented by the quantity and quality of labor power in countryside." Zhen Changzhi (2008) believed that the rural labor capital refers to the population, or the labor power in countryside that have reached the labor age, gained labor ability and are willing to work, among which those rural population who are conducting the producing activity but younger than 16 years old or older than 60 years old are also included.

At the present, the study on rural labor capital is mainly about the following aspects: 1. the features of rural labor capital, discussed by scholars as Liu Wen (2004), Dong Zhiyu (2006), Zhou Jianhua and Zeng Fusheng (2007), and He Xiuling in their articles; 2. the policy and suggestions about rural labor capital investment, which is explored by scholars as Kuang Aiping (2005), Guo Zhiyi and Chang Ye (2007), Liu Hong (2008), etc.; 3. the relationship between rural labor capital and rural laborers' income and economy development. Bai Juhong (2003) used the model of Mincer income function to have a static study on the relationship between education and income of the farmers, concluding that the level of education is on direct proportion to the family income. Xu Nengrui and Fu Guohua (2004) analyzed the function of the labor capital promotion on the income of countryside people in Hainan province with quantitative analysis; 4. the rural labor capital and labor power transfer, proposed by Liu Wen (2004) as to the relationship between labor power transfer and labor capital in countryside; 5. the nurturing and development of rural labor capital.

2.4 The Study on Question of Investment to Human Capital

Once pointed out that "relatively speaking, the policies of China in 1990s attached importance on the realization of capital investment and ignored the labor capital investment as education; at the same time, the policies paid more attention to the urban labor capital investment and ignore rural labor capital investment. Consequently, it is necessary to make new investment policies and achieve the state of balance between countryside and city and various kinds of capitals." The view was further explained in The Policies Promoting Labor Capital (2003). Afterwards, the scholars in China began some relevant studies. Starting from the real situation of China and relating to the higher education, some scholars specifically discussed the employment choices of graduates from universities and risks of labor capital investment (Zhao Hongbin, 2007); some others conducted the researches about the problems of higher education investment and the demanding and supply of graduates from the view of labor capital theory.Zhao Hongbin was interested in the source of the risks of higher education labor

capital investment, while Liu Wen mainly discussed the characteristics of labor capital investment in the shifting period. Mr. Li Yuanbao (2009), on the other hand, analyzed the problems of the "unbalancing development" of labor capital from the aspects of its main body and field, providing valuable support for government to make feasible policies and reference for families and local areas to invest on labor capital. In addition, the scholars further studied the theories of labor capital production, accumulation, transference, realization and decision, thus expanded the domain and views of the application of labor capital theory in China.

Besides, Huang Jingbao (2008) related labor capital theory to graduates employment, and discussed the employment capacity of graduates, but didn't make and further analysis on the investment risks of labor capital. Li Fengliang (2005) compared the labor capital theories and screening theory from the view of Chinese labor power market, believing that the two are controversial but not paradoxical and worth further exploration.

3 Conclusion

Above all, the study of labor capital theory started from western countries, but produced profound influence on developing countries. China proposed the strategic development objectives to "build a country with strong labor resources". With the rapid development and shift of economy and the background of adjusting the industry structure in China, and given the specialty of labor capital as attaching to human beings, the domain of the application of labor capital theory is continuously expanded and extended. Its contents are more and more colorful and have certain practical and guiding significance for policy-making, and have high values of academic research as well.

References

1. Becker, G.: Economic Analysis of behavior of men. Peoples Publishing House, Shanghai (2008)
2. Schultz book, a translation of Wu ZhuHua, The statement of investment in human capital. Economics institute Publishing House, Beijing (1990)
3. Mincer book, a translation of Zhang FengLin, The study of human capital. China Morden Economics Publishing House, Beijing (2001)
4. Zhang, F.: Human Capital Theories and application research. Commercial Press China, Beijing (2007)
5. Li, B.: Human Capital opinion. Beijing Normal University Publishing House, Beijing (2009)

An Introduction on Gary Becker's "Human Capital Theories"

Liu Na*

Beijing University of Civil Engineering and Architecture,
Beijing Normal University, PHD, Financing Projects

Abstract. "human capital theories", written by Gary Becker in 1964 is regarded as one of the classic works in economics. This paper will illustrate the contents and analyze the characteristics of this book after introducing the writer's life and experience. It's academic and referential value will be discussed as well.

Keywords: labor capital theory, labor capital investment, theoretical analysis, empirical analysis.

1 A Review on Gary Becker's Thoughts and Works

As a famous economist, Gary Becker is an important representative of "Chicago School" and has done great contributions to the labor capital theory together with Milton·Riedman、Theodor·Schultz、Ronald·Kos and Jacob·Mincer. The historical contribution of Gary Becker lies in the creative empirical study on economy, enriching the contents of economy and enlarging its study domain.

"Human capital theories"(1964) is Gary Becker's first book on labor capital theory, and is regarded as one of the classics in economy. It was republished in 1975 and 1993. Gary Becker once said, "the concept of labor capital theory was commonly used in such communist countries as Russia, China and countries in Eastern Europe. My theory, as well as that of Schultz and other people's works were widely quoted in these countries. Economists and plan-makers have fully agreed with the concept of investing on human beings." "human capital theories" elaborated the relationship between labor capital and the development of family and society and the influence of capital labor theory on income and the repaying rate of labor capital theory. It also discussed the problems of the repaying rate of college education and middle school education, the relationship among age, profits, wealth and labor capital. At the end of the book, the writer creatively illustrated the relationship between labor capital and family situation caused by the changes of economy.

* Associate Professor
[financing projects] The project of training men for professional outstanding genius of Beijing in 2009 (2009D005017000004); The project of Scientific Research Fund of Beijing, University of Civil Engineering and Architecture(101004906).

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 437–439.
springerlink.com

2 "Human Capital Theories", Its Contents

There are three sections in this book. The first section deals with the theory of labor ca0078pital theory, relevant economic phenomena and its importance. In the first chapter, the writer briefly reviewed the labor capital theory. In the second chapter, the writer mainly discussed the at-post training, believing that it can present the influence of labor capital on income, employment and other economic variables in a more explicit way. The third chapter covers the problem of how to evaluate the total labor capital investment and its changing trend when the expecting profits have changed. The writer put forward the theory that the capacity is in direct proportion to education, namely, those middle school graduates who go to universities can get more repayment from the higher education than those who do not go to college. The second section is empirical analysis, including 5 chapters. The fourth chapter illustrated the repaying rate of college education, holding the idea that those who get higher achievements can have higher profits from higher education. And this is why those intelligent people will choose to go to college. The fifth chapter assessed the social profits and individual profits of higher education and made a comparison between them. In Chapter 6, the author analyzed the social and individual cost and their profits of middle school education. In Chapter 7, the connotation of labor capital theory, that is, the influence of labor capital theory on age-income relation was discussed in detail. The third part of the book introduced the changes in economic field. Chapter 9 is about the relationship between labor capital theory and family. Labor capital can have profound influences on the family economic decision. The parents should not only decide how much they will invest on a child, but also how many children they will give birth. In Chapter 10, the main problem in concern is labor distribution, negotiation costs and knowledge. It discussed the specialization and distribution within and among different department. Chapter 11 is about the relationship between labor capital, birth rate and economic growth, further explaining that with the increasing of labor capital, the investment will increase as well.

Most of the book was from Gary Becker's thesis and teaching script. The chapters are independent with each other, but correlate in some degree. The most worth-mentioning point of this book is that the author chose some sample groups to have empirical study and expanded the study period, thus got more supporting data and inference on education repaying rate.

3 Its Features and Contributions

3.1 The Content Concerns a Vast Field and Is Systematically Organized

The problems concerned various aspects of labor capital theory, including the review, its influence on income and its repaying rate.

3.2 The Empirical Analysis is Creative and Fully Explained

There are many unique and effective economic analytical methods in the book, for example, the age-wealth relation, which can be used to describe the relationship between ages and wealth. Gary Becker also constructed a behavior and skill model to

explain that those underdeveloped countries step out of the shadow of "The balance of Malthus " can help them make efforts toward the low birth rate and high labor capital gains.

3.3 The Results Are of High Feasibility

Because of its high reliability, his study results have important directing meaning for some phenomena. For example, when he is analyzing the education repaying rate, Gary Becker strongly proved that the unadjusted personal repaying rate of middle school graduates is quite high, that why it is a wise action to enlarge the middle school education.

The prominent contribution of the book is that it filled a blank page for the labor capital theory and put forward a new model of economic theory, new theory and ideas on the basis of Schultz and Mincer's study. For example, different from Adam Smith's idea that specialization is decided by the capacity of market, Gary Becker believed that specialization is also decided by some other factors, which are even more important than the market capacity. The book has an overall analysis on the labor capital theory, and provides a universe explanation for various experiential phenomena.

All in all, "human capital theories" is a masterpiece of labor capital theory and its conclusion, ideas, analytical methods, data analysis, theoretical hypothesis and analyzing process have very important academic and referential values.

References

1. Schulz, T.W.: Investment in Human Capital. The American Economic Review 5(15), 1–17 (1967)
2. Becker, G.: Economic Analysis of behavior of men. Shanghai Peoples Publishing House, Shanghai (2008)
3. Min, W.: Economics of Education international encyclopedia, 2nd edn. Higher Education Press, Beijing (2000)
4. Schultz book, a translation of Wu ZhuHua, The statement of investment in human capital. Economics institute Publishing House, Beijing (1990)

Coordinative Operations of Distributed Decision-Making Closed-Loop Supply Chain: A Review*

Jiawang Xu[1,2], Yi Zhang[2], Bin Liu[1], and Liqiang Zhao[1]

[1] School of Economic & Management, Shenyang Aeospace University, Shenyang, China
[2] Shenyang Institute of Automation, Chinese Academy of Sciences, Shenyang, China
ccb867321@yahoo.com.cn

Abstract. The problems and the causes of distributed decision-making in closed-loop supply chain are analysed in this paper. The importance and the significant of the researches for coordinate operations policy design for distributed decision-making closed-loop supply chain are illustrated, and the current research status quo of the issues, such as coordination and operation for closed-lop supply chain are reviewed. Some prospective problems in the coordinate operations for distributed decision-making closed-loop supply chain are discussed.

Keywords: Closed-loop supply chain, Distributed decision-making, Contract coordination, Review.

1 Introduction

Reverse logistics (RL) is the process of planning, implementing, and controlling the efficient and cost effective flow of raw materials, in-process inventory, finished goods, and related information from the point of consumption to the point of origin for the purpose of recapturing value or proper disposal [1]. Moreover, it prevents pollution by reducing the environmental burden of End-of-Life (EOL) at its source [2]. The rise of green concerns makes reverse logistics a time demanding and relevant area of interest. Recycling, remanufacturing, and disposal are the three main factors in this arena for facing the challenges of globalization and sustainability [3]. Closed loop supply chain (CLSC) has also gained an extensive importance today, in the world of increasing environmental concerns and strict regulations on the wastage caused right from inception of a product, through its life period and after it. A CLSC consists of both the forward supply chain, and the reverse supply chain. It has been found that reverse supply chain should be a part of forward supply chain integrated, as it can contribute to lowering overall costs and meeting governmental/environmental regulations. Hence, there is a need to model and analyze closed loop supply chains as a system in total, without splitting into distinct parts of forward and reverse supply chains [4].

* This work is supported by Humanities & Social Sciences Planning Foundation of Ministry of Education Grant #11YJA630165, Special Major Science and Technology Projects of Zhejiang Province Grant #2011C03004, and China Postdoctoral Science Foundation Grant #2008044170.

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 441–448.
springerlink.com

In recent years, there are many theoretical researches related to the closed-loop supply chain coordination mechanism, distributed decision-making, and operation strategy. But it is very few for the research of coordination mechanism for distributed decision-making of closed-loop supply chain management. Here, we will review the current research status quo of the issues， such as coordination and operation for closed-lop supply chain, and put forwards some prospective problems in the coordinate operations for distributed decision-making closed-loop supply chain

2 Distributed Decision-Making in Closed-Loop Supply Chain

Closed-loop supply chain is a complex distributed decision-making system, existing three distributed decision-making problems in strategy level, tactics level and operation level. Strategic level includes product design, recovery and reuse of waste product, and the choices of possible long-term supplier, collector and customers. Tactical level includes long-term (middle-term) manufacturing / remanufacturing, investment of waste materials recycling / reuse, long-term forward / reverse supply chain coordination contract drafting, and market of supply chain upstream and downstream and recovery channels of reverse logistics, recovery efforts designing. The tactical level can be divided into middle-term level and short-term level, middle level (such as 1 year or 2 years) need to make sure that middle production of new product quantity and appropriate production ability, waste product recovery efforts and recovery capacity and other network planning problems. Short-term level need to determine the actual forward logistics, reverse logistics and information flows, especially need to design the short-term operation strategy. Obviously, in these levels and among each level, distributed decision-making problems may be appearing.

As for the researches of distributed decision-making in forward supply chain, one may differentiate between strategic, tactical, and operational tasks in supply chain management. At these strategic level, the over all design of the supply network is achieved [5], while at the tactical level one has contracts that organize the concrete relationship between the supply chain partners [6]. Finally, at the operational level the current material and information flow is actually scheduled [7]. Clearly, the operational interrelation depends on the type of contract one has to observe. On the other hand, the contract one is choosing is usually not independent of the operational setting. More specifically, the operational decision problem has to be anticipated in designing a contract. Schneeweiss et al. [8] investigates the impact of specific contract parameters on the operational performance. It is interesting to observe that the relationship between the design level and the operational planning level is not carefully discussed in the literature. Often one is only considering the logistic interdependencies within a given contract [9], or one is analyzing the formation of contracts mainly within the framework of (economic) contract theory without considering detailed features of the operational level [10]. Furthermore, Schneeweiss et al. [11] investigates the impact of different types of contracts on the operational performance. They are not discussing the actual negotiation of a contract between two supply chain partners but provide rational arguments as to the possible type of contract the partners might choose.

In recent years, there is many other scholars research the distributed decision-making problems and its applications in forward supply chain [12-17], and proposed many algorithms to solve distributed decision-making models [18-21]. However, researches on distributed decision-making in the processing of reverse logistics operation are still fairly new. As we know, negotiation-based coordination in product recovery networks proposed by Walther et al. [22], and the model of decentralized decision-making and the method of protocol design for recycled material flows suggested by Hong et al. [23] are only the most representative two research achievements in this realm. From the present research current situation, the distributed decision making problems with important practical significance in closed-loop supply chain has few other study involved.

3 Coordination of Closed-Loop Supply Chain

The research field through the contract mechanism to coordinate the closed-loop supply chain is develops quickly at present. The existing researches mainly discuss the contract parameters on overall performance, focusing on different decision makers owning private information respectively under the condition of coordinating mechanism, including income/expense sharing contract, quantity commitment contract and buy back contract [24-27].

The price decision is an important mechanism to coordinate supply chain. The price decision coordination methods used by existing researches include quantity discount, price discount, batch discount contracts, and so on [28-31]. In the closed-loop supply chain, manufacturers, retailers and collector can not know other decision maker's private information, but they make policy respectively under the common market demand information. Manufacturers induce retailers to order proper quantity and collectors to adopt appropriate recycling efforts by offering right prices. But in most cases, market demand and waste products recycling information is not to be learned or unstable. On the basis of the existing price coordination mechanism, if the decision-maker to respective master market information as a basis for decisions, it is unable to achieve coordinated operation purpose.

There are leader-follower relationships within closed-loop supply chain. For the practical operations of closed-loop supply chain, the problems of coordination can be come down to Stackelberg game. For instance, Qiu and Huang [32] studied the coordination of closed-loop supply chain with product recycling in stochastic demand. They considered two recycling channels covering both manufactures and sellers to develop a Stackelberg game model for closed-loop supply chain where the manufacturers and sellers play the roles of leader and follower, respectively. Yan and Huang [33] established a multi-echelon closed-loop supply chain with 3PRLP by assuming price-sensitive stochastic demands and effort-sensitive stochastic returns, and designed a target rebate-punish contract between the manufacturer and the 3PRLP. The closed-form analytic expressions for both united optimization strategies in centralized closed-loop system and the one-leader-multiple-follower Stackelberge strategies in decentralized system are deduced, and comparisons are made from the viewpoint of contract coordination.

In distributed decision-making closed-loop supply chain, each channel member owns his private information, and decision-makers may hide the private information for the individual interests. Hence, there exist principal agent problem caused by asymmetric information. Researches in this area generally acquire the characters of corresponding agency problem and corresponding equilibrium solution through the construction of the corresponding quantitative models of principal agent problems, the model deduction and numerical experiments. For example, by comparing the linear incentive contract under symmetric information with that under asymmetric information, Guan et al. [34] analyze the effect of different parameters on the incentive coefficient, manufacturer's decision-making, the expected income of the manufacturer, and the effect of market uncertainty and risk aversion on the manufacturer's profit. Sun et al. [35] established an incentive model based on multi-task principal-agent theory. Through determining qualitative description, quantitative analysis and solution, the optimal incentive intensity coefficient, they obtained a series of meaningful conclusions. Moreover, many other scholars have analyzed the problems of information asymmetry in closed-loop supply chain from conceptions and model level by using game theory models [36]. However, whether principal-agent theory or game theory model in application, there is a certain distance from actual manipulation.

4 Operations of Closed-Loop Supply Chain

Researches of closed loop supply chain operation are the main task and core target for closed loop supply chain management. From the point of view of operations management, researches of closed-loop supply chain operation mainly focus on the production planning, the distribution channel choosing and designing, manufacturing process control, partner selection, performance evaluation, coordination and cooperation of closed-loop supply chain and service management. Xu et al [37] introduced in detail on these specific aspects of the research achievements, here no longer give unnecessary details.

For the static systems, researchers often dealt with the uncertain problems by using robust optimization methods. For instance, Lee and Dong [38] proposed dynamic location and allocation models to cope with the issues of reverse logistics network design. A two-stage stochastic programming model is further developed by which a deterministic model for multi period reverse logistics network design can be extended to an account for the uncertainties. A solution approach integrating is gently proposed sampling method with a heuristic algorithm is also proposed in this research. Qin and Ji [39] employ a fuzzy programming tool to design the product recovery network under uncertainties. Based on different criteria, three types of optimization models are proposed and some properties of them are investigated. To solve the proposed models, they design a hybrid intelligent algorithm which integrates fuzzy simulation and genetic algorithm. Mukhopadhyay and Ma [40] consider the problem of joint procurement and production decisions in remanufacturing under quality and demand uncertainty. They address this issue of a hybrid system where both used and new parts can serve as inputs in the production process to satisfy an uncertain market demand. The yield of the used parts is random. A firm needs to make production and

procurement decisions in the face of this uncertainty. Three different cases are presented depending on the amount of information the firm has about the yield date. They determine the optimal procurement and production quantity for the firm. Xu et al. [41] design a class of closed-loop supply chain with one manufacturer and one supplier, in which re-distribution, remanufacturing and reuse are considered synthetically. They establish a multi-objective robust dynamic operating model for the supply chain under the circumstance of uncertain customers' demands by using the robust linear programming method based on scenario analysis. Zhu et al. [42] consider a class of closed-loop supply chain with re-distribution, remanufacturing and reuse under the environment of uncertain reverse logistics flow, and construct a multi-objective operating model for this closed-loop supply chain. Pan and Nagi [43] Robust supply chain design under uncertain demand in an agile manufacturing setting, and consider the integrated optimization of logistics and production costs associated with the supply chain members.

For the parameter, time lag, and exogenous disturbance uncertainties in dynamic supply chain system, researchers often use robust control theory methods to obtain the system's robust control strategies. In the field of robust control for closed-loop supply chain system, Boakas [44] dealt with the control of production systems that produce many part types with limited capacity. First, he used a simple model to show that the inventory control problem can be solved using modern control theory. You et al. [45] constructed one z-transformation model analyzing the bullwhip effect for multi-echelon inventory system under closed-loop supply chain, and put forward the control strategies. Huang et al. [46] established dynamic models of closed-loop supply chain & robust H_∞ control strategies with time-delay & parameter uncertainty. Guo and Xu [47] presented a H_∞ control method to reduce the bullwhip effect in the remanufacturing system. In this method, a remanufacturing system supply chain system model with a single-item is build, which is based on two types of inventories: the actual product inventory in a manufacturer and the virtual inventory used by a customer.

These studies above are mostly considering the problems of operation strategy design for closed-loop supply chain system, and the problems of uncertainties and robustness appeared in the processing of closed-loop supply chain operation. The researches of present situation is concerned, there are also very few people to study the robust strategy design problems for the coordination operation of closed-loop supply chain under the circumstance of uncertainty based on distributed decision-making.

5 Conclusion

Theoretical researches for closed-loop supply chain are mostly focus on constructing and considering the problems of decision-making in reverse logistics by using microeconomics and game theory methods, but for more practical significance in closed loop supply chain optimization, control and coordination of micro operation aspect question was rarely studied, especially in distributed decision-making model of closed-loop supply chain coordination strategy design. But considering present

research, following problems may be the future research directions for coordinate operation of distributed decision-making closed-loop supply chain.

Closed-loop supply chain is a complex distributed decision-making system. For each decision-maker in the closed-loop supply chain, his decision is not isolated and arbitrary, relationships between his feasible operation strategy and the interests of its upstream and downstream enterprises is very closely. How to operate the closed-loop supply chain system in phase, and achieve interactive win for decentralization decision-makers is very important;

In the operation process of closed-loop supply chain, it's a key problem for the enterprise real and potential implementing closed-loop supply chain management to prompt manufacturers effectively manufacturing/ remanufacturing products (including new products and remanufactured products), retailers effectively selling products (including new products and remanufactured products), retailers or third party recyclers effectively recycling waste products under the asymmetric information condition existing between manufacturer and retailer, manufacturer and the third party recycler.

Closed-loop supply chain system has its inherent high uncertainty characteristics, uncertainties coming from the forward supply chain supply, manufacturing and sales as well as from the reverse supply chain recycling and reuse of waste products and other aspects of the internal workings, from natural disasters, accidents and the external economic environment and other aspects of the external contingencies, will influence the normal operation of closed-loop supply chain system. It will have important theory value and broad application prospects to coordination operation strategy designing for closed-loop supply chain in microcosmic operation level under uncertain environment by using robust optimization / robust control method.

References

1. Rogers, D.S., Tibben-Lembke, R.S.: An examination of reverse logistics practices. Journal of Business Logistics 22(2), 129–148 (2001)
2. Toffel, M.W.: The growing strategic importance of end-of-life product management. California Management Review 45(3), 102–129 (2003)
3. Lambert, S., Riopel, D., Abdul-Kader, W.: A reverse logistics decisions conceptual framework. Computers & Industrial Engineering (2011), doi: 10.1016/j.cie.2011.04.012
4. Kannan, G., Sasikumar, P., Devika, K.: A genetic algorithm approach for solving a closed loop supply chain model: A case of battery recycling. Applied Mathematical Modelling 34, 655–670 (2010)
5. Geoffrion, A.M., Powers, R.F.: Twenty years of strategic distribution system design: An evolutionary perspective. Interfaces 25(5), 105–127 (1995)
6. Cachon, G.P., Lariviere, M.A.: Capacity allocation using past sales: When to turn-and-earn. Management Science 45(5), 685–703 (1999)
7. Henig, M., Gerchak, Y., Ernst, R., Pyke, D.F.: An inventory model embedded in designing a supply contract. Management Science 43(2), 84–189 (1997)
8. Schneeweiss, C.: Hierarchies in Distributed Decision Making. Springer, Berlin (2003)
9. Moinzadeh, K., Nahmias, S.: Adjustment strategies for a fixed delivery contract. Operations Research 48(3), 408–423 (2000)

10. Anupindi, R., Bassok, Y.: Approximations for multi-product contracts with stochastic demands and business volume discounts: Single supplier case. IIE Transactions 30(8), 723–734 (1998)
11. Schneeweiss, C., Zimmer, K., Zimmermann, M.: The design of contracts to coordinate operational interdependencies within the supply chain. Int. J. Production Economics 92, 43–59 (2004)
12. Kelly, J.D., Zyngier, D.: Hierarchical decomposition heuristic for scheduling: Coordinated reasoning for decentralized and distributed decision-making problems. Computers and Chemical Engineering 32, 2684–2705 (2008)
13. Chadwick, D.: Coordinated decision making in distributed applications. Informations Security Technical Report 12, 147–154 (2007)
14. Jung, H., Chen, F.F., Jongju, B.: Decentralized supply chain planning framework for third party logistics partnership. Computers & Industrial Engineering 55, 348–364 (2008)
15. Zangeneh, A., Jadid, S., Rahimi-Kian, A.: A hierarchical decision making model for the prioritization of distributed generation technologies: A case study for Iran. Energy Policy 37, 5752–5763 (2009)
16. Nishi, T., Konishi, M., Ago, M.: A distributed decision making system for integrated optimization of production scheduling and distribution for aluminum production line. Computers and Chemical Engineering 31, 1205–1221 (2007)
17. Dong, M.: An analytic integrated framework for distributed energy supply chain networks. Journal of Shanghai Jiaotong University 42(11), 1772–1775 (2008)
18. Schneeweiss, C.: Distributed decision-making - a unified approach. European Journal of Operational Research 150, 237–252 (2003)
19. Gaudreault, J., Frayret, J.M., Pesant, G.: Distributed search for supply chain coordination. Computers in Industry 60, 441–451 (2009)
20. Li, Y., Yang, S.L.: Distributed multi-level collaborative production - distribution planning model and solution for supply chain. China Mechanical Engineering 19(21), 2589–2594 (2008)
21. Cao, D., Chen, M.: Capacitated plant selection in a decentralized manufacturing environment: A bi-level optimization approach. European Journal of Operational Research 169, 97–110 (2006)
22. Walther, G., Schmid, E., Spengler, T.S.: Negotiation-based coordination in product recovery networks. Int. J. Production Economics 111, 334–350 (2008)
23. Hong, I.H., Ammons, J.C., Realff, M.J.: Decentralized decision-making and protocol design for recycled material flows. Int. J. Production Economics 116, 35–227 (2008)
24. Zhang, Q., Gao, Y.: Coordination of closed-loop supply chain under stochastic demands. Science & Technology Information 27(122), 108 (2009)
25. Chao, G.H., Seyed, M.R., Savaskan, I.R.C.: Quality Improvement Incentives and Product Recall Cost Sharing Contracts. Management Science 55(7), 1122–1138 (2009)
26. Wang, Y.Y.: Closed-loop supply chain coordination under disruptions with buy back contract. Operations Research and Management Science 18(6), 46–52 (2009)
27. Xing, G.J., Lin, X.Y., Da, Q.L.: Study on revenue sharing coordination mechanism of closed-loop supply chain based on flexibility in ordering product. Soft Science 23(8), 30–35 (2009)
28. Xie, C.H., Zhong, W.J., Zhang, Y.L.: Privacy preserving supply cha in quantity discount contract design. Journal of Southeast University (English Edition) 25(1), 132–137 (2009)
29. Jin, Z.Y., Peng, Z.H.: Designing quantity discount with asymmetric demand information. Mathematics in practice and Theory 39(8), 9–18 (2009)

30. Xie, J.X., Zhou, D.M., et al.: Price discount based on early order commitment in a single manufacturer–multiple retailer supply chain. European Journal of Operational Research 200(2), 368–376 (2010)
31. Wang, J.P.: A quantity discount design model for two-echelon supply chain coordination. Journal o f Systems & Management 18(2), 226–231 (2009)
32. Qiu, R.Z., Huang, X.Y.: Coordination model for closed-loop supply chain with product recycling. Journal of Northeastern University (Natural Science) 28(6), 883–886 (2007)
33. Yan, N.N., Huang, X.Y.: Models of closed-loop supply chain with third-party reverse logistics and their applications. Journal of Management Science in China 11(4), 83–93 (2008)
34. Guan, Q.L., Zhou, G.G., Cao, J.: On the incentive mechanism for multi-task principal-agent closed-loop supply chain. Industrial Engineering Journal 12(4), 23–27 (2009)
35. Sun, B.F., Yang, H., Wu, M.N.: Study on incentive mechanism of closed-loop supply chain. Journal of Highway and Transportation Research and Development 26(2), 134–138 (2009)
36. Xiao, T., Shi, K., Yang, D.: Coordination of a supply chain with consumer return under demand uncertainty. Int. J. of Production Economics 124, 171–180 (2010)
37. Xu, J.W., Zhu, Y.L., Huang, X.Y., Qiu, R.Z.: Research progress on closed-loop supply chain management. Systems Engineering, 26(8), 1–10 (2008)
38. Lee, D.H., Dong, M.: Dynamic network design for reverse logistics operations under uncertainty. Transportation Research Part E 45, 61–71 (2009)
39. Qin, Z., Ji, X.: Logistics network design for product recovery in fuzzy environment. European Journal of Operational Research 202(2), 479–490 (2010)
40. Mukhopadhyay, S.K., Ma, H.: Joint procurement and production decisions in remanufacturing under quality and demand uncertainty. Int. J. Production Economics 120(1), 5–17 (2009)
41. Xu, J.W., Huang, X.Y., Guo, H.F.: Robust optimizations model for operating of closed-loop supply chain with uncertain demands. Systems Engineering and Electronics 30(2), 283–287 (2008)
42. Zhu, Y.L., Xu, J.W., et al.: Robust operating strategy design for closed 2loop supply chain with uncertain reverse logistics flow. Control and Decision 24(5), 711–716 (2009)
43. Pan, F., Nagi, R.: Robust supply chain design under uncertain demand in agile manufacturing. Computers & Operations Research 37, 668–683 (2010)
44. Boakas, E.K.: Manufacturing systems: LMI approach. IEEE Transactions on Automatic Control 51(6), 1014–1018 (2006)
45. You, J.X., Sui, M.G., Huo, J.Z.: Bullwhip effect in closed-loop supply chain. System Engineering Theory and Practice 27(12), 111–116 (2007)
46. Huang, X.Y., Yan, N.N., Qiu, R.Z.: Dynamic models of closed-loop supply chain & robust H∞ control strategies with time-delay & parameter uncertainty. Computer Integrated Manufacturing Systems 13, 1313–1321 (2007)
47. Guo, H.F., Xu, J.W.: Remanufacturing system model and H∞ control of its bull whip effect. Systems Engineering and Electronics 30(5), 914–916 (2008)

Identification of Currency Crises: The Empirical Analysis Based on the Binary Regression Tree

Yin Yu

Economics & Management School, Shanghai Second Polytechnic University, Shanghai, China
yuyinnn@yahoo.cn

Abstract. Emerging markets for central, in order to anatomize currency crises, currency crises were decomposed as the macroeconomic-vulnerability crisis, the self-fulfill crisis, the competitiveness- decline crisis, the bank-crises-lead crisis, the debt-default crisis, the capital-flow-reversal crisis and the contagious crisis. According to the choice of identifying indicators, a two-layer identifying system was build. An empirical analysis, which used binary regression tree and the quarter data for the eight countries, identified the six forms of currency crises in emerging markets. Results show that the expected exchange rates and the domestic credit have stronger warning to the currency crises.

Keywords: Currency crises, binary regression tree, indicators.

1 Introduction

Anatomy of financial crises is of advantage to have a good grip of causes of financial crises and defense effectively. Eichengreen and Porters(1986)[**1**] isolated bank failures, debt defaults and foreign-exchange market disturbances from a financial crisis. Based on asymmetric information framework, Mishkin(1992)[**2**] anatomized a financial crisis as a typical financial crisis with increasing in interest rates, stock market declining and increasing in uncertainty and a debt-deflation crisis with a sharp decline in prices and the increased burden of indebtedness. Within the area of study of currency crises, further, Kaminsky(2006)[3] marked six forms: the first-generation model crisis, the second-generation model crisis, the third-generation model crisis, the sovereign crisis, the capital sudden stops crisis and the self-fulfilling crisis which it come to the mature economies with immaculate fundamentals. However, the Kaminsky's classification neglected the contagious characteristics of currency crises. Generally, a currency crisis can trigger a financial crisis or even economic crisis. So, the key reasons of currency crises may contribute to assess the financial system and market volatility. Anatomy of currency crises would not only be propitious to the profound understanding of the causes of currency crises, but also be favorable to predict crises effectively.

In this paper, firstly, I anatomized currency crises as the macroeconomic-vulnerability crisis, the self-fulfill crisis, the competitiveness-decline crisis, the bank-crises-lead crisis, the debt-default crisis, the capital-flow-reversal crisis, and the contagious crisis.

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 449–455.
springerlink.com

Comparing to the Kaminsky's classification, I considered the contagious crisis based on the third-generation model and the contagion theory. Then, a currency crisis was defined as the weighted of Fisher's exchange rates changes and reserve changes. This was a difference in similar literatures. Finally, Brazil, Mexico, Argentina, Korea, Thailand, Indonesia, Philippines, Malaysia were chosen and the quarter data spanned from 1971:1 to 2006:2. The binary regression tree captured the six different forms.

2 Identification of Currency Crises

2.1 Anatomy of Currency Crises

These financial crises, from financial crisis of Eurocurrency system from 1992 to 1993, to financial crisis of Latin America from 1994 to 1995, to financial crisis of southeast Asia from 1997 to 1998, and to financial crisis of America caused by subprime mortgage crisis from 2007, all had one common character: the crises appeared obvious infectivity, one country's financial crises quickly diffused to other countries and areas and evolved into territorial or even global crises[**4, 5, 6**]. So, I anatomized currency crises as the macroeconomic-vulnerability crisis, the self-fulfill crisis, the competitiveness- decline crisis, the bank-crises-lead crisis, the debt-default crisis, the capital-flow-reversal crisis and the contagious crisis[**7**].

2.2 Defining a Currency Crisis

The concept of exchange market pressure was first put forward by Girton and Roper (1977)[**8**]. The model organized the analysis around the demands and supplies of national monies at time *t*.

$$H_t = F_t + D_t = P_t Y_t^{\beta_t} exp\left(-\alpha_t I_t\right) \tag{1}$$

Where H_t was supply of base money issued by the central bank at time t. F_t was base money created against the purchase of foreign assets. D_t was base money created by domes-tic credit expansion. P_t was price level. Y_t was real income. I_t was index of interest rates. β_t was income elasticity, and α_t was interest rates coefficient; α_t, β_t> 0. F_t was determined by

$$F_t = R_t Re_t \tag{2}$$

Where Re_t was stock of international reserves (primary assets) held by the authorities at time t. R_t denoted parity or currency value of primary reserve assets at time t.

Took logarithms of both sides and differentiating the resulting expression, and asterisks denote the foreign country.

$$(re_t - re_t^*) - (i_t - i_t^*) = -(d_t - d_t^*) + (p_t - p_t^*) + \beta_t (y_t - y_t^*) - (1+\alpha_t)(i_t - i_t^*) \tag{3}$$

Using purchasing power parity to substitute the rate of depreciation for the inflation differential, r_t was the domestic price of a unit of foreign exchange. Fisher's exchange rates r_t^e is

$$r_t^e = (i_t - i_t^*) + r_t \tag{4}$$

Where r_t^e was the next exchange rates expected at time *t*. Even though foundational factors influencing the exchange rates changes of a certain monetary the monetary occur nothing change, the monetary will appreciate because of increasing demand based on fluctuating anticipation if everyone anticipates appreciation of the monetary. Therefore, Fisher's exchange rates are able to response attack more availably.

Considered (3) (4), the following equation could be gotten

$$r_t^e - (re_t - re_t^*) = (d_t - d_t^*) - \beta_t(y_t - y_t^*) + (1+\alpha_t)(i_t - i_t^*) \tag{5}$$

Based on ERW[12,13], the exchange risk pressure index (*RPI*) was constructed by Fisher's exchange rates changes.

$$RPI_t = \alpha \Delta r_t^e - \beta(\Delta re_t - \Delta re_t^*) \tag{6}$$

Where RPI_t were weighted average of changes in Fisher's exchange rates and in reserves; Δre_t and Δre_t^* denoted changes in domestic and reference reserves in time t; α and β were the inverses of standard deviations for Δre_t^e and (Δre_t - Δre_t^*), respectively. Equation (6) reflected the interrelation between exchange rates and interest difference. Once currency crisis was defined as those extreme values of the RPI,

$$Crisis_t = \begin{cases} 1, & if \quad RPI_t > \mu_{RPI} + \kappa\sigma_{RPI} \\ 0, & else \end{cases} \tag{7}$$

Where μ_{RPI} and σ_{RPI} were the mean and the standard deviation of total sample, k=1. Therefore, a crisis nation occurred at where its RPIs had been strained to the utmost, moreover, once currency crisis was a period with an extreme value.

3 Empirical Analyses

Indicators used for testing forms of currency crises were constructed by analyzing the indicator frames of the literatures on financial and currency crises. They were mostly quantificational indicators. Considering nonhomogeneous data, I selected the binary regression tree to checkout classifications of currency crises. Whether classifications of currency crises could be identified would be the main issues needing to be addressed.

Regression tree is computer intensive data mining technology starts with the seminal papers by Breiman el at. (1984)[**9**]. The technology has been applied to broad areas, including engineering, market, credit risk assessment, genetics, meteorology, etc. The binary regression tree is mainly used for choosing explanatory variables and critical values and identifying interaction between variables from a tendency of crises.

3.1 Indicators and Data

Based on definition of a currency crisis in this paper, "a crisis quarter" was the quarter with the big RPI value. "A crisis scene" was defined as a crisis quarter plus two

quarters pre-crisis. Samples chosen were Brazil, Mexico, Argentina, Korea, Thailand, Indonesia, Philippines, and Malaysia. Time window was the first quarter of 1971 to the second quarter of 2006. Table 1 showed the time of crises, in which total number of observations in crises were 153, and 755 were in calm.

The indicator set consisted of two layers. The first layer was categories, and the second layer was twenty indicators. They were: (a) the macroeconomic-vulnerability: Financial Deficit/GDP (F/GDP) and M_1 Growth(M_1); (b) the competitiveness-decline: Import(IP), Export(EP), Real Exchange rates(RER), Terms of Trade(TT), Output(OP); (c) the bank-crises-lead: Domestic Credit/GDP(D/GDP), M_2/Foreign Reserves (M_2/Re), M_2 Multiplier(M_2M), Bank Saving(BS), Stork Price(SP), Bank Crises Dumb Variable (BC); (d) the contagious: Current Account Balance /GDP (C/GDP), Financial Market Correlative Effect (FMCE); (e) the debt-default: Foreign Debt/Export (FC/EP), Short-term Debt/ Foreign Reserves (SD/Re); (f) the capital-flow-reversal: World Real Interest rates (WIR) and Foreign Reserves (Re). Data of bank crisis dumb variable were from correlative literatures in the field of crises study, and of other variables were from IMF's international financial statistics database and World Bank's database.

Table 1. Crisis Periods of sample countries

Samples	Periods under observation	Periods under crises(quarters）
Brazil	1980Q1-2006Q2	1983Q1,1985Q4,1986Q4,1989Q3,1990Q4,1991Q4, 1994Q4, 1999Q1
Argentina	1979Q2-2006Q2	1980Q1,1981Q1,1982Q3,1985Q2,1986Q3,1989Q2, 1990Q1, 1994Q4,2002Q1
Mexico	1981Q2-2006Q2	1982Q1,1982Q3,1982Q4,1984Q4,1985Q3,1987Q4, 1992Q4,1994Q4
Indonesia	1978Q2-2006Q2	1978Q4,1983Q2,1986Q3,1992Q4,1997Q3,1997Q4, 1998Q1
Korea	1977Q1-2006Q1	1980Q3, 1983Q1, 1983Q2, 1997Q3, 1997Q4
Malaysia	1971Q2-2006Q2	1975Q3,1985Q3,1997Q3,1997Q4,1998Q2
Philippines	1977Q2-2006Q2	1981Q1,1983Q4,1984Q3,1986Q1,1997Q3
Thailand	1977Q2-2006Q2	1978Q4,1979Q1,1979Q3,1981Q3,1984Q4,1997Q3, 1997Q4,1998Q1,1999Q3,2000Q3,2005Q3

3.2 Analysis and Results

Firstly, data of all indicators were converted into percentage. Processing of the binary regression tree is dualistic because a father-node is divided into two son-nodes in exactly. Each son-node is treated repeatedly like a father-node, so processing of the binary regression tree is also recursive. Besides need to divide a node in two and to consider stopping the regression tree, each terminal node should be designed as a class. In my analysis, each class matched an identifying indicator. That is, the regression tree algorithm designed out twenty nodes to match all observations (See Table 2).

Table 2. Identifying Characteristics of Indicators

	Characteristics	a	b	c	d	e	f
1	r^e_t<0.064; *RER*<19.7; M_2/Re<17.7; *DIR*<11.0	*					
2	r^e_t<0.064; *RER*<19.7; M_2/Re<17.7; *DIR*>11.0		*				
3	r^e_t<0.064; *RER*<19.7; *DIR*<17.5; M_2/Re>17.7		*				
4	r^e_t <0.064; *RER*<14.1; *DIR>17.5*			*			
5	r^e_t <0.064; *DIR>17.5*; 14.1<*RER*<19.7	*					
6	r^e_t <0.064; *RER*>19.7; *FC/EP*<77.5; *WIR*<92.0		*				
7	r^e_t<0.064; 19.7<*RER*<26.4; *FC/EP*<77.5; *WIR*>92.0						*
8	r^e_t <0.064; *RER*>26.4; *FC/EP*<77.5; *WIR*>92.0						*
9	r^e_t <0.064; *RER*>19.7; *FC/EP*>77.5						*
10	r^e_t>0.064; *D/GDP*<60.5; *SD/Re*<17.4; *WIR<81.7*			*			
11	r^e_t>0.064; *D/GDP*<60.5; *SD/Re*<17.4; *WIR>81.7*			*			
12	r^e_t>0.064; *D/GDP*<60.5; *SD/Re*>17.4; *FC/EP<54.2; RER<33.7*					*	
13	r^e_t>0.064; *D/GDP*<60.5; *SD/Re*>17.4; *FC/EP<54.2; RER>33.7*				*		
14	r^e_t>0.064; *D/GDP*<60.5; *SD/Re*>17.4; *FC/EP>54.2; SP<65.8*	*					
15	r^e_t>0.064; *D/GDP*<60.5; *SD/Re*>17.4; *FC/EP>54.2; SP>65.8*						*
16	r^e_t >0.064; *D/GDP*>60.5; *OP*<89.0; *DIR<9.8*		*				
17	r^e_t >0.064; *D/GDP*>60.5; *OP*<89.0; *DIR>9.8; M_2M<27.7*	*					
18	r^e_t >0.064; *D/GDP*>60.5; *OP*<89.0; *DIR>9.8; M_2M>27.7*		*				
19	r^e_t >0.064; 60.5<*D/GDP*<65.5; *OP*>89.0			*			
20	r^e_t >0.064; *D/GDP*>65.5; *OP*>89.0;				*		

The pointer vector was selected from sample set, and took the highest and the lowest values. All values (N=908) was been the only one, and were arranged in sequence. Then, a standardized problem set of the attribute vector consisted of the values taken from total values by median value method. The ten variables used for classifying all observations in indicator set. They were real exchange rates, domestic credit/GDP domestic real interest rates, foreign debt/export, short-term debt/foreign reserves, output, M_2/foreign reserves, world real interest rates, stock price and M_2 multiplier[**3, 10-13**]. Fisher's exchange rates were put in test to test its caution ability in an oncoming crisis. It was interesting that the first separated point was Fisher's exchange rates. This showed that Fisher's exchange rates had a strong caution power

for a crisis. Likelihood of a crisis happening was 71.1% as Fisher's exchange rates were 0.064% or less. However, Fisher's exchange rates presented only to the first separation point. It meant that Fisher's exchange rates did not affect other variables' classification in indicator set.

The next tiers were "real exchange rates" and "domestic credit/GDP". It had been proved that the caution of the real exchange rates was very significant. High "domestic credit/GDP" might foreshow larger pressure on bank and lower level liquidity used for replying attack. "Yes" branch in "real exchange rates" included "domestic real interest rates" and "M_2/foreign reserves", reflected changes in domestic real interest rates, real exchange rates and capital sufficiency rate of central bank. "No" branch in real exchange rates reflected interaction among foreign debt, real exchange rates and world interest rates. "Yes" branch in "domestic credit/GDP" mirrored changes in short-term foreign debt as world interest rates, real exchange rates and stock market were fluctuating, and was able to open out status of capital flow, current account and crises contagion. "No" branch in "domestic credit/GDP" mirrored entity economy operation, degree of financial market development, and whether invest was in open channel or not. Twenty groups were delimited according to node. Characteristics of each group were shown in Table 2. The groups based on Fisher's exchange rates appreciation were 9, as others 11 groups were under Fisher's exchange rates depreciation. According to crises classifications, those groups with similar characteristics were classified into the same group. However, it was also difficult to classify clearly. So, the rank was only a coarse process.

Above analysis demonstrated that a crisis is caused by multifactor together under oscillatory relationship between markets.

4 Conclusions and Implications

This paper identified the forms of crises of eight countries by the binary regression tree. Although boundaries were hazy, it was important to understand currency crises. At the same time, some implications on defending crises were profound.

Firstly, the binary regression tree analysis showed Fisher's exchange rates denoted public expectation to future currency market price. Adjusting policy led directly public expectation, thus it would be important that policy maintain public confidence on economic development. Secondly, the binary regression tree analysis showed domestic credit was an important indicator to forecast financial crises. Level of domestic credit affected directly the bank system capacity and elasticity on replying market pressure. And so, lend entering into capital markets by different forms should be forbidden strictly. Especially, speculative behavior of enterprises and individuals bank bad lent in capital market should be forbidden strictly[**14**].

In addition, besides need to consider fluctuation of exchange market, managers need to consider capital price bubble to cut down impact of bubble breakdown on entity economy.

Acknowledgements. This paper is supported by "Innovation Program of Shanghai Municipal Education Commission" (Project No. 12YS160).

References

1. Eichengreen, B.J., Portes, R.: The anatomy of financial crises, NBER Working Paper, no. 2126 (1986)
2. Mishkin, F.S.: Anatomy of a financial crisis. Journal of Evolutionary Economics 2, 115–130 (1992)
3. Kaminsky, G.L.: Currency crises: Are they all the same? Journal of International Money and Finance 25, 503–527 (2006)
4. Eichengreen, B.J., Rose, A., Wyplosz, C.: Exchange market mayhem: The antecedents and aftermath of speculative attacks. Economic Policy 10(21), 251–312 (1995)
5. Eichengreen, B.J., Rose, A., Wyplosz, C.: Contagious currency crises: First tests. Scandinavian Journal of Economics 98(4), 463–484 (1996)
6. Krugman, P.: Will there be a dollar crisis? Economic Policy 22(51), 435–467 (2007)
7. Yu, Y., Shu-qiong, F.: Analysis and construction of the macro-monitoring indicators based on the currency crisis. Discovery and Innovation 21(1&2) (2009)
8. Girton, L., Roper, D.: A monetary model of exchange market pressure applied to postwar Canadian experience. American Economic Review 67, 537–548 (1977)
9. Breiman, L., Friedman, J., Stone, C.J., Olshen, R.A.: Classification and regression trees. Chapman & Hall, London (1984)
10. Krugman, P.: Building a green economy. New York Times 5 (2010)
11. Cipollini, A., Kapetanios, G.: Forecasting financial crises and contagion in asia using dynamic factor analysis. Journal of Empirical Finance 16(2), 188–200 (2009)
12. Bordo, M.D., Meissner, C.M., Stuckler, D.: Foreign currency debt, financial crises and economic growth: A long run view. Journal of International Money and Finance 29(4), 642–665 (2010)
13. Claessens, S., Ayhan Kose, M., Terrones, M.E.: The global financial crisis: How similar? How different? How costly? Journal of Asian Economics 21(3), 247–264 (2010)
14. Pontines, V., Siregar, R.: Fundamental pitfalls of exchange market pressure-based approaches to identification of currency crises. International Review of Economics & Finance 17(3), 345–365 (2008)

The Issues of Ideal Faith Education for Contemporary College Students and the Countermeasures

Lili Pan[1] and Jia Guo[2]

[1] Beijing Electronic Science and Technology Vocational College, Beijing, China
[2] North China Institute of Science and Technology, Beijing, China
lplb@163.com

Abstract. Education of ideal faith is an essential part of ideological and political education for college students. It appears new characters and problems under the new historical conditions. Taking 528 college freshmen as an example to investigate and analyze ideal faith of contemporary college students in the new period. The result shows that in general the ideal faith of college students is good, but there are some problems with distinctive characteristics of times which should not be ignored. Ideal belief education for college students should be closely related to the reality of their lifestyle and thoughts. Pedagogue should follow the law of formation governing ideal belief, should be integrated into the process of college students' self-development, and should tend to deal with their problems as well. Realization of the social ideal, long-term ideal and lofty ideal should be integrated into the college students' individual ideal, short-term ideal and general ideal.

Keywords: Ideal faith, issues, countermeasures.

1 Introduction

China is going through social changes unparalleled in scope. While releasing tremendous dynamism for China's development and progress, these changes have inevitably given rise to conflicts and problems. Secretary-general Hu Jintao emphasized at a meeting commemorating the 90th anniversary of the founding of the communist party of China that there are always many energetic young people with lofty ideals and great passion in the Party. This is an important guarantee for the Party to remain robust and youthful 90 years after its birth. Young people represent the future of both China and its people. They also represent the future and hope of the Party. The future and density of China depends on the educational level of youth and people. The core of youth education is ideal faith education. Ideal faith is the psychic phenomena and the subjective consciousness of human being. It is the important symbol for human beings to distinguish from animals. Therefore, for college students who being the critical period of life is extremely of consequence to build up ideal faith.

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 457–462.
springerlink.com

2 Issues of Ideal Faith Education for Contemporary College Students

2.1 Inconsistency in between Society Circumstances and College and Family Education

With the rapid development of Chinese economy in the course of globalization, China is being faced with the huge opportunities and severe challenges that she never went through it. The ways of life, values and thinking methods of people are experiencing the trend of diversity. In the globalization age, college students are enjoying the huge changes on their lifestyle and thinking. They are being faced with all-new stage about practice and cognition. However, globalization is the two-edged sword. Some western developed countries are being resorted to globalization by means of all kinds of ways, such as economy, politics and culture to impact on the ideology of college students. The purpose of some western developed countries is the disintegration of china and at the same time the comedown of communism ideal faith. The money worship, extreme individualism and gastronomy reach their zenith at globalization age resulted in the fluctuation of some college students' ideal faith. Consequently, the complexity of social circumstance is more of the necessity of ideal faith as for the inspirit support and power headspring of college students.

According to the questionnaire about "Pursuit of their ideals" of 528 college freshmen of North China Institute of Science and Technology and Beijing Electronic Science and Technology Vocational College, 15% students said they don't think over this question. 25% students said they don't know. They are very confused about the future. 50% students said they hope they will successfully graduate and purse a satisfying job. 5% students said they wish they will devote themselves to the development of society and country. 5% students didn't answer this question. From the whole state, part of college students' value are orientating from the social standard to the individual standard deviation. They more and more depend on the utilitarian and pragmatism. They advocate the enjoyment and consumption. The "psychological phenomenon of deconsecrating" is the key reason for some students to wipe off idealism and ultimate values. Meanwhile, there are surprising phenomena come forth, such as loftiness is be mocked and virtue is be belittled. Therefore, guiding college students to build up the lofty ideal of communism and the common ideal of socialism with Chinese characteristics for ideological-political education operator are experiencing the severe challenges under the background of without ideal faith and being live in a worldly manner.

However, college students' ideal faith education force has not yet formed. The conjunction of school education, family education and social education are usually disconnected from each other and even come forth the conflict and antinomy that make the education effects be counteracted or weaken one another. For example, socialism core value system are being advocated in school education while at the same time they are being ignored or disputed in family education. The New Generation with "Four Qualifications" is being cultivated in school education. Simultaneously, some newspapers and periodicals, broadcasting, internet, TV entertainment, etc spare no effort to publishing, selling, hawking and advocating

mammonism, hedonism and extreme individualism to poison college students' ideology in order to business profits. So the effect of ideal faith education for college students must be badly affected.

2.2 Inconsistency in between Social Ideal Faith Education and Individual Ideal Faith Education

Secretary-general Hu Jintao emphasized on 90th anniversary of celebrated CCP foundation that young people in China must acquire a keen appreciation of the glorious history of tenacious struggles waged by the Chinese people in China's modern history. They should cherish a lifelong love for our great motherland, our great people, and our great nation. They should have lofty ideals and firm convictions, gain knowledge and skills, develop moral integrity and willpower, and prepare themselves for hard work so that they can fully release their talent and prove their value on the big stage of life. Let their youthful vigor shine with brilliance in the course of rendering great services to the Party and the people.

The kernel of ideal faith is the relationship between social ideal faith education and individual ideal faith education. They impact one another and restrict each other. Individual ideal faith is depended on and restricted by social ideal faith while at the same time social ideal faith is cohesion and sublimation of individual ideal faith. Some ideological-political education operators cannot correctly grasp the relationship between social ideal faith and individual ideal faith and cannot deal with the relationship between realism and direction of guiding about ideal faith education for being lack of demonstration research of ideal faith education. Therefore, two extremeness phenomena appear in ideal faith education. First and foremost, ideological-political education operators more emphasize social ideal faith than individual ideal faith so that ideal faith is empty while at the same time is difficult to carry out it. Secondly, ideological-political education operators more emphasize individual ideal faith than social ideal faith so that ideal faith cannot come true because of without the correct direction.

According to the questionnaire about the relationship between social ideal faith and individual ideal faith of 528 college freshmen of North China Institute of Science and Technology and Beijing Electronic Science and Technology Vocational College, 45% college students said individual ideal faith must be achieved under the foundation of accomplishment of social ideal faith. 26% college students said individual ideal faith should be paramount when they conflicted with one another. 20% college students said individual ideal faith should be the first and foremost. At the same time, impairment of the usefulness or value of some people or property is necessary. According to the questionnaire about the relationship between common ideal faith and long-term ideal faith of 528 college freshmen of North China Institute of Science and Technology and Beijing Electronic Science and Technology Vocational College, 57% college students said they identify with the communism ideal faith but they are not sure to realize the communism ideal. 23% college students said for China the most important cause is to achieve common ideal soon. Communism ideal is for away and not within the foreseeable future. 9% college students didn't answer this question.

We can make conclusions that the key problem for ideal faith education is followed some reasons. First and foremost, some college students weaken the consciousness of be in pursuit of lofty ideal faith. Some college students even think individual ideal faith benefits more than communism ideal faith in the long run. Secondly, the moral level of contemporary college students is higher than other communities. They are well-educated people who know what the right is or not while at the same time how to conduct oneself and to be an upright person. But some students are easily impacted on ill factors and western disadvantage trend of thought.

3 Countermeasures against Issues of Ideal Faith Education for Contemporary College Students

3.1 Innovation for Contents of Ideal Faith Education

According to the questionnaire about "what you most concerned the international issues and internal issues are" of 528 college freshmen of North China Institute of Science and Technology and Beijing Electronic Science and Technology Vocational College, 68% college students said they are eagerly concerned about Chinese international status. 54% college students said they pay more attention to the future of Chinese development. 72% college students said a poverty gap is the first and foremost Chinese government must be resolved. 82% college students said they are interested in the issue of education fairness. 97% college students said how to obtain or carve out employment is the most pressing matter for them. 62% college students said democracy processing and legal by rule are the hot problems for China. 97% college students said they are more concerned about combating corruption and building a clean government.

As post-90s contemporary college students, their self-awareness, social concern and participation enthusiasm is high. Nevertheless, they easily have irrational and subjective behavior resulted in their ideology's instability when they are faced with the complicated and multiplex social status quo, which will lead to the fluctuating ideal faith for some college students. On the all-new historical conditions, socialism core value system must be the main contents in ideal faith education so as to construct the harmonious society under the foundation of ideal faith and moral standard.

First and foremost, ideological pedagogues should both teach college students the course or area of development of human being, society and nationality and teach them personal interests in real life so as to achieve the unity of knowledge and practice. Ideological pedagogues can use news comments in class to make college students find out the essence of internal and international status quo by means of describing, analyzing, discuss and decision-making. In the course of survey, they can both discard the dross and select the essential and eliminate the false and retain the true so that they can lessen the internal contradictions of the theory and practice of ideal faith to practice what they preach.

Secondly, contents of ideal faith education should brave in theoretical innovation. That is to say, get along with the times. Ideological pedagogues should press close to the social reality and students' lifestyle under the precondition of theoretical education. Building socialism core value system is a major theoretical innovation on

the ideological and cultural construction. By summing up historical experience and scientific analysis of the current situation, our Party puts forward this theory, which is also a major task. It is of great theoretical and practical significance to construct socialism core value system, which offers theoretical instruction and value standard for ideal belief education on college students. To strengthen ideal belief education on college students, the present situation and causes must be analyzed objectively and an effective way must be explored and guided by socialism core value system. For example, such as emotion and feeling, human relationship, obtain or carve out employment, career layout and psychology problems and so on. Ideological pedagogues should guide the behavior or opinion of college students by right of solving realistic problems to make them build up correct world view, philosophy of life and values. Meanwhile, they should teach college students to be self-reliant, self-confidence and self-mightiness. As a result, contemporary college students are firmly followed the correct ideal faith to take advantage of an opportunity to achieve individual development and success.

Last but not the least, advanced socialist culture must be carried out in the course of ideal faith education. Pedagogue must incorporate improving socialist core values into the entire process of ideal faith education, promoting cultural and ethical progress. Pedagogue should vigorously promote the national spirit centered on patriotism and a spirit of the times with reform and innovation at its core, enhance self-respect, confidence and pride of the Chinese nation. Pedagogue will use the socialist concept of honor and disgrace to guide public conduct, intensify efforts to cultivate public morale, professional ethics, family values, and personal ethics, strengthen moral education among college students, and foster inspiring cultural purists and healthy, civilized lifestyles in the whole society. Moral construction of citizens proves that the great cause of socialism with Chinese characteristics not only need thousands on thousands of a model of morality, but also cultivate thousands on thousands of a model of morality. Moral force is the most important condition of national development, social harmony, and happiness in life. The annual" national moral models" selection activity can guide people to start from the basic moral norms and make people continue to pursue higher moral goals based on the basic moral norms. Ideological pedagogues should carry out moral ideal so that college students can resolutely resist the ideal and belief education of worldly and utilitarian.

3.2 Ideal and Belief Education Carrier Innovation

First and foremost, make full use of new media resources. Chinese Academy of social science literature press and Shanghai Jiao Tong University jointly hosted the "2011 blue book" about “public opinion" conference held in Beijing on July 13, 2011. Experts not only entered into the current Chinese social public opinion, public opinion and the characteristics of crisis communication law but also analyzed the influential public opinion. And" China social public opinion and crisis management report (2011)" was released formally by press of social science literature published. Blue paper pointed out that the new media is increasingly becoming the first media in numerous public opinions. New media has become the mainstream of social media as the same as traditional media.

Secondly, Ideological pedagogue should make full use of new media resources, such as a micro-blog, blogs, SMS, mobile phone, QQ, MSN, red website and so on, based on the traditional media carrier. To some extent, new medias are changing the limitations of traditional education work, which ideological pedagogue can use them grasp students' ideological trends, understand their ideological and psychological needs, enhance the actual effect of ideal belief education. However, at present a lot of college pedagogues only regard the new media as a technical means to survey it not as from the point of view of culture perspective so that college students are extremely vulnerable to the negative impact of new media. Therefore, ideological and political educators should strengthen the college students' information education, cultivate their access to information, process information and create the ability of information so as to make them understand new media correctly and utilize them effectively.

4 Conclusions

Secretary-general Hu Jintao emphasized on 90th anniversary of celebrated CCP foundation that a review of the Party's growth over the past 90 years leaves all of us the same impression, that is, since the very date of its birth, the Party has been representing young people, relying on them and enjoying their support. As a member of the young college students, college students are always responsible for the realization of the great rejuvenation of the Chinese nation task. College students should link their personal destiny with the fate of the individual and the motherland in order to make their ideal faith come true based on the motto of "keep the motherland, serve the people"!

References

1. Public opinion in China and crisis management report. Social Science Literature Press, Beijing (2011)
2. Hu, J.: Speech at a meeting commemorating the 90th anniversary of the founding of the communist party of China. People's Publishing House Press, Beijing (2011)
3. Feng, X.: Contemporary College Students' Ideal Faith Education at the New Temporal and Spatial Situation. Education and Research, Beijing (2011)

A Competence-Driven Staff Assignment Model Based on Stochastic Working Status*

Li Du[a], Yu Zhang[b], and Hua Li[c]

School of Economics and Management, Xidian University
Xi'an, China
[a] duli0919@sina.com, [b] zhangyu_032@163.com,
[c] lihua@xidian.edu.cn

Abstract. Considering that the competences of employees are improved differently when they are in the various working statuses, we developed a new competence-driven staff assignment model by Markov decision process. The model seeks to minimize employee wage and maximize gains of the company from the improvement of desirable competencies. Finally, we proposed the genetic algorithm to solve the model.

Keywords: Staff assignment, knowledge workers, working status, genetic algorithm.

1 Introduction

With the coming of knowledge economy, their competitive and operation mode have many changes in enterprises. Employees' competencies and their development form a key source for competitive advantage in enterprises. There has been considerable literatures are interested in the staff assignment. Billionnet [1] considered a hierarchical workforce problem in which a higher qualified worker can substitute for a lower one. Based on Billionnet's model, Seckiner, Gokcen and Kurt [2] also consider the hierarchical workforce can be assigned to alternative shifts in a day during the course of a week. Hass et. al [3] point out that the companies can implement multiskilling to reduce their costs and increase their advantages. Brusco and Johns [4] present an integer linear programming model of a complex workforce staffing decision when cross-trained employees have different productivity levels in multiple work activity categories. Since then, the multi-skilled staff assignment problem has been concerned by some researchers [5-6].

In fact, due to the dual nature of employees, especially knowledge workers, they possess the ability to learn from performing the task and further develop their competencies. Considering the learning effect, Wu and Sun [7] developed a linear programming model for staff assignment to minimize outsourcing costs and proposed genetic algorithms to solve it. Afterwards, [8-10] have paid attention to competence development of employees from different aspects.

* Supported in part by The Natural Science Foundation of China (70701028), and in part by research grants from Xidian University(K50510060004).

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 463–470.
springerlink.com

Employees are subject to personal preference, external environment and other factors. Their competencies are developed differently when they are in various working statuses. Based on previous research, we present a new model for staff assignment with stochastic working status. The model seeks to minimize employee wage and maximize strategic gains of the company from the increment of desirable competencies. Finally, we propose the genetic algorithm to solve the model.

2 Model

Let us consider a project consisting of *T* tasks, indexed by $t = 1,2,\cdots,T$. Suppose that *T* tasks are in turn, and each task is viewed as a period. The decision on staff assignment has to be made at the start time of period 1, and is conceived as the solution of a dynamic programming problem.

We consider a set of employees, indexed by $k = 1,2,\cdots,K$, which form the staff and is assumed to be fixed during the entire planning horizon. Employees are assumed to possess different knowledge, education, skills and abilities. We refer to the employees' performance at work by the term *competencies* and index competencies by $l = 1,2,\cdots,L$.

Suppose that the degree to which an employee *k* possesses a certain competency *l* can be quantified in the form of a real value. We call this value the *competence score* and denote it by i_{tkl}. Initial values i_{1kl} of the competency scores in period 1 are assumed as known.

As the project progresses, the employees work on a task and their competencies can be improved correspondingly. That is, by learning effects, the competency values increase if employee *k* works on the task; otherwise, the competency values remain the same. Based on their characteristics, knowledge workers is subject to external and internal influences. It is assumed that there are *M* kinds of possible working status, indexed by $m = 1,2,...,M$. His competency value in each competence increases ε_m when he is in working status *m*. Let p_{tkm} represent the probability of employee *k* in working status *m* within period *t*. We represent the decision whether employee *k* is assigned to task *t* by decision variable x_{tk} $(t = 1,2,\cdots,T; k = 1,2,\cdots,K)$, where $x_{tk} = 1$ if employee *k* is assigned to task *t*, and $x_{tk} = M$ otherwise. The competence values remain unchanged with probability 1 if the employee does not work on any task. Thus, for any $t = 1,2,\cdots,T; k = 1,2,\cdots,K; m = 1,2,\cdots,M$, we have

$$p_{tkm}(x_{tk}) = \begin{cases} p_{tkm} & if \;\; x_{tk} = 1 \; and \; i_{(t+1)kl} = i_{tkl} + \varepsilon_m \quad l = 1,2,\cdots,L; \\ 1 & if \;\; x_{tk} = 1 \; and \; i_{(t+1)kl} = i_{tkl} \quad l = 1,2,\cdots,L; \\ 0 & others. \end{cases}$$

One task can be assigned to one or more employees. Because of transfer cost, the employee is not allowed to be substituted in the work once he has been assigned to a task. In addition, we do not consider the outsourcing of any tasks.

Employees with different competence values will generate different work efficiencies. Wu and Sun [7] points out that, work efficiency can be viewed as the share of work performed in one time unit by employee *k* on a task requiring only

competency *l*, if the entire task takes one time unit for an employee with "perfect skills" in competency *l*. [9] addresses that an *efficiency value* of employee *k* in competency *l* during period *t* can be described by using a nonlinear monotonous transformation function ϕ. The function ϕ maps the set of real values into the interval [0, 1]. The specification of an appropriate transformation function is an empirical problem. The class of logistic functions has been used frequently for model organizational learning [8]-[10]. Here, we continue to use this class. Then, the function ϕ is given by

$$\phi(x) = \frac{1}{1 + a\exp(-bx)}, \tag{1}$$

where *x* denotes the competency value, the real parameters $a>0$ and $b>0$. Parameters *a* and *b* are set based on the following rules. It is hard for an employee to increase his competence values if his competence values are very low or very high; otherwise, it is easy to increase his competence values if his competence values are moderate. The reason is that the employee with the low competence values is difficult to finish a task. Similarly, it is difficult for the employee with the high competence values to increase his competence value by working on a task and learning from it, since it has already almost reached its upper limit.

Various competencies are of the different value for the enterprises. From long-term strategic perspective, managers can assign a weight θ_l to each competency *l* that quantifies the relative importance of competency l in comparison to the other competencies, where $\sum_{l=1}^{L}\theta_l = 1$. We consider that the employees apply each competency when they work on the tasks. Then, let $\sum_{l=1}^{L}\theta_l i_{tkl}$ represent the total competence value of employee *k* at time *t*. Thus, equation (1) can be rewritten as:

$$\phi(i_{tk1}, i_{tk2}, ..., i_{tkL}) = \frac{1}{1 + a\exp(-b\sum_{l=1}^{L}\theta_l i_{tkl})}, \tag{2}$$

where $\phi(i_{tk1}, i_{tk2}, \cdots, i_{tkL})$ denotes the efficiency value of employee *k* within period *t*.

Various tasks have the different amount of work. We assume that task *t* requires an overall ideal effort of $d_t (t = 1,2,...,T)$, which is assumed to be both known and deterministic. Similarly with [9], we define the ideal effort d_t as the time needed by an employee with efficiency $\phi(i_{tk1}, i_{tk2}, \cdots, i_{tkL}) = 1$ for completing task *t*. It means that, it takes $d_t / \phi(i_{tk1}, i_{tk2}, \cdots, i_{tkL})$ time units to complete task *t* if employee *k* with efficiency $\phi(i_{tk1}, i_{tk2}, \cdots, i_{tkL})$ is assigned to task *t*. It takes $d_t \Big/ \sum_{k=1}^{K}\left(\frac{M - x_{tk}}{M-1}\right)\phi(i_{tk1}, i_{tk2}, \cdots, i_{tkL})$ time units to complete task *t* if more employees are assigned to task *t*.

Employee wage depends on whether he was assigned to the task, his competency and efficiency values. Let $h(i_{tk1}, i_{tk2}, \cdots, i_{tkL})$ represent the daily wage of employee *k* if he is assigned to task *t* and his competency values are $(i_{tk1}, i_{tk2}, \cdots, i_{tkL})$ within period *t*. Of course, he has no pay if he does not work within period *t*.

Given the notation above, we establish competence-driven staff assignment model by Markov decision process. The period is indexed by *t*, $t = 1,2,\cdots,T$; the state

variables are indexed by $((i_{t11}, i_{t12}, \cdots, i_{t1L}), (i_{t21}, i_{t22}, \cdots, i_{t2L}), \cdots, (i_{tK1}, i_{tK2}, \cdots, i_{tKL}))$, which represent the competency values of employees at time *t*; the decision variables are indexed by $(x_{t1}, \cdots, x_{tK})$, which represent whether employees are assigned to the task at time t; given the competency values of employees and the staff arrangement at time *t*, the probability that the competency values of employees at time (*t*+1) are $((i_{t11} + \varepsilon_{m_1}, i_{t12} + \varepsilon_{m_1}, \cdots, i_{t1L} + \varepsilon_{m_1}), (i_{t21} + \varepsilon_{m_2}, i_{t22} + \varepsilon_{m_2}, \cdots, i_{t2L} + \varepsilon_{m_2}), \cdots, (i_{tK1} + \varepsilon_{m_K}, i_{tK2} + \varepsilon_{m_K}, \cdots, i_{tKL} + \varepsilon_{m_K}))$, i.e., the competence scores of employee *k* in each competency *l* increase ε_{m_k}, is $p_{t1m_1}(x_{t1}) \cdot p_{t2m_2}(x_{t2}) \cdot \cdots \cdot p_{tKm_K}(x_{tK})$. Thus, the optimal equation can be written as:

$$
\begin{aligned}
& g_t\left(\left(i_{t11}, \cdots, i_{t1L}\right), \cdots, \left(i_{tK1}, \cdots, i_{tKL}\right)\right) \\
= \max_{(x_{t1}, \ldots, x_{tK})} & \left\{ -\left(\frac{d_t}{\sum_{k=1}^{K}\left(\frac{M - x_{tk}}{M-1}\right)\varphi\left(i_{tk1}, i_{tk2}, \cdots, i_{tkL}\right)} \right) \sum_{k=1}^{K}\left(\frac{M - x_{tk}}{M-1}\right) h\left(i_{tk1}, i_{tk2}, \cdots, i_{tkL}\right) \right. \\
& + \sum_{m_1 = x_{t1}}^{M} \cdots \sum_{m_K = x_{tK}}^{M} p_{t1m_1}\left(x_{t1}\right) \cdots p_{tKm_K}\left(x_{tK}\right) g_{t+1}\left(\left(i_{t11} + \varepsilon_{m_1}, \cdots, i_{t1L} + \varepsilon_{m_1}\right), \right. \\
& \left. \left. \cdots, \left(i_{tK1} + \varepsilon_{m_K}, \cdots, i_{tKL} + \varepsilon_{m_K}\right)\right)\right\} \\
& t = 1, 2, \ldots, T; \\
& g_{T+1}\left(\left(i_{(T+1)11}, \cdots, i_{(T+1)1L}\right), \cdots, \left(i_{(T+1)K1}, \cdots, i_{(T+1)KL}\right)\right) \\
& = \sum_{k=1}^{K}\sum_{l=1}^{L} \theta_l \left(i_{(T+1)kl} - i_{1kl}\right)
\end{aligned}
\tag{3}
$$

3 Numerical Analysis

On account of nonlinearity and an excessively large number of variables, we hardly obtain the exact solution of this problem by means of general algorithm. Hence, we will apply the genetic algorithm to solve the problem.

Let us consider a project consisting of 5 tasks. The workload of each task is shown in Table 1.

Table 1. The workload of each task

Task ID	1	2	3	4	5
Work load	4.5	8	6.5	10	5

There are 5 employees to be assigned to complete the project. Here, we use 5 competence indicators to describe the characteristics of staffs. In our numerical example, we assume that 10 points is the upper bound of the competence score, and the competence score falls into the interval [0, 10]. Initial competence score of employees in each competency is shown in Table 2.

The competence score of employees in each competency can be increased after they work on a task. However, the amount of the improvement on competencies is

Table 2. Initial competence score of employees in each competency

	Skill 1	Skill 2	Skill 3	Skill 4	Skill 5
Employee 1	5.5	6	7	4	5
Employee 2	4	3.5	3	2.5	4.5
Employee 3	6.5	7	5.5	6	6.5
Employee 4	3.5	3	4	3	4.5
Employee 5	5	5.5	6.5	4.5	5.5

different with their working status. Here, we assume that there are three working status, i.e. ordinary, optimistic and pessimistic status. For example, an employee with competence score "*A*" in some competency is assigned to a task. In an ordinary situation, the competence score in a competency will be increased to $(10+A)/2$ when he finishes the task. In addition, the competence score in each competency will be increased up to 10 if he falls in the optimistic working status. The competence score in each competency remains the same if he lies in the pessimistic working status. We assume that the probability that he is in the ordinary work status is 0.5, and the probabilities that he is in two other working status are both 0.25.

Table 3 shows the weights of competencies in our numerical example.

Table 3. The weights of competencies

skill	1	2	3	4	5
θ_l	0.15	0.3	0.2	0.1	0.25

The logistic function from Eq. (2) is chosen to transform the competence score to an efficiency value. According to the specification of parameters a and b, we set $a=148.4132$ and $b=1$. The change of efficiency values with competence scores is shown in Figure 1.

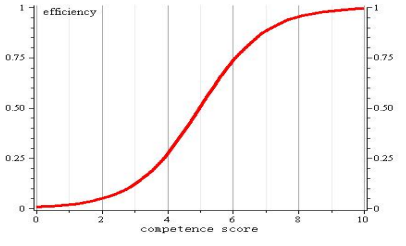


Fig. 1. The change of efficiency value with competence score

For simplicity, we define the following equation as the daily wage of employees.

$$h(i_{tk1}, i_{tk2}, i_{tk3}, i_{tk4}, i_{tk5}) = \sum_{l=1}^{5} \theta_l i_{tkl}$$

A genetic algorithm proceeds by progressively updating a population of candidate solutions, also called chromosomes. The population $P(j)$ is updated by creating new

chromosomes by genetic operators and selecting good quality ones to form the next-generation population $P(j+1)$. A fitness function, which represents the solution quality of a chromosome, should be defined. The updating of $P(j)$ continues until predefined terminating conditions are met.

The proposed genetic algorithm for solving the formulated mathematical model is presented below.

1) Coding

Here, we use integer coding method. According to the number of the decision variables, the number of the code is $T \times K$, where T represents the number of tasks and K represents the number of employees. A chromosome is represented by a row vector $X=[x_{11},\cdots,x_{15},\cdots,x_{51},\cdots,x_{55}]$, in which each element $x_{tk} \in \{0,1\}$ $(t=1,\cdots,5; k=1,\cdots,5)$ represents the task assignment of staff k in stage t.

2) Fitness function

We let $g_1(X)$ denote the value Eq. (3) if the staff assignment is determined by X. Since $g_1(X)$ may be negative, we define the value of the fitness function by $g_1(X)+C$ if X is a feasible solution, where C is a large constant. Otherwise, we define $g_1(X)+C$ by 0 if X is non-feasible solution.

3) Genetic operators

Let N_p represent the number of chromosomes in the population $P(j)$. Three genetic operators, selection, crossover and mutation, are designed to create new chromosomes.

Selection: We use the roulette wheel selection method to select chromosomes. Chromosomes are selected into the next-generation population with a certain probability P_{se}. The probability of each chromosome into the next generation depends on the proportion of the value of its fitness function to the sum of the values of the fitness function in this generation. The chromosome is more likely to be selected if the value of the fitness function is large.

Crossover: We use single-point crossover. Denote P_{cr} the crossover probability. From the population $P(j)$, $N_p \times P_{cr}$ chromosomes are randomly selected and then paired. For each paired chromosomes, a break-point is randomly chosen to interchange some parts of the two chromosomes.

Mutation: Let P_{mu} represent the mutation rate. From the population $P(j)$, $N_p \times P_{mu}$ chromosomes are randomly selected. For some gene x_{tk} in the selected chromosome, a random number ω in [0,1] is generated. If the random number ω is less than P_{mu}, then x_{tk} mutates; otherwise, x_{tk} remains unchanged.

4) Elitist Preservation Strategy

Because of the randomness of the selection, crossover and mutation, the best chromosome may be eliminated. Therefore, the elitist preservation strategy is used to maintain the best chromosome. Denote X^* the best chromosome in the current population $P(j)$, and X^{**} the best chromosome in the next population $P(j+1)$. If X^* is better than X^{**}, then X^* is substituted for the worst chromosome in the population

$P(j+1)$; otherwise, the chromosomes in the population $P(j+1)$ remain unchanged. Thus, the best chromosome in the population $P(j)$ can be kept.

Then, the pseudo code is as follows:

```
Initialization();
Chrom=crtbp(NIND,BaseV);
ObjV;
while gen<= MAXGEN
[y,r]=max(ObjV); J=Chrom(r,:);
Select();
recombin();
Mut();
ObjVSel;
[Chrom ObjV]=reins();
[Y,R]=max(ObjV);[X,s]=min(ObjV);
if y > =Y
    Chrom(s,:)=J(1,:);ObjV(s,1)=y;
end
gen=gen+1;
Output();
```

The GA parameters are set as follows: N_p =40, P_{se} =0.8, P_{cr} =0.6, P_{mu} =0.05, C =1000, MAXGEN=1000. The initial population $P(0)$ is randomly created. The GA program terminates when the pre-determined number of generations has elapsed.

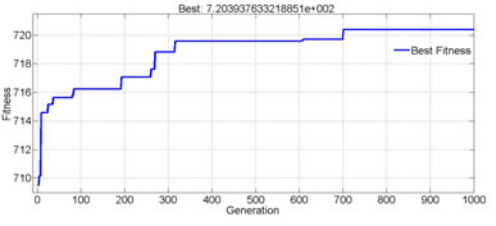


Fig. 2. The change of the fitness function

Table 4. Staff assignment and the corresponding value of the fitness function

	Staff assignment					The value of the fitness function
1	00001	10000	10000	00100	11111	7.203937633218851e+002
2	00001	10000	10001	00100	11111	7.202541535459627e+002
3	00001	10000	10001	00100	11111	7.202541535459627e+002
4	00001	10000	10001	00100	11111	7.202541535459627e+002
5	00001	10000	10001	00100	11111	7.202541535459627e+002
6	00001	10000	10000	00100	11111	7.203937633218851e+002
7	00001	00100	00100	10000	11111	7.191114488366002e+002
8	00001	00100	00001	10000	11111	7.190735786659841e+002
9	00001	00100	00001	10000	11111	7.190735786659841e+002
10	00001	10000	10001	00100	11111	7.202541535459627e+002

We run the program 10 times. The result of the simulation for the first time is shown in Figure 2. The values of the fitness function remain steady after 700 generations. The staff assignment and its corresponding value of the fitness function are shown in Table 4. In table 4, we can see that 10 staff assignments on task 1 and 5 are the same, in which employee 4 performs on task 1, and five employees work together on task 5. It is a little different in 10 staff assignments on task 2, 3 and 4, which lead to the difference of the value of the fitness function. Among these staff assignments, the 1st and the 6th staff assignments are both the best and their values of the fitness function are both the highest.

4 Conclusion

Considering that the competence of employees can be improved differently when they are in the various working statuses, we have developed a competence-driven staff assignment model. The model seeks to minimize employee wage and maximize gains of the company from the improvement of desirable competencies. Finally, we provide the genetic algorithm to solve the model.

Here, we consider the case of the task in turn. In our future research, we intend to consider more complex situations, such as parallel tasks in a project, the selection of projects and so on.

References

1. Billionnet, A.: Integer programming to schedule a hierarchical workforce with variable demands. European Journal of Operational Research 114, 105–114 (1999)
2. Seckiner, S.U., Gokcen, H., Kurt, M.: An integer programming model for hierarchical workforce scheduling problem. European Journal of Operational Research 183, 694–699 (2007)
3. Rodriguez, A.M.: Planning and Scheduling a Multiskilled Workforce. M.S. Thesis, University of Texas at Austin, Austin, TX (1998)
4. Brusco, M.J., Johns, T.R.: Staffing a multiskilled workforce with varying levels of productivity: an analysis of cross-training policies. Decision Sciences 29, 499–515 (1998)
5. Certa, A., Enea, M., Galante, G., Fata, C.M.: Multi-objective Human Resources Allocation in R&D Projects Planning. International Journal of Production Research (2007)
6. Heimerl, C., Kolisch, R.: Scheduling and staffing multiple projects with a multi-skilled workforce 32, 343–368 (2010)
7. Wu, M.C., Sun, S.H.: A project scheduling and staff assignment model considering learning effect. The International Journal of Advanced Manufacturing Technology 28, 1190–1195 (2006)
8. Ngwenyama, O., Guergachi, A., McLaren, T.: Using the learning curve to maximize IT productivity: a decision analysis model for timing software upgrades. Int. J. Production Economics 105, 524–535 (2007)
9. Gutjahr, W.J., Katzensteiner, S., Reiter, P., Stummer, C., Denk, M.: Competence-driven project portfolio selection, scheduling and staff assignment. Central European Journal of Operations Research 16, 281–306 (2008)
10. Gutjahr, W.J.: Optimal dynamic portfolio selection for projects under a competence development model. OR Spectrum 33, 173–206 (2011)

Internal Marketing Establishes the Culture of Market Orientation

Yafang Tsai[1], Shih-Wang Wu[2,*], and Sue-Ting Chang[3]

[1] Department of Health Policy and Management, Chung Shan Medical University, Taiwan.
Department of Medical Management, Chung Shan Medical University Hospital, Taiwan.
avon611@gmail.com
[2] Department of Hospital and Health Care Administration,
Chia Nan University of Pharmacy & Science, Taiwan, Department of Business Administration,
National Chung Cheng University, Taiwan
scottwu101@gmail.com
[3] Department of Technology Application and Human Resource Development, National Taiwan Normal University, Taipei, Taiwan, Department of Administration, Kuang-Tien General Hospital, Taichung, Taiwan

Abstract. By analyzing the antecedent factors that affect market orientation, this study links concepts from these topics with those of internal marketing to provide business administrators with an effective method of building a market orientation culture. The research subjects were presidents and high-level managers at 491 accredited Taiwanese district hospitals or above. For this cross-sectional study, researchers mailed 491 questionnaires in October 2008 and received 104 responses. The data were analyzed with SPSS 17.0, including descriptive statistics and inferential statistical analysis. To examine research hypothesizes, we performed a multiple linear regression to investigate the associations between market orientation and the three sub-constructs of internal marketing. This study finds that internal marketing significantly influences market orientation.

Keywords: Internal marketing, market orientation, service marketing, competitive advantage, health care marketing.

1 Introduction

From a marketing perspective, the most important strategy issues involve understanding and satisfying customer demand [1]. The market orientation strategy claims that by fulfilling customers' needs, organizations create value for customers and achieve customer satisfaction [2], which would ensure their sustainability in a competitive market [3] [4]. Organizations can generate long-term success by satisfying customer needs [5].

The objective of good service is to create value by attending to customer needs [6]. The basis of this value creation lies in satisfying customer needs by providing suitable service quality [7]. Unlike physical products, service products involve direct with

* Corresponding author.

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 471–479.
springerlink.com

customers. Service organizations interact with customers through their employees, who deliver service value for customers. When customers pay for service products, they evaluate service quality according to their predetermined expectations with respect to how service employees should behave. Therefore, the willingness of service employees to deliver quality service plays an important role in an organization's attempts to satisfy customer expectations [8].

The methods of efficiently improving service quality are pertinent to internal marketing because in offering a service, employees not only represent the organization but also provide a crucial so-called "service encounter" [9]. The concept of internal marketing emphasizes that organizational managers should provide a specific organizational vision to all employees through internal education and training to strengthen service provision ability and efficiency among employees, to improve service quality and ultimately to achieve customer satisfaction. For this reason, internal marketing is especially crucial for service organizations [10].

Some internal marketing studies have explored the influence of internal marketing on organizational commitment [11] [12], and still other studies explore how internal marketing influences employee service quality [13] [14] [15]. These studies all focus on the individual level; they explore how internal marketing influences employees' attitudes and behaviors. In contrast, an internal marketing approach considers the entire organization as the target of a competitive strategy [16]. Specifically, internal marketing is concerned with how organizational management should develop educational training, explicitly communicate organizational visions and create reward systems that improve employees' service ability and establish an organizational culture of customer orientation [12].

Most previous studies adopting the internal marketing approach have focused on for-profit sectors, such as the service sector [17], airlines [15] and banks [18] [11]. Notably, there is little research focused on non-profit organizations [12]. This study focuses on the organizational level, exploring how administrators of non-profit organizations apply internal marketing to foster market orientation.

2 Background

2.1 Internal Marketing

The frontline employee plays a vital role in the process of service delivery [19]. Internal marketing should focus on properly educating employees both internally and externally regarding the organizational mission, the strategic mission and customer needs. The organization's long-term vision is communicated to the frontline staff, enabling the frontline staff to identify themselves with the vision while influencing employees' service behavior to improve the service quality provided to customers [10]. Internal marketing is a tool that organizations can use to successfully hire, maintain and motivate employees to provide better customer service [1].

According to the model of a service profit chain [20], when employees' job satisfaction is high, they will provide good service to external customers [21], generating external customer satisfaction. The implementation of internal marketing within organizations not only gives employees a clear working target [17] but also can

be used in resolving employees' work needs. Organizations should thus become more focused on satisfying these needs and contribute significantly to organizational effectiveness [18]. When employees' work needs have been satisfied, their job satisfaction will improve [22]. Internal marketing, that consists of service training, performance-based rewards and a clear organizational vision is a useful management tool to improve employees' organizational commitment and service quality [12].

2.2 Market Orientation

Market orientation is the marketing strategy best suited to the 21st century competitive environment [23]. The concept of market orientation was shaped by Kohli and Jaworski [24] and Narver and Slater [5]. They argued that traditional marketing's heavy emphasis on the execution of marketing activities and planning is too constrained. They recommended that organizations be more active in gathering external market intelligence and disseminating that intelligence to members of relevant departments to induce quicker responses to this intelligence and thus to satisfy customers' product needs [24].

Organization performance can be improved by switching to an organizational culture of market orientation [3], which consists of customer orientation, competitor orientation and inter-functional coordination. In addition to fulfilling the demands of external customers, an organization must also keep abreast of any changes in competitors' marketing strategies, passing information to members of relevant departments. Further, through inter-functional communication, a service can be researched and developed that conforms to customers' needs, permitting a rapid organizational response to customers' needs. Therefore, the concept of market orientation provides organizations with a management philosophy in which the focus of an organization's key coordination activities is on how to satisfy customers' needs with the aim of creating customer value [25].

2.3 The Relationship between Internal Marketing and Market Orientation

A market-oriented organization will pursue its management philosophy of providing guidance to its members regarding behavior, attitude and organizational operations [26] to increase its internal operational efficiency and achieve better results. Such an organization's internal marketing approach is to use education and training to communicate its vision of customer orientation to its employees. This approach will improve employees' attitudes toward serving external customers, ensuring more efficient service delivery [27] and achieving the objectives of market orientation.

The concept of internal marketing is applied to inspire employees to develop new skills and apply these skills within their organization [28], achieving external customer satisfaction through improved service efficiency. To meet the requirements of customer orientation, organizations that value this concept will communicate their vision and mission to their internal customers (i.e., employees) [29]. The organization's commitment to service quality is then passed on to external customers by employees through the provision of value-added services. Therefore, internal marketing should show a positive correlation with market orientation.

Hypothesis 1. Internal marketing has a positive correlation with market orientation.

2.4 The Relationship between Internal Marketing Service Training and Market Orientation

Service organizations emphasize professionalism and continuous technical development in the current information age. Service workers must accumulate practical knowledge through day-to-day problem solving to improve their professionalism [30]. For service organizations, new services may change the nature of customer contacts, which may necessitate employee retraining [31]. Therefore, a service organization's administrators consider the development of employees' knowledge and skills to be an investment rather than a cost [32].

A well-run internal marketing program will encourage employees to develop service skills and knowledge. To enhance professional skills among service employees, a service organization's administrators will often hold training programs, providing an opportunity for their workers to extend their knowledge base and ensuring that they maintain a high level of professional knowledge and skill in providing patients with optimum service.

Hypothesis 2. Internal marketing service training has a positive correlation with market orientation.

2.5 The Relationship between Rewarding Performance in Internal Marketing and in Market Orientation

A growing body of evidence suggests that excellence in service quality is an important driver of customer satisfaction, leading to additional benefits to organizations, such as superior customer satisfaction and customer loyalty. Customer orientation is one of the elements of market orientation. A stronger customer orientation will lead to greater customer satisfaction growth [33]. Services are provided to customers by frontline employees through their interaction with customers. The frontline employee is generally the first person in an organization to notice customers' response to service. One approach to creating customer satisfaction is to enhance employees' service quality through employee recognition and reward programs [34]. Such programs have also been found to improve employee attitudes [35] and to facilitate organizational change [36]. To immediately address problems in service, frontline employees are encouraged to provide suggestions regarding procedural errors or service inefficiencies through an organizational reward system.

Hypothesis 3. Rewarding internal marketing performance is positively correlated with market orientation.

2.6 The Relationship between an Organization Having a Clear Vision of Internal Marketing and Market Orientation

Services are provided through interactions between frontline employees and customers in which value-added services are provided to customers [37]. The ultimate goal of service

organizations is to aim to provide excellent customer service through creating value-added services [38]. To create value-added customer services, service providers seek to customers' needs when they receive feedback upon providing their service. Communicating a vision to employees is one of the methods that organizations have adopted to encourage employees to work towards endpoints that are beneficial to the organization [39]. This allows employees to better understand the role that they are playing and increases the likelihood that they will attain the desired objective. Administrators of a service organization need to provide a clear service vision to employees to help these employees provide excellent customer service.

Hypothesis 4. A clear vision of internal marketing is positively correlated with market orientation.

3 Materials and Methods

Because internal marketing and market orientation are components of overall business strategy [40], which is normally formulated by the president and other senior administrators (including the vice president and managing executives), these senior administrators were the main participants in this research. This study adopted a cross-sectional study using questionnaires to gather research data. Researchers mailed 491 copies of questionnaires in October 2008, and 104 responses were returned, leading to an overall collection rate of 21.18%.

The questionnaire contained a range of closed statements. Respondents were asked to rate their level of agreement on a 5-point Likert Scale (1 = 'strongly disagree' and 5 = 'strongly agree'). Internal Marketing: this item uses the 14-item scale designed by Tsai and Tang [10] and includes three dimensions: service training, rewarding performance and clear organizational vision. Market Orientation: this item uses the 14 questions provided by Slater and Narver [3] over three dimensions: customer orientation, competitor orientation and inter-functional coordination.

All factor loadings were highly significant as a basic requirement for convergent validity. Except for item 7 of the market orientation segment, all standardized factor loadings exceeded the threshold of 0.50, indicating adequate convergent validity of these three scales [41]. The composite reliabilities of all sub-constructs ranged from 0.73 to 0.94, exceeding the threshold of 0.70 [41]. The average variance extraction (AVE) ranged from 0.41 to 0.65, where a value greater than 0.50 is regarded as indicating adequate convergent validity. However, in this study, only one sub-construct (competitor orientation on the market orientation segment) did not reach this criterion. Therefore, almost all items and constructs met the criterion and provided evidence of construct reliability and validity. In addition, an internal consistency reliability estimate was adopted to assess the internal stability of scales. All of the Cronbach's alpha coefficients exceeded 0.70, which is regarded as a respectable standard [42].To examine hypothesis 2, hypothesis 3 and hypothesis 4, we performed a multiple linear regression to investigate the associations between market orientation and the three sub-constructs of internal marketing.

4 Results

The means, standard deviations and bivariate correlations for all observational variables are given in Table 1.

Table 1. Descriptive statistics and correlations for all study variables (N = 104)

No.	Variables	*M*	*SD*	1	2	3	4	5	6	7	8	9	10	11
1	**Internal Marketing**	4.41	0.73	—										
2	Service training	4.46	0.74	0.97	—									
3	Rewarding performance	4.36	0.81	0.91	0.86	—								
4	Clearly organizational vision	4.39	0.77	0.96	0.89	0.80	—							
5	**Market Orientation**	4.22	0.72	0.63	0.63	0.56	0.60	—						
6	Customer orientation	4.68	0.78	0.55	0.55	0.50	0.52	0.90	—					
7	Competitor orientation	3.84	0.79	0.60	0.60	0.53	0.56	0.87	0.68	—				
8	Inter-functional coordination	3.90	0.87	0.53	0.52	0.45	0.52	0.86	0.63	0.69	—			
9	**Organizational Innovation**	3.99	0.76	0.82	0.79	0.76	0.79	0.70	0.60	0.65	0.61	—		
10	Management innovation	3.89	0.83	0.74	0.70	0.69	0.70	0.60	0.51	0.55	0.53	0.91	—	
11	Technical innovation	4.05	0.78	0.80	0.77	0.74	0.77	0.70	0.60	0.66	0.61	0.97	0.77	—

Note: All Pearson's correlation coefficient were significant at $p < 0.001$ level.

The overall association between internal marketing and market orientation was positive and significant ($B = 0.62$, $\beta = 0.63$, $p < 0.001$). The variance explained by overall internal marketing reached 40%, indicating that the effect was substantial. Therefore, hypothesis 1 was supported. To examine hypothesis 2, hypothesis 3 and hypothesis 4, we performed a multiple linear regression to investigate the associations between market orientation and the three sub-constructs of internal marketing. As Table 2 shows, the association between service training and market orientation was significant ($B = 0.38$, $\beta = 0.40$, $p = 0.048$), whereas the coefficients of rewarding performance and a clear organizational vision were not statistically significant when controlling for each other. Therefore, hypothesis 2 was supported, whereas hypothesis 31 and hypothesis 4 were not supported.

5 Discussion

Service organizations depend heavily on the ability of employees to deliver services [43]. Frontline customer-contact employees play the role of boundary spanners at the interface between the firm and the customer and often experience a high level of ambiguity [44]. Service organizations practice internal marketing in communicating employees' roles and functions within the service process, and providing a clear objective and direction can reduce any ambiguity during the service-delivery process. This can also ensure that the service performed is based on a customer-centered perspective with the goal of achieving total customer value.

Effective communication can reduce conflicts between different functions within service organizations [45]. This is extremely helpful for market-oriented organizations that strongly emphasize cross-functional cooperation. Service organizations can take advantage of internal marketing in encouraging inter-functional communication, reducing potential conflicts and thus helping to establish a market-oriented culture. Morgan [46] suggested that internal marketing can also be formulated into a strategy of gradually establishing a market orientation culture. This study justifies the substantial effect of internal marketing on market orientation.

Service organization can increase customer satisfaction by holding service training and establishing customer-focused service behavior, thereby meeting service demands of external customers [33]. Ellinger et al. [47] found that the degree to which frontline service employees receive formal service-related training has a positive influence on market orientation. This study demonstrated that organizations adopt internal marketing as a training exercise for employees, which influences employees' sensitivity to market orientation.

6 Conclusions

Service is a form of behavior that requires interaction between employees and customers. The value of the service lies in the employees' organization's commitment to its service. Besides emphasizing satisfying external customer needs, it is also important to value employees' work-related needs by satisfying these needs and gaining employees' recognition and sense of belonging, creating a happier workforce that will commit to the organization. Organizations that implement internal marketing affect employees' attitudes towards customers.

References

1. Kotler, P.: Marketing Management: Analysis, Planning, Implementation and Control, 7th edn. Prentice-Hall, Englewood Cliffs (1991)
2. Pelham, A.M., Wilson, D.T.: A Longitudinal Study of the Impact of Marketing Structure, Firm Structure, Strategy, and Marketing Orientation Culture on Dimensions of Small-firm Performance. J. the Acad. Mktg. Sci. 24(1), 27–43 (1996)
3. Slater, S.F., Narver, J.C.: Does Competitive Environment Moderate the Market Orientation-Performance Relationship? J. Mktg. 58(1), 46–55 (1994)
4. Zhou, K.Z., Brown, J.R., Chekitan, S.D.: Market Orientation, Competitive Advantage, and Performance: A Demand-based Perspective. J. Business Res. 62(11), 1063–1070 (2009)
5. Narver, J.C., Slater, S.F.: The Effect of a Market Orientation on Business Profitability. J. Mktg. 62, 20–35 (1990)
6. Porter, M.E., Teisberg, E.O.: Redefining Health Care-creating Value-based Competition on Results. Harvard Business School Press (2006)
7. Quinn, G., Ellery, J., Detman, L., Jeffers, D., Gorski, P., Singer, T., Mahan, C.: Creating Patient-centred Healthcare Practices: Social Marketing Tools and Strategies. J. Mgt. & Mktg. in Healthcare 2(3), 221–237 (2009)

8. Humborstad, S.I.W., Humborstad, B., Whitfield, R.: Burnout and Service Employees' Willingness to Deliver Quality Service. J. Human Resources in Hospitality & Tourism 7(1), 45–64 (2008)
9. Tansuhaj, P., Randall, D., McCullough, J.: Applying the Internal Marketing Concept within Large Organizations: as applied to a Credit Union. J. Prof. Serv. Mktg. 6(2), 193–202 (1991)
10. Tsai, Y., Tang, T.W.: How to Improve Service Quality: Internal Marketing as a Determining Factor. Total Quality Mgt. 19(11), 1117–1126 (2008)
11. Caruana, A., Calleya, P.: The Effect of Internal Marketing on Organizational Commitment Among Retail Bank Managers. Intl. J. Bank Mktg. 16(3), 108–116 (1998)
12. Tsai, Y., Wu, S.W.: Using Internal Marketing to Improve Organizational Commitment and Service Quality. J. Adv. Nurs. (2011), doi: 10.1111/j.1365-2648.2011.05696.x
13. Joseph, W.B.: Internal Marketing Builds Service Quality. Mktg. Rev. 16(1), 54–58 (1996)
14. Lings, I.N.: Managing Service Quality with Internal Marketing Schematics. Long Range Planning 33(4), 452–463 (1999)
15. Frost, F.A., Kumar, M.: Service Quality between Internal Customers and Internal Suppliers in an International Airline. Intl. J. Quality & Reliability Mgt. 18(4), 371–386 (2001)
16. Cooper, J., Cronin, J.J.: Internal Marketing: a Competitive Strategy for the Long-term Care Industry. J. Business Res. 48, 177–181 (2000)
17. Boshoff, C., Tait, M.: Quality Perceptions in the Financial Services Sector: the Potential Impact of Internal Marketing. Intl. J. Serv. Ind. Mgt. 7(5), 5–31 (1996)
18. Sargeant, A., Asif, S.: The Strategic Application of Internal Marketing: an Investigation of UK Banking. Intl. J. Bank Mktg. 16(2), 66–79 (1998)
19. Congram, C.A., Friedman, M.L.: The Quality Leadership Connection in Service Businesses. In: Congram, C.A. (ed.) The AMA Handbook of Marketing for the Service Industries, American Management Association, New York, pp. 3–19 (1991)
20. Heskett, J.L., Jones, T.O., Loveman, G.W., Sasser, E.W., Schlesinger, L.A.: Putting the Service-profit Chain to Work. Harvard Business Rev. 72(2), 164–174 (1994)
21. Varey, R.J.: Internal Marketing: A Review and Some Interdisciplinary Research Challenges. Intl. J. Serv. Ind. Mgt. 6(1), 40–53 (1995)
22. Bernstein, S.L.: Internal Marketing of an ED-Based Public Health Initiative. The American J. Emergency Medicine 23(3), 411–413 (2005)
23. Lara, P.R., de Mesa, E.G.: Marketing Strategy Market Orientation: is it Accepted as a 21st Century Marketing Strategy in the Pharmaceutical Industry? An Example: Spain. J. Medical Mktg. 6(4), 260–267 (2006)
24. Kohli, A.K., Jaworski, B.J.: Market Orientation: the Construct, Research Propositions, and Managerial Implications. J. Mktg. 54, 1–18 (1990)
25. Tay, J., Tay, L.: Market Orientation and the Property Development Business in Singapore. Intl. J. Strategic Property Mgt. 11(1), 1–16 (2007)
26. Loubser, S.S.: The Relationship between a Market Orientation and Financial Performance in South African Organizations. South African J. Business Mgt. 31(2), 84–90 (2000)
27. Brady, N.: In Search of Market Orientation: An Experiment in Key Account Management. Mktg. Intelligence and Planning 22(2), 144–159 (2004)
28. Grönroos, C.: Internal Marketing-theory and Practice. In: American Mktg. Association's Serv. Conf. Proc., pp. 41–47 (1985)
29. Hopkinson, G.C.: Stories from the Front-line: How They Construct the Organization. J. Mgt. Studies 40(8), 1943–1969 (2003)

30. Mylopoulos, M., Scardamalia, M.: Doctors' Perspectives on Their Innovations in daily Practice: Implications for Knowledge Building in Health Care. Medical Educ. 42(10), 975–981 (2008)
31. Goffin, K., Mitchell, R.: Innovation Management: Strategy and Implementation Using the Pentathlon Framework. Macmillan, Palgrave (2005)
32. Foreman, S.K., Money, A.H.: Internal Marketing: Concepts, Measurement and Application. J. Mktg. Mgt. 11(8), 755–768 (1995)
33. Gainer, B., Padanyi, P.: The Relationship between Market-oriented Activities and Market-oriented Culture: Implications for the Development of Market Oorientation in Nonprofit Service Organizations. J. Business Res. 58(6), 854–862 (2005)
34. Saunderson, R.: Survey Findings of the Effectiveness of Employee Recognition in the Public Sector. Public Personnel Mgt. 33(3), 255–275 (2004)
35. Cassidy, E., Ackah, C.: A Role for Reward in Organizational Change? Irisb Business and Administrative Res. 18, 53–62 (1997)
36. Kopelman, R.E., Gardberg, N.A., Brandwein, A.C.: Using a Recognition and Reward Initiative to Improve Service Quality: a Quasi-Experimental Field Study in a Public Higher Education Institution. Pub. Personnel Mgt. 40(2), 133–149 (2011)
37. Elmadag, A.B., Ellinger, A.E., Franke, G.R.: Antecedents and Consequences of Frontline Service Employee Commitment to Service Quality. J. Mktg. Theory and Practice 16(2), 95–110 (2008)
38. Paulus, R.A., Davis, K., Steele, G.D.: Continuous Innovation in Health Care: Implications of the Geisinger Experience. Health Affairs 27(5), 1235–1245 (2008)
39. Grazier, K.L., Metzler, B.: Health Care Entrepreneurship: Financing Innovation. J. Health and Human Serv. Administration 28(4), 485–503 (2006)
40. Laforet, S.: Size, Strategic, and Market Orientation Affects on Innovation. J. Business Res. 61(7), 753–764 (2008)
41. Hair, J.F., Black, W.C., Babin, B.J., Anderson, R.E., Tatham, R.L.: Multivariate Data Analysis, 6th edn. Prentice-Hall, New Jersey (2006)
42. DeVellis, R.F.: Scale Development: Theory and Applications, 2nd edn. Sage, CA (2003)
43. Chebat, J.C., Kollias, P.: The impact of empowerment on Customer Contact Employees' Roles in Service Organizations. J. Serv. Res. 3(1), 66–81 (2000)
44. Powpaka, S.: How Market Orientation Affects Female Service Employees in Thailand. J. Business Res. 59(1), 54–61 (2006)
45. Chimhanzi, J., Morgan, R.E.: Explanations from the Marketing/Human Resources Dyad for Marketing Strategy Implementation Effectiveness in Service Firms. J. Business Res. 58(6), 787–796 (2005)
46. Morgan, N.A.: Professional Services Marketing. Butterworth Heinemann CIM, London (1991)
47. Ellinger, A.E., Ketchen, J.D.J., Hult, T.M., Elmadag, A.B., Richey, J.R.G.: Market Orientation, Employee Development Practices, and Performance in Logistics Service Provider Firms. Indus. Mktg. Mgt. 37(4), 353–366 (2008)

CFD Modeling of the Mixing Process in a Fermentation Tank

Lou Peng[1,*], Chen Hui[1], and Tian Bao Long[2]

[1] Tianjin University of Technology, Tianjin, China
[2] Tian Jin University, Tian Jin, China
loupeng19@163.com

Abstract. The impact of different baffles on liquid homogenization in crystallization process was described. In fermentation with crooked baffle, it was found that the flow field can provide more powerful driving force in the axial direction. And combined with the radial driving force provided by the disc turbine, dead zone at the bottom of blades was also eliminated. Due to its certain degree of twist, the flow field in the fermentation with crooked baffle can not only provide axial driving force, but broke the double circulation. Therefore the flow field in crystallization reactor eventually was distributed very well with dramatic reduce of adhesion.

Keywords: Fluent, Fermentation, Process control.

1 Introduction

Mixing is one of the most important unit operations. It is widely used by K.N.G. and N.J.FENTIMAN (1998) in chemical industry, pharmaceutical industry, process industries and so on. In a stirred tank, mixing is a very complex process and is achieved through convection and turbulent exchanges. Several literatures achieved by Wei (2003) on the study of the effect of impeller-to-tank diameter ratio.This complex process can be distinguished and described into two simpler stages of mixing i.e., macro-mixing (mixing at the scale of whole tank) and micro-mixing (mixing at the molecular scale). Chemical reaction is a molecular level process and only micro-mixing can affects its course directly. However, the environment for the micro-mixing is defined by macro-mixing and hence the information related to macro-mixing is also very important to control the performance of a chemical reaction. It is therefore essential to have the knowledge of macro-mixing before studying the micro-mixing. Wang and Liu (2009) has been done to study the subject on the Application of FLUENT Software on Designing Baffle-plate and Deflector.

Crystallization is an important unit operation in chemical industry, and the qualities of the final crystal products are strongly influenced by the mixing processes.Crystallization processes are frequently carried out in straight baffled vessels operating under turbulent flow conditions. Such processes are strongly influenced by the hydrodynamic and mixing

* Corresponding author.

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 481–489.
springerlink.com

characteristics. Over the last two decades, many CFD modeling studies of turbulent flow in agitated vessels with straight baffles have been reported in the literature. However, the baffles mentioned in all of these studies are straight, which have several defects, such as dead zones in some part of the crystallizer and solid particles can not be distributed very well,which is studied by Yang and Gao (2009). Due to these defects of the straight baffles, a newly baffles—Fermentation—were invented. And in the present work, Fermentation with certain degree of twist installed in the crystallizer were investigated. Due to it can provide axial driving force, the dead zones in some part of the crystallizer can be eliminated.

A computational fluid dynamics (CFD) based model was used to understand the mixing process. The sliding mesh method was adopted to simulate the interaction between impeller and baffles. The details of the computational model as well as the results obtained are discussed in the following sections.

2 Modeling

2.1 Governing Equation

Continuity equation

$$\frac{\partial \rho}{\partial t} + div\ (\rho \vec{u}) = 0$$

Momentum equation

$$\frac{\partial (\rho u)}{\partial t} + div\ (\rho u \vec{u}) = div\ (\mu\ gradu\ ) - \frac{\partial p}{\partial x} + S_u$$

$$\frac{\partial (\rho v)}{\partial t} + div\ (\rho v \vec{u}) = div\ (\mu\ gradv\ ) - \frac{\partial p}{\partial y} + S_v$$

$$\frac{\partial (\rho w)}{\partial t} + div\ (\rho w \vec{u}) = div\ (\mu\ gradw\ ) - \frac{\partial p}{\partial z} + S_w$$

Energy equation

$$\frac{\partial (\rho T)}{\partial t} + div\ (\rho \vec{u} T) = div\left(\frac{k}{c_p}\ gradT\ \right) + S_T$$

2.2 Crystallizer Configuration

The crystallizer used in this study was a standard configuration cylindrical vessel of a diameter, T, with four equi-spaced wall mounted baffles of width B=0.1T and the liquid column height is H=T, as shown in Figure 1(straight baffles) and Figure 2(Fermentation). The impeller used in this work was disc turbine, with its diameter of D=T/3.The off-bottom clearance of the impeller was one third of the vessel diameter, C=T/3.Water was selected as the working material.

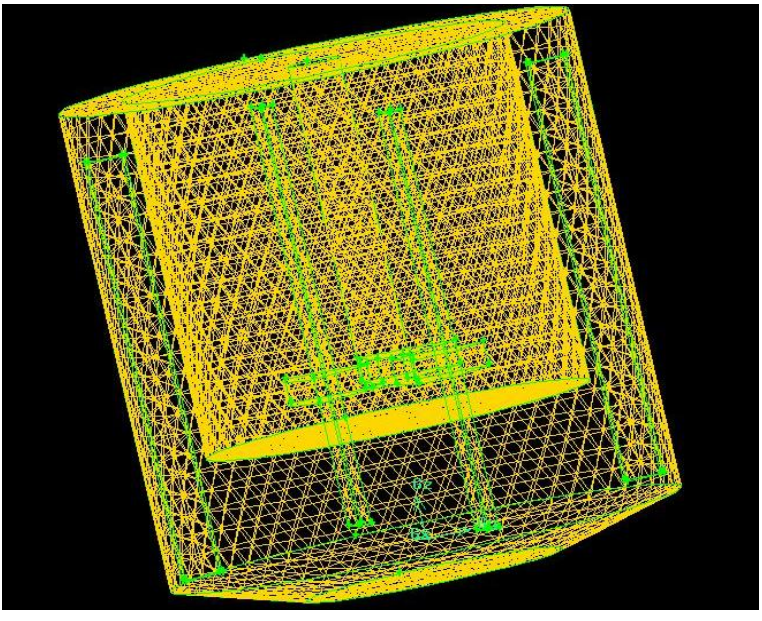


Fig. 1. Straight baffle crystallizer node distribution

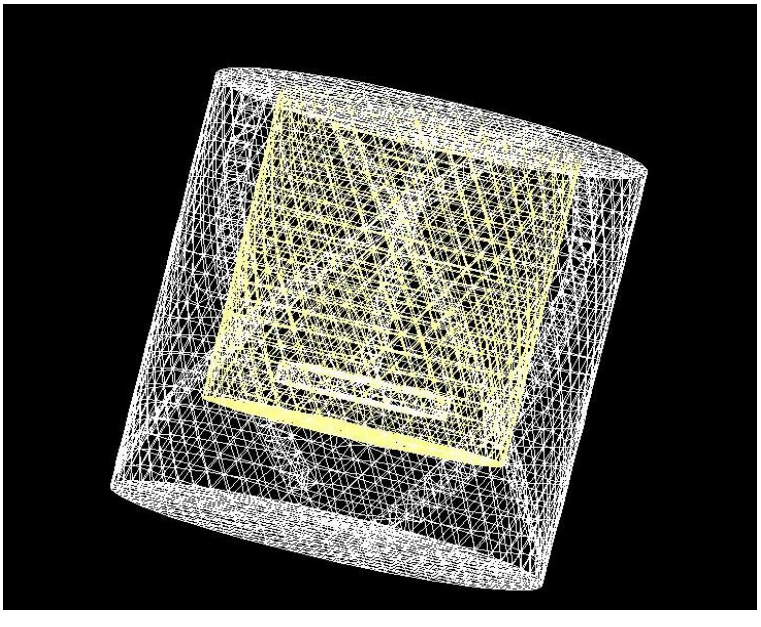


Fig. 2. Surface grid distribution baffle crystallizer

3 Results and Discussion

3.1 Characteristics of the Flow Field

Figure 3-5 show the distributions of the flow field at the same time. As can be seen in these figures above, we can draw our conclusion that: the flow field in the curved baffle vessel distributed better than the other two. The color represents the velocity magnitude in the curved vessel is blue, the low velocity zone is much smaller than the other two. In addition, the vessel without the baffles has dead zone at the bottom.

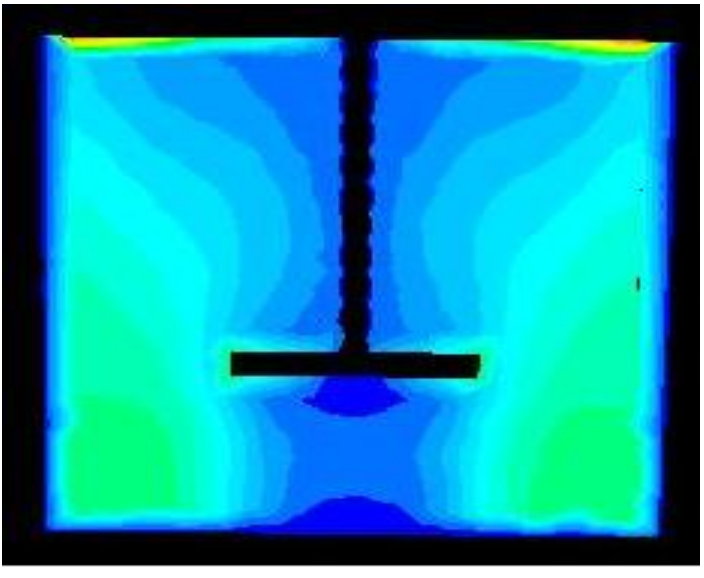


Fig. 3. Flow field distribution distribution netball

3.2 Iso-Surface Analysis

Figure 6 shows the velocity vectors in the straight baffle crystallizer.

In the industrial process, the adhesion effect has a significant impact on the product quality, chemical reaction heat transfer coefficient and productive capability. The existence of the adhesion effect is related to the fluid flow patterns in the crystallizer. Baffle installation method has a great impact on the distribution of adhesion effect. The baffle installation arrangement should prevent the dead zone in the crystallizer. And through the simulation, we can find that the curved baffle with a certain degree of twist can reduce the effect of adhesion.

The read box area is the position which the straight baffle installed, it blocked part of the fluid, resulting in reflux. In the industrial continuous production process, the straight baffle can make part of the fluid , form dead zones, particles are accumulated in the dead zones. The curved baffle, due to its certain degree of twist, fluid flow does not make a major turning point.

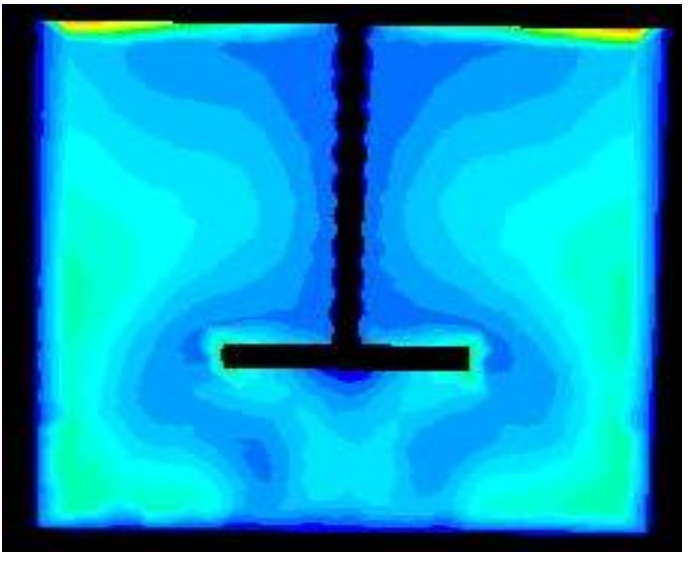


Fig. 4. Flow field distribution using straight baffle

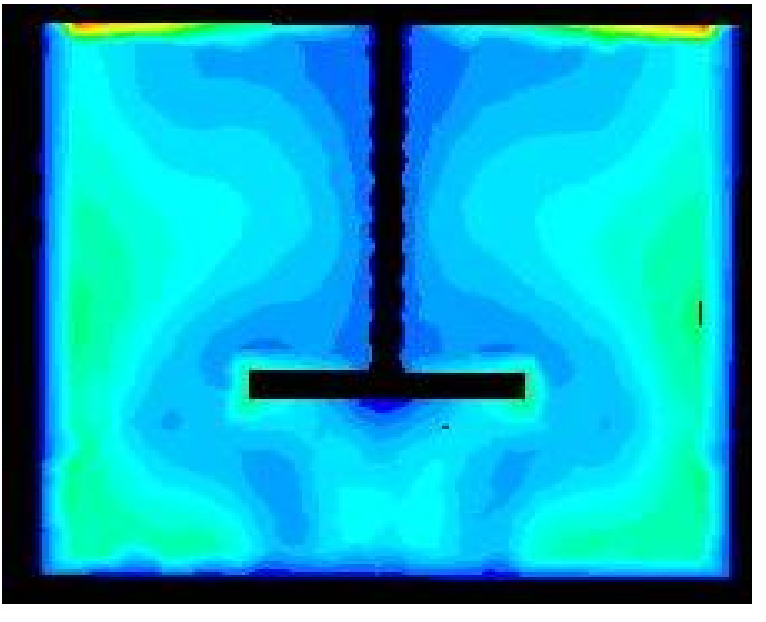


Fig. 5. Flow field distribution map using surface baffle

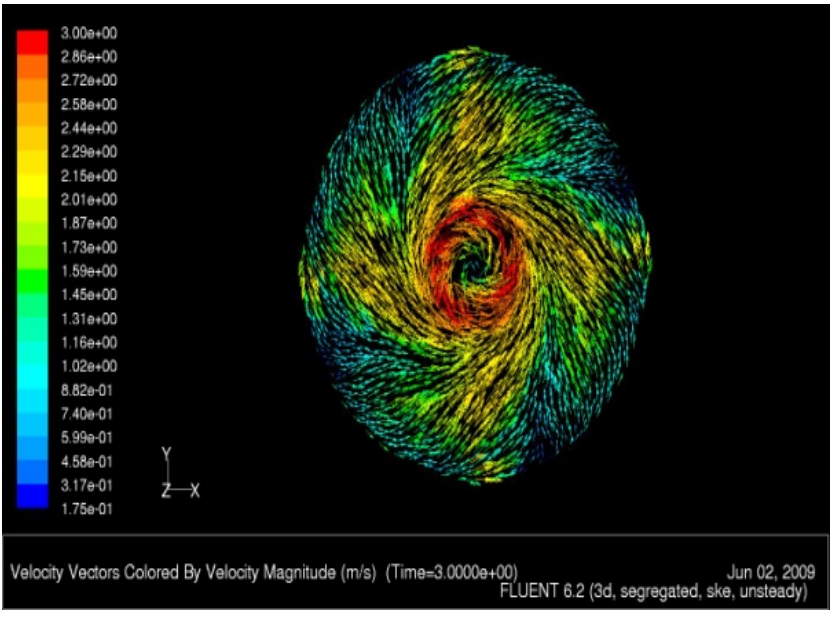


Fig. 6. Z-axis cross-section of velocity vectors using straight baffle（Z=200）

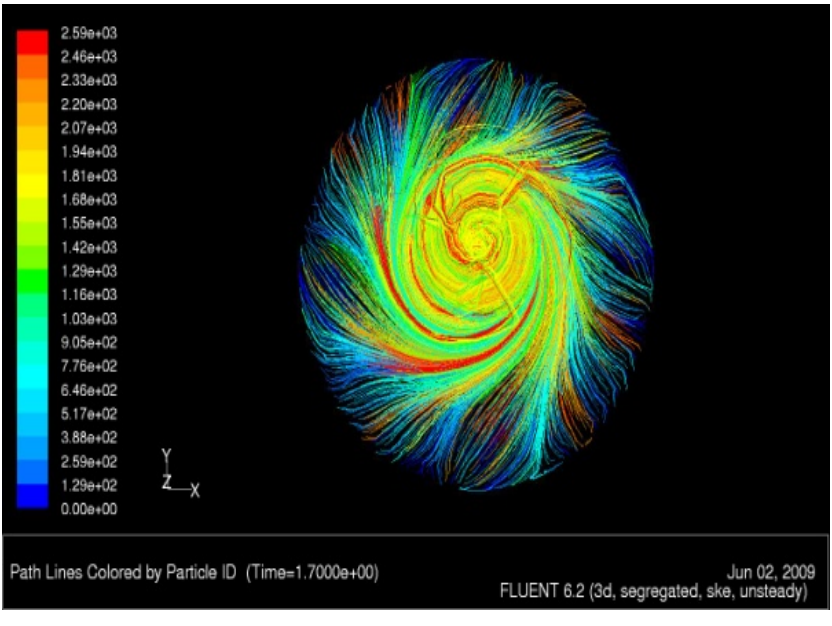


Fig. 7. Z-axis cross-section velocity trace map with curved baffle（Z=100）

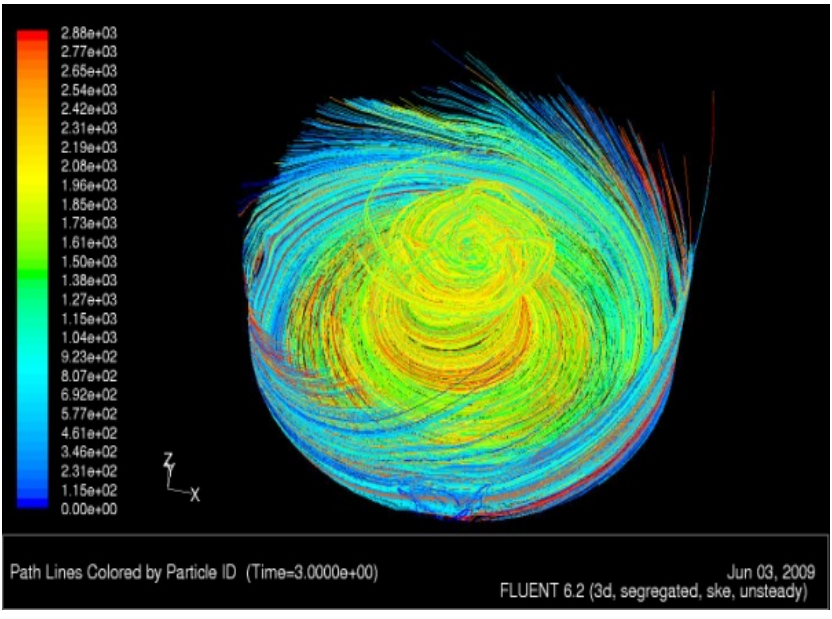


Fig. 8. Z-axis cross-section velocity trace map with curved baffle（Z=200）

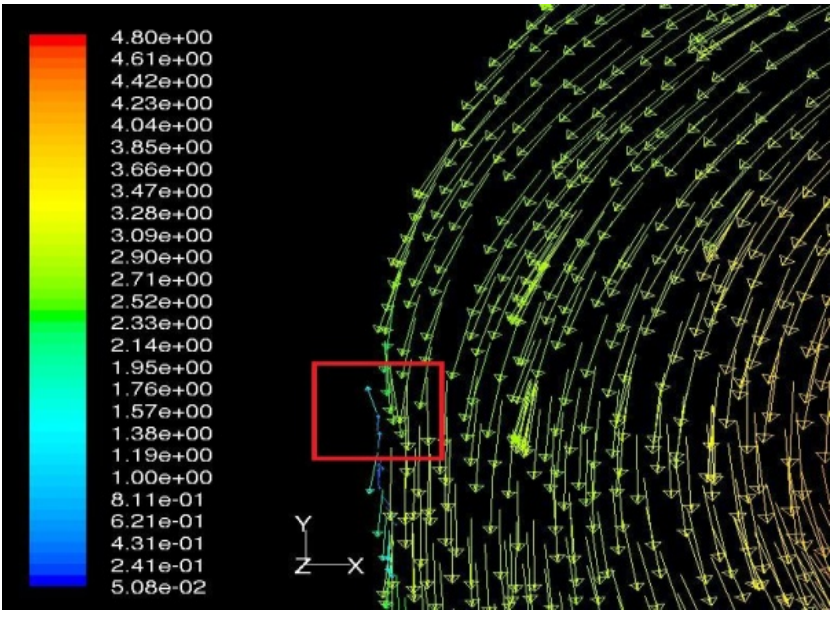


Fig. 9. Straight baffle Local Enlargement

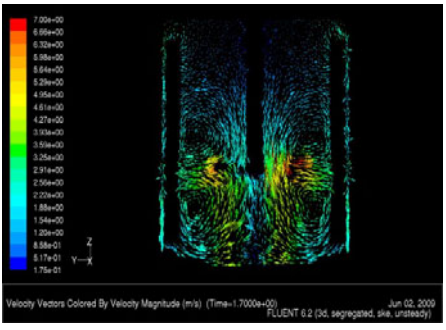


Fig. 10. X-axis cross-section velocity vectors using straight baffle

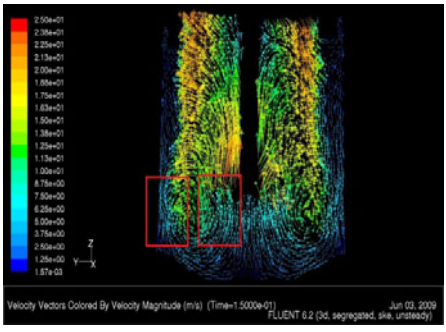


Fig. 11. X-axis cross-section velocity vector using curved baffle

From Figure 10, due to the presence of the straight baffle, its spoiler role in the region makes the fluid disorderly, and small vortices behind the baffle are formatted, witch prevent the fluid swirl phenomenon generation in the vessel. With the rotation, the fluid formed double circulation loops, which played a significant role on the fluid distribution.

We can see that it is necessary to install the baffle, but the bottom circulation loop area is small, because there is not enough driving force. Therefore, there is a dead zone between the two vortex at the bottom.

It is common knowledge that the disc turbine in the straight baffle vessel discharge flows radically, forming two circulation loops, because of the curved baffle can provide a more powerful driving force in the axial direction., two circulation loops break down., and this made the flow field very well. In the red box area, the strong axial driving force break the eddy at the bottom, the fluid was not only discharged by the impeller, but also influenced by the curved baffle, witch leads to an overal axial flow characteristic.

4 Conclusion

In this paper, CFD software was used to modal the flow field in the crystallizer. Compared with the flow field and velocity vectors of straight and Fermentation, it was found that the crystallizer with Fermentation can provide a more powerful driving force in the axial direction. And combined with the radial driving force provided by the disc turbine, it made the flow field in crystallization reactor eventually distributed very well. Therefore, the crystallizer with Fermentation will have a bright future.

Acknowledgements. This study was funded by the project of science and technology support program of Tianjin China (09ZCKFSH00800).

References

1. Fentiman, N.J., Lee, K.C., Yianneskis, M.: Assessment of sliding mesh CFD predictions and LDA measurements of the flow in a tank stirred by a rushton impeller. Chemical Engineering Research & Design 76, 737–747 (1998)

2. Wei, H.Y.: The effect of impeller-to-tank diameter ratio on draw down of solids. Chemical Engineering Science 58, 2011–2022 (2003)
3. Yang, L.B., Gao, Z.S.: Numerical simulation on separation performance of stationary classifier with different opening ratio of baffle plate. Journal of Heilongjiang Institute of Science and Technology 4, 002 (2009)
4. Wang, J.Q., Liu, G.F.: Application of FLUENT Software on Designing Baffle-plate and Deflector in Bag Filter. Petro-Chemical Equipment 5, 024 (2008)

A Ranging Algorithm Based on Binocular Real-Time Video Monitoring

Ying Li and Liquan Mei

School of Science, Xi'an Jiaotong University, 710049, Xi'an, China
leeying840223@163.com, lqmei@mail.xjtu.edu.cn

Abstract. Based on the principle of binocular parallax, this paper designs a simple, practical, fast and convenient passive real-time video monitoring ranging system. The system realizes the real-time video monitoring ranging algorithm of the system and analyzes the accuracy and the error of the ranging system. By doing experiments, the algorithm turns out to be of simple ranging method, good real-time and stability, high accuracy and simple equipments. Another advantage: the system doesn't send signals while working, so it is convenient to conceal yourself, thus having a high safety. Therefore, the system has great application values in the fields like air defense for strategic points, land and sea defense monitoring, real-time monitoring of the battlefield, etc.

Keywords: Stereo vision, binocular parallax, ranging system, digital image.

1 Introduction

Many fields such as air defense for strategic points, land and sea defense monitoring, real-time monitoring of the battlefield, etc., all involves the measurement of the objective parameters in the monitoring area, like the target's distance, azimuth, pitch angle, etc.. Among them, measuring the target's distance is very important.

Many scholars have done some research in this field. Often these systems have a complex structure but simple functions and poor accuracy, such as the two station intersection measurement. It is well know that the triangulation method has been widely used in many applications [1-7] owing to its significant advantages in simplicity and robustness. Zeng et al. have relized the two-dimensional (2D) position measurement by using a two-beam laser triangulation method [5]. But the configuration is also become more complicated due to splitting the laser beams. In addition, many triangulation range finders adopt laser diode and charge-coupled device (CCD) camera, which greatly enhance the cost of the sensor. Huaqiao Gui et al. have realized a simple, low cost real-time 2D positon measurement, they developed a range finder by using linear complementary metal oxidation semiconductor (CMOS) chips and light emitting diodes (LEDs) [8]. Tao Jin et al. have designed a practical method for evaluating the three dimensional (3D) position, but the method need 6 control points which are not in the same plane [9]. However, the actual combat systems and the monitoring systems often need to have simple and portable equipments, and at the same time realize the real-time parameter measurements of the targets in the monitoring area. Therefore, considering the passive

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 491–498.
springerlink.com

real-time monitor positioning and portability, this paper presents a real-time binocular stereo ranging technology based on the digital video, realizes the core location algorithm, and analyzes the precision of the whole system and its effect factors.

2 Principle of the Digital Stereo Image Ranging of the Binocular Parallax

Binocular parallax is also called stereo vision. Its principle is to imitate human's method of perceiving the distance by two eyes. Human's two eyes observe the objective three-dimensional world from two slightly different angles, so projecting them by geometrical optics, different points, having different distance from the observer, will have different positions on the left and right retinas. The positional difference on the two retinas is called binocular parallax, which reflects the object's distance. Therefore, the stereo image ranging sets up two or more detectors, simulates the images of human's visual effect to set up image system, imitates human's method of perceiving the distance by two eyes, and gets the stereo pairs by imaging the same object from different angles, thus measuring the object's distance by the corresponding algorithm.

2.1 Establishment of the Model and Determination of the Objective Parameters

The system's structure [8-12] can be simply shown as Fig. 1(Left), camera *A* and camera *B* are put on the table, their center distance is *b*, the two cameras' optical axes are parallel, and *AB* is perpendicular to the table's axis. All these sets are aim to simulate how human's eyes observe the objects, and also have a easier computation. While observing the object, *A* and *B* will horizontally or vertically rotate as a whole. The system parameters needed are listed as follows.

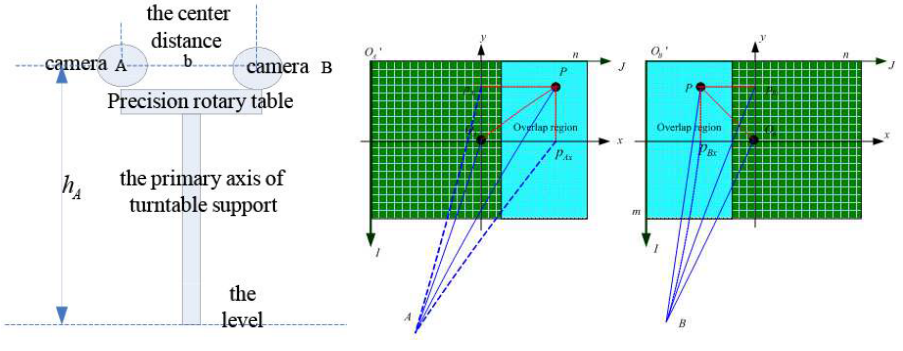


Fig. 1. (Left): the system's structure. (Right): the position of object P in Camera *A* and *B* at time *t* .

At *t*, camera *A* and camera *B* shoot the object *P* simultaneously, but *P* has different positions in A and B. It can be shown as Fig. 1(Right). Set the coordinate system *O'-IJ*, whose origin is the left-hand corner point of the digital photo, to the coordinate

system of the digital image. Set the coordinate system *O-xy*, whose origin is the image's center, to the coordinate system of the image surface. In the coordinate system of the digital image, if the coordinate of a random point *P* is *P(i,j)*, then its coordinate in the image-surface coordinate system *O-xy* can be transformed by the following formula:

$$\begin{cases} x = \dfrac{2j - n}{2} \\ y = \dfrac{m - 2i}{2} \end{cases} \tag{1}$$

Then, in camera *A*'s image-surface coordinate system, the change of the pitch angle in the vertical direction and the horizontal direction is $\Delta\beta_{A,t}$ and $\Delta\alpha_{A,t}$. And, they have the following relationship:

$$\begin{cases} \Delta\beta_{A,t} = \arctan \dfrac{(m - 2i_A) \cdot h_{A,pixel}}{2f_A} \\ \Delta\alpha_{A,t} = \arctan \dfrac{(2j_A - n) \cdot w_{A,pixel}}{2f_A} \end{cases} \tag{2}$$

Similarly, in camera *B*'s image-surface coordinate system, its azimuth change and the pitch angle change are respectively:

$$\begin{cases} \Delta\beta_{B,t} = \arctan \dfrac{(m - 2i_B) \cdot h_{B,pixel}}{2f_B} \\ \Delta\alpha_{B,t} = \arctan \dfrac{(2j_B - n) \cdot w_{B,pixel}}{2f_B} \end{cases} \tag{3}$$

2.2 Distance Measurement

As showed in Fig. 2(Left), *AP* and *BP* are the final distance between the object and the detector [11, 13-15].

According to Fig. 1(Left), the object is always in the overlap viewing region of *A* and *B*. Also the primary optical axes of *A* and *B* are parallel, *AB* is parallel to the level, and the sizes of the images are the same, so the object is in the imaging overlap region of the two cameras. The Fig. 2(Right) can show the relationship between the objective projection and *AB*, we can get:

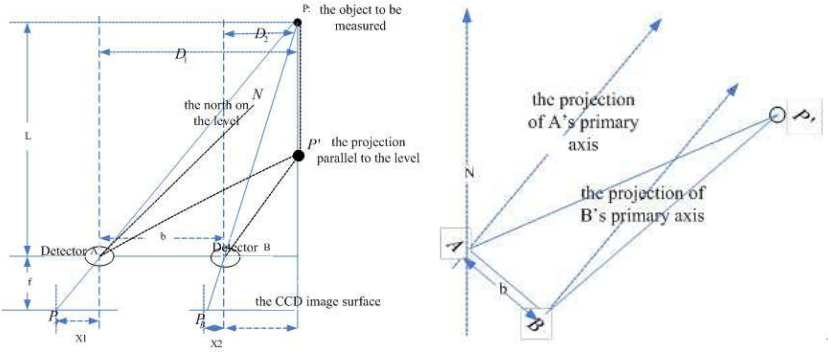


Fig. 2. (Left): the distance between the object and detector. (Right): the horizontal projection of the detector and the object.

$$\begin{cases} \cos\angle PAB = \cos\beta_{A,t}^{'} \cdot \cos(\dfrac{\pi}{2} - \arctan\dfrac{(2j_A - n)\cdot w_{A,pixel}}{2f_A}) \\ \cos\angle PBA = \cos\beta_{B,t}^{'} \cdot \cos(\dfrac{\pi}{2} + \arctan\dfrac{(2j_B - n)\cdot w_{B,pixel}}{2f_B}) \end{cases} \tag{4}$$

So

$$\begin{cases} \angle PAB = \arccos(\cos\beta_{A,t}^{'} \cdot \cos\angle P'AB) \\ \angle PBA = \arccos(\cos\beta_{B,t}^{'} \cdot \cos\angle P'BA) \end{cases} \tag{5}$$

As showed in Fig. 2(Left), it is easy to have the following formulas in $\triangle ABP$:

$$\begin{cases} AP = \dfrac{b\cdot|\tan\angle PBA|}{|\tan\angle PAB| + |\tan\angle PBA|} \cdot \dfrac{1}{|\cos\beta_{A,t}^{'} \cdot \cos\angle P'AB|} \\ BP = \dfrac{b\cdot|\tan\angle PAB|}{|\tan\angle PAB| + |\tan\angle PBA|} \cdot \dfrac{1}{|\cos\beta_{B,t}^{'} \cdot \cos\angle P'BA|} \end{cases} \tag{6}$$

The formulas above are unsuitable for the accuracy analysis of the algorithm. So in order to analyze the accuracy of the algorithm, the formulas are developed by the principle of the optical imaging system. As showed in Fig. 2(Left), point P on the object is imaged to P_A and P_B after passing the two cameras whose center distance is b, and will be received by CCD. Set their positions on CCD to X_1 and X_2 respectively, then according to the imaging principle and the homothetic triangle theory [16-18], when $f_2 = f_1 = f$, we can get:

$$L = \frac{bf}{|X_2 - X_1|} \tag{7}$$

Therefore,

$$\begin{cases} AP = \dfrac{L}{|\cos \Delta\alpha_{A,t}|} = \dfrac{bf}{|(X_2 - X_1)\cos \Delta\alpha_{A,t}|} \\ BP = \dfrac{L}{|\cos \Delta\alpha_{B,t}|} = \dfrac{bf}{|(X_2 - X_1)\cos \Delta\alpha_{B,t}|} \end{cases} \tag{8}$$

Set $x = |X_1 - X_2|$,

$$\begin{cases} AP = \dfrac{L}{|\cos \Delta\alpha_{A,t}|} = \dfrac{bf}{x|\cos \Delta\alpha_{A,t}|} \\ BP = \dfrac{L}{|\cos \Delta\alpha_{B,t}|} = \dfrac{bf}{x|\cos \Delta\alpha_{B,t}|} \end{cases} \tag{9}$$

2.3 The Structure Parameters of System and the Accuracy of Measurement

According to formula Eq.(9), the following formulas can be derived:

$$|\Delta AP| = \frac{bf}{x^2 |\cos(\Delta\alpha_{A,t})|} \tag{10}$$

$$|\Delta BP| = \frac{bf}{x^2 |\cos(\Delta\alpha_{B,t})|} \tag{11}$$

$$x = \frac{bf}{AP|\cos(\Delta\alpha_{A,t})|} \tag{12}$$

$$x = \frac{bf}{BP|\cos(\Delta\alpha_{B,t})|} \tag{13}$$

So we can get the relationship among the system parameters *b,f*, the measured distance *AP*,*BP*, and the relative error (namely accuracy) $\frac{\Delta AP}{AP}, \frac{\Delta BP}{BP}$, that is:

$$\begin{cases} \frac{|\Delta AP|}{AP} = \frac{AP|\cos(\Delta\alpha_{A,t})|}{bf}\Delta x \le \frac{AP}{bf}\Delta x \\ \frac{|\Delta BP|}{BP} = \frac{BP|\cos(\Delta\alpha_{B,t})|}{bf}\Delta x \le \frac{BP}{bf}\Delta x \end{cases} \tag{14}$$

According to the above formulas (14), the longer the distance *b* is, the less the relative error and the higher the accuracy will be. The bigger the focus *f* is, the less the relative error and the higher the accuracy will be. The bigger the measured distance is, the greater the relative error will be. The smaller the imaging position difference is, the less the relative error and the higher the accuracy will be. Taking the accuracy and using convenience into consideration, set the distance *b=1m*.

Table 1. Experimental data of the outfield test

Experimental Data				Objective Distance AP1 (m)	Actual Distance (m)	Relative Error (%)
Object P1	A	azimuth	62°	69.6784	70.5	1.17
		Pitch angle	49°			
		Imaging position	(313,466)			
	B	azimuth	59°			
		Pitch angle	52°			
		Imaging position	(339,517)			
Object P2	A	azimuth	63°	100.9990	98.5	2.53
		Pitch angle	47°			
		Imaging position	(577,512)			
	B	azimuth	59°			
		Pitch angle	52°			
		Imaging position	(704,372)			
Object P3	A	azimuth	59°	88.2452	87.5	0.852
		Pitch angle	47°			
		Imaging position	(233,352)			
	B	azimuth	55°			
		Pitch angle	51°			
		Imaging position	(277,561)			

The receiving system of the digital ranging system based on the binocular parallax is CCD, and the resolution of CCD is determined by the size of the spots and the distance of the spots. If Δx is smaller than the minimum resolution of the receiving system, the object won't be resolved. So make sure that the imaging position difference Δx is bigger than the minimum resolution of the receiving system. If the distance between the two spots is smaller than the center distance, then CCD will think they are one spot. So the minimum of Δx should be $\geq$ the size of CCD spot. The smaller the size of CCD spot is, the smaller Δx will be, the less the relative error will be, and the higher the accuracy will be.

3 Test by Experiments

In order to meet the requirements of accuracy, portability and practicality, in this experiment we set $b=1m$ and the focus of camera A , B, $f=100mm$. The size of each spot is $8.6\mu m \times 8.3\mu m$, and the size of CCD is 867×576. Measure the object's actual distance by the laser rangefinder whose ranging accuracy is $5m$, and solve and compute the algorithm in this paper by MATLAB, then the results listed in Table 1 will be gotten. The experimental results indicate that: the ranging system has a good stability, a high accuracy and a fast response speed. The concrete experimental results are showed as the Table 1.

4 Conclusion

In this paper, by using the principle of binocular parallax and summarizing the former researches, a simple, practical, fast and convenient passive real-time video monitoring ranging system of double CCD is designed. It realizes the real-time video monitoring ranging algorithm of the system and studies the relationship among the system's structure parameters, the ranging areas and the accuracy. It also analyzes the accuracy and the error of the ranging system. By doing experiments, the algorithm turns out to be of simple ranging method, good real time and stability, high accuracy and simple equipments. Another advantage: the system doesn't send signals while working, so it is convenient to conceal yourself, thus having a high safety. Therefore, the system has great application values in the fields like air defense for strategic points, land and sea defense monitoring, real-time monitoring of the battlefield, etc.

Acknowledgments. This work is supported in part by National Science Foundation of China under Grant No. 10971164 and the Fundamental Research Funds for the Central Universities No. xjj20100112.

References

1. Yoshida, K., Hirose, S.: Laser triangulation range finder available under direct sunlight. In: IEEE Int. Conf. Robot. Autom., vol. 3, pp. 1702–1708 (1988)

2. Brenci, M., Mencaglia, A.: Multipoint temperature monitoring in a hazardous environment. In: Proc. SPIE, vol. 158, pp. 97–100 (1992)
3. Wang, H.: Long-range optical triangulation utilising collimated probe beam. Opt. Laser. Eng. 23, 41–52 (1995)
4. Marques, L., Nunes, U.: A new 3D optical triangulation sensor for robotics. In: 10th IEEE International Workshop on AMC, pp. 512–517 (1998)
5. Zeng, L., Yuan, F.: A two-beam laser triangulation for measuring the position of a moving object. Opt. Laser. Eng. 31, 445–453 (1999)
6. Gorti, S.: Triangulation method for determining capillary blood flow and physical characteristics of the Skin. Appl. Opt. 38, 4914–4929 (1999)
7. Lizcano, C., Marquez, M.: Three-dimensional surfaces reconstruction base on laser triangulation. In: Proc. SPIE, vol. 5622, pp. 1322–1327 (2004)
8. Gui, H.: Coarse error analysis and correction of a two-dimensional triangulation range finder. Chin. Opt. Lett. 4, 84–86 (2006)
9. Jin, T.: Evaluating 3D position and velocity of subject in parabolic flight experiment by use of the binocular stereo vision measurement. Chin. Opt. Lett. 8, 601–605 (2010)
10. Jaime, R.: Flat elements on buildings using close-range photogrammetry and laser distance measurement. Opt. Laser. Eng. 46, 541–545 (2008)
11. Francisco, R.M., Wang, Q., Zhang, Q.: Design parameters for adjusting the visual field of binocular stereo cameras. Biosyst. Eng. 105, 59–70 (2010)
12. Xiao, Z., Jin, L., Yu, D., Tang, Z.: A cross-target-based accurate calibration method of binocular stereo systems with large-scale field-of-view. Measurement 43, 747–754 (2010)
13. Hayashi, R.: An integrative model of binocular vision: a stereo model utilizing interocularly unpaired points produces both depth and binocular rivalry. Vision. Res. 44, 2367–2380 (2004)
14. Blake, R., Wilson, H.: Binocular vision. Vision. Res., 1–15 (2010)
15. Bhanu, B., Symosek, P., Snyder, S.: Synergism of binocular and motion stereo for passive ranging. IEEE T. Aero. Elec. Sys. 30, 709–721 (1994)
16. Zhao, P., Wang, N.: Precise perimeter measurement for 3D object with a binocular stereo vision measurement system. Optik 121, 953–957 (2010)
17. Zhao, P., Ni, G.: Simultaneous perimeter measurement for 3D object with a binocular stereo vision measurement system. Opt. Laser. Eng. 48, 505–511 (2010)
18. Bao, X., Li, M.: Defocus and binocular vision based stereo particle pairing method for 3D particle tracking velocimetry. Opt. Laser. Eng. 49, 623–631 (2011)

Operation Mode Selection of Fourth–Party Logistics in China

Liping Chen[1], Luming Yang[1], and Baoquan Sun[2]

[1] Business & Tourism Management School, Yunnan University, Kunming, China
clp911@qq.com, lmyanga@163.com
[2] Public Management School, Yunnan University, Kunming, China
421794817@qq.com

Abstract. Firstly, based on the information at home and abroad, combining the actual conditions in our country, the paper described China's status for development of Fourth–Party logistics. Second, on the basis of that, the paper discusses the four operation modes of Fourth–Party Logistics in China in detail: coordinated operation type, integrated program type, innovative industry type, and dynamic alliance type. Finally, this article analyzes those modes and internal and external logistics business environment, firm size and industry conditions that the four types of models adapt to.

Keywords: Development status, mode of operation, mode selection.

1 Introduction

Fourth–Party logistics（FPL, 4PL) is an integrator and coordinator of supply chain. It deploys and manages all the resources, capabilities and technology that the organization itself and other complementary services organizations have, to provide comprehensive supply chain solutions.[1] However, the logistics situation of our country is that there are many drawbacks, not fully promote the development of 4PL and that the four kinds of Operation mode of 4PL not good distinguished and used, it will be is very important analyzing four kinds of operation mode, selecting out different mode of operation for different enterprises, so as to make better use of 4PL and social resources for the company to serve the community.

2 China's Fourth–Party Logistics Status and Problems of Development

2.1 Development Status of China's Fourth–Party Logistics

4PL is rather late in China and development is very immature. After March 2009, logistics consulting firm gradually increased, but not comprehensive enough to provide services, 4PL functions of the integration of social resources is not well play.

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 499–506.
springerlink.com

Status of Logistics Infrastructure in China: Our current transportation system including rail, highways, waterways, civil aviation and pipelines five modes of transport, have made great development in infrastructure, technical equipment, management, transportation markets. The efficiency is much higher than the U.S., China Ocean Shipping is the world's gross tonnage, and length of highways in China is 3.57 million km, ranked second in the world.[2] Transport routes and operations facilities have been greatly improved.

Table 1. Cargo transported by various means and the growth rate of traffic in 2009

Index	Unit	Absolute	Previous year %
Total	Million tons	278.8	7.5
Railway	Million tons	33.3	1.9
Highway	Million tons	209.7	9.4
Water transport	Million tons	31.4	3
Civil Aviation	Billion tons	445.5	9.3

Source: CCTV, http://news.cctv.com/china/20100225/110706_4.shtml

Third-Party Logistics (the basis of 4PL) Situation: 4PL is to integrate social resources based on the Third-Party logistics (3PL); it is a higher level of development. 3PL in China has the following characteristics: First of all, the efficiency of 3PL providers is very low, and the cost is heavy.

Table 2. China's logistics costs

Annual	Logistics costs (million)	GDP (billion)	Logistics costs/GDP
2008	54,542.0	300,670.0	18.1%
2007	45,406.0	257,305.6	17.6%
2006	38,414.0	211,923.5	18.1%
2005	33,860.0	183,217.4	18.5%
2004	30,057.1	GDP (billion)	Logistics costs/GDP

Source: China Logistics Information Center

Second, function that 3PL providers in China are single, and value-added services are weak. 85% of the proceeds of logistics service providers come from the basic services, such as transportation management and warehouse management. Third, the 3PL market is still quite scattered, small-scale. 3PL companies in the survey, not a logistics service provider with more than 2% market share. Currently, the geographic concentration of China's logistics market is highly, 80% of the proceeds from the Yangtze River Delta and Pearl River Delta.[3]

Logistics Professionals Status: Rapid development of the logistics industry in China will certainly has great demand. It also has increasing requirements for logistics professionals' quality. Scale of training professionals in China about 5,000 a year and training scale of serving officers about 50,000 people. From the demand ratio for all kinds of logistics professionals, demand for top talent is 10%, 30% intermediate talents. Junior professionals and general operator account for 60%. [4]

2.2 Problems in the Development of China Fourth-Party Logistics

China's Logistics Infrastructures are not Fully Utilized. Although China's logistics infrastructure has improved rapidly, large in total, it has not been more fully applied. Main problems: outstanding transportation structural contradictions, and all kinds of transport has not yet formed a reasonable division of labor relations; the scope of market cross seriously and resources of logistics waste seriously; logistics development is uneven, devise from a region and type, the development of logistics enterprises in center have good stability, small gap. Land-based logistics business development with good balance; high waste, high cost is the important reasons to impede improving operational efficiency and so on.[5] Therefore, China should make good use of good logistics resources to enhance logistics efficiency.

Third-Party Logistics Cannot Effectively Support the Development of Fourth-Party Logistics. Generation of 4PL should have a highly developed and strong competitive 3PL market as basement. Above analysis shows, 3PL services have many drawbacks. The third generation of the logistics industry in China, not only in cost and efficiency, but also in size and service, there is great potential for sustained improvement. Only do this, we can better support the development of 4PL industry, or high efficiency and low Cost are empty dream.

Logistics Professionals Are Lack. China's Demand for Logistics is Huge. According to statistics, Logistics professionals' shortfall is about 6000,000 people, needs of undergraduate logistics professionals is about 300,000 to 400,000 people in China, 2010. Logistics planning and consulting professionals, and export-oriented international logistics professionals, logistics researcher are the three professionals in the industry which is the most lack, Logistics professionals have been listed as one of our 12 types of short professionals.[6] As the inadequate professionals of logistics business logistics, without the right people can be introduced, had to employ general staff, resulting in low levels of business, can hardly meet the increasing development requirements of internationalization of logistics industry.

In total, the development of China's 4PL is in the stage of theory beyond than in practical application. Various factors in practice make it quite difficult to apply and accelerate 4PL. 4PL development is still have some limitations.

3 Operation Mode and Analysis of Fourth–Party Logistics in China

For to better use 4PL to provide benefits for the society, we should analysis four kinds of operation mode of 4PL detailed, make a thorough understanding so selecting a suitable operation mode for different enterprises.

The main compositions of 4PL are 3PL providers, IT information servers, management consulting firms, and other value-added service providers. By these four parts together constitute the 4PL, According to the different composition organization; there are four modes of operation: The coordinated operation type, integrated program type, innovative industry type, and dynamic alliance type.

3.1 Coordinated Operation Type

The coordinated operation type is the 3PL and 4PL to jointly develop the market, using the capacity and the market both sides have to supply chain. 4PL provide series services for 3PL, including technology, supply chain strategy, market entrance and project management ability. 4PL service providers working in 3PL Company, not directly contact with the enterprise customer, their ideas and strategies through a 3PL to achieve a specific actor, which aims for customer service.[7] The coordinated operation type's Business area mostly focuses on logistics management, with targeted and flexible business. Outstanding feature of the model is that 3PL (one of the coordinated operation parties) provide strong distribution strengths and that 4PL provide the best solution. A specific example of the pattern is: invested by Andhra Logistics Company, Andhra supply chain Technology Company can use existing client resources of Andhra Logistics Company, provide 4PL services at a lower price or free.

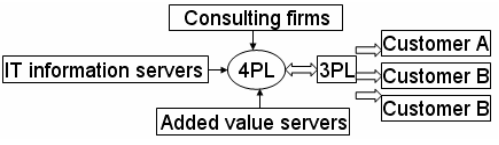


Fig. 1. The coordinated operation type

3.2 Integrated Program Type

Integrated program type is that, 4PL provide the entire supply chain solutions operation and management, and use the resources, capabilities and technology of its members, to provide customers with comprehensive, integrated services. Integrated program type have difference with the coordinated operation type. It not works in 3PL, but as links of business customers and 3PL. So, costumers do not need to contact with 3PL service provider, but directly through the 4PL to implement complex logistics management. The core of this mode is 4PL organizations. As an integrator, 4PL joint venture or partnership with customer, the customer shares the main in the company. A clear focus on Range of clients and the safer customer's business and technology secrets are outstanding advantages of the model. 4PL have stable, close and long-term relationship with customers. But the important premise conditions is the customer's business total will be large enough to make the partners satisfy obtained profit, or they are unwilling to put all resources on a customer. A concrete example: ACCENTURE did a year of consultation and then advised COMPAQ to outsource maintenance and recovery business, implemented Refocusing management.

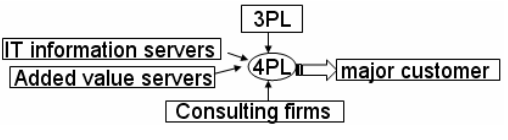


Fig. 2. Integrated program type

3.3 Innovative Industry Type

Innovative industry type is that 4PL provide the supply chain solution program for various industries by collaborating with various service providers who supply resources, technology and capacity. The main difference between Innovative industry types and integrated program type is that the former provides services for customers in various industries, not only a major customer. This model is led by 4PL, jointing other service providers, to provide warehousing, distribution, transportation and other high-end full-service for more than one industry supply chain solutions to customers. In this model, 4PL provide overall industry solutions, enabling the operation of the logistics of the scale more effective expanded, the logistics industry in the operational for more revenue, but also need to improve supply chain solution of wider and more industries. The US carter's logistics company work from the beginning of transporting goods for the head office is to development of providing supply chain solution for many other industries such as Chrysler, Peugeot, Ericsson and other large enterprises.

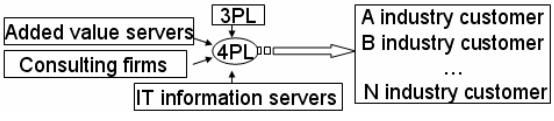


Fig. 3. Innovative industry type

3.4 Dynamic Alliance Type

Dynamic alliance type is that 3PL, consulting companies, suppliers, manufacturers, distributors, some independent service providers and customers connect by information technology then formed a supply chain management alliance in a certain period when a certain market opportunities comes. When market opportunity comes, these organizations quickly organize together to seize the opportunity. As complete a project such an alliance disappeared. This organization is dynamic and flexible. In design, supply, manufacturing, distribution and other areas of the Union, these organizations contribute their core competencies to achieve the profit sharing and risk sharing. Besides the basic features general business, dynamic alliance enterprise also has some special characteristics and here are some special characteristics: organizational process is flexible and dynamic; the process of operation emphasis on explicit and controllable collaboration; information is timely and availably.

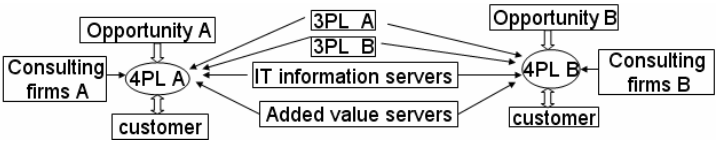


Fig. 4. Dynamic alliance type

4 Operation Mode Selection of Fourth–Party Logistics

4.1 Analysis about the Enterprises Which Coordinated Operation Type Suits to

In the coordinated operation type, 4PL play a role of military adviser to serve customers through 3PL. In 3PL, 4PL joins suppliers, consulting firms and other service providers, serve customers with best ideas about how to make up a supply chain–the ideas serve customers through 3PL. Under this mode, 4PL has a good partnership with 3PL as without 3PL, the former cannot give its ideas to customers directly. 4PL needn't to have leadership; it only provides feasible programs of supply chain. This mode should be in a stable circumstance, because mutations in the internal and external environment are against the cooperation relationship between 3PL and 4PL.

4.2 Analysis about the Enterprises Which Integrated Program Type Suits to

In this mode, the customers which 4PL enterprises serve must have a lot of transactions. What's more, 3PL, Service providers, suppliers and distributors, consulting company and IT service provider also are the members of 4PL enterprises. This fact means that 4PL enterprises need to have strong leadership and management skills and they will become representatives to communicate with major customers. 4PL serve a major customer independently, so it must face the same circumstance as 3PL. In this mode, 4PL enterprises need to have one stable customer, and this customer must have lots of business to be solved. Only in this way, can logistics enterprises function smoothly.

4.3 Analysis about the Enterprises Which Innovative Industry Type Suits to

Innovative industry type provides more chances to 4PL enterprises, and it requires them to enhance the resolution ability of supply chain. Compared with Integrated program type, this one serves more customers in more industries. Different industries require different services. The environment that 4PL frontage is volatile. 4PL have to manage supply chains in lots of industries very well to get profit, which make it face more challenges. So this mode suits to the logistics enterprises that are larger and capable to manage many kinds of industries' logistics.

4.4 Analysis about the Enterprises Which Dynamic Alliance Type Suits to

The most obvious characteristic of dynamic alliance type is that its organizational structure is dynamic and flexible; besides, it is in a complex circumstance. What dynamic alliance type is faced with are not only the various customers, but also partners with different ideas. its cooperation bases on an accidental opportunity, which will be over when project is finished. This kind of cooperation work in a large network environment and it is very opening and competitive. Its dynamics will make it very hard to manage and operate organization. 4PL need to have the ability of innovation, learning, and acceptability and adaptability, only they do this can they make rapid response for the changes in the market and customer demand and grasp chances.

To sum up, from coordinated operation type to dynamic alliance type, Internal and external environment of the four models become more and more flexible and the operation difficulty become harder and harder. Also, 4PL becomes more and more independent; it can realize integrated management of some enterprises; its ability to make up programs of supply chain becomes more and better. All of these reflect that this industry is becoming more perfect. However, the four models which we mention above, do not have very obvious differences, just the reverse, they are in interconnectedness to some extent.

Table 3. Different operation modes of logistics suit to different enterprises

The enterprises characteristic	Coordinated operation type	Integrated program type	Innovative industry type	Dynamic alliance type
Internal and external environment	Stable	Stable	Dynamic	Dynamic
Enterprise independence	Rely on 3PL	Rely on a major customer	Independence	Independence
Level of operation difficulty	Easier	Easy	Difficulty	Highly difficult
Enterprise status	Initial stage and very Small	For a major customer, have some ability	Mature, capable	Highly developed, flexibility, capable

5 Conclusion

No matter which model 4PL adopt, it breaks through the limitation of 3PL. 4PL develops not so well in China today. During this stage, we should operate logistics begin with the coordinated operation type to provide some comprehensive programs of supply chain. With the development of logistics, its model will develop from low level to high level. In future, when we can get more information easily, we will get more opportunity to develop 4PL. Dynamic alliance type will be the better choice.

References

1. Lei, W.: The way to improve the ability of Fourth–Party logistics. Guangxi Journal of Light Industry 1, 115–116 (2009)
2. China railway nets, http://www.railcn.net/news/railway-express/112024.html
3. The pan pearl river delta logistics nets, Trend for Development of 3PL, http://www.pprd56.com/wuliu/zhishi/disanfangwuliu.htm
4. Logistic personnel status in China, http://www.059rc.com/training/news_show.asp?id=10652
5. Tan, B.: Study on the Logistics Performance Evaluation System Based on the Mode of DEA. Zhong Nan University Master thesis

6. Tianjin logistics division training school, http://www.pxloo.com/Special/Special-1-5.html
7. Zhao, G.: Analysis of Fourth–Party Operation mode. Finance &trade Economics 10, 71–73 (2004)
8. Büyüközkan, G., Feyzioğlu, O., Ersoy, M.Ş.: Evaluation of 4PL operating models: A decision making approach based on 2-additive Choquet integral. International Journal of Production Economics 121.1, 112–120 (2009)
9. Zhai, L., Liu, H.: Supply chain operation of virtual enterprise Fourth Industrial. Technology & Economy 4, 114–117 (2009) (in Chinese)
10. Wu, W., Ni, W.: Fourth mode of operation and Core Competence. China Economist 10, 62, 64 (2005) (in Chinese)
11. Li, B.: China's Fourth of development model and strategy. Science & Technology Information 1, 133–134 (2006) (in Chinese)
12. Guo, H., Wang, C., Peng, Z.: Fourth win-win development model. Enterprise Management 03, 94–96 (2009) (in Chinese)
13. Yang, D., Zhao, C.: On the Fourth of China Electronic Commerce. Silicon Valley 12, 198 (2009) (in Chinese)
14. Xia, X., Yang, Y., Liu, Z.: Fourth Development Strategy in China. Modern Business 04, 111 (2009) (in Chinese)
15. Chu, W.: mode selection of logistics management In E-commerce environment. Market Weekly (Disquisition Edition) 8(10), 11–12 (2007) (in Chinese)

A 1.2V Sample-and-Hold Circuit for 14-Bit 250MS/s Pipeline ADC in 65nm CMOS

Wang Xuan[1], Li Fule[1], and Wu Bin[2]

[1] Institute of Microelectronics, Tsinghua University, Beijing 100084, China
[2] Institute of Microelectronics, Chinese Academy of Sciences, Beijing 100029, China
wang-xuan09@mails.tsinghua.edu.cn

Abstract. This paper presents a design of a high speed, high accuracy, low voltage sample and hold circuit used in pipeline analog to digital converter for wireless communication applications in 65nm CMOS technology. Due to the low intrinsic gain of 65nm process, an OTA with simplified class A/B output stage and gain-boosting technique is introduced in the circuit, acquiring a gain of more than 80dB, so as to achieve a 14-bit linearity and a rail-to-rail output swing as well with low power consumption. The linearity issue of the switches is also taken into consideration. The simulation result shows a maximum SFDR of 96.3dB at 11MHz, 95.0dB at 149MHz and 87.1dB at 405MHz input frequency under a sampling rate of 250MS/s.

Keywords: Pipeline ADC, sample-and-hold, linearity, wireless communication application, IF sampling.

1 Introduction

Along with the scaling down of CMOS process and the development of digital signal processing technique, the trend of analog circuits' reduction in SoC is gradually demanded. As a result, particularly in modern wireless communication systems, the removal of mixers in RF receiver has led to superior desired specifications of analog to digital converters (ADCs). As the most essential building block in a high speed and high accuracy ADCs, sample-and-hold (S/H) circuit should be focused extraordinarily by designers. In this work, aspects of high sampling rate and good dynamic performance are taken into consideration. The transit frequency of transistors in 65nm technology is extremely high, which makes it possible to allow S/H circuit to satisfy the necessity of IF sampling. Unfortunately, the linearity of the sampled signal is hardly guaranteed since op-amp's DC gain is mostly limited by poor intrinsic gain during the charge transfer process. The purpose of this work is to balance the trade-offs between the specifications of speed and gain, and to reduce power consumption as well.

This paper is organized as follows. In Section 2, the principle and configuration of S/H circuit is introduced. The circuit design is shown in Section 3. The simulation results of the circuit are given in Section 4. And the conclusion is drawn in Section 5.

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 507–514.
springerlink.com

2 Principle and Architecture of S/H Circuit

A front-end sample and hold circuit is introduced in a well-performed ADC to avoid deterioration of the effective resolution at a high input signal frequency. Thus, the S/H circuit should obtain a performance of high-linearity and low-noise on behalf of not influencing properties of the subsequent stages of quantizer.

In order to achieve a resolution of more than 10-bit, a close-loop sample-and-hold configuration is usually introduced, thus an amplifier with high open-loop gain is required [1]. And for close-loop S/H circuit, the most commonly used architectures are of flip-around and charge-redistribution. Considering that only half of the GBW is needed to obtain the same close-loop bandwidth compared with architecture of charge-redistribution [2], a fully differential flip-around S/H structure (Fig. 1) is chosen in this work to attenuate even-order harmonic, noise and power consumption.

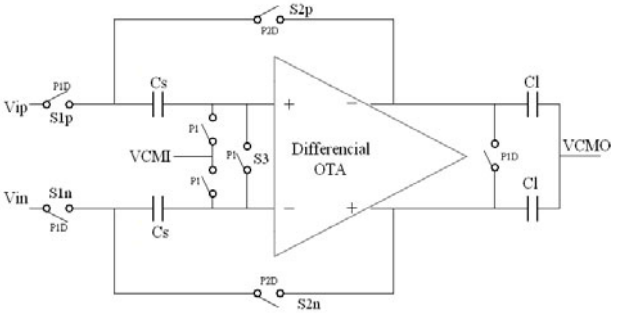


Fig. 1. Flip-around architecture

The circuit operates in two modes of sampling and holding alternately. In sampling mode, S1 and S3 are on while S2 is off, allowing high frequency input signal to be sampled by Cs as the op-amp's input is a virtual ground; when the sampling cycle is over, switch S3 has to be turned off before S1, so called bottom-plate sampling, for the purpose of preventing that charge injection caused by huge sampling switches impacts on sampling accuracy. In holding mode, switches reverse, and OTA works in a close-loop; the output voltage is determined by charges locked on capacitor at the moment S3 being shut off.

3 Circuit Design

3.1 Symmetrical Amplifier with Gain Boosting

Operational amplifier is the core building block in S/H circuit. The charge-transfer process has to be extremely precise to achieve an approving effective resolution and sampling rate, which determines that the op-amp must be designed elaborately to obtain stringent specifications, including DC gain, bandwidth, settling time and output swing.

The most commonly used OTA architectures include: telescopic, folded cascode, symmetrical and two-stages [3]. In this work, 1.2V power supply is too low to guarantee the regular operation of a telescopic topology, and the output swing behavior exhibited by architectures of both folded cascode and symmetrical is also not good enough to conceal noise. Thus, two-stage has to be adopted, since a specialized output stage without cascode-load can present a rail-to-rail output swing. The disadvantage is that Miller compensation must be involved in design to ensure feedback loop's stability, bringing huge compensation capacitor and more power consumption.

Though high gain and high speed characteristics have made folded cascode the most popular architecture in analog designs, in this work, a modified symmetrical topology is chosen as the first stage of the OTA, since no additional inversion amplifier is needed for the both two stages' common mode feedback (CMFB) loop compared with folded cascode, thus to lower the feedback node's parasitic parameters and guarantee the stability. Fig. 2 shows the architecture of OTA used in this S/H circuit.

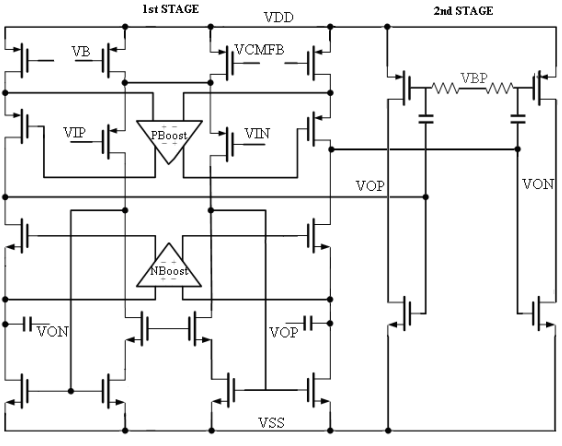


Fig. 2. Symmetrical op-amp with output stage

The output signal of S/H circuit's distortion caused by op-amp nonlinearity has to be controlled within the tolerant margin of the subsequent stage's equivalent resolution:

$$L = \frac{A_{\max} - A_{\min}}{A_{\max} \cdot A_{\min}} < \frac{1}{2} LSB_{next\ stage} \tag{1}$$

Where L measures OTA's linearity; A_{max} and A_{min} are maximum and minimum DC gain respectively. It is estimated that the difference between A_{max} and A_{min} account for 10% of A_{max}. Thus the gain of the OTA being larger than 80dB is a necessity and gain boosting amplifiers (Fig. 3) is used to enhance the effective output impedance, achieving a much higher DC gain accordingly.

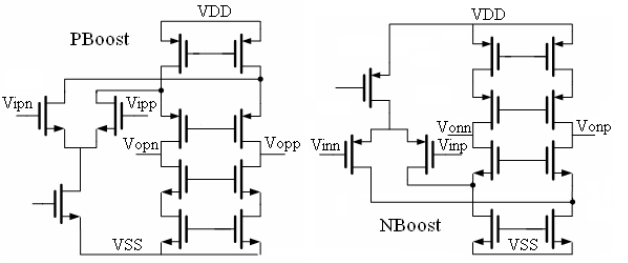


Fig. 3. Gain boosting amplifiers

Gain boosting technique needs not sacrifice much of the power consumption performance, but the existence of boosting amplifiers brings in pole-zero pairs, which may lead to settling problem. To stabilize the feedback loop, a method is discussed in [4], that is, the gain bandwidth of boosting amplifiers has to obey the rule of:

$$\beta\omega_{Main} < \omega_{Boost} < \omega_{Main\_nd} \tag{2}$$

Where β is the feedback factor in the close-loop, ω_{Main} is main op-amp's GBW and ω_{Main_nd} is the frequency of main op-amp's non-dominant pole.

Without constraints of the output swing and CMFB, topology of folded casode is selected for gain boosting amplifiers.

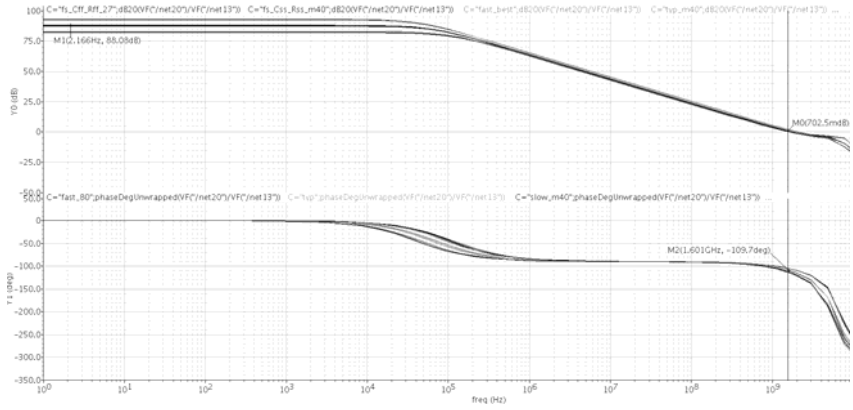
3.2 Class A/B Output Stage

A simplified class A/B output stage is designed in this work. The typical class A/B architecture is shown in [5], and is developed to enhance slew rate behavior and reduce power consumption.

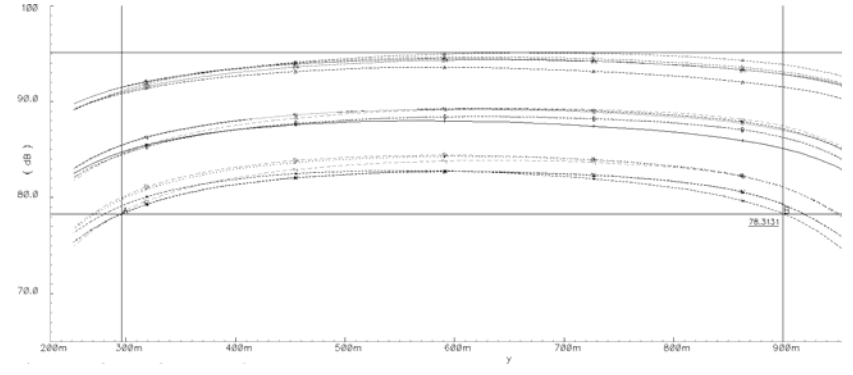
According to the same principle, the bias of output stage will obtain the first stage's output signal by capacitor over VBP shown in Fig. 2, increasing the output transconductance and output current consequently. Besides, the removal of tail current source enables the second stage of OTA to exhibit a rail-to-rail output.

In addition, a differential switched-capacitor CMFB is introduced in this design to stabilize the common mode voltage.

Fig. 4 (a) exhibits the AC response of OTA in corner analyses, and phase margin is 71 degree while GBW reaches 1.6GHz with a 6.4pF capacitor load in typical PVT condition. Fig. 4 (b) shows the sweep result of OTA's DC gain varying with output voltage in 21 corners (including process of transistors and capacitors and temperature corners), and the worst case can still reach a gain of 73.8dB. Besides, in terms of transient simulation, output of the OTA can be settled within 2ns; therefore the S/H circuit's sampling rate is able to reach 250MS/s or even higher.



(a)



(b)

Fig. 4. (a) shows AC response in corner analyses and (b) exhibits DC gain due to output swing

3.3 Switches and Bonding Wire Model

Non-linear contribution due to sampling switches is much more than that caused by a high-gain differential amplifier. The input-depended on-resistance and charge injection lead to severe deterioration of SFDR. In addition, to avoid wideband signal distortion, S/H circuit's input path should have a broad enough bandwidth, which determines that the Ron of S1 (in Fig. 1) ought to be small. According to [6], the sum of resistance of S1 and S3 can be calculated as below:

$$f_{3dB} = \frac{1}{2\pi R_{onsum}(C_S + C_P)} \geq 2^{(N-1)/2} \cdot f_{in} \tag{3}$$

Where f_{in} is input signal frequency, N is converter's resolution, Cs is sampling capacitor, and Cp is parasitic capacitor. As a result, it can be estimated that the on-resistance has to be less than 10-ohm.

In sum, three aspects of small resistance, low relevance to input signal, and small size of switch to attenuate charge injection make the bootstrapped switch technique a necessity. A typical architecture of bootstrapped switch is shown in Fig.5 [7].

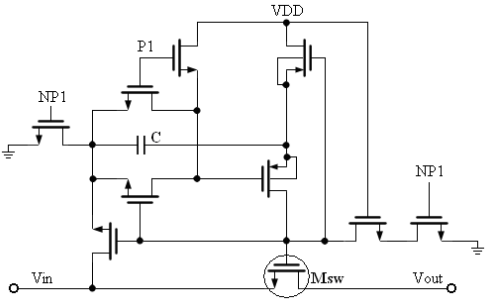


Fig. 5. A simlified architechure of a bootstrapped switch

The bootstrapped gate voltage can preserve a constant of 1.08V higher than input signal as V_{gs} during the sampling phase, enhancing the switch linearity. However, despite of V_{gs} affection, variation of the threshold voltage caused by bulk effect, especially in deep-sub micro process, also influences switches dynamic performance with high frequency input signal. A common solution is increasing the size of M_{sw} in Fig. 5 to decrease the proportion of resistance changes led by V_t, whereas in a low frequency input condition, huge transistor size will aggravate charge injection consequently. Thus a trade-off needs to be made to obtain a satisfied result.

Moreover, the bonding wire model is also considered in this study as the linearity issue. Under an IF sampling condition, impact by bonding inductance cannot be neglected, since the high-frequency current extraction by sampling switches makes parasitic inductance bring in signal-dependent voltage drop. But it will no longer be a problem in the cases of SoC application.

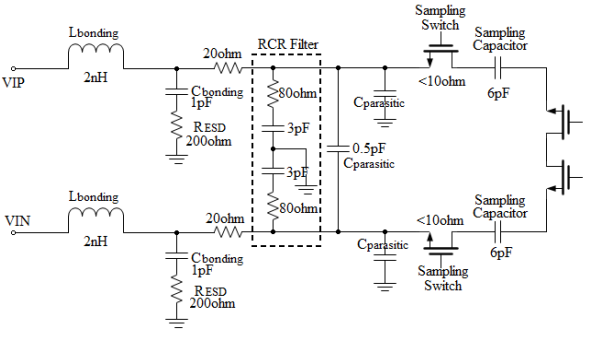


Fig. 6. Bonding wire model

The model is shown in Fig.6 and all the simulation is done with this model in front of the S/H circuit. The series resistance on the signal input path and RCR filter is used to attenuate out-of-band noise.

4 Simulation Results

The performance of S/H circuit is characterized by transient simulation and FFT analysis with a 1.2V power supply and a 250MS/s sampling rate. The total power consumption of the S/H circuit with a 6.4pF capacitor load is 55.2mW.

Fig. 7 shows the spectrum of a 405MHz sampled signal after FFT analysis, and the simulated result reveals an 87.1dB SFDR, reaching the 14-bit linearity specification. Fig.8 exhibits the plot of SFDR versus input frequency while sampling rate keeps 250MS/s. The S/H circuit obtains a maximum SFDR of 96.3dB at 11MHz input frequency.

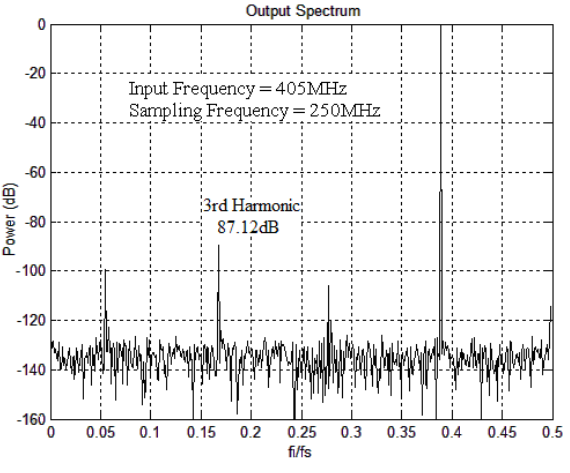


Fig. 7. Output spectrum of an input frequency of 405MHz at a sampling rate of 250MS/s

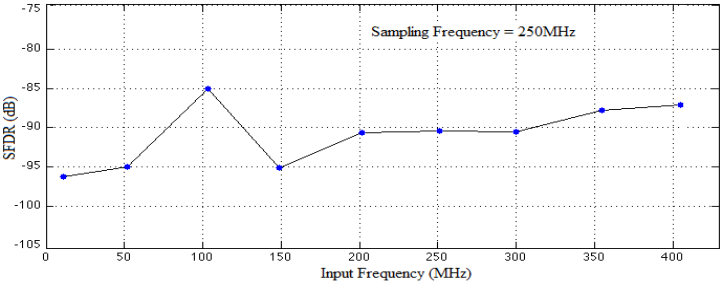


Fig. 8. SFDR versus input frequency

Table 1 summaries the specifications and performance of the designed circuit in this study.

Table 1. Summary of S/H circuit performance

Technology Process	65nm CMOS
Power Supply	1.2V
Sampling Rate	250MS/s
Capacitor Load	6.4pF
SFDR	87.1dB @ 405MHz fin
Power Consumption	55.2mW

5 Conclusion

A 1.2V S/H circuit used in 14-bit 250MS/s pipeline ADC is presented in this study. The modified S/H circuit is designed in 65nm CMOS process. This design uses several methods to improve the performance including higher sampling rate, better dynamic behavior, and lower power consumption, so as to make this design suitable for IF sampling applications in communication systems.

Acknowledgments. This work was supported by the Major National Science & Technology Program of China under Grant No. 2009ZX03007-002. /National Natural Science Foundation of China (No 60976022).

References

1. Boni, A., Pierazzi, A.: A 10-b 185-MS/s Track-and-Hold in 0.35-umCMOS. IEEE J. Solid-State Circuits 36(2), 195–203 (2001)
2. Kelly, D., Yang, W., Mehr, I.: A 3V 340mW 14b 75MSPS CMOS ADC with 85dB SFDR at Nyquist. IEEE J. Solid-State Circuits 36, 1931–1936 (2001)
3. Sansen, W.M.C.: Analog Design Essentials, pp. 144–154. Springer (2008)
4. Bult, K., Geelen, G.J.G.M.: A Fast-Settling CMOS OpAmp for SC Circuits with 90-dB DC Gain. IEEE J. Solid-State Circuits 25, 1379–1384 (1990)
5. Chuang, C.T.: Analysis of The Setting Behavior of An Operational Amplifier. IEEE J. Solid-State Circuits 33, 74–80 (1982)
6. Waltari, M., Halonen, K.: A 220-MSample/s CMOS sample-and-hold circuit using double-sampling. AICSP 18(1), 21–23 (1999)
7. Waltari, M., Sumanen, L.: A self-calirated pipeline ADC with 200 MHz IF-sampling frontend. AICSP 37(3), 201–213 (2003)

The Effect Evaluating System of VMS Information Guidance on Urban Road Network

Chengxiang Zhuge[1], Chunfu Shao[1], Changqing Zheng[2], and Liang Qiao[2]

[1] MOE Key Laboratory for Urban Transportation Complex Systems Theory and Technology, Beijing Jiaotong University, Beijing, China
[2] Traffic Management Bureau, Beijing, China
zgcx615@126.com

Abstract. With the VMS gradually extensively used in major cities, it is necessary to evaluate its guiding result scientifically and systematically. This paper studies on the effect evaluating system of VMS, mainly includes evaluation index and comprehensive evaluation two parts. First, the evaluation index is made up of five indexes which are network, economy, environment, safety and driver. Second, compare the common comprehensive evaluation method and the Grey relational analysis is finally chose. Finally, take one region network of Beijing as an example to apply the evaluation system and prove its efficiency.

Keywords: VMS (Variable Message Signs), Guiding Result, Evaluation Index, Comprehensive Evaluation.

1 Introduction

As a part of the Intelligent Transport System (ITS), the real time condition of the upcoming roads is provided to divers using VMS, to help them choose the appropriate route and get rid of the congested road sections or where traffic incidents occur, so as to balance the traffic flows over the network and enhance the utility of it. Since VMS has been widely built up and utilized in big cities over the country, it is becoming necessary to evaluate the guiding result in a scientific way comprehensively.

Hongyun Yao (2005) presented the framework of index system for VMS evaluation, including internal physical evaluation and operation evaluation [1]. Baofeng Sun (2005) evaluated the effect of highway VMS projects on social economy, giving out the appraisal of VMS's guiding result from the point of benefit [2]. Hui Zhang (2009) compared the condition of before and after building up VMS from economic benefits, traffic benefits and social benefits [3]. Srinivas Peeta(2000) evaluated the Advanced Information Systems' effect, which include fluency, safety and air quality[4]. Although the research on VMS's guiding result is carried out for a few years, the evaluation system has not been formed.

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 515–522.
springerlink.com

2 Build Up the Evaluation Index System

2.1 Principle for Building Up the Evaluation Index System

In order to establish a scientific and complete index system, this paper obeys the following principles:

(1) Scientificalness. Every index should scientifically reflect the guiding result.
(2) Globality. The indexes should reflect VMS's guiding result roundly.
(3) Operability. The evaluation indexes' data should be collected easily.
(4) Independence. Each index should be independent and make sure no an index belong to its same level index.

2.2 Process of Building Up Evaluation Index

According to the above principles, the evaluation index system, which is showed by Table 1, is built up. It contains five first-class indexes and each first-class index is made up of second-class indexes. In order to describe the following equations conveniently, some rules will be made as follows:

I represents the total number of the road section; α_i represents the section i's weight, $\alpha_i = L_i / L$, in where, L_i represents the length of section i, L represents the total length of the whole road network. Q_i represents the traffic volume, V_i represents the space mean speed, C_D represents the design traffic capacity.

Table 1. The evaluation index system of VMS's guiding result on urban road network

VMS's Guiding Result on Urban Road Network	
First-Class Index	Second-Class Index
Network	V/C Ratio
	Average Speed
Economic	Time Saving
	Fule Consumption
Environment	Noise Pollution
	Exhaust Emission (Include CO, HC, NO)
Safety	Mortality Rate per Million Vehicle Kilometer
	Ponderance Index of Accident
	Mortality Rate per Ten Thousand Vehicles
	Equivalent Accident Number per Ten Thousand Vehicles
	Injuries and Deaths Ratio
Driver	VMS Utilization Ratio
	Drivers' Anxiety

Network. Network index includes V/C Ratio and Average Speed.

V/C Ratio. V/C Ratio is an index can reflect the road congestion, the equation to calculate it is showed as follows:

$$\mu_1 = \sum_{i=1}^{I} \alpha_i \times Q_i / C_D \tag{1}$$

Road Network's Average Speed. Road Network's Average Speed can also reflect the VMS's guiding result. The equation to calculate it is showed as follows:

$$\mu_2 = \sum_{i=1}^{I} \alpha_i \times V_i \tag{2}$$

Economic. Economic index includes Time Saving and Fule Consumption.

Time Saving. The network's time benefit is used to evaluate the value of drivers' time loss on road network, which evaluates the VMS's guiding result from the point of view of the economy. The calculation process is as follows:

Value of drivers' time loss on road network (B_{time}):

$$B_{time} = \sum_{i=1}^{I} Q_i \times \frac{L_i}{V_i} \times vot \tag{3}$$

vot represents unit time value (Unit: Yuan/Hour).

Fule Consumption. In the actual road traffic condition, drivers can avoid taking the congestion road under the VMS's guiding, which can avoid the vehicles moving hardly and reduce the fule consumption. Thus, fule consumption is an important index reflects the VMS's guiding result. The detailed calculation process is as follows:

Firstly, use the second steady fule consumption model, which is proposed by Post.etc, to calculate each section's fule consumption.

In unit time, section i 's fule consumption function (F_i):

$$F_i = L_i \times Q_i \times (a + b \times V_i + c \times V_i^2) \tag{4}$$

In where, $a = 170 \frac{ml}{km}$, $b = -4.55 \frac{ml \cdot h}{km^2}$, $c = 0.049 \frac{ml \cdot h^2}{km^3}$.

In unit time, the road network's fule consumption function (B_{Flue}):

$$B_{Flue} = \sum_{i=1}^{I} F_i \times P_{Flue} \tag{5}$$

In where, P_{Flue} represents the current fule price (Unit: Yuan/Liters) .

Environment. Environment index includes Noise Pollution and Exhaust Emission.

Noise Pollution. Traffic noise pollution degree is mainly affected by traffic flow, vehicle composition, traffic speed, road conditions, etc. So the VMS's guiding can affect it through affecting its factors.

Road network's noise function(μ_3):

$$\mu_3 = \sum_{i=1}^{I} \alpha_i \times L_{eqi} \tag{6}$$

In where, L_{eqi} represents section i 's equivalent A-weighted sound pressure level.

Exhaust Emission (Include CO, HC, NO). Drivers can not only reduce the fule consumption, but also reduce the exhaust emission under the VMS's guiding. So exhaust emission is another important index, this paper mainly focuses on three gases: CO, HC, and NO.

The quadratic polynomial fitting result between speed and exhaust emission and their related parameters are showed as follows:

$$y_{CO} = 0.0113x^2 - 1.7768x + 81.57 \quad R^2 = 0.9532 \tag{7}$$

$$y_{HC} = 0.0009x^2 - 0.1527x + 9.03 \quad R^2 = 0.9871 \tag{8}$$

$$y_{NO} = 0.025x + 2.8 \quad R^2 = 0.9871 \tag{9}$$

The road network's exhaust emission is the sum of every section's exhaust emission; it can be calculated as follows:

$$S_{Gas} = \sum_{i=1}^{I} L_i \times \beta_{Gas} \times Q_i \qquad Gas = CO、HC、NO \tag{10}$$

Safety. The urban congestion can be improved under the VMS's guiding, as a result, the distribution of road network's traffic flow tend to be more reasonable and drivers' lane changing behavior is reduced. So the rate of the traffic accidents will be reduced. The road's safety can be improved. This paper provide two methods to evaluate road's safety under two opposite conditions: having or not having accident data.

(1) Having accident data

When having accident data, the following four indexes can be chose to evaluate the safety, they are mortality rate per million vehicle kilometer (K_1), ponderance index of accident (K_2),mortality rate per ten thousand vehicles (K_3), equivalent accident number per ten thousand vehicles (K_4) [5]. Section i 's each index can be calculated by equation (11) to equation (15).

$$k_{1i} = \frac{F_i}{Q_i L_i} \times 10^8 \tag{11}$$

$$k_{2i} = \frac{F_i}{J_i} \tag{12}$$

$$k_{3i} = \frac{F_i}{Q_i} \times 10^4 \tag{13}$$

$$k_{4i} = \frac{(ETAN)_i}{Q_i} \times 10^4 \tag{14}$$

$$(ETAN)_i = (TAN)_i + 9.5F_i + 3.5J_i \tag{15}$$

In equation (15), F_i represents section i 's accident death toll (Unit: person); $(ETAN)_i$ represents section i 's the total number of equivalent accident (Unit: time). $(TAN)_i$ represents section i 's the number of accident (Unit: time). J_i represents section i 's wounded number in the accident (Unit: person).

Road network's each safety index can be calculated as follows:

$$K_a = \sum_{i=1}^{I} \alpha_i \times k_{ai} \qquad a = 1, 2, 3, 4 \tag{16}$$

(2) Not having accident data

When not having accident data, this paper quotes Canada's urban and rural two-lane road survey analysis results: there is following relationship among the section i 's average speed, velocity difference and the rate of casualty (h_i):

$$h_i = 0.01802V_{85i} + 0.01884V_{di} - 1.94294 \quad R^2 = 0.95984 \tag{17}$$

In where, V_{85i} and V_{15i} separately represents the section i 's summation curve at 85% point speed and 15% point speed; V_{di} represents section i 's velocity difference, namely $V_{di} = V_{85i} - V_{15i}$ [6].

Driver. Driver index includes VMS Utilization Ratio and Drivers' Anxiety.

VMS Utilization Ratio. The VMS's guiding result can be evaluated from the view of drivers, if the VMS utilization ratio is big, then VMS play a part in guiding driving route. Under the reasonable guiding information, the bigger VMS utilization ratio, the better guiding result.

Drivers' Anxiety. Traffic congestion not only causes economic losses and environmental damage, but also makes drivers produce irritating mood. The VMS's guiding information let drivers understand surrounding condition, which can increase their controllability and decrease the anxiety. Thus, drivers' anxiety can be one of the indexes to evaluate the VMS's guiding result.

3 Comprehensive Evaluation Method of VMS's Guiding Result

Because there is no comprehensive evaluation methods completely match to VMS's guiding result evaluation. Thus, this paper compare different methods which are used

for ITS's evaluation and chose a better one. There are three factors for choosing: 1) considering evaluation index of VMS's guiding result; 2) considering the number of evaluation object; 3) considering evaluation result. Based on considering three factors above, the Grey relational analysis is chose.

The Grey relational analysis judge the connection's tightness according to the similarity of sequence curve's geometric shape, the tighter the curve close, the bigger the correlation among corresponding sequence is, the smaller the vice is. The evaluation includes 6 steps and each step's detailed content is as follows:

Step1: Ensure the analysis sequence. For this paper use the method of comparison between before and after, thus the number of evaluation is 2, and the number of evaluation index of each evaluation object is decided by decision maker, here assume n indexes and the original array matrix is as follows:

$$X\left[x_{i1}, x_{i2}, ...x_{in}\right] \qquad i = 1, 2 \tag{18}$$

Step2: The original data's dimensionless processing. In order to facilitate analysis and guarantee the indexes' equivalent, the original data should be handled with dimensionless and normalization.

Step3: Calculate maximum differential and minimum differential.

$$\Delta(\max) = \max_i \max_j (\left| x_{ij}' - 1 \right|) \qquad i = 1, 2;\ j = 1, 2, ..., n \tag{19}$$

$$\Delta(\min) = \min_i \min_j (\left| x_{ij}' - 1 \right|) \qquad i = 1, 2;\ j = 1, 2, ..., n \tag{20}$$

Step4: Calculate relational coefficient. Input the step 3's result into equation (21) to calculate relational coefficient.

$$\xi_{ij} = \frac{\Delta(\min) + \beta\Delta(\max)}{\Delta_{ij} + \beta\Delta(\max)} \qquad i = 1, 2;\ j = 1, 2, ..., n \tag{21}$$

In where, ξ_{ij} represents the relational coefficient of the i th evaluation object's j th index with the j th optimal index value in the optimal index muster. β ($0 < \beta < 1$) represents the differentiated coefficient. The value of β is 0.5.

Step5: Chose the reasonable method to calculate the weight according to practical condition. This paper chose AHP to calculate the weight.

Step6. Calculate relational coefficient according to equation (22).

$$E_i = \sum_{j=1}^{n} \omega_j \xi_{ij} \qquad i = 1, 2 \tag{22}$$

Obviously, the bigger value of E_i , the better VMS's guiding [7].

4 Example

This paper evaluates the VMS's planning project from the view of its guiding result. Take one area of Beijing's network for an example, VMS's layout is showed by Figure1.

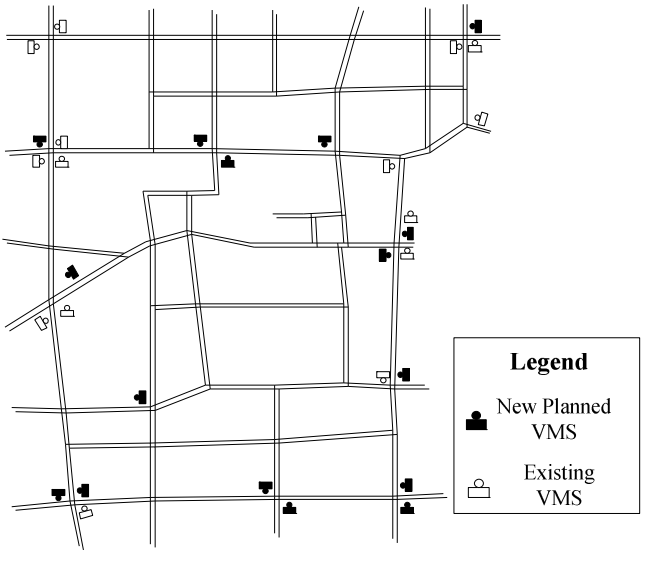


Fig. 1. The map of VMS layout in the planning project

Because of limited collected data, only 10 second-class indexes are calculated. The calculation result is showed by Table 2 and the results can be used to evaluate VMS's guiding result from single view; Based on calculation result, AHP is used to calculate the index's weight and Grey Relational Analysis is used to give a final evaluation to VMS. The comprehensive evaluation result is showed by Table 3. The relational coefficient increases from 0.952 to 0.999 because of building up VMS, namely, road network's condition is improved about 5%.

Table 2. The calculation result of evaluation indexes of VMS's induct result

First-Class Index	Second-Class Index	Before VMS	After VMS
Network	V/C Ratio	0.548	0.536
	Average Speed (km/h)	51.2	53.7
Economic	Time Saving (Yuan/h)	119835	106279
	Fule Consumption (Yuan/h)	44930	43842
Environment	Noise Pollution (DB)	77.1	77.2
	Co Emission (kg/h)	5569	4720
	HC Emission (kg/h)	929	839
	No Emission (kg/h)	974	975
Safety	Injuries and Deaths Ratio	2.7	2.9
Driver	Drivers' Anxiety (0-100)	68	52

Table 3. The calculation result of evaluation indexes of VMS's induct result

First-Class Index	Second-Class Index	Before VMS	After VMS	Changed Value	Changed Rate
Network	V/C Ratio	0.446	0.456	0.01	2.25%
Network	Average Speed (km/h)	0.088	0.092	0.004	4.93%
Economic	Time Saving (Yuan/h)	0.14	0.158	0.018	12.74%
Economic	Fule Consumption (Yuan/h)	0.052	0.053	0.001	2.46%
Environment	Noise Pollution (DB)	0.075	0.075	0	-0.20%
Environment	Co Emission (kg/h)	0.008	0.009	0.001	18.06%
Environment	HC Emission (kg/h)	0.065	0.072	0.007	10.74%
Environment	No Emission (kg/h)	0.016	0.016	0	-0.10%
Safety	Injuries and Deaths Ratio	0.035	0.032	-0.003	-7.70%
Driver	Drivers' Anxiety (0-100)	0.028	0.036	0.008	30.72%
Comprehensive Evaluation		0.952	0.999	0.047	4.98%

5 Conclusion

This paper build up evaluation index system of VMS's guiding result and the system is made up of network, economic, environment, safety and driver. The Grey relational Analysis is used to give out a comprehensive evaluation of VMS. The mainly research achievement is as follows: (1) Build up an evaluation system which is suitable for evaluating VMS's guiding result. (2) Take Beijing for an example and put the evaluation theory into practice.

Henceforth, the mainly research content will be carried out is to develop an evaluating software of VMS's guiding result.

References

1. Yao, H.Y., Liu, X.Y.: Research on VMS Evaluation Index System. Management of Municipal Facilities, 3–6 (2005)
2. Sun, B.F., Jun, Z.C., Wu, W.J.: Reviews of Socio-Economic Impact Evaluation for VMS Project on Freeway. Journal of Transportation Systems Engineering and Information Technology, 18–21 (2005)
3. Zhang, H.: Research on Planning and Design of Variable Message Signs in Beijing. Beijing Jiaotong Universitys, Beijing (2009)
4. Peeta, S., Poonuru, K., Sinha, K.: Evaluation of mobility impacts of advanced information systems. Journal of Transportation Engineering, 212–220 (2000)
5. Pei, Y.L., Dai, T.Y.: Fuzzy Evaluating Method to Distinguish Black Spot. Journal of Highway and Transportation Research and Development, 121–125, 138 (2005)
6. Cui, H.J., Wei, L.Y.: Research on Road Condition and Transport Accident Based on Speed. Journal of Hebei University of Technology, 95–98 (2001)
7. Lin, Z., Chui, Y.P.: Research on Evaluation for Expressway Network Based on Information Entropy-Grey Relative Degree, Traffic and, pp. 122–125 (2008)

A Transitive Trust Chain in Ad Hoc Network

Chongzheng Zhai and Jing Liu

School of Computer Science and Software,
Tianjin Polytechnic University, Tianjin, China
zhaichongzheng@163.com

Abstract. To further enhance the safety of Ad Hoc Network, the paper presents a new trust transfer model with Trusted Platform Module (TPM) in Ad Hoc Network. A scheme which takes a trust chain from the node of Ad Hoc Network to Ad Hoc Network as its design objective is proposed, and the trust relationships between peers can be evaluated with a trust model in Ad Hoc Network. Authenticated routing for Ad Hoc Network (ARAN) is improved by introduction of trust level and then selects the routing of highest trust level. At last the trust transfer model in Ad Hoc network is analyzed.

Keywords: Trust chain, Ad Hoc Network, trust model, ARAN, trust level.

1 Introduction

At present, the computer network represented by Internet develops rapidly and the security issue which people face is more and more severe .Most of information security system consists of passive defense ways such as firewall or invasion detection or virus defense etc, but they can't withstand the attack from the system effectively, and also can't prevent the inexhaustible virus and vicious attack action. For this reason, a new security architecture which can puts its hand to terminal and defends initiatively towards network attack is badly needed. Trusted computing which is researching provides a new trajectory to the solution of this issue. ISO/IEC15408 standard [1] define trust as follows: a trusted module or operation or the action of process is forecasted in random operation condition, and defends effectively the destroy from application virus or definite physical interference.

TCG [2] (Trusted Computing Group) introduces the trust concept to computing platform and solves the security issue of platform from the angle of systemic integrity. Delves into trusted computing, a new trust chain transfers in Ad Hoc network scheme which combines trusted model is proposed and implemented.

2 Remote Attestation

Remote attestation includes the measurement and report of integrity. The main purpose of remote attestation is to verify the system state of current computing environment to distant verifier. In this way, we can make the verifier believe that the system behavior is credible and interacts with the local system.

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 523–529.
springerlink.com

The measurement of integrity includes the measure model of integrity and the storage of the credibility. The measure model of integrity measures the system components which affect the integrity state of the platform. The trusted storage stores the digest which the hash function hashed the integrity values of components measured in the PCRs (Platform Configuration Registers), the PCRs ensure the safety of measure values, system will also store the corresponding SML (Stored Measurement Log) which contains the related information of measurements in PCRs, and list in the order, together with the PCR value to send to the verifier.

The function of the report of integrity is that remote verifier sends remote attestation request to obtain the PCR value of prover. The prover sends the integrity measurements of the platform to the verifier, remote verifier verifies the trust of platform, and the verifier will refuse to interact or provide services to the platform if the trusted environment of platform is destroyed.

3 ARAN Secure Routing Protocol

ARAN [4] is introduced by Kimaya Sanzgiri, it's based on public key cryptography technology, it can achieve the identification of the routing information and message integrity and non-repudiation, it can resist the tampering, forfery and replay attacks to routing messages, but relatively non-secure routing protocol, there are several shortcomings[5]: the high computational complexity, the complexity of routing lookup, the large memory capacity of the node. Taking into account the shortcomings of ARAN, a new trust chain transfer in Ad Hoc network scheme which combines trust level is proposed on the basis of ARAN, the scheme improves ARAN, it minimizes computational complexity, the complexity of routing lookup protocol as possible.

4 A Transitive Trust Chain in Ad Hoc Network

Each Ad Hoc network node is based on TPM security chip. In this scheme, each Ad Hoc network node needs a AIK certificate to show its identify. Then assesses the trust level of neighbor node and selects a highest trust level routing on basis on trust level.

4.1 Ad Hoc Node Initialization

Each Ad Hoc node applies for an AIK certificate to Privacy CA out of line. Privacy CA quits while issues the AIK certificate to each Ad Hoc network node, keeping Ad Hoc network without a central feature. Ad Hoc network node can use the remote attestation mechanism to prove its identity to verifier after it has an AIK certificate.

4.2 Measurement of Trust Level of Neighbor Node

Because in the Ad Hoc network, source node must find a communication path through neighbor nodes, so we need to measure the trust level of the neighbor nodes. The trust level is defined as the sum of trust level of integrity and behavior in Ad Hoc network. The trust level of integrity is defined as follows:

$$T_{integrity}(A, B) = \frac{\alpha EBIOS+\beta ELoader+\gamma EOS+\eta EApplication}{\alpha+\beta+\gamma+\eta} \quad (1)$$

E_{BIOS} is the ratio between integrity measurement and the reference of BIOS, E_{Loader} is the ratio the integrity measurement and the reference of OS loader, E_{OS} is the ratio of the integrity measurement and the reference of OS, $E_{Application}$ is the ratio of integrity measurement and reference of application.α,β,γ,η are positive integer.

The trust level of behavior is based on the direct trust of neighbor nodes. According to the historical record which interacts with its neighbor nodes and context information, we can assess the trust level of neighbor node. Node sets itself as monitor mode so as to monitor weather the next node to send and modify the packet after sending the datagrams. The trust level of neighbor node is set to T_{send}, the trust level without any tempering of the packet is set to T_{right}, we can use the simple Bayesian model to measure the trust level of behavior. Both the trust level are calculated as follows:

$$T_{send}=\frac{1+Ssend}{2+Ssend+Fsend} \quad (2)$$

S_{send} is the number of packets sended by neighbor node; F_{send} is the number of packets not transmitted by neighbor node.

$$T_{right}=\frac{1+Sright}{2+Sright+Fright} \quad (3)$$

S_{right} is the number of packets sended correctly; F_{right} is the number of packets tempered by neighbor node.

The source of the behavior trust level is from different types of experience, we should give different weights to the experience of different types when we calculate the direct trust level. The trust level of behavior is defined as follows:

$$T_{Behavior}=W_{send} \times T_{send} + W_{right} \times T_{right} \quad (4)$$

W_{send} ,W_{right} are the weights of the two trust level.

1) Assess the trust level of integrity: The node sends a random number to begin the measurement of the trust of integrity. Neighbor node receives the request of measurement, it submits the random number to the local TPM.TPM signs the random number and the PCR value of current platform by the AIK private key, and returns a whole event log. The node sends the random number and the PCR values and SML to the requestor, then the requestor uses the AIK public key from AIK certificate to verify the signature, and compares the received random number to the random number sended, this can prevent replay attack. Then calculates the integrity values of the platform from SML, if the integrity values received equals with the integrity values of current platform which calculated. If they match, we can set the integrity value received as the integrity value of current platform of neighbor node. According to PCR value received, we can calculate the integrity trust level of neighbor node by the model of integrity trust level.

2) Assess the trust level of behavior: According to the number of packets sended and send correctly by neighbor node, we can assess the behavior trust level of neighbor node by the model of behavior trust level.

3) Assess the trust level of node: The trust level is defined as follows:

$$T(A,B)=\alpha T_{integrity}(A,B)+(1-\alpha)\, T_{Behavior}(A,B)$$

Because the trust level of integrity reflects the safe state of terminal node, so set $0.5<\alpha<1$, $T_{integrity}(A, B)$ is the integrity trust level of neighbor node, $T_{Behavior}(A, B)$ is the behavior trust level of neighbor node.

4.3 Select Routing

Select a highest trust level route to transmit data. Each node need to maintain a table about the trust level of neighbor node. Node A begins to discover the routing when the source node A want to communicate with destination node X, the steps as follows:

1) When a packet arrives, if node A can't found the effective routing to destination in routing table or the routing is invalid, it will seek a routing. Node A broadcasts a routing request RREQ packet to the neighbor nodes, RREQ as follows:{REQ, IP_X, $Cert_A$, ID_{AX}, t, T_{REQ}}K_{AS}, $Cert_A$. REQ is the packet type identification, IP_X is the destination node's IP address, $Cert_A$ is the certificate of the source node A,ID_{AX} is the identified of routing lookup, it will automatically increase 1 when the node A initiates a new request, t is the current time of the source node A lunches a request for routing, T_{REQ} is the original trust level of routing, values the defaults. The RREQ packet is signed by the AIK private key of node A.

2) Neighbor node B receives the RREQ packet from the source node broadcasted, then it validates the signature by the public key of certificate in node A, and verifies if the certificate is expire. The node should judge the freshness of (IP_X,ID_{AX}) pair, if the routing identifier has the same, checks the time of node A's request. It's possible the route identifier turns after itself overflowing. Then check the trust level of last node, if its trust level is lower, pass the routing request, otherwise compares T_{REQ} and A, if lower than T_{REQ} , updates A's the trust level as T_{REQ} .The intermediate node can be directly returned the routing response to node A. If B is the destination or it have the route, then reply. If not, adds the certificate to B and continue broadcasting after signed by the AIK private key [{REQ, IP_X, $Cert_A$, ID_{AX}, t, T_{REQ}}K_{AS}]K_{BS}, $Cert_B$.

3) Neighbor node receives the packets of node B, then certificate and discards it if it's incorrect. After that checks the trust level of last hop node, it will not reply request if trust level lower, otherwise compares the trust level of node B to T_{REQ}, and gets the lower value. Abandons the certificate and signature of node B, adds IP address of node B, and then uses its AIK private key to sign, adds its certificate. The message as follow :{REQ, IP_X, $Cert_A$, ID_{AX}, t, T_{REQ}}K_{AS},IP_B]K_{CS}, $Cert_C$.

To the destination node X like this, the destination node X validates the certificate and signature, if it's correct, compares the trust level of several routing requests, selects a highest trust level routing to source node A. Selects the route with least number hops when the trust level of the routing is same.

4) The destination node unicasts response packet along the reverse path of the source. Suppose that neighbor node D is the first node which receives the routing response, the message as:{REP, IP_A, ID_{AX}, t, $Cert_X$, T_{REP},IP_B,IP_C,IP_D} K_{XS}, $Cert_X$.

REP is the packet type identification; $Cert_X$ is the certificate of destination node X, T_{REP} is the trust level of routing response. Node D should validate the signature when it receives the message. If it's correct, and then compares the trust level of current routing

response to the trust level of last hop node, gets the lower. Abandons the signature and certificate, signs by the private key of node D, and adds its certificate, passes down next node until source node.

5) Source Node A receives unicast message back from the destination node. At first it validates the signature. If it's correct, then compares the trust level of routing response TREP whether to reach the threshold value of trust level, it decides to trust the routing response directly or not. If the trust level of routing response is higher than threshold value, and then receives, otherwise discards.

5 Performance Comparison

5.1 Comparison of Safety

Because both the trust chain transfer mechanism and ARAN adopt public key cryptography, so the intensity of their security is almost the same. They can prevent unauthorized nodes to join by their certificate, and the trust chain mechanism can also avoids deception, black hole attacks, replay attacks, tampering the routing message and serial number attack etc. However, ARAN secure routing protocol can't prevent the attack of internal malicious nodes, and also can't do anything for the selfish node, even if the selfish nodes are authenticated by the trust server, because ARAN secure routing protocol can't distinguish the selfish node in the network. ARAN is improved by introduction of the trust chain transfer mechanism proposed in the passage. Other nodes will drop the trust level to the malicious component of node or selfish node in Ad Hoc network. Ad Hoc network will be separated from them if their trust level lower than threshold value. Other nodes don't deal with the message sended by the malicious nodes or selfish nodes, also other nodes won't transmit any message to them, and the route of malicious nodes will be canceled.

5.2 Comparison of Routing Strategies

In this instantaneous topology, S is source node, we set its trust level of routing request is 0.8.D is a destination node, its trust level is 0.7,other nodes such as A、B、C、E、F、G、H are the Ad Hoc network nodes passed in the routing discovery process. In the routing protocol, the destination node maybe receives multiple routing paths to reach the source node after the route discovery. ARAN will choose a minimum delay path, under the same load, that is the least path hops. However, the trust chain mechanism will choose a path of maximum reliability; it will choose a path of minimum hops when the trust level is same. As shown in Figure 1,dashed line identifies the route map in the request process, each node also identifies the trust level of neighbor nodes, for example, the trust level of node S to node A and B are 0.6,0.7 respectively. Suppose that node F is a malicious node or selfish node, its trust level is only 0.3 because it has malicious components or does not send packets. Suppose the threshold trust level of each node is 0.5.ARAN will choose a path with least hops, that is S→A→E→F→D ,the trust level of this route is 0.3.But because node F is a malicious or selfish node, its trust level is lower than threshold value, so the trust chain transfer mechanism doesn't choose the path with node F, however, the trust level of the two route such as S→A→E→G→H→D and S→B→C→E→G→H→D is 0.6, the

trust chain mechanism will choose a route of relatively less hops, S→A→E→G→H→D. The trust level of the route is higher than that of ARAN chooses, despite its longer path length compared to ARAN. The selection of longer path will cause certain effect to the response time of the routing, but reliable data transmission can substantially increase the transmission efficiency.

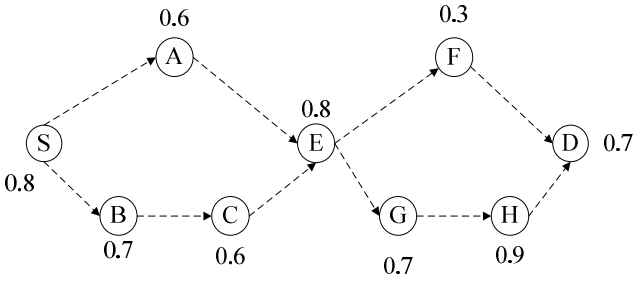


Fig. 1. Instantaneous Topology and routing

6 Conclusion

As the TCG's TNC architecture does not apply to Ad Hoc Networks. The passage goes into the trust computing and the features of Ad Hoc network, utilizes the measurement and report mechanism of trust computing, a new scheme which combines with the trust model is proposed. At first each Ad Hoc network node applies for an AIK certificate, and then according to the trust model to assess the trust level of neighbor node, selects a maximum trust level of routing.

However, the scheme will utilize the integrity value of platform of neighbor node when calculates the integrity trust level of neighbor node, but the integrity value of platform of neighbor node is variational, the integrity trust level of neighbor node is also variational, it can reflect the current state and configuration of nodes at any time. The following work should be further explored the dynamic collection of integrity value of platform, we need to design an optimized trust chain transfer mechanism, it can reflects the variation of trust level of nodes.

Acknowledgments. This work is supported by National Nature Science Foundation of China (No.60970016) and research foundation (No.08ZHTGCG01400 and No.06YFGZGX17500) from Tianjin Science and Technology Committee.

References

1. Standard ISO/IEC.Information Technology Open Systems Interconnection Evaluation criteria for information technology (1999)
2. Trusted computing group. Trust platform module main specification Version 1.2 (2005)

3. Trusted Computing Group.TCG TPM Specification part 1 Design Principles.USA: Trusted Computing Group, pp. 16–26 (2006)
4. Sanzgiri, K., La Flamme, D., Dahill, B., Levine, B.N.: Authenticated Routing for Ad Hoc Networks. IEEE Journal on Selected Areas in Communications 23(3), 598–610 (2005)
5. Tie, M., Li, J., Huang, Z.: The identification routing protocol of Ad Hoc network and performance analysis. Computer Applications 25(9), 2066–2069 (2005)

The System Architecture of Intelligent Monitoring for Broadcasting TV Based on Web Services

Chunlai Zhou

Automation Dept., Information Engineering School Communication
University of China, 100024 Beijing, China
clzhou@cuc.edu.cn

Abstract. The complex systems are usually equipped with monitoring systems to ensure the security and the reliability of theirs. Monitoring systems are usually designed according to main systems, and are closed ones. Therefore it is hard to communicate or cooperate to do a job among them, which is needed when many systems cooperate with each other to perform a job, such as associated live broadcasting of multi-TV stations. This paper presented a kind of system architecture of intelligent monitoring based on web services, which combines the monitoring systems together and by means of mobile devices to support system mobile maintenance and fault diagnosis. The system architecture is very powerful and convenient to implement the different system's cooperation.

Keywords: Intelligent Monitoring, System Architecture, Web Services, Mobile Computing.

1 Introduction

The various monitoring systems, such as the program studio monitoring system, the broadcasting signal monitoring system and the transmitter monitoring system etc., are developed in order to guarantee the security and reliability of the broadcasting TV system in many TV broadcasting stations. For the broadcasting industrial management, different level monitoring centers such as a county, a city and a province center have been set up to supervise the signal quality and contents of programs. The architectures of these monitoring systems are diverse and functions are different, usually each one is a stand alone and closed. Therefore, it is a challenge to realize the communication and cooperation among such monitoring systems.

In the broadcasting TV industry, the requirement to integrate such monitoring systems is urgent needed. The high level management departments need to know the program broadcasting effects of all stations; the technicians need to master the system working status; the associated live broadcasting of multi-TV stations need the monitoring systems to perform concurrent actions according to the unified controls. If the information and functions of the monitoring systems are merged together, that will be very convenient for management staff and technicians and will dramatically guarantee the whole system's reliability and stabilization.

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 531–538.
springerlink.com

In consideration of such similar applications, a kind of system architecture of intelligent monitoring based on web services and mobile computing devices is presented in this paper. The paper described the system framework, key technologies used in the architecture and the part of our works.

2 Key Technologies

2.1 Monitoring System Based on the Network

There are two modes for a monitoring system based on the network: Client/Server (C/S) and Browser/Server (B/S). The C/S mode is a dominant one and is considered as an ideal one before. The B/S mode also has its advantage and is taking a big share with the internet spreading. There is also the mixed mode with two modes as in Fig.1.

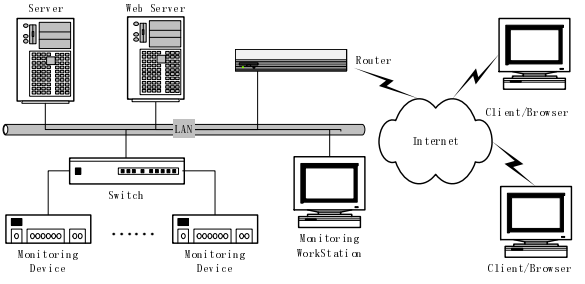


Fig. 1. Monitoring system with mixed mode

Fig.1 is a typical structure of monitoring system with mixed mode. The site monitoring devices acquire the data from monitored objects and communicate with the monitoring host. The host deals with the data, analyses the system working status and makes the decisions. It also has the functions to diagnose the system and provide the problem-solving. The server with the database stores the historical data and dynamic data acquired, and also answer the access from clients. Web server provides the function of user's accessing in remote by a browser [1][2].

2.2 Web Services

There are various software languages and software models used in developing the monitoring systems. Therefore, the control patterns and communication protocols of monitoring systems are not the same. It is difficult to colligate them and implement communicate each other. The common way to combine such inhomogeneous systems is to find a mechanism with standard norms, definitions, unify interfaces, and being accepted by all the OS platforms and software languages. The Web Service is regarded as to be such a mechanism considering of the popularity of the web applications in network times. Web services promise the creation of loosely coupled application to application communication over the WWW [3][4], and are designed according to the idea to take the internet as a virtual computing space. The interfaces of Web services are described with XML and communicated with clients through XML packages. The basic web service

stack includes following three layers: transmission, communication protocol and description of service contents as in Fig. 2.

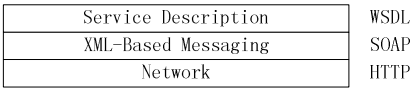


Fig. 2. Basic web service stack

Fig.3 is a common implementation of web services, where HTTP is used to realize the network transmission; it is very natural to take HTTP as a transmission protocol for it is de facto standard of internet. The SOAP is used for communication protocol. The SOAP has a standard encapsulation mechanism, and can carry out the information orientated to file and fulfill remote call by means of XML. WSDL describes the services provided by Web Services, which is an encapsulated package in SOAP format with information of functions, procedure call and the address.

The SOAP is an extensible protocol based on XML message, and has defined the message structure and message procedure protocol, which is independent to the transmission protocol. A SOAP package is consisted of a XML document, one message is a specified XML document. The elements of a SOAP message involves following elements:

<Envelop…/>, root element;
<Header…/>, optional header element;
<Body…/>, obligatory element, included calling and response information.

The SOAP provides the methods to encapsulate the information, and message package is transmitted by HTTP protocol, therefore if only both systems exchange data by means of the SOAP, it can be called each other. In this way, the inhomogeneous systems can be integrated together. The communication method between the systems is depicted as in Fig. 3.

In Fig. 3, one piece of request information from service requestor is sent out with a SOAP enveloped package via network protocol HTTP, and reached service provider. The service provider receives the package, analyzes the information and sets up the calling stacks, then activates the interface of application server. The application server finishes asked job and envelopes the results with SOAP package, then returns it.

The SOAP protocol is just a mechanism to realize the specified task and is needed to realize by program language when implementing web services. Common programming languages support this function, for example, Delphi provides abundant units including SOAP to support developing Web services.

2.3 Android OS

The Android open source OS, based on LINUX kernel (version 2.6), is primarily designed for mobile platform by Google, and widely accepted in the mobile and portable computing market, and for certain, it will be used in embedded real-time applications such as monitoring system as we did. The Android OS Software architecture is shown in Fig 4 [5][6].

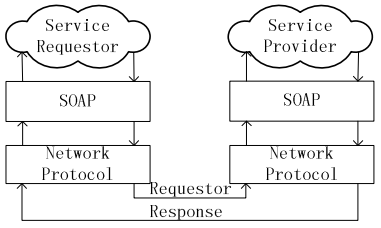


Fig. 3. Communication method of SOAP calling

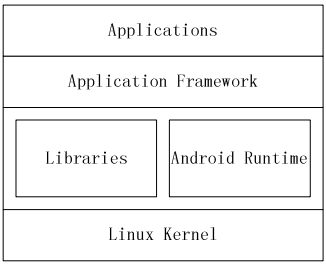


Fig. 4. Android OS architecture

3 System Architecture Based on Web Services

3.1 System Architecture

The system architecture presented in this paper based on the technologies aforementioned is in Fig. 5, which integrates available intelligent monitoring systems and realizes the communication and cooperation cross different districts, further more, provides the mobile computing ability to the system diagnosis and maintenance.

The architecture is composed of three parts: intelligent monitoring systems available, web services and mobile computing terminals.

The data acquired by monitoring devices at sites are sent to monitoring host (depicted as Server in the Fig. 5) via the network in real time. The monitoring devices have two ways to transmit data, one is by private cable network with TCP/IP, and the other one is through the telecom wireless network with Android OS. The server deals with the data from the devices, implements monitoring functions such as controlling and alarming etc, and restores the data to the dynamic database. The web server provides the web services to the users in the internet and also provides the access ability of browser users. The software of mobile devices based on Android OS accesses web services by HTTP and SOAP protocols, and also receive the information from local monitoring devices. Therefore, the manager with the mobile device can process the accident event in time.

3.2 Design of Web Services

In the system described in Fig.5, the monitoring systems have different modes and the business logics such as algorithms, control logics, diagnosis rules etc. The original system structures are not changed excepted that a web server is added if the system has not one. The software of web services is developed according to the unified specifications that describe the functions, such as:

- *Definition of the interface types*
- *Specifications of various units of the systems*
- *Measures of test and diagnosis for the interfaces*
- *Access control private*
- *Business logic encapsulation*
- *Normalized the diagnosis techniques*

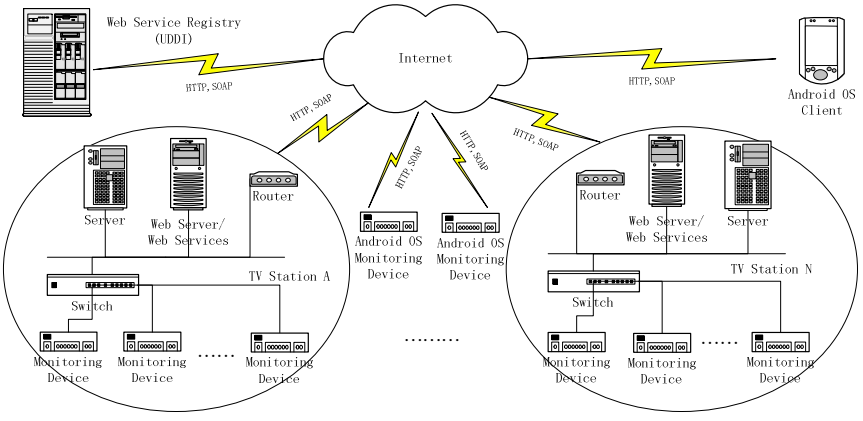


Fig. 5. System architecture of intelligent monitoring based on web services

Web services usually provide two services: one is for station's clients, such as:

- *Broadcasting plans of advertisement clients*
- *Program broadcasting schedules*
- *Programs searching and download service*

The other one is to monitor the quality of the programs broadcasted, diagnose fault equipment and control remotely, such as:

- *Dynamic indexes of the programs of different broadcasting channels*
- *Status of the key equipments of the broadcasting stations*
- *State predictions of the equipments*
- *Fault analysis and diagnosis*
- *Remote control privilege*

The functions of Web Service are very abundant and complicated, which need to be elaborate designed and planed. The services should be divided into several sectors, such as information provider, management data and maintenance guides, and oriented respectively to clients, managers and engineers. The information provider mainly faces to the publics who have business relationship with the TV stations or interesting in the information published by the stations. The management data are for administrations which are statistical indexes used to measure the service levels of the broadcasting systems. The maintenance guides are the intelligent parts of the systems that provide the solutions about system analysis, problem fixing, maintenance and fault diagnosis. The different users have different access rights.

The web service is key factor to integrate the monitoring systems of different districts together and to form a wholly more powerful monitoring system.

3.3 Web Service Registry

The publish mechanism is needed for web services so that the users can find required services, which is Universal Description, Description and Integration (UDDI). UDDI is a set of registry norms including web service publish. UDDI provides a mechanism

for clients to dynamically find other web services. Using a UDDI interface, businesses can dynamically connect to services provided by external business partners. A UDDI registry has two kinds of clients: businesses that want to publish a service (and its usage interfaces), and clients who want to obtain services of a certain kind and bind programmatically to them [4].

The interaction of a client with web services is described as Fig. 6 that shows the relationship among service requestor, service provider and web service registry as well as the processing procedure.

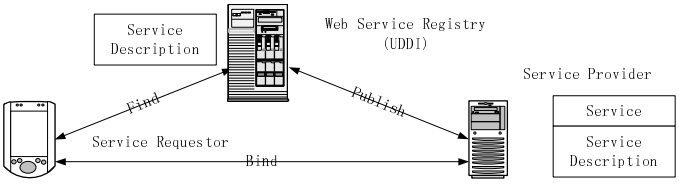


Fig. 6. Interaction of a client with Web Services

3.4 Mobile Computing

The main feature of the architecture presented in the paper is to integrate inhomogeneous monitoring systems together and realize the mobile computing that is supported by the mobile terminal and software. The monitoring device based on Android OS has been developed to support mobile computing. The mobile terminal can be developed based on Android OS or just adopted an intelligent mobile telephone with Android OS available from the market. If the mobile telephone is adopted as a mobile device, the only thing to do is to develop the software according to the special applications. The mobile computing provides ubiquitous computing ability such as information access, system management and equipment maintenance and diagnosis etc. For example, the engineers who are in charge of the broadcasting transmitters, the terminal should install the software to manage the transmitters, so that who can receive the a alarm message from the monitoring device of the transmitter, and deal with the alarm with proper action, say, to issue the control command remotely in time wherever he is; when there are some things he isn't sure, he can search the answers from web services and get the solutions on-line. Such a way for equipment maintaining isn't affected by the long distance between the equipment and the staff, and is very convenient to deal with the urgent events, and that can guarantee the system's reliability powerfully.

The device with Android OS has limited hardware deployment, it is not suitable to process a large volume data and carry out too complicated algorithm, so that in the system architecture of this paper, it is only used as client terminal as the web services or the monitoring device to support the mobile computing.

Google has provided the package--ksoap2-android for Web Service application development, which can be downloaded and installed in the Android OS [7]. The main consideration to adopt the Android OS in the our designing is for its openness.

The main steps to create a client application of the web service with ksoap2-android in the mobile devices are as following:

Step 1. Create an object *HttpTransportSE* that is used to call the function of web services;

Step 2. Create an object *SoapSerializationEnvelope* that is used to envelope the SOAP protocol;

Step 3. Create an object*SoapObject*;

Step 4. If there are parameters to be past, call the function of *addProperty*(*string name, Object value*) of *SoapObject*;

Step 5. Call the function *setOutputSoapObject*() of *SoapSerializationEnvelope* to envelope the message package;

Step 6. Call remote web services with *SoapSerializationEnvelope* as argument and *call*() as function;

Step 7. Access the attribute of the *bodyIn* of *SoapSerializationEnvelope*, which returns a *SoapObject* object represented the information from web services.

Step 8. Analyze the *SoapObject* and get the result from web services.

The feature of web services is whatever the business logic complicated is, it is transparent to a client, as only as the client sent out a request, the server will return a SOAP message, therefore it can be used to implemented powerful system functions in the mobile computing.

4 Questions to Be Solved

Security. Web services are an open service, they will encounter inevitably the attack from the internet, it may be dangerous to the systems although the security measures have been set up. How to prevent and predict the risks will be a key factor to successfully deploy the web services.

Self-Learning Power of the System. One of goals of our system is to integrate the distributed monitoring systems, and collect all intelligent diagnosis technologies and maintain experiments in the broadcasting industry, as well as assembly them on-line, so as to provide powerful fault diagnosis functions. But the theory and methods about fault diagnosis are constantly developing, therefore, the web services should have the ability of self-learning with the systems running, that is to say new knowledge could be added into the web services smoothly, which is a challenge for web service design.

Optimum Service Selection. It will be a crucial problem for users how to select a suitable service from a great deal web services, which will determine the efficiency of web services [8].

5 Conclusion and Further Works

We designed the system architecture of intelligent monitoring based on web services for broadcasting TV system, and developed a prototype system. The system can implement intelligent monitoring and integrate different system modes, realized the ubiquitous mobile computing for the fault diagnosis and system maintenance, and

also provides a plenty of services for clients. The architecture is suitable for other industrial applications too.

The next work need to do is to implement specialized the mobile device based on Android OS and perfect the software functions of web services to fully reach the target of system designing target.

Acknowledgements. This work is supported by "211" Project Fund (No.21103050104) and Beijing Cultural Innovation Fund (HG0842).

References

1. Zhou, C.: Automatic monitoring technology of broadcasting TV. China Broadcasting Press (July 2009) (Chinese)
2. Zhou, C., et al.: Design and Implement of Monitoring Controller of TV Transmitter based on ARM-Linux. In: The 2nd International Conference on Industrial and Information Systems, Dalian, China (July 2010)
3. Kreger, H.: IBM Software Group, Web Services Conceptual Architecture, WSCA 1.0 (May 2001)
4. A Web Services Primer By Venu Vasudevan (April 04, 2001), http://www.xml.com/pub/a/ws/2001/04/04/webservices/index.html
5. Google Android SDK, http://developer.android.com/sdk/index.html
6. Sbaa, A., El Bejjet, R., Medromi, H.: A System Architecture for Industrial Application based on the Android Mobile OS. International Journal of Research and Reviews in Computer Science (IJRRCS) 2(2), 232–327 (2011)
7. http://code.google.com/p/ksoap2-android/
8. Michael Maximilien, E., Singh, M.P.: Agent-based Architecture for Autonomic Web Service Selection, http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.86.4587

A New Nonlinear Integrable Couplings of GJ Equations Hierarchy and Its Hamiltonian Structure

Dandan Liu and Tie-cheng Xia*

Department of Mathematics,
Shanghai University, Shanghai, 200444, China
liudandan8489@yahoo.com.cn, xiatc@yahoo.com.cn

Abstract. In this paper, based on the rudimentary knowledge of the nonlinear integrable couplings, we establish a scheme for constructing nonlinear integrable Hamiltonian couplings of soliton hierarchy. Variational identities over the corresponding loop algebras are used to offer Hamiltonian structures for the resulting continuous couplings. As an application, we use this method to obtain a nonlinear integrable couplings and Hamiltonian structure of the GJ hierarchy by using Tu scheme and the quadratic-form identity.

Keywords: Zero curvature equations, Integrable couplings, The GJ hierarchy.

1 Introduction

Integrable couplings [1] are coupled systems of integrable equations, which has been introduced when we study of Virasoro symmetric algebras. It is an important topic to look for integrable couplings because integrable couplings have much richer mathematical structures and better physical meanings. In recent years, many methods of searching for integrable couplings have been developed [1 - 10], especially authors have also used the enlarged matrix spectral problem method[2] to end out integrable couplings. For example, they ever used the following the spectral matrix to obtain integrable couplings

$$\overline{U} = \begin{bmatrix} U(u) & U_a(v) \\ 0 & U(u) \end{bmatrix},$$

where the sub-spectral matrix U is associated with a given integrable equation $u_t = K(u)$. However, soon afterwards, researcher find out that obtained integrable couplings is a relatively simple. So in order to get better integrable couplings to the known integrable system, we need to introduce an enlarged relatively complex spectral matrix

$$\overline{U} = \begin{bmatrix} U(u) & U_a(v) \\ 0 & U(u)+U_a(v) \end{bmatrix}. \tag{1}$$

* Corresponding author.

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 539–548.
springerlink.com

Therefore, from zero curvature equation

$$\overline{U}_t - \overline{V}_x + \left[\overline{U}, \overline{V}\right] = 0, \tag{2}$$

Where

$$\overline{V} = \overline{V}\left(\overline{u}\right) = \begin{bmatrix} V(u) & V_a\left(\overline{u}\right) \\ 0 & V(u) + V_a\left(\overline{u}\right) \end{bmatrix}, \tag{3}$$

And $\overline{u}$ consist of u and v, we can give rise to

$$\begin{cases} U_t - V_x + [U, V] = 0. \\ U_{a,t} - V_{a,x} + [U, V_a] + [U_a, V] + [U_a, V_a] = 0. \end{cases} \tag{4}$$

This is an integrable couplings of Eq.(1), due to Eq.(2), and it is a nonlinear integrable coupling because the commutator $[U_a, V_a]$ can generate nonlinear terms.

Let us further take a solution $\overline{W}$ to the enlarged stationary zero curvature equation

$$\overline{W}_x = \left[\overline{U}, \overline{W}\right]. \tag{5}$$

Then, we use the quadratic-form identity or variational identity.[5,6]

$$\frac{\delta}{\delta \overline{u}} \int \left\langle \overline{W}, \overline{U}_\lambda \right\rangle dx = \lambda^{-\gamma} \frac{\partial}{\partial \lambda} \lambda^{\gamma} \left\langle \overline{W}, \overline{U}_{\overline{u}} \right\rangle, \tag{6}$$

Where γ is a constant, to search for the Hamiltonian structures for the integrable couplings [6]. In the variational identity(6), $\langle \cdots \rangle$ is non-degenerate, symmetric and ad-invariant bilinear form over the Lie algebra:

$$\overline{g} = \left\{ \begin{bmatrix} A & B \\ 0 & A+B \end{bmatrix} \middle| A, B \in g \right\}. \tag{7}$$

In what follows, we will make above idea to apply the GJ hierarchy.

2 The Integrable Couplings of the GJ Equations Hierarchy

2.1 GJ Hierarchy

For the GJ spectral problem[11]:

$$\begin{cases} \phi_x = U\phi = U(u,\lambda)\phi, \phi_t = V\phi, \lambda_t = 0 \\ U = \begin{pmatrix} -\lambda + s & q \\ r & \lambda - s \end{pmatrix}, u = \begin{pmatrix} q \\ r \\ s \end{pmatrix}. \end{cases} \tag{8}$$

Setting

$$V = \begin{pmatrix} a & b \\ c & -a \end{pmatrix} = \sum_{i\geq 0} V_i \lambda^{-i} = \sum_{i\geq 0} \begin{pmatrix} a_i & b_i \\ c_i & -a_i \end{pmatrix} \lambda^{-i}. \tag{9}$$

The stationary zero curvature equation $V_x = [U,V]$ yields that:

$$\begin{cases} a_{nx} = qc_n - rb_n \\ b_{nx} = -2b_{n+1} + 2sb_n - 2qa_n \\ c_{nx} = 2c_{n+1} + 2ra_n - 2sc_n \end{cases} \tag{10}$$

Choose the initial data:

$$a_0 = -1, b_0 = c_0 = 0, \tag{11}$$

Then we have

$$\begin{cases} a_1 = 0, b_1 = q, c_1 = r, a_2 = \frac{1}{2}qr, b_2 = qs - \frac{1}{2}q_x, c_2 = rs + \frac{1}{2}r_x, \\ a_3 = qrs + \frac{1}{4}(qr_x - q_x r), b_3 = -q_x s - \frac{1}{2}qs_x + \frac{1}{4}q_{xx} + qs^2 - \frac{1}{2}q^2 r, \\ c_3 = r_x s + \frac{1}{2}rs_x + \frac{1}{4}r_{xx} - \frac{1}{2}qr^2 + rs^2, \end{cases} \tag{12}$$

From the compatibility conditions of the following problems

$$\begin{cases} \phi_x = U\phi, \phi_t = V^{(n)}\phi, \\ V^{(n)} = (\lambda^n V)_+ + \Delta_{1n}, n \geq 0, \end{cases} \tag{13}$$

where $\Delta_{1n} = \begin{pmatrix} a_{n+1} & 0 \\ 0 & -a_{n+1} \end{pmatrix}$.

We can determine the GJ hierarchy of soliton equations[15]

$$u_{t_n} = \begin{pmatrix} q \\ r \\ s \end{pmatrix}_{t_n} = K_n = \begin{pmatrix} -2b_{n+1} + 2qa_{n+1} \\ 2c_{n+1} - 2ra_{n+1} \\ a_{n+1,x} \end{pmatrix} = J \begin{pmatrix} c_{n+1} \\ b_{n+1} \\ 2a_{n+1} \end{pmatrix} = JL^n \begin{pmatrix} c_1 \\ b_1 \\ 2a_1 \end{pmatrix}, n \geq 0. \quad (14)$$

With the Hamiltonian operator J , the hereditary recursion operator L and the Hamiltonian functions H_n, respectively, as follow:

$$J = \begin{pmatrix} 0 & -2 & q \\ 2 & 0 & -r \\ -q & r & \partial \end{pmatrix}, L = \begin{pmatrix} \frac{1}{2}\partial + s & 0 & -\frac{1}{2}r \\ 0 & -\frac{1}{2}\partial + s & -\frac{1}{2}q \\ \partial^{-1}q\partial + 2\partial^{-1}qs & \partial^{-1}r\partial - 2\partial^{-1}rs & 0 \end{pmatrix}, H_n = \int \frac{2a_{n+2}}{n+1}dx, n \geq 0. \quad (15)$$

2.2 Integrable Couplings

Let us now begin with an enlarged spectral matrix

$$\overline{U} = \begin{pmatrix} U & U_a \\ 0 & U + U_a \end{pmatrix}, U_a = \begin{pmatrix} p_1 & p_2 \\ p_3 & -p_1 \end{pmatrix}, \quad (16)$$

$$\overline{V} = \begin{pmatrix} V & V_a \\ 0 & V + V_a \end{pmatrix}, V_a = \begin{pmatrix} e & f \\ g & -e \end{pmatrix}, \quad (17)$$

With the help of the corresponding enlarged stationary zero curvature equation $\overline{V}_x = [\overline{U}, \overline{V}]$, we have:

$$V_{ax} = [U, V_a] + [U_a, V] + [U_a, V_a] \quad (18)$$

Which equivalently generates

Setting

$$V = \begin{pmatrix} e & f \\ g & -e \end{pmatrix} = \sum_{i\geq 0} V_i \lambda^{-i} = \sum_{i\geq 0} \begin{pmatrix} e_i & f_i \\ g_i & -e_i \end{pmatrix} \lambda^{-i}. \quad (19)$$

And then, the eq.(18) can be transformed into

$$\begin{cases} e_{nx} = qg_n - rf_n + p_2c_n - p_3b_n + p_2g_n - p_3f_n, \\ f_{nx} = -2f_{n+1} + 2sf_n - 2qe_n + 2p_1b_n - 2p_2a_n + 2p_1f_n - 2p_2e_n, \\ g_{nx} = 2g_{n+1} + 2re_n - 2sg_n + 2p_3a_n - 2p_1c_n + 2p_3e_n - 2p_1g_n, \end{cases} \quad (20)$$

We choose the initial data

$$e_0 = -1, f_0 = g_0 = 0, \tag{21}$$

Then we have

$$\begin{cases} e_1 = 0, f_1 = q + 2p_2, g_1 = r + 2p_3, e_2 = \frac{1}{2}qr + qp_3 + rp_2 + p_2p_3, \\ f_2 = -\frac{1}{2}q_x - p_{2,x} + qs + 2sp_2 + 2qp_1 + 2p_1p_2, g_2 = \frac{1}{2}r_x + p_{3,x} + rs + 2p_3s + 2rp_1 + 2p_1p_3. \\ e_3 = \frac{1}{4}(qr_x - q_xr) + \frac{1}{2}(p_{3,x}q - p_3q_x) + \frac{1}{2}(p_2r_x - p_{2,x}r) + \frac{1}{2}(p_2p_{3,x} - p_{2,x}p_3) + qrs + 2p_3qs \\ + 2p_1qr + 2p_1p_3q + 2p_2rs + 2p_2p_3s + 2p_1p_2r + 2p_1p_2p_3, \\ f_3 = \frac{1}{4}q_{xx} + \frac{1}{2}p_{2,xx} - q_xs - \frac{1}{2}qs_x - p_2s_x - 2p_{2,x}s - p_{1,x}q - p_{1,x}p_2 - 2p_1p_{2,x} + qs^2 + 2p_2s^2 \\ + 2p_1qs + 4p_1p_2s - \frac{1}{2}q^2r - p_3q^2 - 2p_2qr - 2p_2p_3q + 2p_1qs - 2p_1q_x + 2p_1^2q + 2p_1^2p_2 - p_2^2r - p_2^2p_3, \\ g_3 = \frac{1}{4}r_{xx} + \frac{1}{2}p_{3,xx} + r_xs + \frac{1}{2}rs_x + 2p_{3,x}s + p_3s_x + p_{1,x}r + p_{1,x}p_3 + 2p_1p_{3,x} - \frac{1}{2}qr^2 - 2p_3qr - p_2r^2 \\ - 2p_2p_3r + rs^2 + 2p_3s^2 + 4p_1rs + 2p_1p_3s - p_3^2q - p_2p_3^2 + 2p_1r_x + 2p_1p_3s + 2p_1^2r + 2p_1^2p_3. \end{cases}$$

Using the zero curvature equation,

$$\begin{cases} \overline{U}_{tn} - \overline{V}_x + [\overline{U}, \overline{V}] = 0, \\ U_{at_n} = V_{ax}^{(n)} - [U, V_a^{(n)}] - [U, \Delta_{2n}] - [U_a, V^{(n)}] - [U_a, V_a] - [U_a, \Delta_{2n}], \end{cases} \tag{22}$$

We can rewrite Eqs. (22) as

$$\begin{pmatrix} p_1 \\ p_2 \\ p_3 \end{pmatrix}_{t_n} = \begin{pmatrix} e_{n+1,x} \\ -2f_{n+1} + 2qe_{n+1} + 2p_2a_{n+1} + 2p_2e_{n+1} \\ 2g_{n+1} - 2re_{n+1} - 2p_3a_{n+1} - 2p_3e_{n+1} \end{pmatrix}, \tag{23}$$

Also, we have the following results:

$$\overline{u}_{t_n} = \begin{pmatrix} q \\ r \\ s \\ p_1 \\ p_2 \\ p_3 \end{pmatrix}_{t_n} = \begin{pmatrix} -2b_{n+1} + 2qa_{n+1} \\ 2c_{n+1} - 2ra_{n+1} \\ a_{n+1,x} \\ e_{n+1,x} \\ -2f_{n+1} + 2qe_{n+1} + 2p_1a_{n+1} + 2p_2e_{n+1} \\ 2g_{n+1} - 2re_{n+1} - 2p_3a_{n+1} - 2p_3e_{n+1} \end{pmatrix}, n \geq 0. \tag{24}$$

Obviously, when $p_1 = p_2 = p_3 = 0$ in eq.(24), the above results become eq.(14). So we can say eq.(24) is integrable couplings of the GJ hierarchy. By reducing eq.(24), we have

$$\begin{cases} q_{t2} = 2q_x s + qs_x - \frac{1}{2}q_{xx} + \frac{1}{2}q^2 r_x - \frac{1}{2}qq_x r - 2qs^2 + q^2 r + 2q^2 rs, \\ r_{t2} = 2r_x s + rs_x + \frac{1}{2}r_{xx} - \frac{1}{2}qrr_x + \frac{1}{2}q_x r^2 - qr^2 + 2rs^2 - 2qr^2 s, \\ s_{t2} = q_x rs + qr_x s + qrs_x + \frac{1}{4}(qr_{xx} - q_{xx} r), \\ p_{1,t2} = e_{3,x} \\ p_{2,t2} = -2f_3 + 2qe_3 + 2p_2 a_3 + 2p_2 e_3, \\ p_{3,t2} = 2g_3 - 2re_3 - 2p_3 a_3 - 3p_3 e_3. \end{cases} \tag{25}$$

where

$$a_3 = qrs + \frac{1}{4}(qr_x - q_x r), e_3 = \frac{1}{4}(qr_x - q_x r) + \frac{1}{2}(p_{3,x} q - p_3 q_x) + \frac{1}{2}(p_2 r_x - p_{2,x} r) + \frac{1}{2}(p_2 p_{3,x} - p_{2,x} p_3)$$
$$+ qrs + 2p_3 qs + 2p_1 qr + 2p_1 p_3 q + 2p_2 rs + 2p_2 p_3 s + 2p_1 p_2 r + 2p_1 p_2 p_3,$$

$$f_3 = \frac{1}{4}q_{xx} + \frac{1}{2}p_{2,xx} - q_x s - \frac{1}{2}qs_x - p_2 s_x - 2p_{2,x} s - p_{1,x} q - p_{1,x} p_2 - 2p_1 p_{2,x} + qs^2 + 2p_2 s^2 + 2p_1 qs$$
$$+ 4p_1 p_2 s - \frac{1}{2}q^2 r - p_3 q^2 - 2p_2 qr - 2p_2 p_3 q + 2p_1 qs - 2p_1 q_x + 2p_1^2 q + 2p_1^2 p_2 - p_2^2 r - p_2^2 p_3,$$

$$g_3 = \frac{1}{4}r_{xx} + \frac{1}{2}p_{3,xx} + r_x s + \frac{1}{2}rs_x + 2p_{3,x} s + p_3 s_x + p_{1,x} r + p_{1,x} p_3 + 2p_1 p_{3,x} - \frac{1}{2}qr^2 - 2p_3 qr - p_2 r^2$$
$$- 2p_2 p_3 r + rs^2 + 2p_3 s^2 + 4p_1 rs + 2p_1 p_3 s - p_3^2 q - p_2 p_3^2 + 2p_1 r_x + 2p_1 p_3 s + 2p_1^2 r + 2p_1^2 p_3,$$

So, we can say that the system in (24) with $n \geq 2$ provide a hierarchy of nonlinear integrable couplings for the GJ hierarchy of the soliton equation. Now, we proceed to search for the Hamiltonian structure of the equations hierarchy (24).

3 Hamiltonian Structures of the Integrable Couplings of the GJ Equations Hierarchy

To construct Hamiltonian structures of the integrable couplings obtained, we need to compute non-degenerate, symmetric and invariant bilinear forms on the following Lie algebra:

$$\overline{g} = \left\{ \begin{bmatrix} A & B \\ 0 & A+B \end{bmatrix} \middle| A, B \in gl(2) \right\}. \tag{26}$$

For computations convenience, we transform this Lie algebra $\overline{g}$ into a vector from through the mapping

$$\delta : \overline{g} \to R^6, A \to (a_1, a_2, \cdots, a_6)^T, A = \begin{bmatrix} a_1 & a_2 & a_4 & a_5 \\ a_3 & -a_1 & a_6 & -a_4 \\ 0 & 0 & a_1 + a_4 & a_2 + a_5 \\ 0 & 0 & a_3 + a_6 & -a_1 - a_4 \end{bmatrix} \in \overline{g} \tag{27}$$

The mapping δ induces a Lie algebraic structure on R^6, isomorphic to the matrix Lie algebra $\overline{g}$ above. It is easy to see that the corresponding commutator $[\cdots]$ on R^6 is given by

$$[a,b]^T = a^T R(b), a = (a_1, a_2, \cdots, a_6)^T, b = (b_1, b_2, \cdots, b_6)^T \in R^6, \tag{28}$$

Where

$$R(b) = \begin{pmatrix} 0 & 2b_2 & -2b_3 & 0 & 2b_5 & -2b_6 \\ b_3 & -2b_1 & 0 & b_6 & -2b_4 & 0 \\ -b_2 & 0 & 2b_1 & -b_5 & 0 & 2b_4 \\ 0 & 0 & 0 & 0 & 2b_2 + 2b_5 & -2b_3 - 2b_6 \\ 0 & 0 & 0 & b_3 + b_6 & -2b_1 - 2b_4 & 0 \\ 0 & 0 & 0 & -b_2 - b_5 & 0 & 2b_1 + 2b_4 \end{pmatrix}. \tag{29}$$

Define a bilinear form on R^6 as follows:

$$\langle a,b \rangle = a^T F b, \tag{30}$$

where F is a constant matrix, which is main idea by Zhang and Guo presented in 2005[5].

Then the symmetric property $\langle a,b \rangle = \langle b,a \rangle$ and the ad-invariance property under the Lie product

$$\langle a,[b,c] \rangle = \langle [a,b],c \rangle \tag{31}$$

requires that $F^T = F$ and

$$(R(b)F)^T = -R(b)F \qquad \forall b \in R^6. \tag{32}$$

So we can obtain

$$F = \begin{bmatrix} 2\eta_1 & 0 & 0 & 2\eta_2 & 0 & 0 \\ 0 & 0 & \eta_1 & 0 & 0 & \eta_2 \\ 0 & \eta_1 & 0 & 0 & \eta_1 & 0 \\ 2\eta_2 & 0 & 0 & 2\eta_2 & 0 & 0 \\ 0 & 0 & \eta_2 & 0 & 0 & \eta_2 \\ 0 & \eta_2 & 0 & 0 & \eta_2 & 0 \end{bmatrix} \tag{33}$$

where η_1 and η_2 are arbitrary constants.

Therefore, a bilinear form on the underlying Lie algebra $\overline{g}$ is defined by

$$\begin{aligned}\left\langle a,b\right\rangle_{\overline{g}} &= (2b_1\eta_1 + 2b_4\eta_2)a_1 + (b_3\eta_1 + b_6\eta_2)a_2 + (b_2\eta_1 + b_5\eta_2)a_3 \\ &+ (2b_1\eta_2 + 2b_4\eta_2)a_4 + (b_3\eta_2 + b_6\eta_2)a_5 + (b_2\eta_2 + b_5\eta_2)a_6.\end{aligned} \tag{34}$$

where

$$A = \begin{bmatrix} a_1 & a_2 & a_4 & a_5 \\ a_3 & -a_1 & a_6 & -a_4 \\ 0 & 0 & a_1 + a_4 & a_2 + a_5 \\ 0 & 0 & a_3 + a_6 & -a_1 - a_4 \end{bmatrix}, B = \begin{bmatrix} b_1 & b_2 & b_4 & b_5 \\ b_3 & -b_1 & b_6 & -b_4 \\ 0 & 0 & b_1 + b_4 & b_2 + b_5 \\ 0 & 0 & b_3 + b_6 & -b_1 - b_4 \end{bmatrix} \tag{35}$$

It is non-degenerate if and only if

$$(\eta_1 - \eta_2)\eta_2 \neq 0. \tag{36}$$

Based on (37) and eqs.(16) and (17), we can easily compute that

$$\begin{aligned}&\left\langle \overline{V}, \overline{U}_{\lambda} \right\rangle = -2\eta_1 a - 2\eta_2 e, \\ &\left\langle \overline{V}, \overline{U}_{\bar{u}} \right\rangle = (\eta_1 c + \eta_2 g, \eta_1 b + \eta_2 f, 2\eta_1 a + 2\eta_2 e, \eta_2 (c + g), \eta_2 (b + f))^T, \\ &\gamma = -\frac{\lambda}{2}\frac{d}{d\lambda}\ln\left|\left\langle \overline{V}, \overline{V} \right\rangle\right| = 0\end{aligned}$$

By applying the operator Γ_{n+1} to both sides of variational identity (6) we deduce that

$$\begin{aligned}&\frac{\delta}{\delta \bar{u}}\int \frac{2\eta_1 a_{n+1} + 2\eta_2 e_{n+1}}{n} dx = \\ &(\eta_1 c_n + \eta_2 g_n, \eta_1 b_n + \eta_2 f_n, 2\eta_1 a_n + 2\eta_2 e_n, 2\eta_2 a_n + 2\eta_2 e_n, \eta_2 (c_n + g_n), \eta_2 (b_n + f_n))^T,\end{aligned} \tag{37}$$

So we obtain that equation hierarchy (24) possess the Hamiltonian structures:

$$\overline{U}_{tn} = \overline{K}_n(\bar{u}) = \overline{J}\frac{\delta H_n}{\delta \bar{u}}, n \geq 1. \tag{38}$$

where the Hamiltonian operator and the Hamiltonian functionals are given by

$$\bar{J} = \frac{1}{\eta_1 - \eta_2}\begin{bmatrix} 1 & 0 & 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 & \frac{\eta_1}{\eta_2} & 0 \\ 0 & 1 & 0 & 0 & -\frac{\eta_1}{\eta_2} & 0 \\ 0 & 0 & -1 & \frac{\eta_1}{\eta_2} & 0 & 0 \end{bmatrix} \tag{39}$$

and

$$H_n = \int \frac{2\eta_1 a_{n+2} + 2\eta_2 e_{n+2}}{n+1} dx, n \geq 0. \tag{40}$$

With the help of (22), we can see a recursion relation

$$\bar{L}\frac{\delta H_n}{\delta \bar{u}} = \frac{\delta H_{n+1}}{\delta \bar{u}}, n \geq 0. \tag{41}$$

with

$$\bar{L} = \begin{bmatrix} L & 0 \\ L_a & L + L_a \end{bmatrix}, \tag{42}$$

where

$$L_a = \begin{bmatrix} p_1 & 0 & -\frac{p_3}{2} \\ 0 & p_1 & -\frac{p_2}{2} \\ \partial^{-1} p_3 \partial + 2\partial^{-1} p_1 p_2 + 2\partial^{-1} p_1 q + 2\partial^{-1} p_2 s & \partial^{-1} p_3 \partial - 2\partial^{-1} p_1 p_3 - 2\partial^{-1} p_1 r - 2\partial^{-1} p_3 s & 0 \end{bmatrix}$$

Up to now, we have already obtained Hamiltonian structure eq.(40) of integrable couplings of the GJ hierarchy.

Acknowledgments. The Project supported by the Natural Science Foundation of China (Grant No.61072147, 11071159), the Natural Science Foundation of Shanghai and Zhejiang Province(Grant No.09ZR1410800 and and Y6100791) , the Shanghai Leading Academic Discipline Project(NO.J50101) and the National Key Basic Research Project of China(KLMM0806).

References

1. Ma, W.X., Fuchssteiner, B.: Integrable theory of the perturbation equationa. Chaos Solitons Fract. 7, 1227–1250 (1996)
2. Xia, T.C., Zhang, G.L., Fan, E.G.: New Integrable Couplings of Generalized Kaup Newell Hierarchy and Its Hamiltonian structures,Commun. Theor. Phys. Commun. Theor. Phys. 56, 1–4 (2011)

3. Zhang, Y.F.: A hierarchy of integrable Hamiltonian equation. Acta Mathematica Applicata Sin. 23(2), 181–187 (2002)
4. Ma, W.X., Xu, X.X., Zhang, Y.F.: Semi-direct sums of Lie algebras and discrete integrable couplings. Phys. Lett. A 351, 125–130 (2006)
5. Guo, F.K., Zhang, Y.F.: J. Phys. A: Math. Gen. 38, 8537 (2005)
6. Ma, W.X.: A Hamiltonian structure associated with a matrix spectral problem of arbitrary-order. Phys. Lett. A 367, 473–477 (2007)
7. Guo, F.K., Zhang, Y.F.: The integrable coupling of the AKNS hierarchy and its Hamiltonian structure. Chaos, Solitons and Fractals, 321898–321902 (2007)
8. Ma, W.X., Gao, L.: Coupling integrable couplings. Modern Phys. Lett. B 23, 1847–1860 (2009)
9. Ma, W.X., Zhu, Z.N.: Constructing nonlinear discrete integrable Hamiltonian couplings. Computers and Mathematics with Application 60, 2601–2608 (2010)
10. Ma, W.X.: Nonlinear conitnuous integrable Hamiltonian couplings. Applied Mathematics and Computation 217, 7238–7244 (2011)
11. Tu, G.Z.: The trace identity, a powerful tool for constructing the Hamiltonian structure of integrable systems. J. Math. Phys. 30, 330–338 (1989)

Applied Research on Tower Crane Safety Supervising System Based on Internet of Things

Leqiang Bai[1], Yuejiao Sun[1], and Xin Guo[2]

[1] Information & Control Engineering Faculty
Shenyang Jianzhu University
Shenyang Liaoning China
[2] International Business Development Division
Neusoft Corporation
Shenyang Liaoning China
baileqiang@sjzu.edu.cn, jiayou.syj@163.com, guoxin@neusoft.com

Abstract. This paper analyses current situation of tower crane safety supervising system and puts forward a scheme of a new intelligent system: with the network architecture of the Internet of Things as the foundation, using Wireless Sensor Networks technology, combining CAN bus technology with Zigbee technology, which provides industrial control with a simple, real-time, low cost, high reliability and efficient methods. In this paper, the structure and function of the intelligent system are introduced in detail.

Keywords: Internet of Things, Tower Crane, Wireless Sensor Networks, Zigbee, CAN.

1 Introduction

Internet of Things (IoT) is based on radio frequency identification, infrared sensors, global positioning systems, laser scanners and other information sensing device, according to the agreement agreed to anything connected to the Internet, the exchange of information and communication to achieve intelligent identifying, locating, tracking, monitoring and managing a network [1][2]. From the technical perspective, the IoT is the intelligent network to achieve automatic interchange and process of information between objects and objects, people and objects, by intelligent sensor devices, through transport network, arriving at specified information processing center. In 2010, Chinese Premier Wen Jiabao in his government report, for the first time mentioned "speed up the development and application of Internet of Things", which led to the upsurge of the research of IoT in China [3].

Furthermore, modern construction increasingly rely on tower cranes, and tower cranes are operated crossover frequently as the pursuit of construction efficiency. It is impossible to avoid malignant accidents of tower cranes completely. Tower crane accident happens in the process of installation, disassemble and application. Thus, it is of vital significance to design tower crane safety supervising system based on the Internet of Things by using intelligent information technology, so as to realize

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 549–556.
springerlink.com

real-time monitoring of the tower crane operation and then strengthen supervising the whole process of the production safety effectively.

2 Analyse Current Situation of Tower Crane Safety Supervising System

Early tower crane safety supervising products mainly contain lifting weight overload limitator, lifting torque limitator, range limiting stopper, wind speed alarm device, start alarm, hoisting height limiting stopper, etc. Some tower crane protection device collects some data, such as the machine vehicle height, range, loadage, intersection angle to judge whether they are beyond limiting range. Tower crane anti-collision system has been designed by some research institutions, which can prevent collisions between tower cranes. But the function of these systems is still need to be improved, there are many risky factors hide in application of tower crane. In addition, potential safety hazard still exists if the operator doesn't obey the operating norms.

Currently, there have been many researches in the field of tower crane safety. But, some shortcomings are still exist, such as a lack of integration degree, too much equipments leading to complex installation; using PLC technology, processing speed is not high enough; poor man-machine interaction; small monitoring capacity; high cost and so on. Some systems have complex wiring without using the bus technology; inconvenience installation and removal; being hard to maintain and extend their functions; when need to increase monitoring parameters, they can only be redesigned; not facilitate standardization, standardized. Therefore, it is necessary to design a scheme of tower crane safety supervising system which based on the wireless network communication, having data processing ability, being able to adapt to the task of measurement and control. This paper aims at studying the new intelligent system based on IoT by using Wireless Sensor Networks (WSN) technology.

3 The Architecture of Intelligent System

Internet of Things makes all kinds of perception technology, modern network technology, artificial intelligence and automation technology polymerization and integrated application. The intelligent system based on IoT means machine to machine (M2M) through the various of the wireline/wireless, long/short communication network to provide management and service with safety controllable on-line monitoring, alarm linkage, scheduling command, plan management, remote control, security guard and so on, in order to realize the "manage, control, operate" integration of tower crane efficiently and safely.

The architecture intelligence system based on IoT is shown in Figure 1. Currently, the industry recognized that IoT architecture consists of three layers: from the bottom up, the perception layer, the network layer and the application layer. In accordance with concrete functions: the bottom layer is the sensing equipment for information acquisition; the middle layer is the network for data transmission, while the top layer is designed for applications and middleware [4].

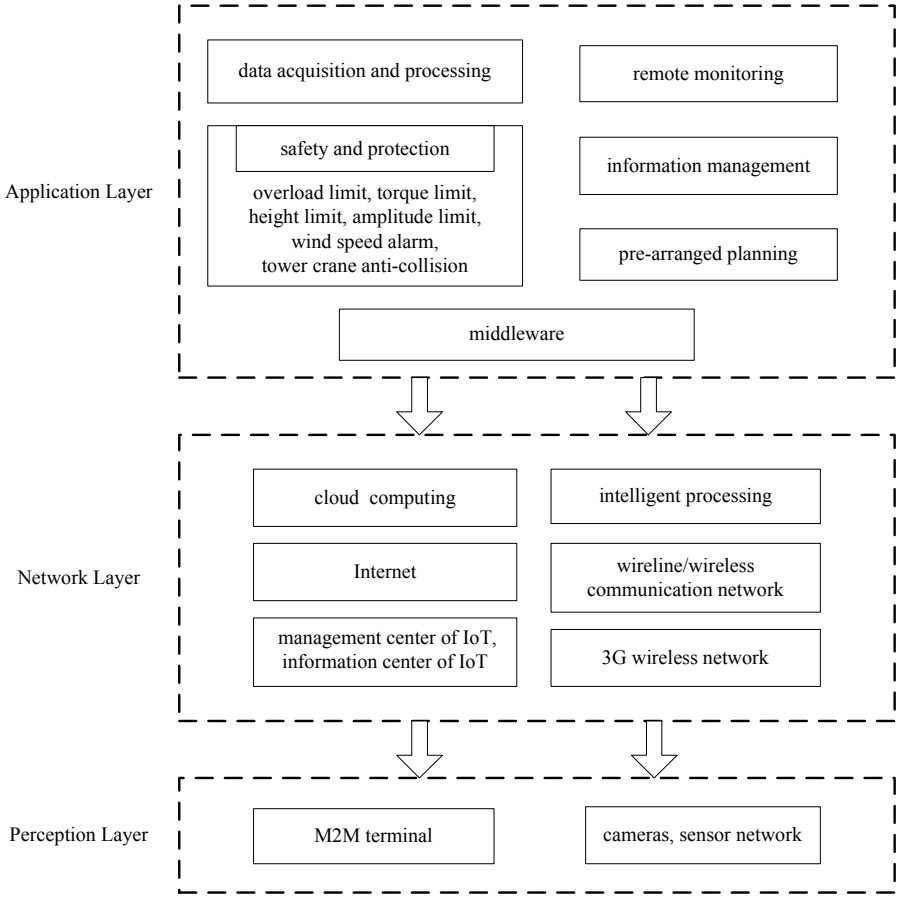


Fig. 1. The Architecture of Intelligent System

3.1 Perception Layer

The perception layer is the core competence in realizing the comprehensive perception of IoT. This layer is a significant part in IoT of key technology, standardization and industrialization. The perception layer consists of the sensor information collection and video acquisition equipment, including perception terminals as follows: altimetric sensor, displacement transducer, angular transducer, air velocity transducer, camera, etc. The main function of the perception layer is collecting information, it reflects the characteristics of recognition and network communication of IoT, that is, the objects brought into IoT must possess the ability of automatic identification and M2M.

3.2 Network Layer

The network layer consists of various private network, Internet, wireline/wireless communication network, network management system and cloud computing platform. It is responsible for transmitting and intelligent processing the information get from the perception layer. It makes the objects interconnection which needs networking. The information collected regularly from the sensors on the IoT transmitted through the network, while transmission, it must adapt to various heterogeneous network and agreement to ensure the correctness and timeliness of the data. In the network layer the broad coverage of mobile communication network is the infrastructure of IoT. This layer is the most standardized the most industrialized, and the most mature part in IoT.

Cloud computing business gradually becomes one of the most important applications of the Internet of things. Cloud computing data center is used to enhance its IT infrastructure operational efficiency. Through the cloud center to integrate scattered in various branches of the small data centers. By building a cloud computing-based data center, IT infrastructure of these enterprises will become more flexible and elastic.

3.3 Application Layer

The application layer is the interface between IoT and user (including people, organization and other systems), it combined with industrial demands to realize the intelligent application of IoT. Middleware is an independent system software or service program, distributed applications share resources in different technology with the help of this software. Applications can work in multiple platform or OS environment through the middleware, to realize monitoring and alarming, information management and other functions. The application layer is aimed at providing various applications by intelligent computing and intelligent processing, combines IoT technology with the requirements of industry [5].

4 Model and Functions of the Intelligent System

The sensor technology and intelligent embedded technology have been widely used in the IoT. The IoT have intelligent characteristics, that is, network system should have automation, self feedback and intelligent control characteristics. The main steps developing the IoT mainly takes as follows: firstly, mark properties of object, the properties including the static ones and dynamic ones, they are real-time detected by sensors; secondly, identification device reads off the properties of object, and then transform data format of the information to adapt to network transmission; thirdly, transmit the information of the objects through the network to the information processing center, where related computing of object communication is completed. The model of the intelligent system is shown in Figure 2, which has the function of data acquisition and processing, remote monitoring and information management.

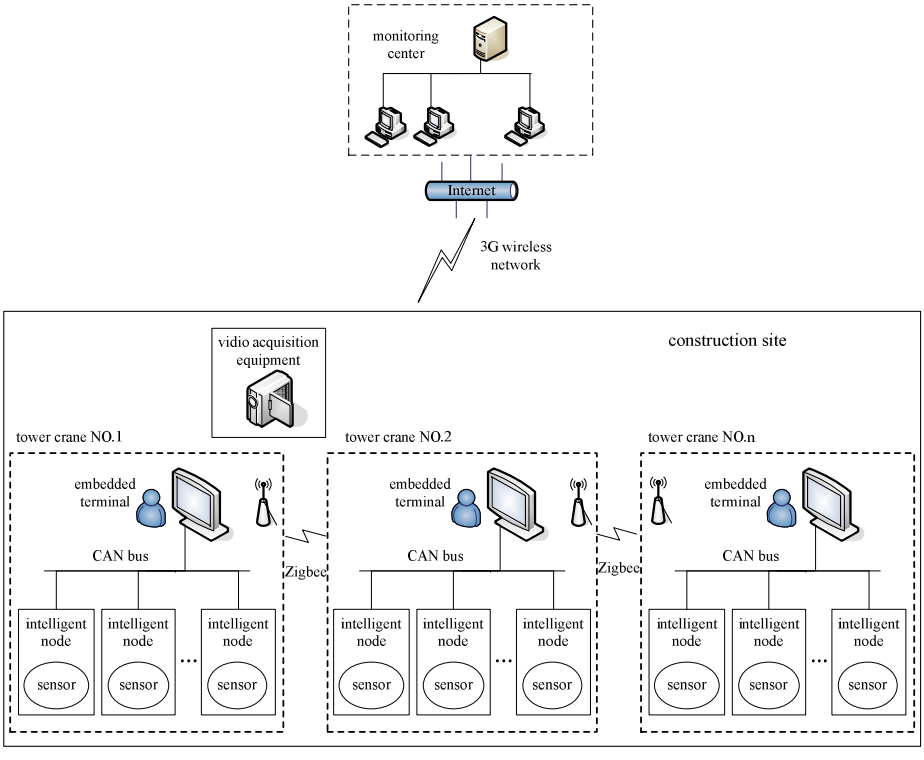


Fig. 2. Model of the Intelligent System

4.1 Data Acquisition and Processing

At the construction site, the whole system is based on Wireless Sensor Networks which is responsible for detecting, sending and receiving signals. WSN synthesizes sensor technology, modern network and wireless communication technology, embedded computing technology and distributed information processing technology. It is a novel technology about acquiring and processing information.

The node of sensor network in the system consists of data acquisition unit, process unit, and data transfer unit. The types of the sensors are decided by the form of monitored physical signals, while the sensor technology is designed to identify the dynamic properties [6]. ARM9 is used as the embedded CPU processors. The system chooses the miniaturization operating system, embedded Linux operating system, to carry out complex task scheduling and management.

Data Acquisition Unit. A large number of various sensors are set on the tower crane, and each one is an information source, different kinds of sensors catch information of different content and different format. The information include static data of tower crane: length of loading jib, length of counter-jib, height of fixed tower, height of cat head; dynamic data of tower crane: turning angle, machine vehicle radius, sling tightly length, loadage, dip angle of tower body; environment monitoring data: wind

speed at the construction site (set the anemograph at the top of the highest tower crane), etc. The sensor acquires real-time data, according to certain frequency, and then updates the data constantly.

Data Processing Unit. ARM module is used for embedded terminal set in the crane cab as client, because ARM control chip has the incomparable speed compared with single-chip microcomputer. ARM module is responsible for managing the information acquired from sensors. It has the following functions: anti-collision alarming; estimating whether the environment of the construction site (such as wind speed) satisfy the operation conditions; overload protection; collapse protection and so on.

Data Transfer Unit. The real-time information is transmitted accurately through wireline/wireless communication network so as to connect to Internet. In this paper, data transfer unit is put forward, combining CAN bus technology and Zigbee technology.

CAN Bus Technology. CAN bus technology is used in this system to realize control and data collection. Compares with the general communication bus, the data communication of CAN bus has the prominent reliability, timeliness and the flexibility, and the strongpoint of Field bus [7]. By using the bus technology, some drawbacks can be overcome, such as complex wiring; inconvenience installation and removal; expansibility sent. When need to increase monitoring parameters, only increase the node parallel in the bus rather than design the system again, it facilitates standardization, standardized. The actual communication media can be twisted-pair wiring, when monitored objects distributed in the concentrated area, cost will be reduced obviously, and capacity of resisting disturbance will be enhanced.

Zigbee Technology. Real-time and short distance wireless communication network are needed in industrial control environment. In construction site, by using Zigbee network in charge of communication between tower cranes, it has the following advantages: wireless network topology is more flexible for industrial network; no wiring, avoiding troubles during the construction; extensive cover range, from dozens of meters to hundreds of meters can be satisfied. Zigbee network, as one of the technology of short-range wireless network, is established mainly for data transmission in automation control. Zigbee is the primary criteria of sensor network and widely applied in all kinds of automatic control field, specially directing at communication in low speed and short distance, possessing the advantages of low power consumption, low time delay, low cost, high reliability, high security, large capacity, good compatibility, etc. While the IEEE 802.15.4/Zigbee protocol stack is being considered as a promising technology for low-cost low-power Wireless Sensor Networks, several issues in their specifications are still open.

The whole structure design. The main responsibility of CAN node is to acquire the state of controlled object and to send the status information in CAN bus, While, execute the command from ARM module transmitted through CAN bus. ARM module transmits status information to the related tower crane through Zigbee network. This system possess the advantages of high reliability, simple structure,

strong anti-interference capability, low power consumption, low cost, small delay and the advantages of wireless network, it is a very economic and practical measurement and control system.

4.2 Remote Monitoring

At the construction site, both embedded terminal and camera are internally installed with 3G mobile communication module. There are two ways to join the transmitting link to the monitoring center, one is to join the terminal directly by 3G wireless network, and the other is to send the data through the 3G wireless link and the Internet. The latter one has been used here. Tower crane operation status in different area is monitored by the monitoring center. In addition, once illegal operations happened, the monitoring center will brake and alarm immediately. If there is a safety accident, intelligent system can get the feedback information immediately and find the reasons. This can be an effective way to prevent illegal operations, and analysis the cause of the accident.

Wireless communication, together with video compression technology brings more convenience to remote surveillance. The mobile telecom operators have built a wide-spread data communication platform. So,using the public wireless platform is the first choice to transmit the data. Nowadays, the 3G wireless network has been set throughout the country from city to city, providing faster speed and wider bandwidth than that of the 2.5G and 2.75G, running a speed of at least 144 kbps and a bandwidth of 2Mbps[8], making it possible to set a high-quality remote wireless surveillance system.

The data acquisition equipment captures the video data from the camera set in construction site. After being compressed and coded, the video data is sent to the 3G wireless network by the 3G transmission module, and then gets to the monitoring center through the wireless transmission link. The original video data can be monitored in the monitoring center after decoded. Manager use video to monitor the information of construction site through 3G wireless network at anytime.

The design of the monitoring center is based on C/S mode. The sever gets the video coding flow decoded as soon as the data reaches the monitoring center. The client visits the server by the LAN, monitoring the image.

4.3 Information Management

A large number of various sensor are set at the construction site, they acquire real-time dates, and then update the data constantly according to certain frequency. Thus, integrated business management platform is used to collect and classify the sensor information, and then transmit them with directivity. Cloud computing platform and intelligent network of the IoT can make decisions based on the data acquired from sensor network and control the behavior of the object. Management system assigned unique ID number to each tower crane, relevant information is gathered and recorded. Management system deals with the information, at the same time data are stored and transmitted to the special management center.

5 Conclusion

In this paper, the scheme of tower crane safety supervising system based on IoT is introduced in detail, which can supervise the whole process of the tower production safety effectively. The scheme has the following essential functions: the first is to acquire, process and transmit data; the second is safety monitoring and alerting; the third is to manage the enormous information. This system fuses informatization and industrialization; it is economical and easy to apply in practice and has good practical value and application prospect.

Acknowledgments. This research was financially supported by the Open Research Fund from the State Key Laboratory of Rolling and Automation, Northeastern University, Grant No.: 2011002.

References

1. Wang, Y.: Reviews of the Internet of Things development. Science and Technology Information 03, 7–37 (2010)
2. Gan, Z.: Research on origin and development background of the Internet of Things. Modern Economic Information 01, 157–158 (2010)
3. Wei, Z.: Research on Architecture of the Internet of Things. In: ICCDA, 3rd edn., Xi'an, vol. 9, pp. 400–402 (2011)
4. Shen, S.-B., Fan, Q.-L., Zong, P., Mao, Y.-Q., Huang, W.: Study on the Architecture and Associated Technologies for Internet of Things. Journal of Nanjing University of Posts and Telecommunications. Natural Science, 1–11 (2009)
5. Guo, S.: Primary Exploration of Internet of Things. In: 3rd ICCDA, Xi'an, vol. 3, pp. 555–558 (2011)
6. Zhang, F.-S.: Internet of Things: open a new life of intelligent era, pp. 175–184. Shan Xi People's Publishing House (2010)
7. Wang, Z.-X., Cheng, W., Zhang, X.-B., Yu, L.-J.: The Research of In-Car Audio Entertainment Node Based on CAN. In: 3rd ICCDA, Xi'an, vol. 6, pp. 434–437 (2011)
8. Bai, L., Zhou, C., Hu, B., Shen, L.: A Research of the Wireless Video Surveillance of the Bus Based on 3G and H.264 Standard. Road Transfer Technology 4, 1 (2009)

An Enhanced Instruction Tracer for Malware Analysis

Zheyuan Liu

Control and Network Institute,
Northwestern Polytechnical University,
127 Youyi Xilu, Xi'an, China 710072
liuzheyuan@mail.nwpu.edu.cn

Abstract. Modern malicious applications use advanced anti-debugger, anti-virtualization, and code packing techniques to obfuscate the malware's true activities and divert security analysts. Malware analysts currently do not have a simple method for tracing malicious code activity at the instruction-level in a highly undetectable environment. This paper presents an enhanced instruction tracer as an extension to the Xen Ether virtualization framework to aid in the task of malicious software analysis. It places a malware binary into a virtualized environment and records the contents of all processor general register values that occur during its execution. Evaluation shows its new level of introspection for advanced malware that was not available with the previous state-of-the-art analysis tools.

Keywords: Virtualization, malware analysis, instruction trace.

1 Introduction

Today's malware can spread globally to thousands of computers in only a matter of days and performs much more sophisticated and damaging attacks to systems [7]. Defending against malicious computer attacks has become a tremendous and critical industry and research market and the need for comprehensive analysis and rapid response grows every year.

In order to protect their anonymity from security analysts and increase the difficulty of defense, malware authors generally obfuscate their code and communication channels. Common methods for hiding data are for the malicious code to detect if analysis tools are being used and correspondingly change its behavior. Analysts commonly run malware inside of a virtual machine (VM) in order to sandbox the code from infecting the host machine and to enable the ability to rewind the state of the system. Malware can detect that it is running inside a virtual machine by checking for virtualized hardware, differences in certain instruction return values, and backdoor I/O ports used by the hypervisor to communicate to the VM [9]. As virtualization is increasingly applied today to consolidate physical server hardware into virtual hardware, not performing malicious activity when inside of a VM would eliminate a large percentage of potential victims.

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 557–564.
springerlink.com

The technique of recording the set of all instructions executed during a program execution presented in this article is closer to the early HP 2100A [10] tracer back in the early 1970's than modern tracing applications. It runs on the Xen privileged dom0 and communicates solely with the hypervisor. This similarity allows our tool to monitor and trace any operation of a guest operating system including the operating system's execution itself. Moreover, as each instruction is recorded, the values of currently accessed variables are also recorded in its instruction trace log.

In this paper, we present an enhanced instruction tracer as an extension to the existing Xen Ether malware analyzer. The Ether additions allow for an instruction tracer running outside of a virtual machine to monitor and transparently record the variable values of a program while it executes. We believe that this new tool will contribute additional knowledge to malware analysts and provide a starting point for subsequent research in malware behavior analysis. The rest of the paper is organized as follows. Section 2 discusses related work. In Section 3 we present the design decisions and implementation details of our tool. Experimental results conducted on a physical machine are shown in Section 4. Section 5 concludes the paper.

2 Related Work

The work completed in this paper integrates new functionality into previous research done on Xen virtualization framework [5] and the Ether analysis framework [6]. The Xen Ether framework runs solely on the Xen hypervisor and dom0 without requiring any modification to the virtualized guest operating system in domU. The separation of the analysis framework and the guest operating system is realized by extensively utilizing the Intel VT virtualization extensions. Ether requires the use of Xen 3.0 and hardware-assisted virtualization (HVM) guests which allows for the highest level of transparent malware analysis without the guest being able to detect analysis. It implements instruction tracing of a guest operating system by setting the trap flag after each instruction which causes a debug exception and a VM-exit event that the Xen hypervisor handles [3]. Then it allows our tool to perform its analysis and instruction trace logging in between each guest instruction.

A similar project called Rotalume [8] uses protected instruction tracing as the foundation for malware analysis. Our work shares some of the same ideas about process introspection from an overlaid system, but focuses on the angle of internal algorithm analysis. While Rotalume focuses more on the automation of determining emulator protection semantics, our tool was developed to provide run-time variable information in a static analysis environment with a guarantee of transparency. Future work combining the variable identification techniques of the two in a second-pass analysis could provide more easily understood information about what values were stored in program variables at run-time.

3 Methodology

3.1 Intended Malware Analysis Workflow

Our tool was designed for use by a malware analyst who wants to gain insight into the inner-working of advanced malicious software as it would behave on a physical system. The proposed solution utilizes the common software package Xen for virtualization. It provides a simple workflow that requires some major tasks to complete. Firstly we take a binary file (that may or may not be confirmed to be malicious) and execute the file while under the environment of our tool. As the binary is executed, a log file is produced which contains all instructions that were executed, a list of all register values referenced in the instructions, and a list of all possible string array values referenced. Generally after the above operations malware analysts would probably transfer the specialized instruction trace file to a workstation, which will then load an unpacked version of the original binary file.

3.2 Xen Ether Memory Manipulation Analysis

The essential of our tool consists of modifications to the Xen Ether malware analysis framework in order to record additional data about the malicious program's state during execution. Specifically, the plug-in provides logged access to the run-time values of all referenced registers and uses a simple heuristic to also record all potential string array references by pointers. The following subsection provides an insight into how the virtual guests, Xen and Ether communicate and code flow between them during development of our tool.

Retrieving Instruction Trace from Malware Sample. The goal of the Xen portion of our tool is to single-step a single virtualized guest program and to record the register values during each instruction. In order to enable single-stepping of a processor, the x86 architecture provides a single bit Trap Flag (TF) that can be set in the FLAGS register. The FLAGS register is a status register on x86 processors that contains various bit-flags that control or describe the state of the processor. If the Trap Flag is set to true then the processor will make a call to interrupt 1 (INT 1) after each instruction is executed. This interrupt is typically used by kernel debuggers that would set this flag on the physical processor when running in a bare-metal environment. In the Xen virtualization environment, however, the hypervisor creates a Virtual Machine Control Structure (VMCS) for each virtual processor of each virtual machine. The VMCS is a structure defined in the Intel VT-x documentation that the hardware virtualization extensions use to store the state of each virtual processor. In addition to virtual registers and flags that the hardware processor has, the VMCS contains VM-specific control fields to set which operations should trap into the hypervisor, where the processor should start executing during a VM-exit exception, and metadata about the

cause of a VM-exit [3]. The Xen hypervisor can trap on specific interrupts and has been modified by Ether to allow analysis code to be run while the virtual processor is paused at each instruction.

The process that Ether uses to enable single-stepping of a virtual guest is presented on the leftmost of Fig.1. The code flow originates in the original Xen hypervisor, specifically the call to *vmx_properly_set_trap_flag*() which is used by Xen to properly handle returning execution to the guest during a VM-enter. If single-stepping was selected during the Ether configuration, then the original Xen code diverges into a section which sets the Trap Flag for the virtual processor and also sets the VMCS exception bitmap such that the Xen hypervisor will trap on the next exception. This causes the guest to return and execute a single instruction which triggers an INT 1 exception due to the Trap Flag being set. The Xen hypervisor then catches this exception and sends code execution to *vmx_handle_debug_exception*(). Still inside the Xen hypervisor (code modified by Ether), the code receives the debug exception and if Ether has enabled single-stepping then it makes a call to the next Xen Ether hypervisor function *ether_handle_instruction*(). It is in this instruction handler that the Xen hypervisor finally gives control to the Ether dom0 code by sending an ETHER_NOTIFY_INSTRUCTION message to it via a shared memory page. In the Ether main loop (and code section where modifications by our tool were made), the program awaits messages sent from the hypervisor - when an ETHER_NOTIFY_INSTRUCTION message arrives, it calls a final function named *disasm_instruction*() which handles the disassembly of a virtual guest's current instruction.

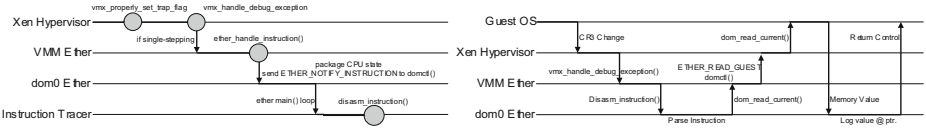


Fig. 1. Control and data flow during single-stepping and a tracer call

The previous actions were required in order to force the virtual guest operating system to essentially pause after each instruction it executes, allowing our code to execute and analyze the virtual guest's state while it is paused. The rightmost of Fig.1 describes the data and code flow used to parse individual machine instructions and capture the register and memory values. First, an additional step is required before *vmx_handle_debug_exception*() is called; the algorithm does not need to see all instructions executed on a virtual processor but those executed by a single process within the virtual guest. When the user executes our analysis program, one of the command arguments is the malware filename, N_m, which the user wishes to filter by. Every operating system uses its own method for loading a process into memory and executing it and fortunately the Ether developers determined how Windows XP SP2 loads processes and can determine the current executing process by filename. Once Ether detects that Windows has switched to a process whose filename matches the filter provided by the user,

Ether will determine the value of Control Register 3 (CR3), $CR3_m$ and instruct the hypervisor to trap whenever CR3's value is changed to $CR3_m$. This is useful since CR3 is a processor register (virtualized to the guest) used for virtual memory addressing and always contains the page directory base register (PDBR) which is unique for each process on an operating system [4]. Therefore, if another process P_x with the same filename as the malware (N_x=N_m) is executed after the malicious program P_m has started, then Ether will only single-step execution for P_m since $\forall x, CR3_x \neq CR3_m$.

At this point, the Ether code originally called *disasm_instruction*() which would use libdisasm [1] to parse the x86 opcode into a human-readable string and log the disassembled instruction to a file. Our tool modifies this code section to also record the values of all registers which appear in the operands of the instruction and also follow pointer values to possible strings in memory. Algorithm 1 describes the process to log instruction register values and the potential string arrays that the reference values point to in memory. Ether already uses libdisasm's *x86_disasm*() function to generate a structure containing all information about each instruction, so Algorithm 1 begins after this call. For each trapped instruction of the guest operating system, our tool first uses the libdisasm library to parse the instruction's operation code and determine if one of the x86 general registers is used: EAX, EBX, ECX, EDX, ESP, EBP, EDI, or ESI. Next, for each operand o_m that contains a register value, the value of that register is stored into *val* by accessing internal Xen state structures for the virtual machine under analysis. If *val* is of non-zero value then the code proceeds in an attempt to treat o_m as a pointer and finds the value of the memory it references. To do this, a call is made to the internal Xen function *domain_read_current*(*val*) which does a mapping from the virtual machine's virtual memory into the host machine's virtual memory and returns the value stored at the address *val*. Next, the algorithm uses a basic heuristic to determine if the register in o_m is pointing to a string array (run-time strings will be useful in analysis) by searching for ASCII characters. The current value in memory that the pointer points to is checked to see if it is within the printable ASCII characters (0x20 to 0x7F); if it is, then the value is copied into a buffer, the address is incremented by one byte, and the loop repeats. Once either a pre-defined maximum number of characters have been copied from the string or a non-ASCII character is reached then the loop completes and all processing for the current instruction is finished other than logging the results out to the instruction trace file.

The worst-case runtime complexity of this algorithm is $O(IRS)$ where I is the total number of instructions that are traced, R is the average number of operands containing register values per instruction, and S is the maximum string length constraint set by the analyst. Ether already requires $O(I)$ time to single-step and to perform an instruction trace and our tool adds to Ether's instruction trace runtime by $O(RS)$. The x86 instruction set only contains instructions with a possible zero to three operands, so in the worst-case scenario all instructions will contain three operands, all of which contain register values (extremely unlikely). Therefore, the worst-case runtime complexity that our tool adds to the

Algorithm 1. Register and C String Array Logging

```
var: string[MAX_SIZE] {limited to 32 bytes in prototype}
var: ∀i, reg_i ∈ ∑
var: ⋀ = all printable ASCII characters
for instruction I_m in P_m do
  for operand o_m in I_m do
    if o_m ∈ ∑ then
      val = xen_state.register_values[o_m] {store value of all registers used in
      operands}
      if val ≠ 0 then
        ptr_val = domain_read_current(val) {if value is not null, assume it is a
        pointer and load memory at value}
        j = 1
        while ptr_val[j] ∈ ⋀ and j < MAX_SIZE do
          string[j] = ptr_val[j] {if memory value contains ASCII characters, store
          them in a buffer}
          j = j + 1
        end while
      end if
    end if
  end for
  print I_m : (∀i, o_m) : (∀j, ptr_val_j : string[j]) {log all register values and their
  possible string references for each instruction}
end for
```

previous Ether instruction trace operation is $O(3S)$. Realistically, the call to *domain_read_current*() takes the most amount of additional time since it requires a hypercall into the Xen hypervisor and must wait for data to be returned.

Retrieving Memory Values During Execution. For each operand that contains a general purpose register, the code of our tool parses Xen's internal structure for the virtual machine under analysis and stores the current value of the register used in the operand. If the register's value is 0, then no further action is taken for that specific operand. This is a simple heuristic to determine if a register has a useful value. The algorithm next assumes that whatever value the register holds could potentially be a pointer to a string array. A call to *domain_read_current*(*val*) is made which is an internal Xen function that maps a guest domain's virtual memory address into the host machine's physical memory address and allows our tool to read memory at a specific location in the guest. If register holds a pointer value to a location in the guest's memory which contains printable ASCII characters, then the string array is copied out of memory into a buffer until either a non-printable ASCII character is reached or 32 bytes have been copied. Note that the static size limit was an implementation limitation and not a limitation of the methodology. Finally, the instruction mnemonic, the value of registers referenced in the instruction operands, and any possible string arrays pointed to are logged into a text file for each instruction in the malware's trace.

4 Evaluation and Discussion

4.1 Hardware and Software for Development

A physical machine built with an Intel Core 2 Duo E8400 processor (with Intel VT-x) and 4GB of DDR2 RAM was used during the development of our Ether extension. The Debian Lenny 5.0.6 with Linux kernel 2.6.26 was installed as the dom0 operating system running on top of Xen 3.1. Finally, Xen Ether 0.1 was installed via the official instructions [2] with a Windows XP SP2 domU guest used for analysis which was allocated 256MB of RAM.

Table 1. Tracing performance analysis between Ether and our tool

Program	Ether Size	Our Tool Size	Ether Time	Our Tool Time
notepad.exe	145 KB	415 KB	53.1s	56.2s
ipconfig.exe	272 KB	654 KB	91.4s	97.8s
winmine.exe	355 KB	802 KB	45.3s	46.2s
calc.exe	921 KB	2189 KB	234.4s	227.8s
iexpress.exe	1104 KB	2203 KB	70.3s	71.2s
SpyEye.exe	127,593 KB	347,358 KB	257.2s	285.4s

Table 2. Percentage difference in time and speed of Ether and our tool

Program	Additional Size	Additional Time
notepad.exe	+186.2%	+5.84%
ipconfig.exe	+140.4%	+7.00%
winmine.exe	+125.9%	+1.99%
calc.exe	+137.7%	-2.82%
iexpress.exe	+119.4%	+1.28%
SpyEye.exe	+172.2%	+10.96%
Average	+147.0%	+4.04%

4.2 Evaluation Results

To examine the experimental performance of our tool's additions to the Xen Ether code, multiple binaries were traced using each system and the amount of time taken was recorded. Since our instruction tracer runs independently from the virtual machine (cannot induce programs to be executed inside of the guest), the best method for recording the time taken was to start the timer with the launch of it then manually start the program to be traced in the guest. This timing method is not particularly accurate, but should be able to at least give an estimate to the amount of overhead incurred by the code additions. Since

each timing had to be completed manually, each program was just run once with the timer for each condition: tracing with Ether and tracing with our tool. Additionally, the size of the trace files created by both Ether and our tool are compared in order to present the storage overhead of using our tool. The timing and storage comparisons are shown in Table 1. The percentage increases incurred on the storage and time requirements for using our tool are displayed in Table 2.

5 Conclusion

This paper has described a research tool for malicious software security analysts. By applying our tool to a malicious binary program, a researcher is able to statically view what values variables and memory addresses the program held during its actual runtime. Furthermore, since the instruction tracing module is built upon the Xen Ether framework, it is transparent to the execution of the malware for all anti-debugging attacks and most anti-VM attacks. The insight that will be obtained by using this novel extension will hopefully allow future malware analysts to better detect and defend against the increasingly advanced malicious attacks on computer systems.

References

1. libdisasm. x86 Disassembler Library, http://bastard.sourceforge.net/libdisasm.html
2. Ether Source Code (April 7, 2009), http://ether.gtisc.gatech.edu/source.html
3. Intel 64 and IA-32 Architectures Software Developer's Manual, vol. 3B, Intel Corporation (2008)
4. Intel 64 and IA-32 Architectures Software Developer's Manual, vol. 3A, Intel Corporation (2008)
5. Barham, P., et al.: Xen and the art of virtualization. In: Proceedings of the 19th ACM Symposium on Operating System Principles, pp. 164–177. ACM, New York (2003)
6. Dinaburg, A., Royal, P., Sharif, M., Ether, L.W.: Ether: Malware analysis via hardware virtualization extensions. In: Proceedings of the 15th ACM Conference on Computer and Communications Security, pp. 51–62. ACM Press, USA (2008)
7. Chen, Z., Gao, L., Kwiat, K.: Modeling the spread of active worms. In: Proceedings of the 22nd Annual Joint Conference of the IEEE Computer and Communications, vol. 3, pp. 1890–1900. IEEE, San Francisco (2003)
8. Sharif, M., Lanzi, A., Giffin, J., Lee, W.: Automatic reverse engineering of malware emulators. In: Proceedings of the 30th IEEE Symposium on Security and Privacy, pp. 94–109. IEEE, Oakland (2009)
9. Ferrie, P.: Attacks on more virtual machine emulators. Symantec Technology Exchange (2007), http://ivanlef0u.fr/repo/windoz/attacks2.pdf
10. Barnes, D.H., Wear, L.L.: Instruction tracing via microprogramming. In: Proceedings of Conference Record of the 7th Annual Workshop on Microprogramming, pp. 25–27. ACM Press (1974)

An Efficient Scan-to-Scan Integration/Correlation Algorithm for Sea Surveillance Radar

Byung-Doo Kim and Byung-Gil Lee

Electronics and Telecommunications Research Institute(ETRI)
218 Gajeong-ro, Yuseong-gu, Daejeon, South Korea, 305-700
{bdkim,bglee}@etri.re.kr

Abstract. In most sea surveillance radars based on non-coherent integration, the scan-to-scan integration which improves signal to noise ratio and mitigates unwanted echo with a small correlation time is widely used to extract stationary and slow moving targets. The problems of the conventional scan-to-scan integration are: (a) the target tail is generated through integrating previous scans and current scan with respect to a fast moving target which has a high radar cross section; (b) it is difficult to detect a small and fast moving target due to the attenuated small target echo. In this contribution, the problems of the conventional scan-to-scan integration are investigated carefully by using time series analysis, and then an efficient scan-to-scan integration/correlation algorithm is proposed to overcome the problems of the existing techniques. In the proposed algorithm, the signal selection logic eliminating the unwanted target tail and recovering the attenuated target echo has been operated according to the difference between the current and previous scans with two thresholds. Simulation results demonstrate the proposed algorithm's performance of detecting a small and fast moving target without the target tail effect.

Keywords: radar signal processing, scan-to-scan integration/correlation, target tail, sea clutter.

1 Introduction

Radar unfortunately receives not only desired target echo but also unwanted echo caused by the natural environment. Clutter is the term used to denote unwanted echoes and can be classified into rain, cloud, ground, sea, and others [1], [2]. It makes difficult the detection of wanted targets.

Especially, sea clutter depending on the shape of the sea surface is not so easily canceled as ground clutter since the former is moving and fluctuating while the latter is almost fixed. Sea clutter becomes a more serious problem when small, slowly moving boats have to be detected within sea waves.

A large proportion of radars use only the envelope of the received signal in their processing. Since they do not use the signal phase, these systems can be non-coherent from pulse to pulse, and in consequence can use simple transmitters such as pulsed magnetrons and their receivers may employ a logarithmic detector that provides wide dynamic range and facilitates the analog calculation of products, ratios, and powers.

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 565–572.
springerlink.com

The envelope of sea clutter after logarithmic detector is distributed roughly according to a Rayleigh probability distribution [3], [4], [5].

In most sea surveillance radars with non-coherent integration, Scan-to-scan Integration (SI) is used to extract stationary and slow moving targets with a small signal to noise ratio and suppress unwanted echo with small correlation time. Since the scan rate of the sea surveillance radar is about 3 to 30 revolutions per minute, echoes of sea clutter at the identical position from two consecutive scans are obviously uncorrelated to each other [6].

The problems of conventional SI are that a target tail is generated by integration of the current scan and previous scans for the fast moving target with high Radar Cross Section (RCS), and the detection of a small, fast target is difficult due to attenuated small target echo.

A Scan-to-scan Integration/Correlation (SIC) algorithm which can detect small, fast targets has proposed in [7]. In the SIC algorithm, SI process is divided into two branches. One branch provides optimum clutter attenuation by means of more signal weighting to prior scan amplitude while another branch ensures that targets are detected even if they are fast and three thresholds are required in Scan-to-scan Correlation (SC) process. These are referred to clutter distribution before and after SI weighting.

In this study, a conventional SI algorithm is investigated carefully by means of time series analysis, and an efficient SIC algorithm with only two thresholds is proposed. The basic idea of the proposed SIC algorithm is moving target indication by utilizing the difference between two consecutive scans such as a single delay line canceller. The proposed SIC algorithm ensures the detection of a small, stationary target as well as small, fast target without target tail.

2 Radar Signal Processing

Fig. 1 shows the general signal processor structure for non-coherent sea surveillance radar systems that is considered in this paper.

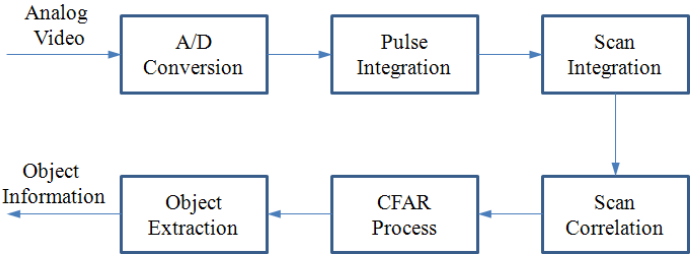


Fig. 1. Block diagram of the radar signal processing

The reflected echo at the receiver input is represented with the consideration of two cases in which the input has three (i.e., the target echo, the clutter echo and the receiver noise) and two terms (i.e., only the clutter echo and the noise) [8]. Hence, the input can be written as:

$$r_n(t) = \begin{cases} s_n(t) + c_n(t) + v_n(t) & t = t_n \\ c_n(t) + v_n(t) & t \neq t_n \end{cases} \quad (1)$$

where n is the scan index, t_n is the elapsed time corresponding to the target range, $s_n(t)$ is the target echo, $c_n(t)$ is the clutter and $v_n(t)$ is the receiver noise.

The envelope is quantized by an Analog to Digital (A/D) converter for digital signal processing, and the quantized video signal can be expressed as follows:

$$q_n(k) = \begin{cases} q_{n,1}[k] & k = p_n \\ q_{n,0}[k] & k \neq p_n \end{cases} \quad (2)$$

where k is the range cell index and p_n is the range cell index with respect to the target range.

A/D converted video data are stored during several Pulse Repetition Intervals (PRIs), and those are processed by a pulse to pulse integration algorithm to obtain improved Signal to Noise Ratio (SNR). This process is called the non-coherent integration, and the pulse-to-pulse integrated video signal can be expressed as follows:

$$a_n(k) = \begin{cases} h_{n,1}[k] & k = p_n \\ h_{n,0}[k] & k \neq p_n \end{cases} \quad (3)$$

where $h_{n,1}[k]$ is the integrated video with the target echo at the scan n and $h_{n,0}[k]$ is the integrated video without the target echo at the scan n.

2.1 Conventional Scan-to-Scan Integration

Scan-to-scan integration is performed to allow increasing Signal-to-Clutter Ratio (SCR) for the sea clutter with short time correlation.

Generally, the sea clutter spike occurs at a random range cell for consecutive scans. By recursive integration of several scans, the clutter level is reduced while stationary target amplitude remains. In this way, small stationary targets emerge from the clutter.

The conventional scan-to-scan integration can be expressed by recursive equation given by:

$$A_n[k] = \alpha A_{n-1}[k] + \beta a_n[k] \quad (4)$$

where $A_{n-1}[k]$, $A_n[k]$ are the integrated video up to the (n-1)-th and the n-th scan, respectively, $a_n[k]$ is the current scan video at the n-th scan and α, β is the weighting factor for the priori integrated video and the current video, respectively.

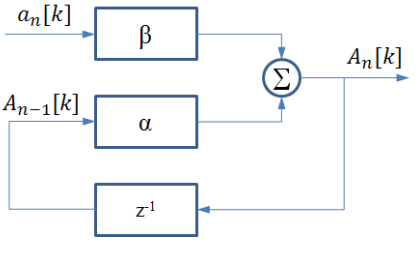


Fig. 2. Principle of scan-to-scan integration

The recursive equation in (4) can be rewritten as follows:

$$A_n[k] = \beta \sum_{m=0}^{n-1} \alpha^m a_{n-m}[k]. \tag{5}$$

Substituting (3) into (5) results in:

$$A_n[k] = \begin{cases} \beta \sum_{m=0}^{n-1} \alpha^m h_{n-m,1}[k] \\ \beta \sum_{m=0}^{n-1} \alpha^m h_{n-m,0}[k]. \end{cases} \tag{6}$$

2.2 Analysis of Conventional Scan-to-Scan Integration

In the case of a stationary target, the cell index of the target range at the current scan is assumed to be almost identical with those at the previous scans. In addition, if the target fluctuation is according to the Swerling model 0, the amplitude of the quantized target echo is analogous with those at the previous scans. Therefore, the following assumption would be made:

$$h_{1,1}[p_1] \cong h_{2,1}[p_2] \cong \cdots \cong h_{n,1}[p_n]. \tag{7}$$

From the above assumption, the expectation of (6) is derived by using the geometric series and the constraint the weighting factor of SI, which can be written as:

$$E\{A_n[k]\} = \begin{cases} \bar{h}_1[k] & k = p_n \\ \bar{h}_0[k] & k \neq p_n \end{cases} \tag{8}$$

where

$$E\{A_n[k]\} = \beta \sum_{m=0}^{n-1} \alpha^m E\{h_{n-m,1}[k]\} \cong \frac{\beta}{1-\alpha} \bar{h}_1[k] = \bar{h}_1[k]. \tag{9}$$

The expectation of integrated scan without target echo after SI can be obtained by same approach.

As one can see, the radar video with target echo and non-target echo are attenuated to mean level. Generally, because the amplitude of the radar video with target echo is greater than the amplitude of the radar video with non-target echo, the Constant False Alarm Rate (CFAR) processor can detect the target.

In the case of a fast moving target, its cell index of the target range at the current scan can be assumed to be different from those at the previous scans. From these factors, the following assumption would be made:

$$p_1 \neq p_2 \neq \cdots \neq p_n. \tag{10}$$

According to (10), the expectation of (6) can be derived as the following equation.

$$E\{A_n[k]\} = \begin{cases} \beta h_{n,1}[k] + \alpha\Delta h[k] & k = p_n \\ \bar{h}_0[k] & k \neq p_n \end{cases} \tag{11}$$

where

$$\Delta h[k] = h_{n-1,1}[k] + \beta\alpha h_{n-2,1}[k] + \cdots + \beta\alpha^{n-2} h_{1,1}[k] \tag{12}$$

$$E\{A_n[k]\} = \beta h_{n,1}[k] + \beta\alpha h_{n-1,1}[k] + \cdots + \beta\alpha^{n-1} h_{1,1}[k]. \tag{13}$$

As one can see from (11), not only the video amplitude contained target echo at the current scan is attenuated by signal weighting, β, but the target tail denoted by $\alpha\Delta h$ is also generated as a side effect of SI, and $\Delta h\Delta$is described as the result of SI until previous scan by (12).

The averaged video with non-target echo after SI can be detected as a new target by CFAR processor if the amplitude of the target tail is greater than the averaged video with non-target echo.

3 A New Scan-to-Scan Integration/Correlation Algorithm

A proposed SIC is configured with signal weighting part and signal selection part such as conventional SIC technique. But signal selection logic has been operated according to the difference of current and prior integrated scan with two thresholds.

The signal selection logic eliminates an unwanted target tail and recovers the attenuated target echo by signal weighting.

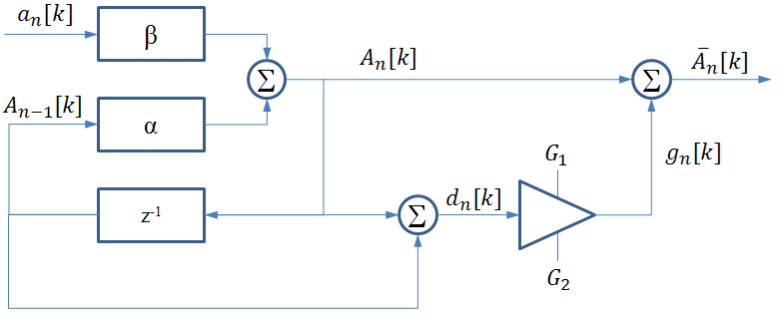


Fig. 3. Principle of the proposed SIC algorithm

For example, if the difference is less than G_1, the current range cell is regarded as a target tail and is removed from the current integrated scan by signal selection logic. On the contrary, if the difference is greater than G_2, the current range cell is regarded as a reduced target signal and is recovered to the current integrated scan by signal selection logic, and if the difference is greater than G_1 and less than G_2, current range cell pass into the next step without any processing.

From the results of (11) and (12), the signal selection logic is simply designed by the subtraction of consecutive scans and the mathematical descriptions are as follows:

$$d_n[k] = A_n[k] + A_{n-1}[k]. \tag{14}$$

Equation (14) can be rewritten as follows:

$$d_n[k] = \begin{cases} \beta h_{n,1}[k] - \beta^2 \sum_{m=0}^{n-2} \alpha^m h_{n-m-1,1}[k] & k = p_n \\ \beta h_{n,0}[k] - \beta^2 \sum_{m=0}^{n-2} \alpha^m h_{n-m-1,0}[k] & k \neq p_n. \end{cases} \tag{15}$$

From (7) and (8), for the stationary target, the difference of consecutive scans can be expressed as follows:

$$d_n[k] = \begin{cases} \beta(h_{n,1}[k] - \bar{h}_1[k]) & k = p_n \\ \beta\big(h_{n,0}[k] - \bar{h}_0[k]\big) & k \neq p_n. \end{cases} \tag{16}$$

From (10) and (11), for the fast moving target, the difference of consecutive scans can be expressed as follows:

$$d_n[k] = \begin{cases} \beta(h_{n,1}[k] - \Delta h[k]) & k = p_n \\ \beta\big(h_{n,0}[k] - \bar{h}_0[k]\big) & k \neq p_n. \end{cases} \tag{17}$$

We consider a first order term of target tail to obtain two thresholds that determine whether attenuated current target echo or target tail.

First-order term of target tail:

$$d_n[p_{n-1}] = -\beta^2 h_{n-1,1}[p_{n-1}] < G_1 = -\beta^2 G_0. \tag{18}$$

Attenuated current target echo:

$$d_n[p_n] = -\beta h_{n,1}[p_n] > G_2 = \beta G_0. \tag{19}$$

G_0 can be obtained from the specified false alarm probability and variance of clutter voltage. Hence, the signal selection logic can be configured as follows:

$$g_n[k] = \begin{cases} -\alpha \Delta h[k] & \text{for } d_n[k] < G_1 \\ 0 & \text{for } G_2 < d_n[k] < G_1 \\ \alpha a_n[k] & \text{for } d_n[k] > G_2. \end{cases} \tag{20}$$

Finally, the video signal after the proposed SIC algorithm is obtained as follows:

$$\bar{A}_n[k] = A_n[k] + g_n[k]. \tag{21}$$

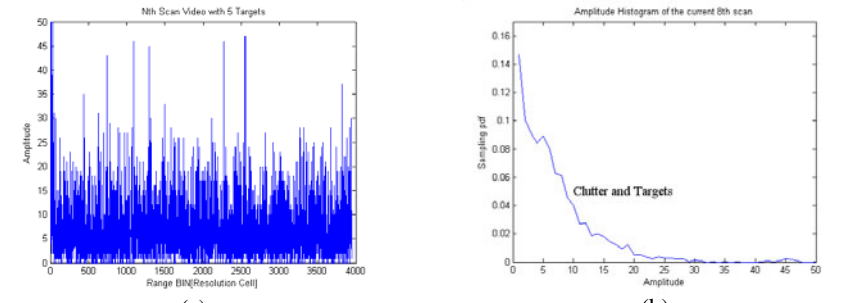
4 Computer Simulation

A computer simulation is performed to demonstrate the proposed algorithm's performance of detecting a small and fast moving target without target tail effect.

The simulated radar signal has two stationary targets and three moving targets with different speeds that are 40, 60, 80 knots of radial velocity, respectively. The sea clutter with Rayleigh distribution is generated utilizing Morchin's sea clutter model [9]. The simulated radar video signal which is quantized by 8-bit A/D converter

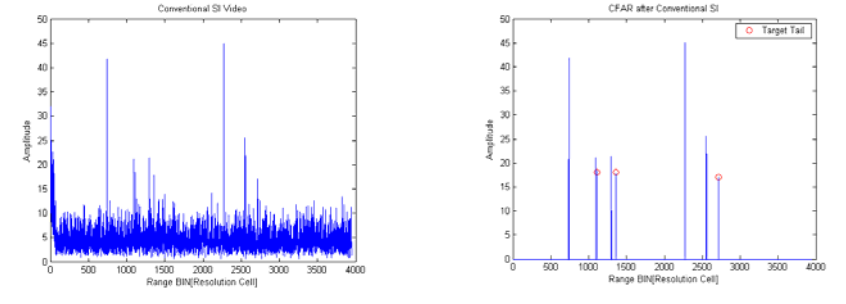
includes 5 target echoes with target RCS 2(m^2) and sea clutter at sea state 3. The sampled radar video is depicted in Fig. 4 (a) and amplitude histogram of sea clutter and targets echoes is depicted in Fig. 4 (b).

The sampled radar video is processed in pulse integration, scan integration, and CFAR in order of precedence. The results of conventional SI and CFAR processor are depicted in Fig. 5. As one can see, a target tail is detected as a real target by CFAR processor in Fig. 5 (b).



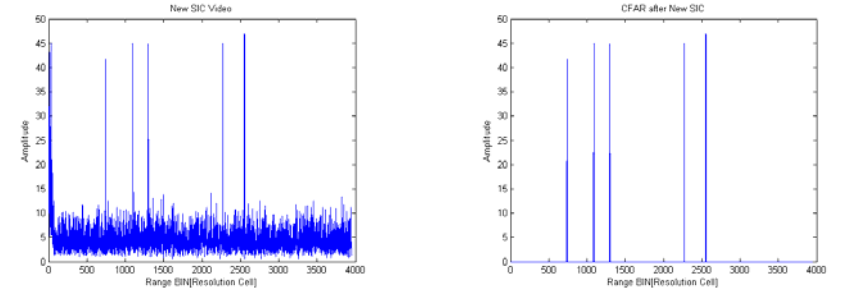
(a) (b)

Fig. 4. Sampled Radar video and amplitude histogram



(a) (b)

Fig. 5. Radar video after Conventional SI and target echo after CFAR processing



(a) (b)

Fig. 6. Radar video after Proposed SIC and target echo after CFAR processing

The results of the proposed SIC algorithm and CFAR are depicted in Fig. 6. The target tail is removed and the attenuated moving target echo is recovered by signal selection logic. The result of CFAR after SIC processing is depicted in Fig. 6 (b). Only the current target echo is detected without target tail by CFAR processor. Not only the stationary target but also the moving target echo remains but sea clutter is attenuated by the proposed SIC algorithm.

5 Conclusion

In this study, the conventional scan-to-scan integration algorithm has been investigated by using a time series analysis, and an efficient SIC technique which utilizes the difference of consecutive scans has been developed to detect a small and fast moving target through removing its tail.

A computer simulation is carried out to demonstrate the proposed SIC algorithm's performance. The simulation results show that the proposed algorithm is effective to detect a small and fast moving target without the target tail. In addition, it is revealed that the radar signal processing results strongly depended on the type of radars and the operational parameters. Hence, the two thresholds used in the scan-to-scan correlation should be calculated in the consideration of the variance of sea clutter and the false alarm probability.

Acknowledgments. This work was supported by ETRI through Maritime Safety & Maritime Traffic Management R&D Program of the MLTM/KIMST (D10902411 H360000110, Development of u-VTS for Maritime Safety).

References

1. Skolnik, M.I.: Introduction to Radar Systems, 3rd edn. McGraw-Hill, New York (2001)
2. Barton, D.K.: Modern Radar System Analysis. Artech House, Massachusetts (1988)
3. Ward, K.D., Baker, C.J., Watts, S.: Maritime surveillance radar Part 1: Radar scattering from the ocean surface. IEE Proceedings F Radar and Signal Processing 137(2), 51–62 (1990)
4. Green, B.A.: Radar detection probability with logarithmic detectors. IEEE Transactions on Information Theory 4(1), 50–52 (1958)
5. Croney, J.: Clutter on radar displays-reduction by use of logarithmic receivers. Wireless Engineer 33(4), 83–95 (1956)
6. Shyu, H.C., Sun, Y.S., Shen, W.H.: The Analysis of Scan-to-Scan Integration Techniques for Sea Clutter. In: IEEE Radar Conference, pp. 228–223 (1994)
7. Hofele, F.X.: Scan to Scan Integration Correlation for the Detection of Small Fast Targets. In: IEEE International Radar Conference, pp. 380–384 (2001)
8. Spafford, L.J.: Optimum radar signal processing in clutter. IEEE Transactions on Information Theory 14, 734–743 (1968)
9. Edde, B.: Radar Principles, Technology, Applications. Prentice Hall PTR (1993)

Frequency Conversion Control Module Design for Electric Cars Based on SVPWM

Wei Feng[1], Changán Di[1,2], Peng Bian[3], Deren Kong[2], and Zhenwei Duan[2]

[1] Xián Modern Chemistry Research Institute
[2] Nanjing University of Science and Technology
[3] North Huaán Industry Group Co., Ltd
210094, Nanjing University of Science and Technology,
School of Mechanical Engineering
dica3@sina.com

Abstract. In order to increase electric cars power efficiency when control frequency conversion, and reduce the interference of high harmonics, we propose a design method of voltage space vector (SVPWM) frequency conversion control module. Based on TMS320LF2407 the PWM signal producing module design, IPM module as the core, we design the inverter circuit and Developed SVPWM real-time motion control program. This program use the manner of segmentation modulation, achieve frequency conversion control a range of 5 ~ 100Hz frequency. Variable frequency speed-governing experiment shows that, The control method can reduce the switching losses, meanwhile suppress the temperature rise.

Keywords: Frequency Conversion Control, SVPWM, DSP, IPM.

1 Introduction

Electric cars are effective urban transportation vehicles for clean. They absolve people's worry from shrinking of the oil resources and so on. As clean and energy-efficient new vehicles, they are the only vehicles which run as "zero discharge" and no pollution. Motor frequency conversion control technology of electric cars is a hot research currently. Among them, the use of voltage space vector (SVPWM) control methods to control the motor has good power efficiency. This paper presents frequency conversion control module design method for electric bicycle based on SVPWM.

2 Frequency Conversion Control Module Hardware Design

In this paper, the frequency drive module uses DC-AC inverter circuit. Its DC side is storage battery. It's a three-phase alternating current after IPM Module drive circuit inverter for the output frequency conversion [1] [2] [3]. Block diagram is shown in Fig.1.

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 573–578.
springerlink.com

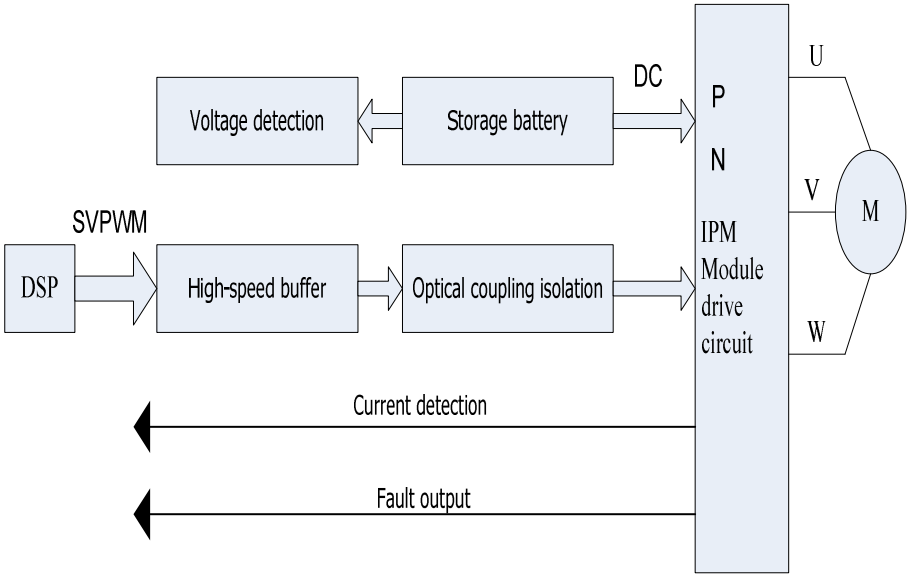


Fig. 1. DC - AC inverter principle based on the IPM

2.1 SVPWM Signal Input

DSP uses dedicated motor control chip TMS320LF2407.SVPWM signal exported from the DSP module is a low-voltage high-frequency signal, has limited drive capability. Therefore, before the SVPWM signal entering the optical coupler, first use the tri-state buffer 74HC540 with enable control to increase signal drive capability. After SVPWM signal through the tri-state buffer, by the current-limiting resistance, connect with LED cathode of high-speed optical coupler 6N137. SVPWM signal through optical coupling isolation into the IPM module, control IGBT's (integrated within the IPM) one pair of bridge arms Vp, Vn, the other bridge arms are the same. The use of isolated power measures, through the optical coupler, completely isolate SVPWM signal generation circuit from the SVPWM signal of inputting IPM . 1200V DIP-IPM input signal logic drives using the high level. Circuit is shown in Fig.2.

2.2 Frequency Conversion Main Circuit

The inverter module of the system uses Mitsubishi intelligent power module (IPM) PS2205x. Internal integrated IGBT drive circuit. It can endure high pressure. Its highest current can reach 25A. It is designed as the device for low-power frequency inverter. The interior of PS2205x integrate high pressure HVIC in 1200V, it makes the PWM signal can be directly entered without additional drive circuit. PS2205x's supply voltage is 15V, through a circumscribed bootstrap capacitor and diode, can realize single power supply. The maximum carrier frequency is 15kHz, can meet the system requirements.

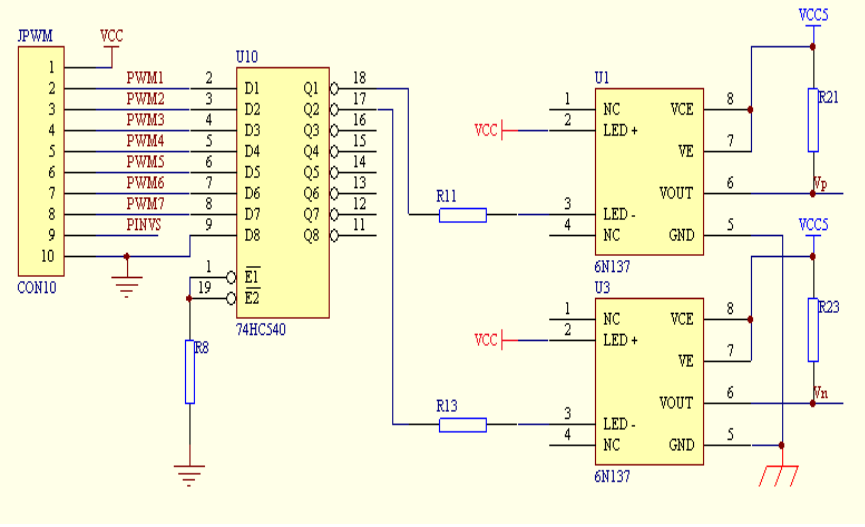


Fig. 2. SVPWM signal input circuit (One group)

3 DSP Software Implementation of SVPWM Algorithm

In TMS320LF2407, the built-in event manager EV module has space vector PWM waveforms generated built-in circuit. In order to output space vector PWM waveforms, need to set the following register [4] [5] [6].

1) Set ACTRx register define the output mode of compare output pin;
2) Set COMCONx Register to enable the compare operation and space vector PWM mode, and set reloading conditions of the CMPRx as underflow;
3) Set general timer working mode, and start the timer.

In the program, we need according to external demands to ensure the component values U_α, U_β of the inputting motor voltage U_{out} and ensure the following parameters of each PWM period:

1) Two adjacent vectors, that determine the sectors number of U_{out};
2) According to SVPWM modulation period T to calculate the respectively acting time T1, T2 and T0 of two basic vector space vector and 0 vector;
3) Write the corresponding opening way of U_x to ACTRx.14~12 bits, and write 1 to ACTRx.15; or Write the opening way of U_{x+60} to ACTRx.14~12 bits , at the same time write 0 to ACTRx.15;
4) Write the T1 / 2 value to the CMPR1 register, and Write the (T2+T1) / 2 value to the CMPR2 register.

Using interrupt mode loads the value of the timer in the software. In the main program based on user settings, determine the frequency of operation of the motor,

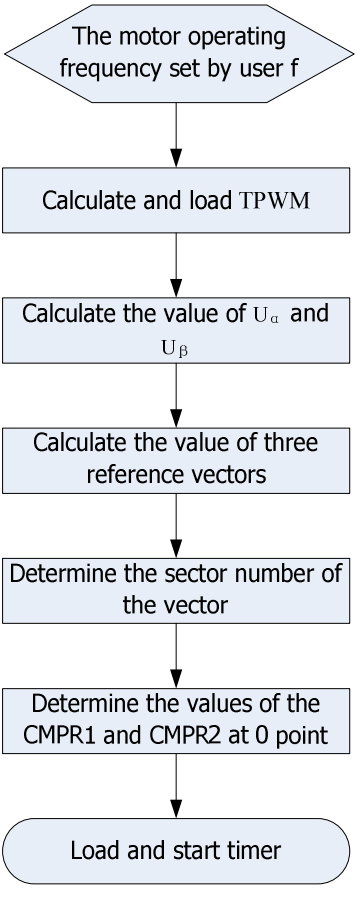


Fig. 3. Calculation flow chart

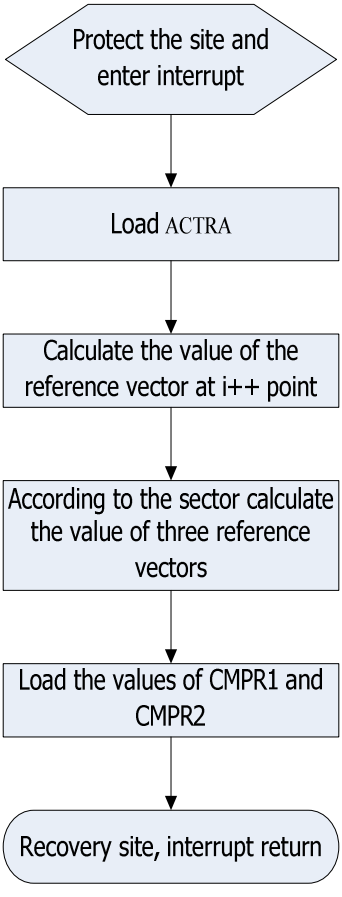


Fig. 4. Timer 1 interrupt handling of main program

and calculate the corresponding value of the timer. We take n points from the angle of 2π. The calculation flow charts show in Fig.3 and Fig.4.

Programs define two arrays Ua[n], Ub[n], they respectively store α、β component value of U_{out}. Points determine according to motor operating frequency, segmentation synchronous modulation table shows in able 1. Based on the reference quantity value of three sectors to determine the sector of each U_{out}, and stored in the sector array SECTOR[n]. The timer comparison value is calculated in the interrupt, it can achieve real-time calculation of SVPWM. When control the motor, set the motor operating frequency by the user, by the main program to calculate the α, β components of U_{out} and the sectors. The remaining tasks are handled by the interrupt subroutine, it enables the processor to have more time to deal with other control tasks.

When control the motor frequency, motor operation frequency changes from 5 ~ 100Hz. In this, select the segmentation synchronous modulation, the segmentation

method shows in table 1[4] [5] [6]. The use of sub-modulation can eliminated when the motor is running in the low-frequency instability. While at high frequencies, the reduction of Points doesn't affect the normal operation of the motor.

Table 1. Sub-synchronous modulation table

Frequency interval of modulation wave	Highest carrier frequency	Points
5-10Hz	7.2kHz	720
10Hz-20Hz	7.2kHz	360
20-36Hz	7.2kHz	200
36-50Hz	6kHz	120
50-72Hz	7.2kHz	100
72-100	6kHz	60

4 The Result and Analysis of Frequency Conversion Speed Control Experiment

The experiment of frequency conversion speed control is primarily used to verify the Results and performances of electric motor speed control when system frequency conversion speed control. The System uses the open-loop speed control SVPWM, set the parameters of SVPWM are as follows:

1) Modulation cycle TPWM is 6000*TCLK, carrier frequency is 6.144 kHz;
2) Output power frequency is 8.53Hz, Sampling points of per cycle are 720;
3) Modulation depth TP is 0.7;

In the experiment, the measurement of SVPWM output signal 1and the output waveform is shown in Fig.5. Fig.5 shows that each way of switches with 1 / 3 of the time is not open. Hence, compared to control mode of SPWM, at the occasion of the higher carrier frequency, it can obvious reduce the switching losses. At the same time, the line voltage of output power are also measured, the waveform is shown in Fig. 6.

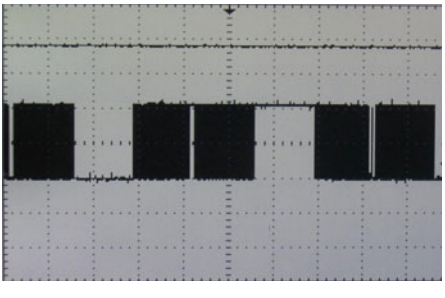


Fig. 5. PWM control signal of channel

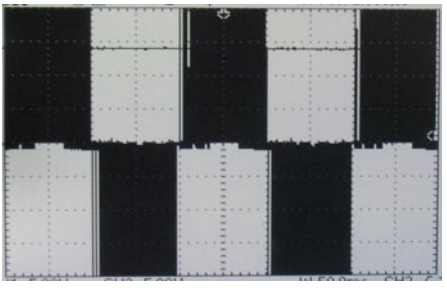


Fig. 6. Line voltage wave of output power

In the actual control, system uses the way of sectioned spectrum modulation to control the motor at frequency conversion. In the experiment, for different working frequency ranges have carried out experiments and the results are well. Experiments show that, the system which uses the frequency conversion control method of SVPWM achieves motor speed control in the range of 5~100H. The effect is well. Meanwhile, using the way of sectioned spectrum modulation reduces motor the unstable phenomenon at low frequencies, noise is relatively small.

5 Conclusion

With the development of the power electronics, microelectronics and advanced control theories, motor control technology has made substantial progress. According to frequency conversion speed control required of environmental electric cars, using intelligent power module (IPM) PS2205x and high-performance DSP chip TMS320LF2407, designed a set of fully digital frequency conversion speed control system of AC asynchronous motor based on DPS, and the DSP implementation of software algorithms based on SVPWM has a detailed study. Finally, carry out experiments on the design of the prototype, Respectively give the experimental results under the different of parameters. Studies have shown that, SVPWM control method improve the voltage utilization, reduce the switching loss and reduce the temperature rise. There is reference value of this paper studies on improving the utilization of electrical energy.

References

1. Li, F., Wang, Y.: The Motor and Drag Technology. Tsinghua University, Beijing (2006)
2. Jiang, S.: TMS320C24X DSP Hardware Development Tutorial. Mechanical Industry Press, Beijing (2005)
3. Texas Instrument. TMS320LF2407 DSP controller (2002)
4. Liu, H., et al.: DSP Theory and Motor Control Applications: Series based on TMS320LF240x. Beijing Aerospace University Press, Beijing (2006)
5. Li, B., Fu, X.: Frequency speed control vector control system research application TMS320LF2407DSP. Power Electronics (4) (2005)
6. Wang, X., et al.: DSP control of electric motor——DSP applications of TI Company, 2nd edn. Beijing Aerospace University Press, Beijing (2009)

Research and Design of Data Integration Architecture for Rural Planning

Weipeng Cheng, Wenxue Wei, and Feng Zhang

College of Information Science and Engineering,
SDUST, Qingdao, Shandong 266590, China

Abstract. Rural planning is an important step in the urbanization construction and urban system perfection. Data integration of rural planning is the base of constructing a data center and application platform all over the country. A 7-layer framework of data integration of rural planning is designed in the paper on the base of analyzing the features of data of rural planning, such as distributed, of different frequency, from multiple sources and multi-semantic, referencing to the SIG 7 layers architecture and adopting the idea of service based of OGSA. Also the service mechanism and mixed deployment mode of data integration are studied in the paper in detail. The architecture has a characteristic of "general management, local autonomy", so it has the heterogeneous data integrated with great efficiency and high extensibility.

Keywords: Rural planning, data integration, architecture, Web service, service oriented.

1 Introduction

At present, our country has entered an accelerating period of the urbanization. In order to integrate the space of villages and town and improve the level of establishment ceiling, building a platform for data integration of rural planning is absolutely essential. The data integration of rural planning can solve many problems which appeared in the process of planning of rural village and town, for instance, loss of basic data, weakness in application capacity, lack of harmony and timeliness, thus providing basic conditions for rural planning and construction management.

Data integration is combining the data into together, which are geographically widespread, self-governing, and heterogeneous in mode so that users can acquire data and functions and methods of handling data transparently. There have been many achievements about the data integration of geographic information, but they are not well suited for the planning of village and town[1]-[4]. The data integration of rural planning is still in the initial stage of analyzing, and the whole structure has not been studied[5]-[9]. Now the data integration is mainly about using XML to shield the heterogeneity of recourses[10]-[11]. SOA(Service-Oriented Architecture) is an idea of architecture to realize the data integration of distributed heterogeneous system and resource sharing. SIG (Spatial Information Grid) and OGSA (Open Grid Service Architecture) are exactly the result of this idea[12]-[15]. The OGC (Open Geospatial

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 579–586.
springerlink.com

Consortium) defines series of technical standards about geo-spatial data sharing. At the present time, OSGA is chiefly used in the information integration of enterprises and SIG is at the stage of conceptual study, however neither OSGA nor SIG are applied in mass data integration. In the paper, a 7-layer framework of data integration of rural planning is designed and a mixed deployment mode combining real-time integration with data warehouse.

2 Analysis of the Characteristic

Before integrating the data, we should analyze the data appeared in the rural planning and its characteristics. Planning of village and town is not only a politically governmental management, also a geographically plan, so the process involves two applications, MIS and GIS. As a result, the data appeared is especially complicated, which can be divided into three groups: geographic data, attributive data and document data, as shown in the table 1.

Table 1. Rural planning data category table

Data Name	Data Category	Instructions
Geographic Data	Basic Geographical Data	The data that are used for background display, determining the position, include administrative districts data, topography data, aerial photographs and satellite photograph data.
	Rural Planning Geographical Data	The GIS data that serve for rural planning and construction, include planning result data(for example rural overall planning data, controlling planning data, construction data, project planning data and so on),and planning approval data(for example red line data, pipe line data, road data and so on).
Attribute Data	Basic Management Data	The relational data that are used for describing district itself, including public facilities data, ecological environment data, resources and assets data, economic industry data, and general history data.
	Rural Planning Management Data	The relational data that are produced and used in rural planning process, include business examination and approval data, planning management data.
Document	Basic Document	The document include all kinds of different types of pictures and area videos document that are used for displaying map in GIS, the laws and regulations used in planning management.
	Rural Planning Document	The dynamic documents that are produced in rural planning process include image document and planning management approval document.

Rural planning embodies the traditional culture, as well as the new plan. Therefore there are many population and socioeconomic data. The information has a characteristic of distributed, of different frequency, muti-semantic and from multiple sources.

1) Distributed: The resources are in different locations. There are so many administrator apartments and each of them has its own geographic database and government management database. Also the level of the data is different caused by different division, together with the "data island", brought by different departments. Recognizing these problems, we need to link the data with each other according to the geographic features of geographic entities.
2) Of different efficiency: From the table 1,we can see that the data can be divided into too parts: one is basic data, which have a long renewing cycle, complicated data sources, large amount, frequent applications, stable data and the problem of data sharing with other systems; the other is the planning data, which is generated and used in the planning process, and the features of this part is strictly requirement of effectiveness, a short renewing cycle, small amount, frequent changes, low share degree and in the scope of rural planning; also there are use data, authority data etc.
3) Multi-semantic: On the same entity, the geometric features are same, however with many different semantics. There are geographical characteristics, together with social and economical information. Consequently, the problem of multi-semantic arises. The space-time feature makes the data sequences at the same time but in different spaces, or in the same space but at different time.
4) From multiple sources: As in data sources, mobile equipments and mapping equipments have different ways to memory, extract and manipulate data. The spatial data are stored in graphic files in all kinds of formats, and different GIS software adopts different storage format; attribute data are mainly stored in different relational databases.

3 Architecture Design

3.1 Integral Structure Design

The data integration of rural planning has the features of effectiveness and large amount, has the sharing of GIS data and integration of management data. SIG is information resource of assembling and sharing space, SIG is Spatial Information Infrastructure, which can assemble and share spatial information resources, can provide integration of organization and process, and can supply on-demand services. A 7-layer framework of data integration of rural planning is designed in the paper referencing to the SIG 7 layers architecture and adopting the idea of service based of OGSA, as shown in figure 1.

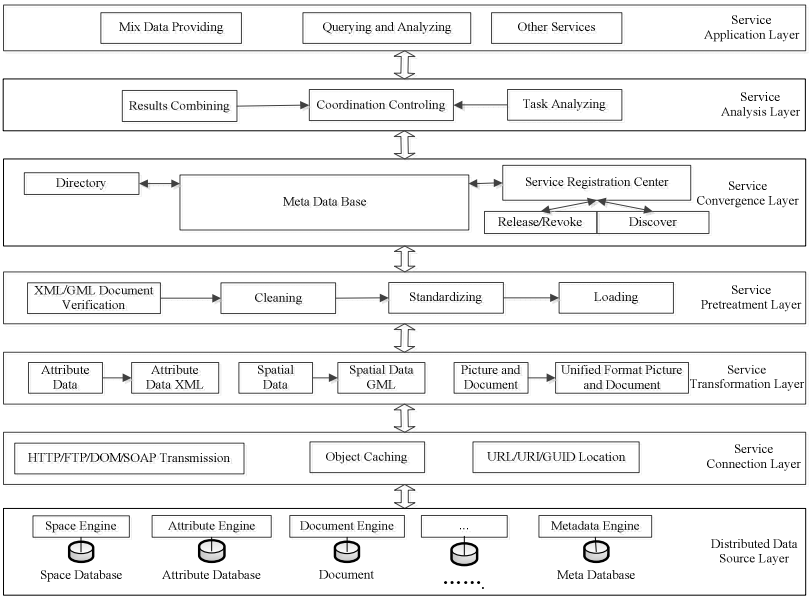


Fig. 1. The general framework of rural planning data integration

1) Distributed data source layer: This layer is at the bottom of the architecture and is the fundamental service provider. It not only includes the supply of many data service, also is executors of retouching, deletion and addition etc.
2) Service connection layer: Connect service providers in the distributed network, use URI and GUID to locate, employ HTTP and FTP to transport, and offer a buffer for the upper layer.
3) Transition Layer: Transform the data from the lower layer to a format recognized by the data center. In the paper the attribute data use XML, which is a standard format of information interchange in MIS system and is suitable for information exchange between data centers in two levels. The spatial data adopt GML format and generate GML documents based on relevant regulations. However a detailed standard need to be established for the document exchanges.
4) Preprocessing layer: Clean and standardize the data from the lower layer according to some rule and provide data for the upper layer.
5) Assembling layer: The layer is the core layer in the architecture, and the center of registry, publishing and binding services.
6) Analyzing and transaction layer: Analyze the service requests from the upper layer, exchange the requests to many subtasks, assemble the processing results and turn over it to the upper layer.
7) Application layer: Interact with users, present the results and requests.

3.2 Detailed Design of Service Assembling Layer

To make the data integration framework to be a service oriented architecture, we can use Web Service technology and the spatial data service launched by OGC, then

abstract all the resources(such as spatial data resources, computational resources) into service. All services use the follow standard: the browsing part using the WMS (Web Map Service) interface standard launched by OGC, the inquiring part using the WFS (Web Feature Service) interface standard launched by OGC, the spatial data transmission using GML standard, the attribute data transmission using XML standard. The realization mechanism of service orientation is as shown in figure 2.

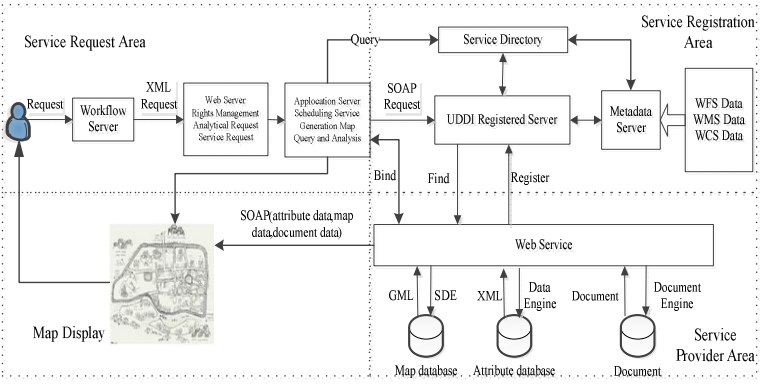


Fig. 2. The realization mechanism of service orientation in rural planning data integration

Web Service system structure contains three basic entity: service provider, service requester and service agency (service register).In service requested area, the users request the Web server, choose the service chain to perform through the workflow service and then use server to call service; In service provided area, the data provider in the bottom encapsulates itself to Web service and registers in UDDI; In service registered area, which include service registry and service list, Web service just provides service as stateless application and there is no relationship during data services in data center, so meta data server is used to communicate, which uses space service standard, including WFS, WMS, WCS(Web Coverage Service); Eventually, all requested services commit to users using SOAP, then the results are showed to the users.

4 Deployments

A multi-tier distributed architecture is designed in the paper: the national data center of rural planning, the provinces and autonomous regions, the cities and counties. This distributed data center has a strict hierarchical relationship so as to avoid override update.

The ways of integrating data mainly include two kinds, data integration and data warehouse. Building data warehouse can accelerate the response but with a bad real-time performance. If a real-time integration is used, the data can not be called up in time from all the country, let alone need to be transformed and loaded many times during the call.

The data integration of rural planning, as a large integration process, contains planning data and basic data. In order to improve the efficiency, the features of the data must be taken into consideration. So we designed a mixed deployment mode, with a combination of the two integration methods, as shown in figure 3.

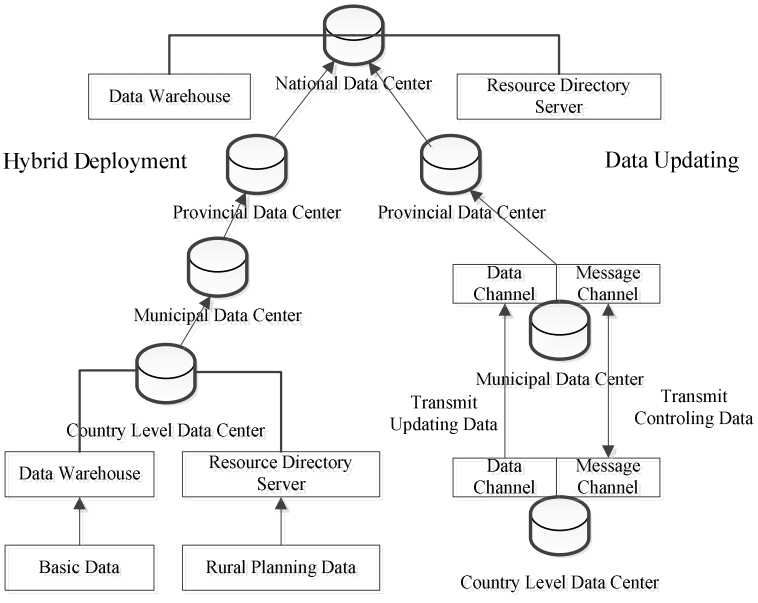


Fig. 3. The hybrid deployment of rural planning data integration

The middle and upper layer manages the lower layer logically in this way of deployment, and it manages the data which are collected by the lower layer physically, the lower layer manages its data autonomously, namely realize "global management, local autonomy". The characteristics of mixed deployment are as follows.

1) "Baisc data". We can use the way of data warehouse for the big, frequent and stabile data, and call data from data warehouse of this layer directly when use. In resource management department of each layer, we establish integral data center, and it realizes update of date through the data exchange mechanism between layer and layer, namely, integrate once, use more than once.
2) " Rural planning data". This data are produced by this system itself. It generally has small amount except for using GIS, which needs high efficiency and can use real-time data integration methods. The data center of each layer establishes resource list and Meta server to the data of lower layer so that it can visit the latest data server provided by data center of lower layer through the remote call. Use and integrate at any time.
3) Grassroots data center is established on the base of LAN or private network to ensure the efficiency of business processing. It connects through Internet network among multilayer data center. It is loosely coupled among different data center in grassroots and easily to maintain and extend.

The graphical data set, which is produced by rural planning process, is the basic source of rural planning data. The users of the bottom update data to the data center of county level, and then update every level data center from bottom to top. There are two transmission channels in the data updating process between every two layers: message channel, response for transferring control message, and data channel, response for data exchange, as is shown in the right of figure 3.

The national data center realizes rural planning of remote geographic collaborative and visual interaction based on "one map". The provincial rural planning data center inserts all historical digital maps and text into its data center according to the relevant regulations and standards to establish rural planning historical database; Data sharing mechanism through the national network interconnection reports data and plan documents for the superior center and provide guidance for the regulation of the city/county level rural planning; The provincial rural planning data center is equipped with large networking data server and space database management system, space and planning data entry and processing platform, and data sharing interface and planning information releasing platform. Different planning departments plan activities and create and update database using unified planning data layer and standards.

The city/country data center is equipped with rural planning information platform terminal, which connects with the provincial one. It searches and updates the date of living environment and ecological in its regulatory scope regularly, and puts on record and makes detailed investigation to illegal building and the change of land use that is harmful to the environment according to the situation provided by management personnel in village or town, then it reports to the provincial rural planning data center through this platform.

5 Conclusions

Data integration of rural planning provides a significantly technical support in coordinating urban-rural development and integration of city and countryside, and supplies data support for setting out the rural plan and the constructing the village and town. The service mechanism and mixed deployment mode of data integration are studied in the paper at length based on service centered and the SIG 7 layers architecture and the architecture has a feature of "general management, local autonomy", so it has the heterogeneous data integrated with great efficiency and high extensibility. The architecture not only meets the requirement of data integration, also can be combined with distributed data computation, so as to complete the rural planning efficiently.

References

1. Xu, K., Wu, H., Gong, J.: Research on Mechanism of Geographic Information Service Supported by Multi-tier Heterogeneous Geospatial Database. Geomatics and Information Science of Wuhan University 04, 402–404 (2008)
2. Cui, B.-L., Xu, L.-L., Cao, K.-B.: GIS Data Organization for Planning and Land Management of Village and Town. Geomatics World 03, 6–10 (2010)

3. Hu, Q.-L., et al.: Application of GIS in New Rural Planning. Journal of Anhui Agricultural Sciences 01, 611–614 (2011)
4. Cai, C., Cui, T.-J., Ge, W.: Development Research of Geographic Information Web Service. Geomatics World 01, 35–39 (2011)
5. Wu, X.: Data center integrated development platform, pp. 5–9. Publishing House of Electronics Industry, BeiJing (2010)
6. Huang, Z., Dai, X.: Study of Multi-Source GIS Data Integration Method. Urban Geotechnical Investigation & Surveying 02, 22–25 (2010)
7. Gong, Q.: Research on the geospatial information grid computing mode and architecture design under the condition of heterogeneity. Science of Surveying and Mapping 04, 59–61 (2010)
8. Wang, F.-X., Hou, Y.-Z., Yang, J.: Geospatial Data Sharing and Interoperability Technologies in Grid Environment. Computer Science 36(01), 96–100 (2009)
9. Gong, Q.: Review to grid programming model and the corresponding technology of geographic spatial information grid. Science of Surveying and Mapping, 11–10, 1–6 (2011)
10. Huang, J., Wu, X.-C., Xing, T.-Y.: A study of multilevel heterogeneous land resources data exchange based on GML. Science of Surveying and Mapping 34(1), 115–117 (2009)
11. Fang, Y., Cao, J., Zeng, C.: New Implementation Mechanism of Interoperable WebGIS Based on GML. Computer Engineering 07, 44–46 (2007)
12. Liu, Y.-X., Jing, N., Chen, L.: Spatial Information Grid: A New Generation of Spatial Information Infrastructure and Service Framework. Geomatics World 07, 22–27 (2007)
13. Hu, M.: Research on Seamless Integration Technology of Distributed Heterogeneous Spatial Data Based on Data Center Patterns. China University of Geosciences, Wu Han (2009)
14. Zhao, J.: Design and Applied Research of Regional Planning MIS Based on GIS. ChongQing University, ChongQing (2007)
15. Huang, C.L., Yang, Y., Zhao, X.: Research on the Service oriented Architecture of Space Information Grid. Journal of the Academy of Equipment Command & Technology 17(5), 48–50 (2006)

The Design and Implementation of the Control and Monitoring System for a Multi-pot Reactor

Kai Zheng, Su Wang, and Jun Zheng

Computer Center, East China Normal University, Shanghai, China, 200062
{kzheng,swang,jzheng}@cc.ecnu.edu.cn

Abstract. Reactors are widely used in chemical industry, pharmaceutical industry and many other industries. Advanced reactors aided with computer software can fully control the reaction process and record the whole reaction parameters. Compared with conventional software design, reactor control software design has its own specialty. In this paper, we introduce a design and implementation of the control and monitoring system for a multi-pot reactor. Also, we focus our attention on exploring the development of industrial software, discussing some typical problems encountered and providing corresponding solutions.

Keywords: Industrial Software, Control and Monitoring system, Reactor.

1 Introduction

Reactors are chemical devices, which are used to mix two or more types of liquids or gases for reaction. It's possible to observe reaction process, analyze and obtain the necessary data during reaction. Such devices are widely needed by chemical industry, pharmaceutical laboratories and many other industries.

With the development of science and technology, research is becoming increasingly sophisticated. Highly intelligent devices are needed to provide precise control and accurate data. Manual control is gradually replaced by intelligent control with the aid of computer. Advanced reactors can use computer software to fully control the reaction process and record the whole reaction parameters.

Not long ago we developed a control and monitoring system of the four-pot reactor for Japanese customers. The reactor has four separate reaction vessels, each reaction vessel can be used for two-liquid (gas) reaction. The device is fully controlled by computer, e.g., pre-setting reaction process and the parameters; recording and computing the parameters during the reaction; viewing the trends of the corresponding parameters; using user interface to control the reaction process if necessary.

Reactor control and monitoring software belongs to industrial software. During development phase, it was found that there were some differences between the development of industrial software and conventional software. Understanding these differences is helpful for developing high-quality industrial software. Here we introduce the system design and implementation, discuss some typical problems encountered in development, and also provide corresponding solutions.

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 587–593.
springerlink.com

2 System Design

2.1 Hardware Architecture

The system consists of three parts: the mechanical parts, control modules and computers. Mechanical parts mainly consists of four reaction vessels, each reaction vessel is equipped with the two micro-pumps, which are used for dropping reaction liquid. There are also related heating, cooling and mixing devices. Control modules include a temperature controller and some I/O control modules. Temperature controller can monitor the current temperature and set the target temperature of the Reaction vessel (After setting the target temperature, temperature controller will start heating or cooling device to make the temperature of reaction liquid closer to target temperature). I/O control modules are bridges between the mechanical parts and computers, which are used to change computer signals to peripheral device signals and the reverse. There are two types of I/O control modules, digital ones and analog ones. Mechanical parts directly connected with I/O control modules, while the latter connected to the computer through the network. The reactor generally equipped with two computers, one computer with control software is for setting and controlling the reaction process, another with monitoring software is for monitoring reaction progress or observing the trend graphics of each parameter.

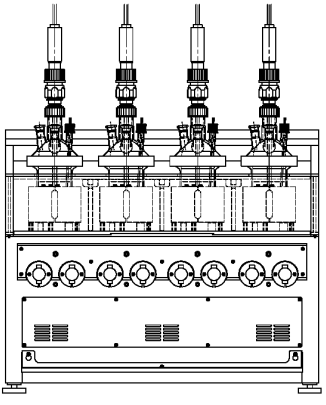


Fig. 1. Image of 4-pot reactor

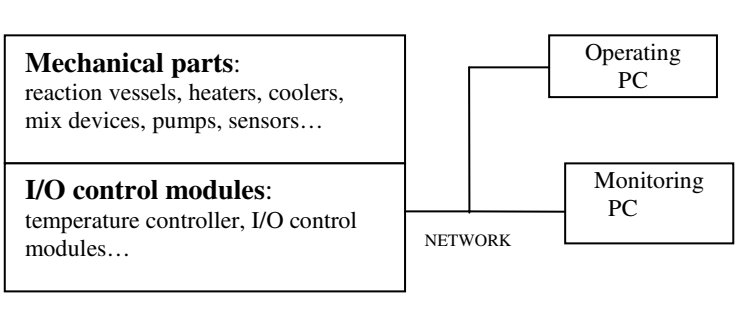


Fig. 2. System hardware structure

2.2 Software Architecture

There are two kinds of software, the software for operating PC and the software for monitoring PC.

Operating PC software includes user interface, hardware control software, data collection, trending software and etc. It can pre-set conditions and control the entire reaction process, collect related data and draw trends based on these data. Monitoring PC software is mainly composed of user interface and trending software, which can monitor the reaction process, view trends of the various parameters, and terminate the reaction process if necessary.

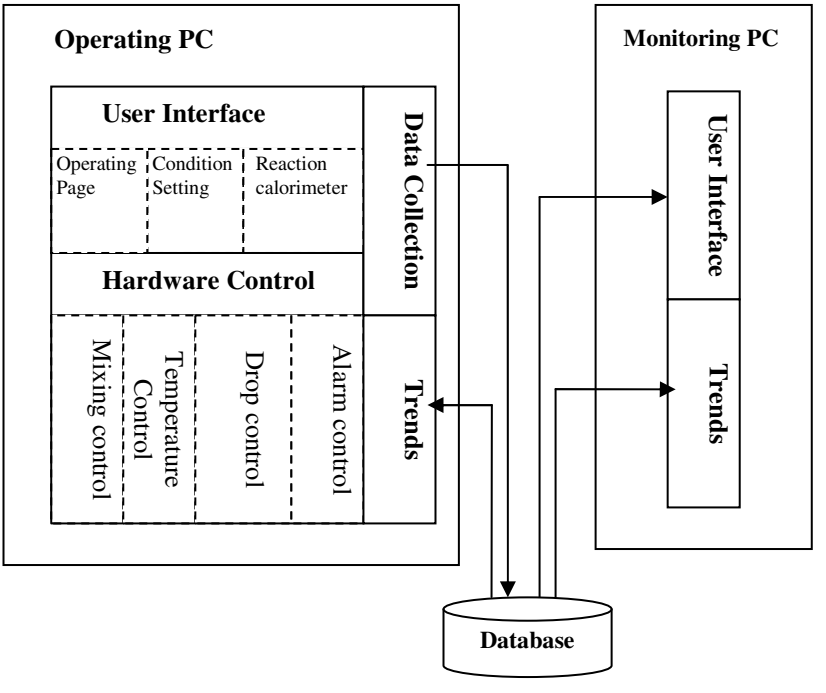


Fig. 3. System Software Architecture

2.3 Database Design

The system uses SQL Server database, there are three types of data stored in the database:

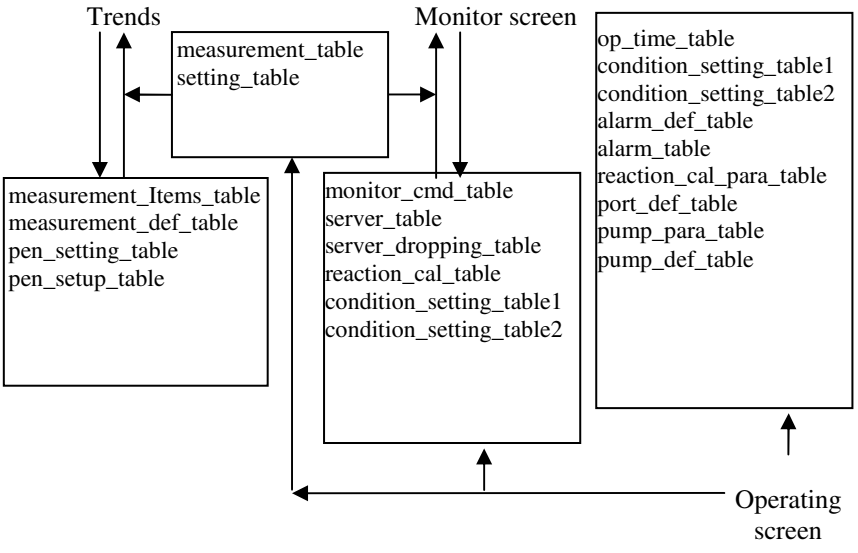


Fig. 4. Main Database Table

The first one is measured values and setting values recorded in reaction process, which can be used to draw trends as need. In addition, the trend graphics also needs some tables to store the pen color values and attribute values of trend graphics.

The second one is various types of data needed by operating PC software, including the condition parameters which are set before reaction, the reaction calorimeter parameters, alarm definition, and other display-related data.

The third one is the data for monitoring PC software. Monitor screen has a similar appearance to operating screen, there are only a few functional differences between them (e.g., in monitor screen there are no other control functions but only stop function). As the user interface is much complex, it's difficult to use Web elements to construct the UI, so we use C/S mode which regard the database as a bridge instead of using the Web CGI mode. First, we save various parameters and values in database on operating PC(server side), and then on monitoring PC(client side), we fetch the necessary data from the database. On the server side, data is generated every second, on the client side, data is fetched every second for screen refresh.

3 Problems and Solutions

Compared with the conventional computer program, industrial control software has its particularity. In development, a number of problems are encountered, the corresponding solutions are given through analysis.

3.1 Real Time Control

At present, Windows is not a real time operating system, some programming languages (such as Java, C and C#) call Sleep method of Thread class to simulate a specified time passed (The function of sleep method is to make program sleeping a specified number of milliseconds), it is assumed the program will pause for a second if instruction sleep(1000) is called. Although the pause time caused by sleep(1000) instruction may be not actually one second, the main problem is: in traditional programming, conventionally instruction's execution time is not considered because execution speed is very quick, which means the execution time for most types of instruction will be regard as 0 except sleep instruction. But in actual environment, we found some instructions, especially I/O instructions, are relatively time consuming. If no action is taken, after several cycle time error will accumulate soon, finally has great influence upon control accuracy.

As a resolution, all instruction's execution time is taken into account. In a cycle, cycle start time is recorded at beginning, and cycle end time is recorded in the end, the difference of them is regard as actual cycle time. So, time error cause by instruction's execution time will not accumulate.

3.2 Competing for Shared Resources

The system has 4 reaction vessels (4 pots), each pot is controlled independently. Each pot has its own parameters. For example, temperature of pot1 is different from that of pot2. All temperatures are read from one temperature controller. This temperature controller has 8 temperature sensors, two for each pot: one is for temperature of reaction liquid, another is for temperature of the surface of the reaction vessels.

In this system, one temperature control is shared by all pots, and it should be accessed through RS232C interface, the communication protocol of it is similar to wait-stop protocol, PC send a data request package to temperature controller, then temperature controller reply a data acknowledge package with requested data. So temperature controller is a serial device, and it's a shared resource, we should take methods such as semaphore to prevent competition for resources.

To solve this problem, in c# language, Mutex class which is in System.Threading name space is used, before entering resource sharing area, WaitOne method of the resource related Mutex Object is called, process will check the Mutex Object first, if it’s free, lock the Object and enter the sharing area; if it’s occupied, just wait. At the end of resource sharing area, ReleaseMutex method of the Mutex Object is called to release the Mutex Object, so that other process can access the shared resource.

It can also use Enter and Exit methods of the Monitor class to solve the problem of competing for shared resources.

3.3 Hardware Access

System operating screen and trends need the latest values of the various measured items.

These measured items are divided into two groups, items in first group can be gotten via direct access to peripherals, such as pressure, PH value, etc.

Items in second group require a PC to send data requests to the peripheral through the port controller, the peripheral reply the required data after receiving request package. Such items include temperature and measured balance values.

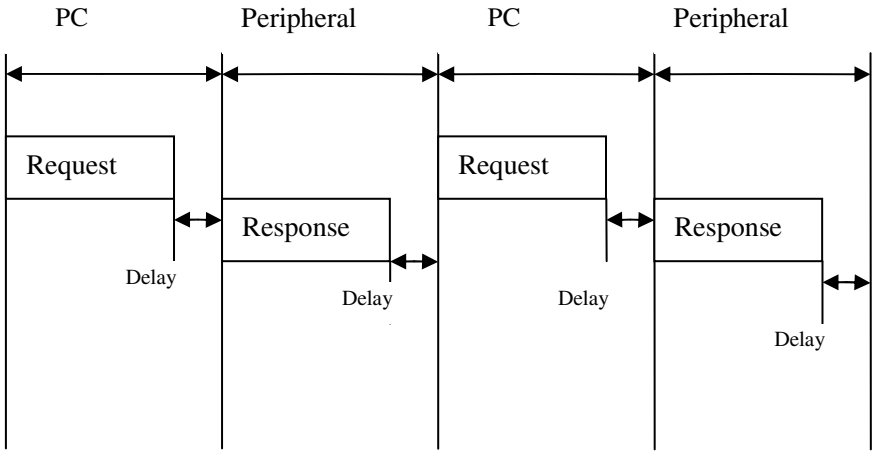


Fig. 5. Communication diagram for PC and peripheral(2nd group)

If the program directly using the I/O instruction to get the measured values, there will be a considerable delay before latest measured values can be fetched, especially for items in second group.

To solve this problem, asynchronous way is used to fetch data. We check the peripherals regularly(e.g. using one second as the interval), get the latest measured values and save them to memory variables, then the applications can obtain these measured values directly from memory variables. There is almost no delay in this way.

4 Conclusion

The paper discusses a design and implementation of the control and monitoring system for a multi-pot reactor. Compared with conventional software design, reactor control software design has its own specialty, a lot of attention should be paid to it. The paper discusses several problems encountered in development and presents their solutions, hoping that it can benefit to other developers who have similar problems.

Acknowledgments. This research is supported by the natural science foundation of Shanghai, China. No. 10ZR1410400.

References

1. TOHO ELECTRONICS INC. TTM-00BT Detailed Manuals (2009), http://www.toho-inc.com/
2. CONTEC Co., Ltd. Isolated Digital Output Module DO-32(FIT)GY Detailed Manuals (2006), http://www.contec.co.jp/
3. CONTEC Co., Ltd. Isolated Analog Output Module DAI12-4(FIT)GY Detailed Manuals (2006), http://www.contec.co.jp/
4. CONTEC Co., Ltd. Isolated Analog Input Module ADI12-8(FIT)GY Detailed Manuals (2006), http://www.contec.co.jp/
5. Christian Nagel, Bill Evjen etc. Professional C# 4 and.NET 4 (Chinese version). Tsinghua University Press (2010)

Study on the Measurement Method of the Impact Force of Projectiles with High Kinetic Energy

Changán Di[1], Lilong Gao[2], Yanjuan Li[3], and Deren Kong[1]

[1] Nanjing University of Science and Technology
[2] Xián Modern Chemistry Research Institute
[3] North Huaán Industry Group Co., Ltd
210094, Nanjing University of Science and Technology,
School of Mechanical Engineering
dica3@sina.com

Abstract. In order to make it simple to evaluate the vulnerability of explosives, the mechanical properties of the explosive loading source needs to be measured. By taking PVD films as sensitive elements, the design of the sensors, which can test the impact force, has been accomplished. The test system of impact force based on PXI has been constructed by means of charge integrators. In the end, the system got tested by 82 mm loading devices. And the pressure curve of the loaded 82 flat nose bullets were got as well. A method of impact force test is provided for the measurement of mechanical properties of loading source.

Keywords: Explosive, Vulnerability, Impact force, PVDF.

1 Introduction

The vulnerability of explosives and propellants is one of the important features to evaluate its properties. Besides the security system in the design, manufacture and transportation, the elements that may affect explosives and propellants vulnerability also include the security system in the service of its productions and its own. The method to assess the vulnerability of explosives and propellants in its design, manufacture and transportation are listed as followings: the diameter experiment in critical initiation, sensitivity experiment of drop hammer(percentage), friction sensitivity experiment, cap-sensitivity test(including the sensitivity test of adding booster explosives), drop sensitivity test. The evaluation method in service include rapid burning test, sensitivity test of shock wave, the sensitivity test of efflux , bump test of rapid bullets(including large west mulberry test and small shoot test of bullets), slow hot baked test(SCB) and launch sensitivity test.

During rapid projectile hit test, the stress and strain of explosives in the hit of high-speed projectile. The test is kind of destructive because of the rapid projectile's large mass, high speed and large kinetic energy. Meanwhile, the strain gauge will not be robust when directly pasted onto explosives because of the special characteristic of its materials. As a result, it will be sort of difficult to measure the stress and strain in the impact by means of photoelectric and strain measurement. It needs to measure the

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 595–599.
springerlink.com

loading signals and response signals to study the loading characteristic of explosives. The properties of explosives will be studied according to the captured signals. Based on these, an exclusive mechanical test device of explosives has been designed in this paper, with the purpose of accomplishing the impact force and pressure in the hit of high-speed projectiles and offering a test method for bullets' impact test.

2 Design of Measurement System

The schematic of impact test of piezoelectric film are shown in the fig.1. The system is composed of target assembly unit, charge converter 8 and data acquisition unit 9. The target assembly unit consists of front cover 4,piezoelectric film 5 and end cover 6. The piezoelectric film 5 is put between front cover 4 and end cover 6.

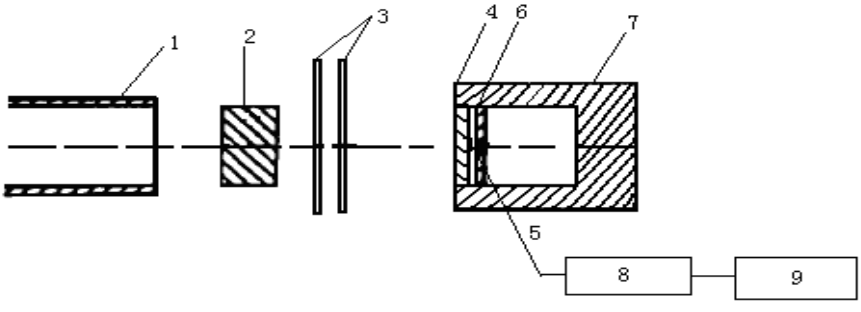


Fig. 1. Schematic diagram of device

1.load device 2. flat nose projectile 3.measurement device of speed 4.front cover 5.piezoelecric film 6.end cover 7.skewback 8.charge converter 9.data acquisition device

2.1 Piezoelectric Film

The piezoelectric film is light and soft. It is easy to be designed according to the demands from tests. It is of high piezoelectric properties. Namely, when we put pressure upon it in certain directions, there will be change in the opposite direction on its surface. Its piezoelectric properties are about 3 times to 5 times that of quartz. Because of the special characteristic of explosive load device's structure, the piezoelectric film of Polyvinylidene fluoride（PVF2） is taken to measure the impact force of high-speed projectile. The specifications of piezoelectric film are listed as followings: PVF2 11-.125-EK X 3 Inch. The range is 0-30GPa. The specific parameters are shown below:

- Size: 3.18×3.18 mm
- Piezoelectric film: PVF2, The thickness of piezoelectric film: 0.028mm
- Packaging materials: kapton, The thickness of packaging materials: 0.025mm
- Type: monopodium
- Effective area Ag(cm2): 0.1

2.2 Electric Charge Converter

The sensitivity of PVDF piezoelectric film is greater, if the follow-conditioning circuit and the electric charge converter are connected, it will be beyond the range of the electric charge converter input signal and causing signal overload, here using a charge integrator to conversion and measurement, directly supply the voltage waveform acquired by integral method to data collector, the range of measurement up to 0-300 Kbar. The electric charge converter model CI-50-0.1. The working principle of electric charge converter is shown in Fig.2.

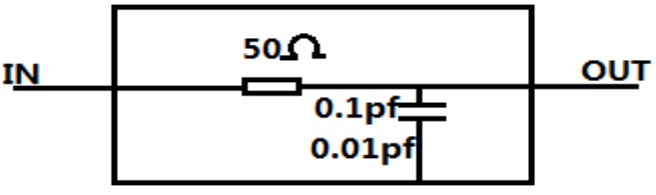


Fig. 2. The operation schematic diagram of electric charge converter

2.3 The Calibration of Piezoelectric Films Impact Force Sensor System

The system of piezoelectric films impact force sensor has been calibrated by the manufacturer, the dynamic calibration curve shown in Fig.3, it is conducting on the pneumatic fragment generating device, the calibrating range of pneumatic fragment generating device: 0 ~ 350kbar. In the factory, re-use of immediate switch device to conduct quasi-static calibration for each piezoelectric films to obtain the sensitivity of the coincident individual characteristics and correct its output signal amplitude.

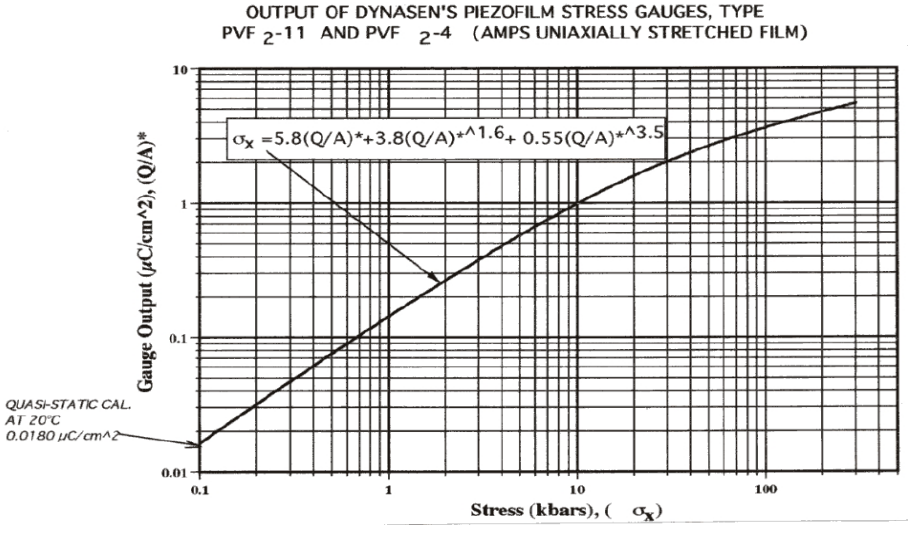


Fig. 3. The quasi-static calibration curve of piezoelectric films impact force sensor

Quasi-static calibration is conducting in the immediate switch device, using standard sensor (carbon resistance meter) compared with the piezoelectric film sensor to calibrate, calibration pressure 0.1kbar, pulse width 10ms. According to the calibration data and temperature influence, correct for each piezoelectric film, the revised results:

$$(\frac{Q}{A})^{*}(\mu C/cm^2) = (\frac{Q}{A}) \times \frac{0.0180}{C_S}\left[\frac{(1.0+0.01(20-T_t)(1-Pt/10))}{(1.0+0.01(20-T_c))}\right] \quad (1)$$

Among them: $(\frac{Q}{A}) = (\frac{V_m}{A_g}C_c)$, produced the charge number per unit area, $(\mu C/cm^2)$

C_S : sensitivity by single quasi-static calibration $(\mu C/cm^2)$

Tt: test temperature
Tc: calibration temperature
Pt: estimate of the measured pressure
Vm: output voltage of charge converter
Cc: capacitor value of charge converter μC
Ag: sensitive unit area of strain gage cm2

Output voltage can be calculated by the equations below:

$$\sigma_x = 5.8(\frac{Q}{A})^{*} + 3.8(\frac{Q}{A})^{*\wedge 1.6} + 0.55(\frac{Q}{A})^{*\wedge 3.5}\ (\text{kbar}) \quad (2)$$

We can work out the impact force after we multiply the pressure with the action area.

3 Test Situations

Flat-nose bullets are used to strike the bearing carrier which has powder insider. The speed is 450m/s. When the powder gets exploded, the impact pressure acquired is 18.5GPa. The curves of the data in this experiment is shown in fig.4.

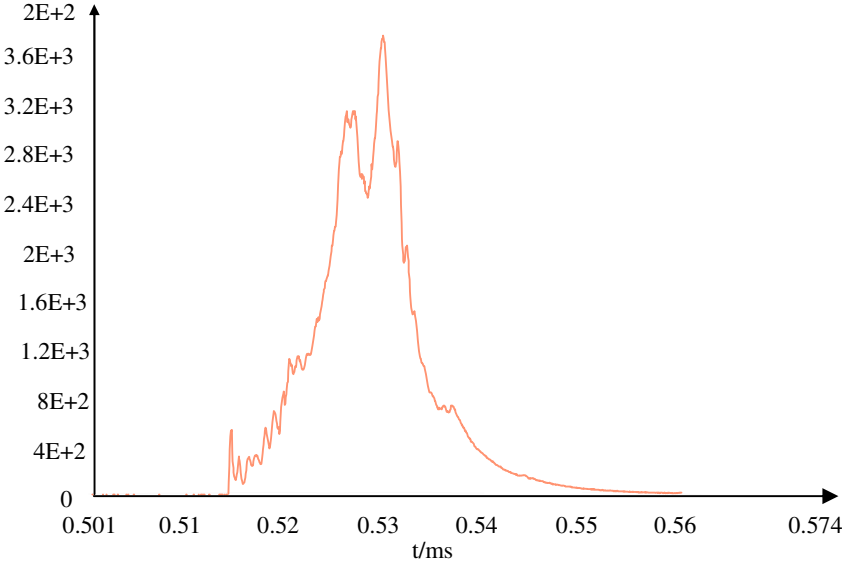


Fig. 4. The load force acquired by PVDF in experiment

From the figure above, we can get the information below:

a. The peak value of impact force is about 3700KN;
b. The rise time of load is 16.3 μs . When other conditions are stable, the rise time of load reflects the value of load speed. Namely, the shorter the rise time is, the quicker the load speed.
c. The impact wave reflects between the piezoelectric film and target materials for a while. As a result, there appear the high-frequency vibration processes at the wave top. After this signal, the impact wave pressure tends to balance. Moreover, the high-frequency vibrations may contain the interference signals brought out by the stray inductance in the circuits.

4 Conclusions

The high-speed load of bullets (82mm in diameter) has the characteristics of huge impact loading and high speed. In this paper, the exclusive sensor for impact force is worked out by means of PVDF film. Meanwhile, we accomplish the measurements of load force in certain experiments, and acquire the test curves of load force, which can provide measurement methods for impact force with high speed and huge loading.

References

1. Han, B., Zhang, X., Xing, Y.: Study on LOVA Performances Test of a New Propelling Charge. Chinese Journal of Explosives & Propellants 31(1), 53–55 (2008)
2. Jie, L., Zhu, X., Wang, G., Yuan, Y.: The Research of Sensing Mechanism of PVDF in Large Strain Condition. Chinese Journal of Sensors and Actuators 23(6), 896–898 (2010)
3. Wang, S., Hu, H.: The Effect of Different Charging Processes on the Launch Safety. Chinese Journal of Explosives & Propellants 26(1), 20–22 (2003)
4. Zhu, M., Zeng, S.: Finite element analysis of flat-nose bullets impacting rigid plane. Protection Engineering (1), 8–14 (1994)

An Efficient Method of License Plate Location Based on Structure Features

Fan Ye and Mei Xie

Image Processing and Information Security Lab, School of Electronic Engineering, University of Electronic Science and Technology of China, Chengdu, China
yefan19871108@163.com, xiemei@ee.uestc.edu.cn

Abstract. License plate recognition system is an important part of intelligent transportation system, while the vehicle license plate location is the key of it. In this paper, a new method of license plate (LP) positioning is proposed. In this method, firstly, the input image is preprocessed. Then, the texture and color features are used to remove most of disturbance. After that, the regions which have the structure and texture that is similar to the license plate are extracted. Finally, pseudo-region is eliminated and true LP is precisely located. The working principle is taking full advantages of plate texture, color characteristic and structure features to choose candidate regions. Experimental results illustrate that the method is robust and can increase the location accuracy distinctly.

Keywords: Vehicle license plate location, texture, color, pseudo-region, structure features.

1 Introduction

Automatic license plate recognition system has already been used in traffic surveying and monitoring, automatic toll, etc. Of all these application areas, license plate location plays an important role, and also it is the most difficult step.

Nowadays, the main methods of vehicle license plate location are as follows: the method based on projective invariance [1], the method based on Hough transform [2], the method based on the color feature of the vehicle license plate [3], the method based on morphology [4], the method based on the texture feature of the vehicle license plate [5, 6], and so forth. The presently available methods turn out to be not effective enough in the case of complicated background, varying illumination, and low-quality images.

In order to improve the location accuracy, this paper presents a new approach for vehicle license plate location based on structure features. The rest of this paper is organized as follows. In section two, the license plate location method is discussed in detail. Experimental result and analysis are presented in section three. Concluding remarks are given in section four.

2 Method for License Plate Location

This method consists of four stages which will be discussed in the following subsections.

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 601–606.
springerlink.com

2.1 Pre-processing

Original images often suffer from various types of degradation such as noises and low contrast. Therefore, the input images need to be preprocessed. In this proposed method, only the image brightness information is used. So, the original image should be converted into grayscale according to formula (1).

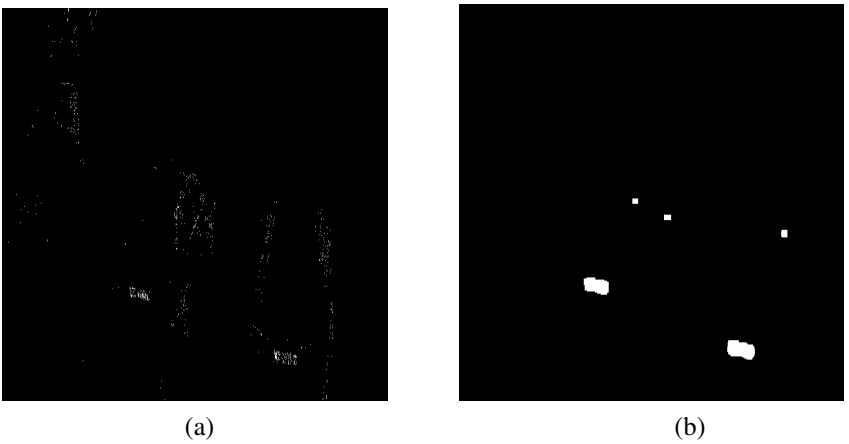
$$Gray = 0.299R + 0.587G + 0.114B \tag{1}$$

In order to filter out the small particle noises which maybe affect the positioning effect, median filter is chosen. Due to various reasons, the license plate is not prominent against the background, here, top-hat transform is employed to restrain the background and give prominence to the license plate region. Top-hat is a morphological operator. It can suppress the background of number plates and enhance plate regions, so it is very favorable to plate location.

2.2 Getting Candidate Regions

There is lots of texture information in the plate region, especially the vertical edge. This feature is employed for locating the plate in an image. The following steps which describe in details can get the candidate regions.

Step 1: Using the operator "sobel" to detect the vertical edge of the license plate, pay attention to choose the appropriate operator.
Step 2: Using the color feature to eliminate some disturbance of the edge detection image, because of the fixed color collocation of the license plate. Before using this feature, the gray image should be transformed into HIS image. After do that, the de-noise edge detection image is shown in Fig.1 (a), which will be based on to do mathematical morphology.
Step 3: Using mathematical morphology method to get the candidate regions. The structure element is 3*3, which must be selected appropriately; otherwise it will influence the final positioning accuracy. And the detail processes are: erosion, erosion, close operating, dilation, dilation, dilation. The results of these processes are connective regions. Fig.1 (b) shows these regions.



(a) (b)

Fig. 1. (a) The De-noise Edge Image (b) Results after Morphological Processing

2.3 Coarse Location

Chinese license plate has an obvious structure feature, namely the width height ratio is 3.14.Taking into account possible inclination of license plate, the width height ratio could be set between 1.5~5.The connected regions without the proper width height ratio are eliminated. If only one connected region remains, it could be regarded as the region of the vehicle license plate. Otherwise, the other three conditions which are listed as follows will be used.

Condition 1: Density, the density of the license plate is defined as follows:
From the vertical edge detection image, it shows that the density of the license plate region is bigger than that of non-plate region. The proportion of the white pixels (the character brackets after the edge detection) in the whole candidate region is defined as the density. Let N be the number of the white pixel in the whole candidate region, L is the width of the candidate region, H is the height of the candidate region, and D is the density, which can be calculated as the formula(2):

$$D = N/(L \times H) \tag{2}$$

The density (D) of the license plate is between two specific values, which can be a structure feature to exclude the pseudo-plate regions.

Condition 2: Using "ostu" threshold method to transform the gray image into the binary image *f(x,y)*.And then get the subtraction image *g(x,y)* according to the following formula(3):

$$g(x, y) = |f(x, y+1) - f(x, y)| \tag{3}$$

where, if $g(x,y)=0$, then $f(x,y)=0$, else $g(x,y)\neq 0$, then $f(x,y)=255$, x is the number of the row and y is the column. So the number of the white pixel of each row is equal to the gray level jump times in the gray image. The number of this gray level jump times is between two specific values when the area is the real license plate.

Condition 3: The variance of the plate region is smaller than that of pseudo-plate region in the image after edge detection. Calculating the variance (dF) of these candidate regions. The region that its dF is the minimal is the real license plate region. The formula (4) to calculate the variance is as follows:

$$dF = \sum_{i=1}^{L \times H} \frac{(X-a)^2}{L \times H} \tag{4}$$

where, X is the grayscale value of the candidate region, a is the mean of the grayscale value of the candidate region, L is the width of the candidate region, H is the height of the candidate region.

After using these structure features and the three conditions, it can exclude the pseudo-regions which not fit above features and get the coarse location of the license plate. The results are shown in Fig.2.

(a) (b)

Fig. 2. Results of Coarse Location

2.4 Precisely Location

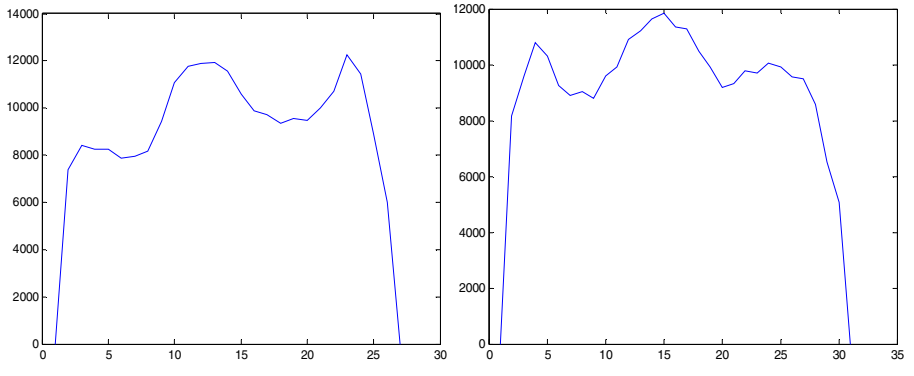
The results show that the coarse location is not precise. So the images need to be further processed. Using Hough transform to detect the tilt angle and make them paralleled before the precisely location. This precisely location method is divided into two parts, horizontal location and vertical location.

1) Level Location
In this paper, using the formula (5) below for the level projection, which is

$$T_h(i) = \sum_{j=1}^{n} f(i, j) \tag{5}$$

where, n represents the number of the column of the binary image *f(x,y)*, i is the row and j is the column.

The results are shown as follows (Fig.3):



(a) (b)

Fig. 3. Level Projection of the Image

From the level projection of the image, setting a threshold and using the first row that its level projection is bigger than the threshold as the up and down border of the license plate, so candidates of the license plate can be got. The level precisely location results are shown in Fig.4.

(a) (b)

Fig. 4. Candidates of the License Plate after the Level Location

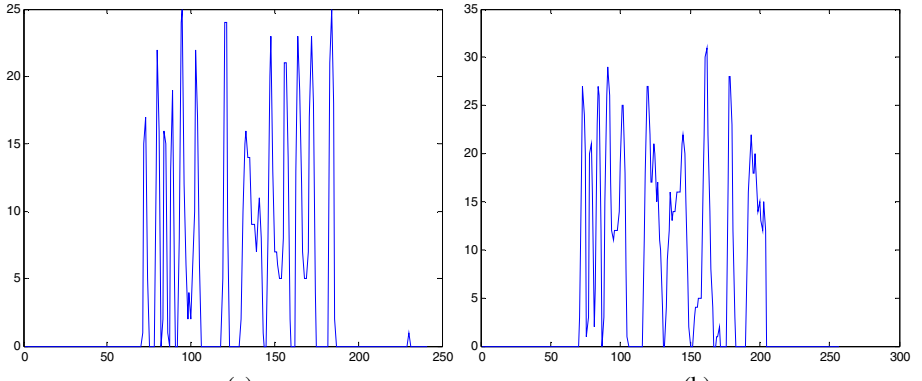
2) Vertical Location

During the vertical location, the integrated location method based on edge detection and vertical projection is taken. When converting the gray-scale image into binary image, selecting the "ostu" method and taking the formula (6) below to calculate the vertical projection, which is:

$$T_v(j) = \sum_{i=1}^{m} f(i, j) \tag{6}$$

where, m represents the number of the row of the binary image *f(x,y)*, i is the row and j is the column.

The results are shown as follows (Fig.5):



(a) (b)

Fig. 5. Vertical Projection of the Image

As the result, the left and right border of the license plate can be got that using the first column which its vertical projection is bigger than the specific threshold that we set according to different plate images. The finally located images are shown in Fig.6.

(a) (b)

Fig. 6. Final Location of the License Plate

3 Experiment Result and Analysis

A data base of 250 images, which are taken from many kinds of conditions, such as different angles and different lightening conditions, corrupted and stained license plates is used in the experiment. The experiment is oriented to Chinese LPs, and the result is illustrated in Table 1.

Table 1. Experimental result

Different condition	Total image number	Location number	Location rate
day	150	146	97.3%
night	100	95	95.0%

This experiment result reveals that the proposed method is robust and has high accuracy.

4 Conclusions

As indicated above, we can get good result using this method. It fully uses the special structures of the license plate and employs the fixed structure features of the license plate, which is very stable, so comparing to other method, it has a good performance. As a whole, our experiment shows that we can get good location accuracy with this method in different conditions, and the time for location is very short. So, this method can be applied to high real-time requirements of intelligent system.

Acknowledgements. This paper is supported by Image Processing and Information Security Lab.

References

1. Yang, W.P., Li, J.C., Shen, Z.K.: Technique of Automatic Orientation of Car License Plate Targets. Journal of Image and Graphics, pp. 835–839 (2002)
2. Yanamura, Y., Goto, M., Nishiyama, D., Soga, M., Nakatani, H., Saji: Extraction and Tracking of The License Plate Using Hough Transform and Voted Block Matching. In: Proceedings of IEEE Intelligent Vehicles Symposium, pp. 243–246 (2003)
3. Li, W.J., Liang, D.Q., Zhang, Q., Fan, X.: A Novel Approach for Vehicle License Plate Location Based on Edge-Color Pair. Chinese Journal of Computers 27(2), 204–208 (2004)
4. Suryanarayana, P.V., Mitra, S.K., Banerjee, A., Roy, A.K.: A Morphology Based Approach for Car License Plate Extraction. In: IEEE Indicon 2005 Conference, Chennai, India, December 11-13, pp. 24–27 (2005)
5. Liu, G.Q., Zhang, X.S., Zhang, X.B.: License Plate Location Based on Texture Characteristic of Image. Journal of Image and Graphics 10(11), 1420–1423 (2005)
6. Yang, J., Qi, F.H.: A License Plate Locating Approach Based on Shape and Texture Characteristics. Computer Engineering 32(2), 15–22 (2006)

The Multithreading Parallel ACA Algorithm Based on OpenMP

Yulei Zhao, Chuangming Tong, and Zhiqi Ju

The Missile Institute, Air Force Engineering University,
Sanyuan, Shaanxi Province, 713800
ylzhao1983@yeah.net, cmtong@126.com, no1jzq@yahoo.com.cn

Abstract. The Memory of the computers has been not able to satisfy large scale scientific calculation. With the coming of multicore era of the processor CMOS chip, the technique of multithreading has been an effective way to improve the efficiency of the program implementing. OpenMP is an application programmers interface (API) designing for the programming in the shared-storage multiprocessor. This paper introduces OpenMP to a new algorithm-the Adaptive Cross algorithm (ACA). And the efficiency of the program is improved greatly through the numerical experiments.

Keywords: The Adaptive Cross Algorithm, Multicore, OpenMP.

1 Introduction

The adaptive cross algorithm (ACA) which present firstly by Bebendorf in documentation [3]-[5] compresses the boundary integral operator by progressive smooth kernel. The ACA algorithm takes use of the character of rank-deficient of the coupled matrix block. The beauty of the ACA algorithm is its algebraic nature. That is to say its computational speed-up is carried out through the liner algebra manipulation, such as QR-decomposition, singular value decomposition (SVD), LU decomposition and so on. So the development and implementation of the ACA algorithm do not dependent on the complete knowledge of the integral kernel, primary function or the integral equation itself. And because of its pure algebra nature, it has high commonality, and it can be modular and easy to be transplanted to all kinds of MoM code. In general, The ACA is a multilevel matrix partition technique which implements rank-revealing LU decomposition to rank-deficient matrix to reduce the computational cost.

As all known, the impedance matrix based on the MFIE and EFIE is nonsingular full rank matrix. But because of the character of the Green's function, this full rank matrix is made of lots of rank-deficient matrix blocks, such as, the matrix represent the coupling of separate targets is rank-deficient. That is, the coupled matrixes which represent the coupling relation of two groups can be expressed by a few column vectors accurately. To this condition, the ACA groups all elements as shown in Fig. 1. The way of group is similar to [4] and [7]. In this way, the original nonsingular full rank matrix is decomposed to lots of rank-deficient sub-matrix of all kinds of size.

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 607–614.
springerlink.com

The diagonal block sub-matrix which represent self-coupling and the sub-matrix be adjacent to the diagonal will be solved with MoM. While the remains which represent the coupling of far groups can be compress by ACA because of its nature of rank-deficient. By this way, these sub-matrixes of all kinds of size can be reduced effectively.

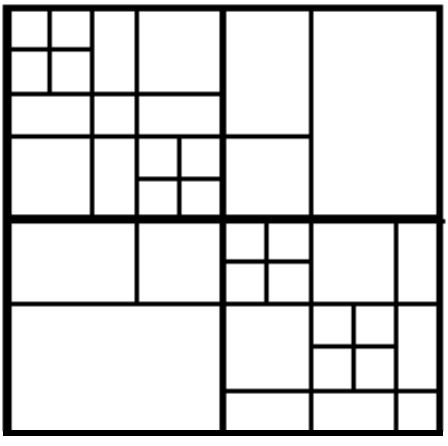


Fig. 1. Matrix division

The character of the ACA algorithm indicates its good parallelism. So it can performance computing by different processor core simultaneously to improve the efficiency of the program, which is the so-called multithreading.

Now the processor chip is in the era of multicore, so that multiple process or threads can be performed simultaneously but not like the single core environment in which the CPU can only be used by turns. But the traditional programs are written for single core, the most of which can not be accelerated directly on the multicore computers.. In order to improve the efficiency of the program, it needs to divide the single computing task to multiple parallel sub-tasks, and perform these sub-tasks on different processor cores. So studying the Implementation technology of parallel computing on multicore PCs has The important practical significance. OpenMP is an application programmers interface (API) designing for the programming in the shared storage multiprocessor. It is an Industrial standard that Support for shared-memory parallel programming [15]-[18].

OpenMP consists of a set of compiler directives and library routines that extend FORTRAN, C, and C++ codes to express shared-memory parallelism. For the synchronization of shared variables, reasonable distribution of load and other tasks, it provided effective support with the characteristics of simplicity and rapid development. OpenMP is an industry standard of portable multi-threaded applications developping, with high efficiency at the fine-grained(circulating levels) and coarse-grained (function level) thread technology. For the serial application program converting into parallel applications, OpenMP instruction is a powerful tool and it is

easy to operate. It has the potential to substantially improve the performance of the application because of the OpenMP implementing synchronously on the symmetric multiprocessor or multi-core systems. OpenMP automaticly turns a loop to threads on multiprocessor systems to improve the performance of applications. The user does not have to deal with iterative partitioning, data sharing, thread scheduling and synchronizing, and so on.

It offers a simple way of exploiting parallelism without interfering with algorithm design; an OpenMP program compiles and operates correctly in both parallel and serial execution environments. Using OpenMP's directive-based parallelism also simplifies the act of converting existing serial code to efficient parallel code.

The most important part of OpenMP is the compiling directive sentence. In C/C++, the format of the compiling directive sentence is

```
#pragma omp directive [clause[clause]…]
```

Such as "`parallel for`" is a piece of directive, the clause after which is selectable. For example, `#pragma omp parallel for private(j)`, where `private` is a clause.

The main directive of OpenMP is:

`Parallel`: using before a code segment to indicate that this code segment will be performed by multiple threads at the same time;

`for`: using before the loop of `for` to distribute the loop to multiple threads to be performed at the same time. It must be independent between each loop;

`parallel for`: the combination of `parallel` and `for`, also being used before the loop of `for` to indicate that the code of `for` loop be performed by multiple threads at the same time;

`sections`: using before the code segment that may be performed synchronously;

`barrier`: for the thread synchronization of the code in the parallel area. All of the threads performing will not stop until to the `barrier`.

OpenMP is a parallel compiling model based on threads, Fig 2 is the Fork-Join model that OpenMP usually used. A OpenMP program executes from a single thread, which needs to be executed at some points the program will spawn additional threads, and makeup a thread group. These threads execute synchronously in a code area called parallel area until to next parallel area(or to the end of the program).

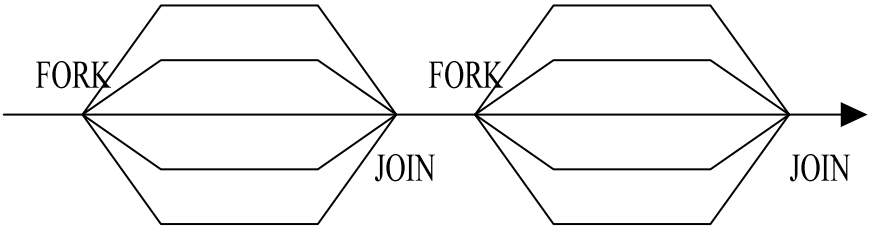


Fig. 2. FORK-JOIN model

2 The ACA Algorithm Based on OpenMP

For a matrix $Z^{m\times n}$ with size of $m\times n$ ($Z^{m\times n}$ is a sub-matrix of the impedance matrix), the purpose is to find two matrix $U^{m\times r}$ and $V^{r\times n}$ to compress the $Z^{m\times n}$ with the ACA algorithm. That is

$$Z^{m\times n} \approx Z_{app}^{m\times n} = U^{m\times r} V^{r\times n} = \sum_{i=1}^{r} u_i^{m\times 1} v_i^{1\times n} \tag{1}$$

$U^{m\times r}$ and $V^{r\times n}$ are two full rank matrix, where *r* is the rank of $Z^{m\times n}$. The bigger the *r* is, the better the $Z_{app}^{m\times n}$ approach to $Z^{m\times n}$. The condition to end the iteration can be set as

$$\frac{\left\|u_r\right\|_F \left\|v_r\right\|_F}{\left\|Z_{app}^{m\times n}\right\|_F} \le \varepsilon \tag{2}$$

Where ε is the normalization error, which in general takes 10^{-4}. If $\varepsilon = 0$,

$$\frac{\left\|u_r\right\|_F \left\|v_r\right\|_F}{\left\|Z_{app}^{m\times n}\right\|_F} \le \varepsilon \tag{3}$$

It only needs part of the knowledge of the original matrix to approximate it with the ACA algorithm. The memory complexity is $O(r(m+n))$. The operation times per iteration of the 1st step to 4th step of the kth iteration is $O(r(m+n))$. There are r iterations, so the complexity of the CPU time is $O(r^2(m+n))$. What is interesting is if the $Z^{m\times n}$ is full rank, the ACA algorithm is column Gaussian elimination in nature. The algorithm is pure algebraic, which predicate that it doesn't need to make clear every element of the matrix like the FMM.

Documentation [1] is the detail of the ACA algorithm.

The calculation of row and column of the original matrix are independent when The sub-matrix of the ACA algorithm approximating the original matrix. There is no data exchange, so it is conveniently to use OpenMP to program. Fig. 3 is the model of the ACA algorithm based on OpenMP.

In the parallel area of the OpenMP, all threads' data in default are mutual shared, so it is easy to cause read-write error of thread data. To solve this problem, it needs to declare the data scope explicitly. The scope is used to control whether a variable in different threads is shared or private. A compiling directive statement may contains more than one data scope clause, but a variable can appear only in one scope clause. That is a variable can not be both shared and private. What is needs to pay attention is that in the parallel area of nest loop, if the parallelization compiling directive sentence in the outer loop, the circulation index of inner loop must be declared to be private.

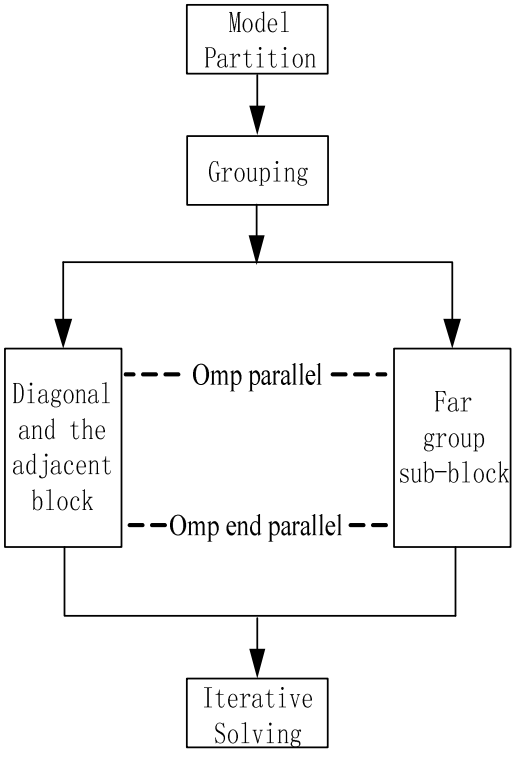


Fig. 3. OpenMP+ACA model

3 The Numerical Results

Fig.4 is a model made of two PEC spheres the separation of which is 3λ, and their radius are both 1λ. Subdivide the model with triangle surface element. The maximal length of side takes $\lambda/8$, accordingly, the triangle surface element number of sphere1(sphere2) are 867. The matrixes generated by the interacting of the two spheres are A_{12} and A_{21}. Take the A_{12} for example, A_{12} is a $N_{12} \times N_{12}$ scale matrix (N_{12} = triangle surface element number $\times 3$, is the number of the unknown). Every element is a complex made of two double type data. Table 1 represents the memory comparison between the ACA algorithm and the traditional MoM.

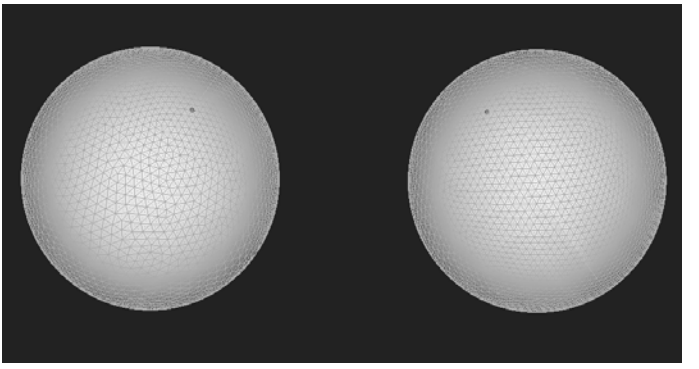


Fig. 4. Two PEC Spheres (the radius is 1λ , the separation is 3λ)

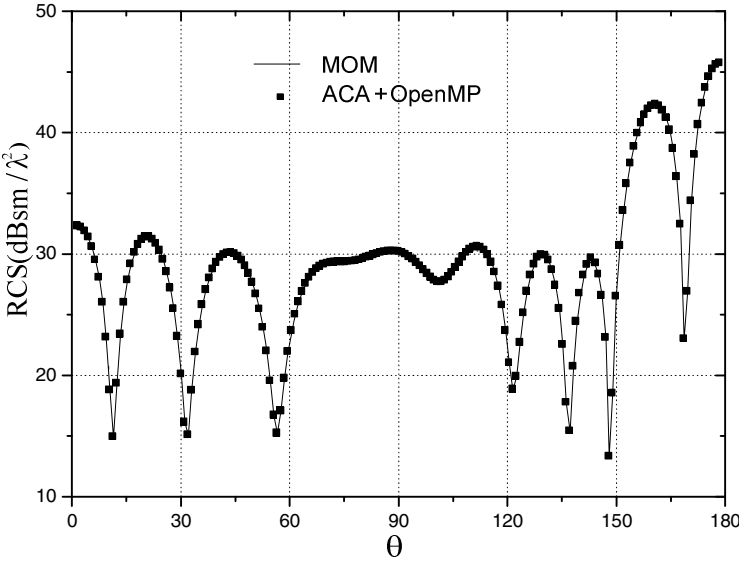


Fig. 5. The RCS of the two PEC spheres

Table 1. The comparison of times, and memory overhead

	Time(s)	Memory overhead	speedup
MoM	954	432.8 MB	-
ACA	446	6.32 MB	-
OpenMP+ACA	235	6.32 MB	1.90

4 Conclusions

OpenMP is a tool designed for multicore parallel computing, in other words, it is more suitable for parallel computing on single computer shared memory structure.

Due to the use of threads shared memory between the coordinate parallel computing. It has high efficiency, small memory overhead as well as simple and intuitive programming statements on multiprocessor systems. So it is easy to program (now the latest version of C, C + +, Fortran compiler basically have built-in OpenMP support). In this paper, OpenMP is combined with the ACA algorithm. Because of the good parallelism of the ACA algorithm and the advantage of the OpenMP on multithreading development, their combination reflects huge advantage at the execution efficiency of the program. (From the numerical results, the speedup of the parallel code is about 1.9 times of the serial code).

References

1. Zhao, K., Vouvakis, M.N., Lee, J.-F.: Accelerated Method of Moments Computations of EMC problems. IEEE Trans. Electromegnetic Compatibility 47(4) (November 2005)
2. Geng, N., Sullivan, A., Carin, L.: Multilevel Fast-Multipole Algorithm for Scattering from Conducting Targets Above or Embedded in a Lossy Half Space. IEEE. Trans. Geoscience and Remote Sensing 38(4), 1561–1573 (2000)
3. Bebendorf, M.: Approximation of boundary element matrices. Numer. Math. 86(4), 565–589 (2000)
4. Kurz, S., Rain, O., Rjasanow, S.: The adaptive cross-approximation technique for the 3-D boundary element method. IEEE Trans. Magn. 38(2), 421–424 (2002)
5. Bebendorf, M., Rjasanow, S.: Adaptive low-rank approximation of collocation matrices. Computing 70(1), 1–24 (2003)
6. Ostrowski, J., Andjelíc, Z., Bebendorf, M.: Fast BEM-Solution of Laplace Problems With H-Matrices and ACA. IEEE Trans. Magn. 42(4), 627–630 (2006)
7. Lu, C.C., Chew, W.C.: A multilevel algorithm for solving a boundary integral equation of wave scattering. Microw. Opt. Tech. Lett. 7(10), 456–461 (1994)
8. Song, J.M., Chew, W.C.: Multilevel fast multipole algorithm for solving combined field integral equation of electromagnetic scattering. Microw. Opt. Tech. Lett. 10(1), 14–19 (1995)
9. Rokhlin, V.: Rapid solution of integral equations of classic potential theory. Comput. Phys. 60(1), 187–207 (1985)
10. Rokhlin, V.: Rapid solution of integral equations of scattering in two dimensions. Comput. Phys. 86(2), 414–439 (1990)
11. Coifman, R., Rokhlin, V., Wandzura, S.: The fast multipole method for the wave equation: A pedestrian prescription. IEEE Antennas Propag. Mag. 35(3), 7–12 (1993)
12. Bleszynski, E., Bleszynski, M., Jaroszewicz, T.: AIM: Adaptive integral method for solving large-scale electromagnetic scattering and radiationproblems. Radio Sci. 31(5), 1225–1251 (1996)
13. Wang, C.F., Ling, F., Jin, J.M.: Adaptive integral solution of combined field integral equation. Microw. Opt. Tech. Lett. 19(5), 321–328 (1998)
14. Phillips, J.R., White, J.K.: A precorrected-FFT method for electrostatic analysis of complicated 3-D structures. IEEE Trans. Comput. Aided Des. Integr. Circuits Syst. 16(10), 1059–1072 (1997)

15. Chow, E., Hysom, D.: Assessing Performance of Hybrid MPI/OpenMP Programs on SMP Clusters. R. Tech. Report UCRL-JC-143957 (2001)
16. SuChuan Dong, M., Karniadakis, G.E.M.: Dual-level parallelism for high-order CFD methods. In: Supercomputing Conference (2002)
17. Karniadakis, G.E., Kirby II, R.M.: Parallel Scientific Computing in C++ and MPI. Cambridge University Press (2003)
18. Chapman, B., Jost, G., Pas, R.V.D.: Using OpenMP. The MIT Press (2007)

Algorithms with Restrictive Input Constrains on Genome Exemplar Distance Caculation

Yehong Chen[1], Daming Zhu[2], Weifang Tang[1], Qi Chen[1], and Jie Zhang[1]

[1] Shandong Polytechnic University, Information School, Shandong Jinan, China
[2] Shandong University, Computer Science and Technology Institute, Jinan, China
chenyh@spu.edu.cn

Abstract. The comparative study of gene order rearrangements has been restricted to the case when the genes in one genome are homologous to at most one gene in the other genome. Sankoff (1999) proposes the estimation of true exemplars, a selection of one gene from each gene family in both genomes such that the distance between the resulting exemplar strings is minimized, which is called exemplar distance. David Bryant in his paper show that the calculation of the exemplar distance between two genomes is NP-hard for both the signed reversals distance and the breakpoint distance, even with quite restrictive conditions on the input. Limite the input according to strictive conditions used by David Bryants' proof, here we provides two algorithm to solve problem under some input conditions. one is branch and bound algorithms ,the other is Greedy Algorithm for both distances .different DATA can use different Algorithm to get better performance.

Keywords: Exemplar Distance, NP-HARD, Branch And Bound Algorithm, Greedy Algorithm, vertex cover problem, Set cover.

1 Introduction

Recently, Sankoff (1999) has proposed a method for estimating which of the multiple copies of a gene in two genomes G, and H best reflects the position of the ancestral gene in the common ancestor genome of G and H. The selection must result in the less distances between both genomes.

David Bryant showt hat both of the exemplar distance(EBD and ERD)problems posed by Sankoff (1999) are NP-hard, even with quite restrictive conditions on the input.

Theorem 1. The calculation of either the EBD or the ERD between two pegged genomes G and H is an NP-hard problem, even the biggest gene family only have two numbers.

In David Bryants' paper ,he proof Theorem 1 and Lemma.

Although, David byant proof the problem is HP-hrad, we all so need to Design practical algorithm to solve the actual problem.

Accord to Theorem 1,we show a branch and bound algorithms for both distances, the input genomes must be pegged one, and every gene family have numbers no more than two.

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 615–622.
springerlink.com

2 Definitions

We will use the same notation as Sankoff (1999). Given an alphabet A, let G and H be two strings (genomes) of signed (+ or -) symbols (representing genes) from A, of lengths lG and lH, respectively. For each a ∈ A, let kx(a) be the number of occurrences (+ or -) of symbol a in genome X. All occurrences of the symbol a in both genomes are said to constitute a gene family, the "a family". For our purposes, that the genes in a family are not exact copies is immaterial; we simply assume that the families have been constructed correctly.

A gene is singleton in a genome if it is the only member of its family in that genome. A genome is pegged if every pair of genes from the same gene family is separated by at least one singleton.

Theorem 1. The calculation of either the EBD or the ERD between two pegged genomes G and H is an NP-hard problem, even when $k_G(a) <= 2$ and $k_H(a) <= 2$ for all a ∈ A.

Lemma. The calculation of the EBD between two pegged genomes is NP-hard even when $k_G(a) = 1$ and $k_H(a) <= 2$ for all a ∈ A.

Then gives two restricted examples in this format, the problem is simplified format problem but still is NP-hard problem, from these perspective formats of input, this two kinds of case is low complexity, firstly, we try to give a more restricted algorithm on these formats, at the future, we should gradually pursuit of less restricted algorithm.

Case 1: assume that the two genome, repeating the same gene , in the same length, considering only breakpoints and reversals distance.

In order to make the problem simplify, we give some constraints on input data。

Let

$$G = V_1 £_1 V_1' V_2 £_2 V_2' \ldots Vn £n Vn'$$

and

$$H = V_1 - £_1 V_1' V_2 - £_2 V_2' \ldots\ldots Vn - £n Vn'$$

Here H is corresponding to G, we only need to consider the input of G.G is a string have pegged gene form, Vi and Vi'(1=<i<= n) is singleton, corresponding to A vertex of the graph. £ = {el,e2,... ,em}, ei is gene copies which corresponding to a edge of the graph. $£_i$(1=<i<= n) is a subset of £,which include all the edge linked to V_i These pairs insert in the string as number chars. Every edge only two copies showed between the two linked vector area.

David Bryant claim that:

(1) G corresponding to an vertex cover of sizeλ and only if the EBD between G and H is at most 2λ。
(2) G corresponding to an vertex cover of size λ if and only if the ERD between G and H is at most λ.

Case 2: The two genomes now become

$$G=V_1V_1'V_2V_2'\ldots V_nV_n'X_1\text{-}e_1X_2\text{-}e_2X_3\ldots X_m\text{-}e_mX_{m+1}$$

and

$$H = V_n £_n V_n' V_{n-1} £_{n-1} V_{n-1}' ..... V_1 £_1 V_1' X_{m+1} X_m ... X_2 X_1$$

From David Bryants' paper, we get limited input format and the relevant undirected graph

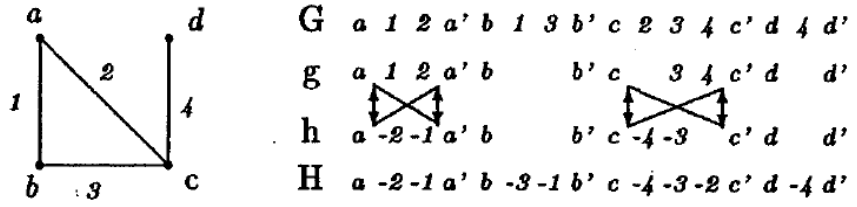


Fig. 1. An example of the reduction from VERTEX COVER. On the left a graph g with four vertices and four edges. On the right we have the genomes G and H, and the exemplar strings corresponding to the vertex cover {a, c}. We represent the ,breakpoints by vertical arrows and the two reversals required by dotted lines.

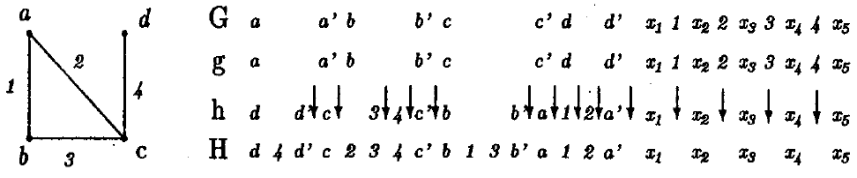


Fig. 2. An example of the second reduction from VERTEX COVER. On the leftb a graph $\mathcal{G}$ with four vertices and four edges. On the right we have the genomes *G* and *H*, and the exemplar strings corresponding to the vertex cover {a,c}. The breakpoints of *h* with respect to *g* are marked by arrows.

Case 1, or in case 2, the main problems are found the minimum vertex set corresponding, to eliminate the extra genes。 only the vertex In covered set are not empty, the algorithm is made to the deleted empty vertex as more as possible. For k_G (a) < = 2,and k_H (a)< = 2,according these two conditions, we can design the branch and bound algorithm, make a good solution, when the search tree branch more than than 2, reduce the search efficiency, then we will explore the effective the approximate solution of the optimal solution, make a good solution of the algorithm.

3 Branch and Bound Algorithm

We have designed a algorithm, Build a search tree use string G as root node . The tree of each node contain data: genome strings and the ei what need of reduction, the current on the edge of the vertex cover number vc.

The root node reduction, delete the first appearance ei, get left node, building the left subtreethe of root node; delete the second ei, get right node, structure right nodes.

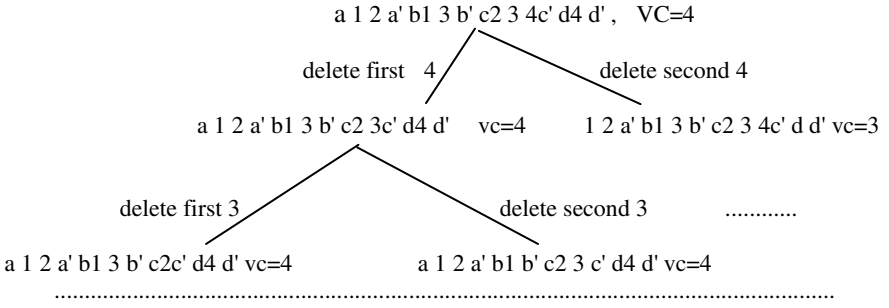


Fig. 3. Search tree building

As So on, the high is Equal to edge counts, every lay down, the vc at most Reduced by 1.so when we build to the leaf node we get the exemplar string. We found the search tree height is is m, which is the pair numbers of all the recurring genes that contained in the root node , and search tree each drop a layer, may only a vertex may be emptyed, vc is most minus 1.

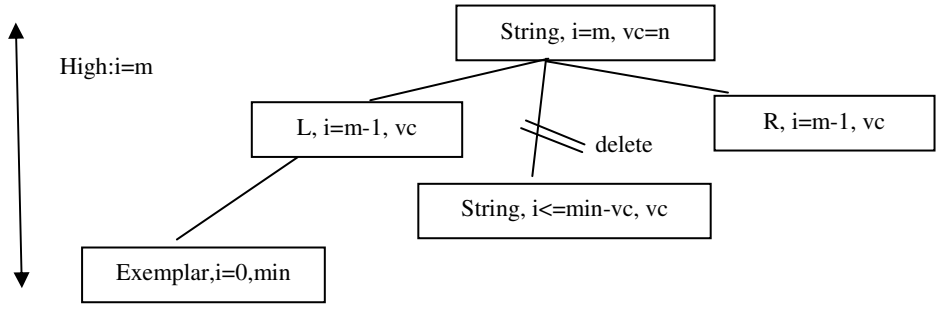


Fig. 4. Branch cutting

We use the strategy of depth first search tree structure, the most the left branch of the tree from root to leaves was firstly constructed . when the first leave have been made , we get a possible solution, the structure work of the later subtree can be contrasted beyond the min solution, only might come to get the more small distance of the branch can continue to down structure. Set out variable Min equal to vc of the first leave , because the height of tree is fixed, each layer down vc at most minus 1.consider the subnode, need to reduction i, sub tree height is i, due to the current vc, if i<=Min-vc, this sub branch cannot to get the more less possible leaves which vc small than min, we can give up this one node building work, thus greatly improved the search tree structure time and space of the price. Finally the leaves of the node is built for simplifying the optimal results, the optimal string is not the only, but the sample is the same distance, can produce minimum sample distance.

4 Pseudo-Code

```
  // Global variables defined
  /*deposited the final reduction String, at first an
empty String is empty.*/
  String Exem;
  /*record the current minimum vertex cover, a number of
initial value ∞.*/
  int Min=32767;
  Main()// main program
  {input: string G,  Int m, Int n ;
   /*G is string for reduction , m is the number of genes
in G twice,n is associated with the edge of the
vertices;*/
  //G is roots, establish reduction search tree.
  Buildingsearchtree (G, m,n);
  /*EBD(exemplar breakpoint distance) is 2*Min,
ERD(exemplar reversals distance) is Min.*/
  Output Min;
  output Exem; }//reduction of samples。

  //structural search tree
  Void Buildingsearchtree (string G, int i, int vc)
  /*G :the root node of the bunch, i: edge to reduction ,
vc : covered vertices*/
  {if (i <=0) // as the leaves
  // onlyleaves get minimum node vc can be go to.
  {   Min=vc;
   Copystring(Exem,G);
   Return;
  }
  String L,R; // left and right the son of root.
  Int Flag=0;
  /* said into the son tree, whether or not vc have
decrease*/
  CopyL(G,L,i,flag);
  /*from G get L, delete the first appearance of i, flag
said whether or not decrease vc */
  If i-1>(vc+flag-min)
  /*according to the cut conditions ,one son tree can get
more small vc was built.*/
  Buildingsearchtree (L, i-1, vc+flag);
  CopyR(G,R,i,flag);
  /*from G get R, delete the last appearance of i, flag
said whether or not decrease vc */
  If i-1>(vc+flag-min)
  /*according to the cut conditions ,one son tree can get
more small vc was built. */
  Buildingsearchtree (R, i-1, vc+flag);
  }
```

```
  / * from root string of G get left the son of L, delete
the tree first appearance of i, flag said If or no
decrease vc */
  Void CopyL(string G,string L,int i,int flag)
  {flag=0;Int j=0;
  While(((g[j]-'0')!=i))
   {L[j]=G[j]; j++;}
  If ([j-1], g[j+1] both are not number char)
//pegged area empty
  Flag=-1;
  While(g[++j]!='\0')L[j]=G[j];
  L[j]='/0';
  }
  / *from root string of G get right the son of R, delete
the tree second appearance of i,flag said if or no
decrease vc */
  Void CopyR(string G,string R, int i, int flag)
  {flag=0;
  Int Len=lenth(G);
  Int J=len-1;
  While((g[j]-'0')!=i)
   {L[j]=G[j]; J--;}
  If ([j-1], g[j+1] both are not number char)
   //pegged area empty
   Flag=-1;
  While(j>=0) {L[j]=G[j];j--}
  }
```

5 Problem Extends

Vertex cover problems can be extended as the set covering problem, when figure of edge is not related to the top of two side, but more than two, vertex cover problems became the set covering problem. When the "vertex" is not associated to pairs of vertex, but a set of vertices, the union of set which include all vertices in it is the set covering, so vertex cover can be regarded as a kind of the set covering as |set|=2. Also, theorem 1 can also expand to theorem 2.which prove the gene repeat number of sample larger than 2 of EBD and ERD problem are also HP-problems .

Keep the case 1 and case 2 described input form in the same , problems can be expanded to $k_G(a) > 2$, $k_H(a) > 2$ for all a $\in$ a. this problem is still HP-HARD, at this point, the corresponding hp-c problem is the set covering problem. Vertex cover problem can be seen as a set number of elements in the set covering for 2.

Theorem 2. The calculation of either the EBD or the ERD between two pegged genomes G and H is an NP-hard problem, even when $k_G(a) > 2$ or $k_H(a) > 2$ for $\forall$ a, a $\in$ A.

Proof. We provide a reduction from SET COVER:

SET COVER

Instance: g = (V,F).Integer λ.

Question: Is there ,V' In V, such that |V'| = A and each elements in F is adjacent to at least one element in V'?

Let g = (V, F) and λmake up an arbitrary instance of SET COVER with V = {V_l,V_2,... ,Vn}, and F = {f_l,f_2,... ,fm}. We construct an alphabet A of size 2n + m given by

$$A = V \cup \{V'_i : V'_i \in V\} \cup £.$$

For each i = 1,..., n let $£_i$ be a string of the symbols e_j for the set that are adjacent to V_i. Let - $£_i$ denote the string e_i with order reversed and opposite, signs-the signed reversal of e_i.

Let

$$G = V_1 £_1 V_1' V_2 £_2 V_2' \ldots Vn £n Vn'$$

and

$$H = V_1 - £_1 V1' V2 - £_2 V2' \ldots Vn - £n Vn'$$

Here H is corresponding to G, we only need to consider the input of G.G is a string have pegged gene form, Vi and Vi'(1=<i<= n) is singleton, corresponding to A vertex of the graph . £ = {el,e2,... ,em}, ei is gene copies which corresponding to set of set of the V, $£_i$(1=<i<= n) is a subset of £,which include all the set linked to V_i, These sets insert in the string as number chars. Every set have more than two copies showed between the two linked vector area.

We claim that

(1) g has a set cover of size λ and only if the EBD between the G and H is at most 2λ
(2) g has an set cover of size λ if and only if the ERD between G and H is at most λ.

For case 2,the proof is ease to give.

Then, Our problem is for the set cover efficient algorithm, because here the tree branch can be greater than 2, search tree become Multi-abifurcation tree , even though branch and bound algorithm can work, but the running efficiency is not very good, so we explore below a approximate solution of the problem of the set covering.

6 Approximate Solution

Above gives is general optimal solution method, according to the observed in some distinctive instances we can get some quick approximate solution method. For example, by observation , we found in the genome appear some gene duplication continuous string, that length is bigger. we can choose to keep these big word string, then as more as possible to delete empty the other strings. we assume that choice of the largest pegged substring , the time complexity can achieve polynomial finish.

Solving the sample genomeal algorithm is described.

1. To traverse G, find out the longest a substring" ViEiVi ' ", |EI | have maximum
2. To traverse G, delete all not in Ei, belong to the other pegged sub_stringnot .
3. If there is remain unempty pegged string need to reduction , choice the largest one, then goto 2.

Example:

a 1 2 a' b1 3 b' C2 3 4C' d4 d' ------>A1a'b1b'c234c'dd'------>A1a'bb'c234c'dd'

Above for kG (a) > 2 or kH (a) > 2 is also applicable. Instead, if the use of the branch and bound problems will rise, great complexity for 2 bifurcation tree into many bifurcation tree, but used the above approximate solution for, sample reduction algorithm complexity not increase. So the approximate algorithm is more general.

Above the approximate algorithm ,the approximate performance ratio is related to the length of pegged sub string distribution , if the length distribution is more balance, use the approximate algorithm is reduced, the efficiency of approximation is lower, but if the clause length difference is very big, first to the longest clauses do gene retained, approximate degrees and efficiency are ideal.

We have a finally law: For a given G string, considering only EBD and ERD.

1. First to G of all the clauses in length pegged sort, from large to small in processing.
2. The clauses for large, with approximate algorithm is reduction
3. Remaining clauses length gaps, with the branch and bound method reduction.

References

1. Sankoffd: Genome rearrangements with gene of families. J. Bioinformatics. 15, 909–917 (1999)
2. Sankoff, D., Nadeau, J.H. (eds.): Comparative Genomics. Kluwer Academic Publishers, Netherlands (2000)
3. Zhu, D., Ma, S.: Algorithm design and analysis. Higher Education Press, China (2009)

The Research Based on Self-adaptive Filter and Kalman Filter on Multi-robot

Lidan Gao[1], Zhonghua Huang[2], Nu Zhang[3], Shunhua Li[4], and Yan Feng[5]

[1] School of Mechatronical Engineering, Beijing Institute of Technology, Beijing, China
[2] Machinery Manufacture LTD.CO.Northeast Industries Group Corporation, Jiangbei Jilin, China
[1] douningmeng@163.com, [2] huangzh@bit.edu.cn,
[3] crossbow.zn@gmail.com, [4] lishunhua999@163.com,
[5] fengyan0211@126.com

Abstract. The paper presents a method of ultrasonic and wireless network to the issue of co-location on multi-robot, and proposes a method that makes use of robot network and self-adaptive filter to achieve to process the signal of ultrasonic. Simulation results show that the proposed method can detect the signal of ultrasonic while there exists the complex ultrasonic noises in the environment which have the same frequency with the transmitting ultrasonic, and complete the task of robot positioning in the case of multi-robot collisions.

Keywords: Self-adaptive filter, ultrasonic, multi-robot, co-location.

1 Introduction

In recent, the technology of co-location on multi-robot has been gaining more and more attention and has become an active area of research. If the robots were able to detect their peers and can exchange information with peers, you can use the robot relative observation information to improve their positioning accuracy, especially for heterogeneous robots, it is more obviously advantages in the area of co-location[1].

Our research proposed a method to the problem that use ultrasound positioning and wireless networks in multi-robot co-location, and to make use of robot networks and adaptive filtering method to solve the issue of using ultrasonic signal processing and kalman filter to complete co-location and integration of self-positioning.

2 The Principle of Multi-robot Co-location

The technology of co-location is the use of multi-robot observation group information signs in an unknown environment to communicate with each other, and through information exchange, it can obtain and share information on environmental measurements with other robots. According to the robot system, it sends ultrasound and takes the controlling robot as a base station in multi-robot co-location system[2].

Suppose a group of at least two executive robots, When the robots both have ultrasonic receivers and transmitters, they can measure the distance between base station and their respective, and the distance between two robots, each step with at

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 623–629.
springerlink.com

least one robot at rest, and with the base station connection point of deviation from the x-axis is known, you can calculate the location of another robot.

Multi-robot co-location shown in Figure 1,Small circles represent robots. Assumed that the multi-robot system composed of three robots, one is controlling robot, the second one is for environmental exploration, and the last robot is to aid positioning, their respective tasks as followed:

K—level 2 controlling robot, used for receiving ultrasound, calculating the distance.

R_1—environment detecting robot, send ultrasound, used for positioning.

R_2—assisted positioning robot, used for receiving ultrasound, calculating the distance.

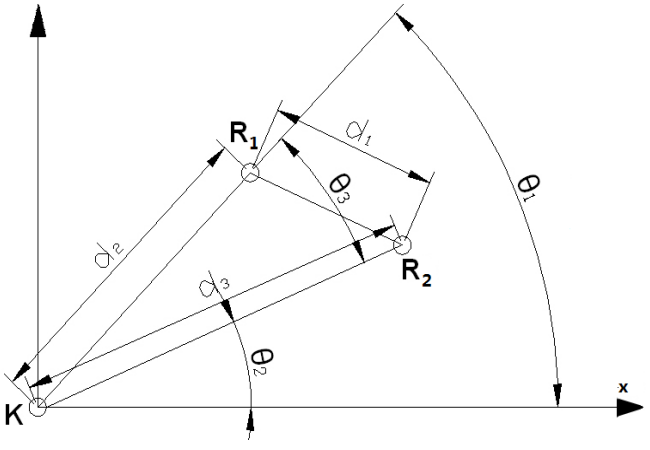


Fig. 1. The diagram of ultrasonic positioning

Suppose K and R_2 keep still at first, the distance between them can be obtained by ultrasonic sensors, and angle θ_2 between R_2 and the x-axis is known, then angle θ_1 between R_1 and the x-axis is:

$$\theta_1 = \theta_2 + \theta_3 \tag{1}$$

$$\theta_3 = ar\cos\frac{(d_3)^2 + (d_2)^2 - (d_1)^2}{2d_2d_3} \tag{2}$$

Where:

d_1—the distance between R_1 and R_2;
d_2—the distance between K and R_1;
d_3—the distance between K and R_2.

3 The Application of Adaptive Filter in Multi-robot Co-location

First, assuming each robot on multi-robot systems has the function to communicate with other robots. According to the principle of multi-robot co-location, Out of the ultrasonic the environment detecting robot also issued communication signals, the communicational signals contain information of frequency of environment rover piezoelectric excitation and size of the excitation voltage, as well as the parameters of piezoelectric sensors features. When the robot receives a communication signal, take the following work:

(1) Start timer immediately;

(2) Detect the robot's environment to calculate the characteristics, frequency and amplitude of ultrasonic signals;

(3) According to the characteristics of ultrasonic sensors receiving circuit, work out the characteristics of frequency and amplitude of the sampled signal, and then get the expected value.

4 The Integration of Robot Co-location Positioning and Autonomous Positioning of Robots

Assuming a single robot equipped with encoder and electronic compass, it can be self-positioning. But when a single robot roles by external forces, it is forced have to change position and orientation, the speculate at this time caused by the error is very large from the real data.

The algorithms of Kalman filtering[3] predicts the location, it can be a good solution in the actual process due to the error code disk impact. Therefore by integrating Kalman filter and self-observing position and orientation information to realize the cooperation in multi-robot positioning. Their cooperation diagram as follows:

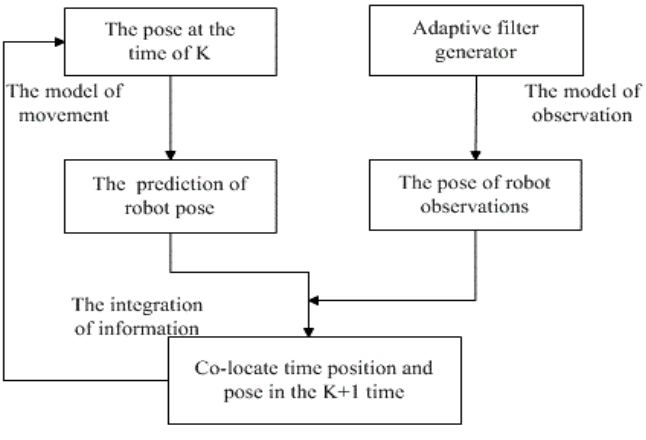


Fig. 2. The framework of positioning robot cooperation

Kalman filter is a recursive form of state estimation, Kalman filter based on the formula:

$$\hat{\bar{x}}_k = A\hat{x}_{k-1} + Bu_k \tag{3}$$

$$\hat{x}_k = \hat{\bar{x}}_{k-1} + k(z_k - H\hat{\bar{x}}_{k-1}) \tag{4}$$

$$\bar{p}_k = Ap_{k-1}A^T + Q \tag{5}$$

$$p_k = (I - K_k H)\bar{p}_k \tag{6}$$

$$K_k = \bar{p}_k H^T (H\bar{p}_k H^T + R)^{-1} \tag{7}$$

Where:

$\hat{\bar{x}}_k$ —the next step predictive value;

u_k —control value;

$\bar{p}_k$ —pre-observation covariance;

p_k —after the observation covariance;

$\hat{x}_k$ —update the value of using observational data.

Then establish the state equation and observation equation model of the system. Use Kalman filter to integrate self-bit data information and relative observation information of the robots, for the equation of motion and observation equation of the robots are nonlinear functions, At first, process the nonlinear system linearization, that is, solve the equations of the Jacobi matrix, and then use the Kalman filter equations to derive the co-location algorithm of the robot. Use the value x which obtains using robot's co-location algorithm to replace in (5), the use of new observational data to update, then you can guess the true value of the robot.

5 Simulation

The algorithm of Pattern matching[4] extract the amplitude as characteristics, the sampling period is 2.5×10^{-5} s, the sampling frequency is 4×10^{8} Hz. Extract the average amplitude, that is to say that remove the symbol sampling point divided by the sum of the number of sampling points. If greater than 4.5 less than 5.5 is considered a match. Otherwise, the match fails.

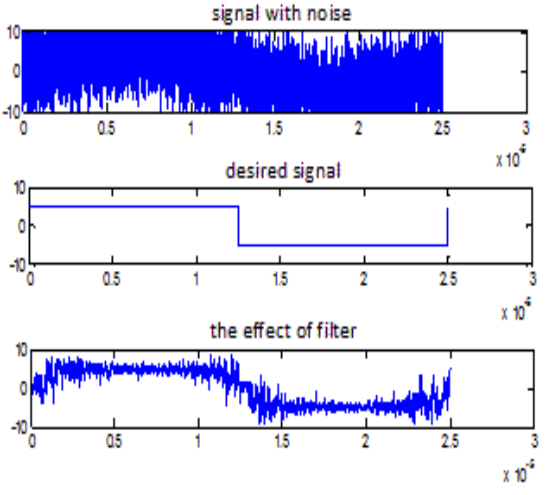


Fig. 3. Signal with ultrasonic transducer filtering effect

Figure 3 shows the received signal in ultrasonic transducer signals, the frequency of 4×10^4 Hz, the square wave amplitude is 5 v, the noise frequency is 4×10^4 Hz, 5 v amplitude sine wave and Gaussian white noise, filter through self- adapt the filter. After calculation, the mean amplitude is 4.956 v, the pattern is matched.

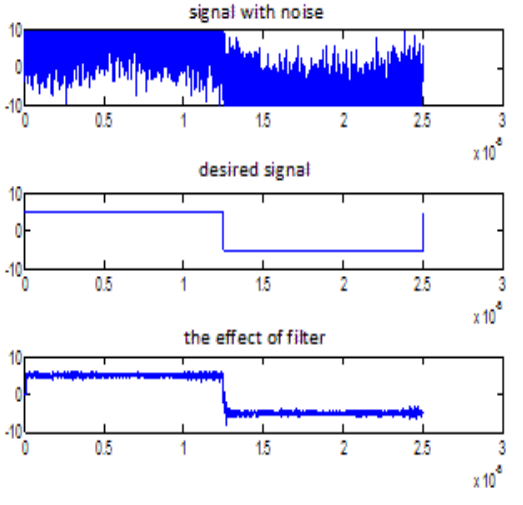


Fig. 4. Signal without ultrasonic transducer filtering effect

Figure 4 shows the signal was not received ultrasonic transducer signals, the noise frequency 4×10^4 Hz, 5 v amplitude sine wave and Gaussian white noise, filtered

through the adaptive filter. After calculation, the mean amplitude is 4.126 v, pattern matching fails.

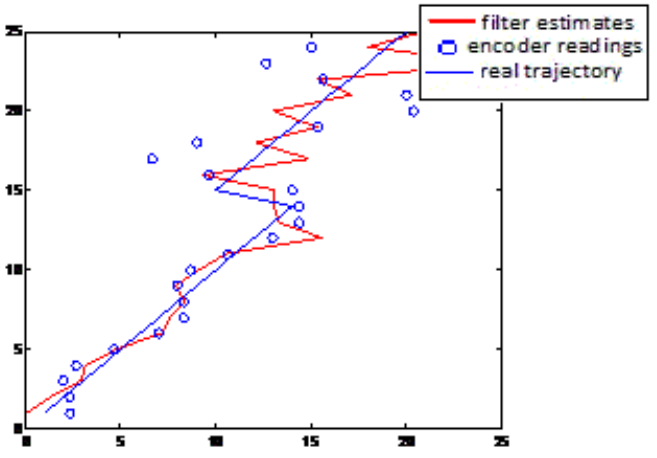


Fig. 5. The updated trajectory using the value of co-location

Figure 5 is the use of the value of the updated path of co-location. Assuming the robot cause a collision in the coordinates (15:15). The solid line is the robot trajectory, the dashed line is the use of co-location corrected trajectory, seen from the figure, and the robot can complete the case of collision orientation.

6 Conclusion

Simulation results show that the proposed method that can detect ultrasonic signals to complete the task of positioning of robot in case of having complex interference and collision between the robots.

Because of the properties of ultrasonic, the receiver is likely to not receive ultrasound signals when there is obstacle between the ultrasonic transmitter and ultrasonic receiver. Also, the positioning method is a measurement-based approach, although ultrasonic distance measurement error is small, however, the error may be larger along the movement of robot system.

Therefore, this article as a supplementary method of positioning is feasible. While robot can be positioned by the law of inertia on the autonomous multi-robot system, the use of this method in this article can correct the error caused by self-positioning.

References

1. Lucibello, P.: On the Role of High-Gain Feedback in P-type Learning Control of Robot Arms. IEEE Transactions on Robotics and Automation 12, 602–605 (1996), Francesco Bullo

2. Pietruszkiewicz, W.: A Comparison of Nonlinear Kalman Filtering Applied to Feed-Forward Neural Networks as Learning Algorithms. In: 2010 9th IEEE International Conference on Cybernetic Intelligent Systems (CIS), pp. 1–6. IEEE Press, Guilin (2010)
3. Sung-Min, H., Sang-Keon, P., Jae-Hag, J., Kang-Woong, L.: Mobile Robot Navigation by Circular Path Planning Algorithm Using Camera and Ultrasonic Sensor. In: 2009 IEEE International Symposium on Industrial Electronics, pp. 1749–1754. IEEE Press, Beijing (2009)
4. Jang-Jong, F., Keh-Yih, S.: The Design of Efficient Algorithms for Two-Dimensional Pattern Matching. 1995 IEEE Transactions on Knowledge and Data Engineering 7(2), 318–327 (1995)
5. Luo, R.C., Yu-Chih, L., Ching-Chung, K.: Autonomous Mobile Robot Navigation and Localization Based on Floor Plan Map Information and Sensory Fusion approach. In: 2010 IEEE Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI), pp. 121–126. IEEE Press, Fort Douglas (2010)

Research on the Trunk Line's Transmission Performance of Multi-hop WMN Based on 802.11n

Zhu Quan[1,2], Jiang Xinhua[1], and Zou Fumin[2]

[1] School of Information Science and Engineering, Central-South University, Changsha Hunan 410075, China
[2] Research Center for Next-Generation Internet Tech & Apps, Fujian University of Technology, Fuzhou Fujian 350014, China

Abstract. This paper had a research on the main factors which affect the transmission performance. Based on 802.11n, It proposed the multi-radio node structure of multi-hop Mesh backbone network, achieved the prototype system of wireless mesh router, and solved the two problems of multi-hop transmission: multi-hop attenuation and basis bandwidth bottlenecks. The experiment showed that it had more than 165 Mbps basis bandwidth, and under the limited 60Mbps environment, the bandwidth attenuation of per hop is less than 1%, which basically satisfied the application of on-board Internet.

Keywords: 802.11n, batman-adv, multi-radio mesh, multi-hop transmission performance.

1 Introduction

With the rapid development of mobile Internet[1], the rail and road vehicle users are increasing demand for broadband access to the public information network during travel. The existing information network has been difficult to meet the demand of mobile users for accessing. So on-board broadband Internet applications has brought wide attention in the World[2-4]. Obviously, the wireless broadband coverage for the trunk lines of highway/railway is one of essential conditions to extend broadband Internet to fast moving vehicles. Reference [5] proposed the banyan-tree topology is fit for the wireless broadband coverage of the traffic trunk lines, whose backbone transmission link is using Wireless Mesh Networks. Reference [6] proposed the train broadband access prototype system based on Wireless Mesh Networks technology. The backbone transfer based on WMN technology has a series of advantages, such as less investment, flexible networking[7], and the Railway Wireless Mesh Network (RWMN) shows a low transmission delay, high bandwidth and a strong network robustness[5], which is a better technology solution for the on-board broadband wireless access.

However, there is a key problem of Wireless Mesh networks: the basis bandwidth is not high and with the increase of wireless hops n, the transmission bandwidth will have a sharp decline at least for 1/n while the transmission delay has a rapid growth. Although multi-radio Wireless Mesh technology dramatically improves the transmission performance of multi-hop WMN, with the increase of wireless hops, the

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 631–636.
springerlink.com

transmission bandwidth will still fall significantly. It is difficult to improve the multi-hop WMN backbone link transmission performance effectively.

Currently, the WMN technology has become very hot in industry applications, as an important carrier of on-board broadband Internet applications, it has a very important academic and engineering value to have a research on backbone transmission technologies of multi-hop WMN. Therefore, this paper will integrate banyan-tree topology, multi-Radio Mesh Node Structure, and the 802.11n based on Reference [5,8], to realize the prototype system of wireless mesh router based on 802.11n, improving the multi-hop basis bandwidth and reducing its multi-hop transmission decay, which can achieve the T3 level communication system standards and meet the needs of on-board Internet applications.

2 Multi-hop Backbone Transmission of WMN

2.1 Multi-hop WMN Architecture

Shown in Figure 1, the multi-hop WMN backbone link along transportation was formed by a series of Mesh Routers (Mesh Router, MR) based on the banyan-tree topology[5], achieving full wireless broadband coverage and providing seamless wireless broadband access for the Mesh terminal (Mesh Client, MC) which deployed in the transport. When the MR directly connected to the wired network, it becomes a Mesh Gateway (Mesh Gateway, MG) which brought user data together and exchanged with Switch Control Center (SCC). SCC is the core of multi-hop WMN, it is achieved centralized control of network management, and had a real-time adjustment of the network topology according to dynamic changes of node failure probability.

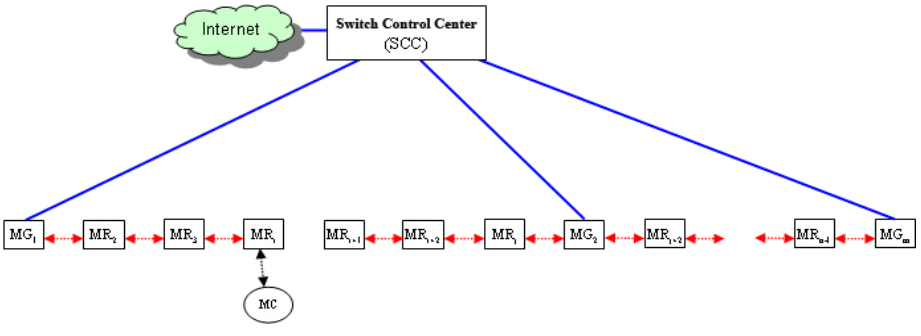


Fig. 1. The figure of multi-hop WMN architecture

Multi-hop WMN contains multiple loops, and we can get one of the wireless multi-hop chains by dynamic minimum spanning tree algorithm (DMSTA) [5]. As the topology shown in Figure 2, it will be a linear multi-hop WMN backbone transmission link along the trunk line in the actual deployment. It should be noted that the content discussed in this paper is suitable for the entire multi-hop WMN, and in order to facilitate the description of the problem, we just consider only one wireless multi-hop chain here.

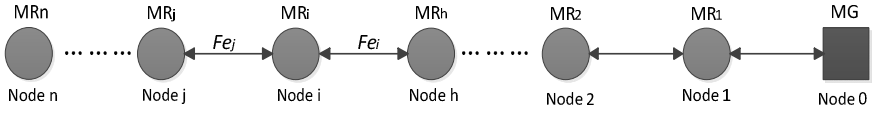


Fig. 2. Linear Network Topology

Figure 2 shows a wireless multi-hop chain, with an MG node and N (N≥1) MR nodes, which totally has n-hop wireless transmission. This multi-hop WMN is still faced the two key problems of traditional WMN.

Reference [8] shows that some of the key factors which affect the performance of backbone link include the interference between Mesh nodes, data forwarding and CPU processing capability. Through the use of hybrid channel multiplexing schemes, asymmetric bandwidth allocation mechanism, and the capacity enhancement of mesh router nodes, we can partly solve the problem that multi-hop transmission bandwidth declining dramatically, and reduce the bandwidth decay rate of multi-hop WMN transmission. Using 802.11n technology can improve the basis bandwidth of backbone link, and then improve overall network performance of multi-hop transmission.

Therefore, this paper will be based on 802.11n technology[9], and will focus on the need of on-board broadband Internet applications and have a research on backbone transmission link of multi-hop WMN, for further enhancement of its multi-hop backbone link bandwidth and reducing the bandwidth decay rate of its multi-hop transmission.

3 Multi-radio Mesh Node Structure Based on 802.11n

In a multi-hop chain, MR is the main body of wireless transmission, which bears the MC access, ingress and egress data transmission. In the traditional WMN, these three tasks are serial to work. In order to improve the performance of the WMN, it should be made that these three tasks can be executed concurrently. Meanwhile, MG is needed for data transmission of all these MR in this link, but the return link and access link bandwidth is always equal in the traditional structure. It would clearly be a bottleneck of multi-hop WMN backbone transmission, so it is needed to maximize the basis bandwidth. To this goal, the Multi-Radio Mesh Node Structure is shown in Figure 3.

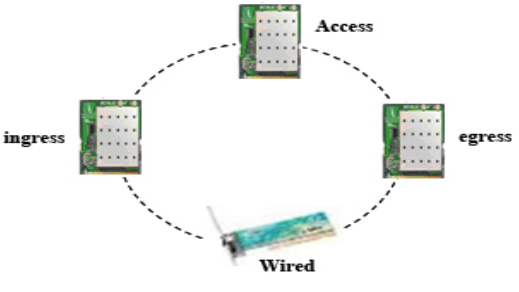


Fig. 3. Multi-Radio Mesh Node Structure

Multi-Radio Mesh Node consists of three wireless modules and a wired module. The three modules are responsible for the MC wireless access, ingress and egress data transmission mission. When the node is as a gateway node, the wired module is responsible for public network data transmission. As this structure is using multi-radio wireless module and the independent wireless module is responsible for their respective tasks, the concurrent execution of three items become possible.

For the convenience of description, this paper considered the former three modules as wireless access, ingress and egress module. To avoid the wireless interference in and between the nodes, each wireless module needs to operate in orthogonal channels. As the current wireless users usually work in the 2.4GHz band with 802.11g wireless card, the access module must support the 2.4GHz band. The ingress and egress module is only responsible for data transmission between Mesh nodes, so it can work in the 5GHz band using the Wi-Fi standard, such as 802.11a or 802.11n. The wireless modules in adjacent mesh nodes need to work on orthogonal channels. Through the use of hybrid channel multiplexing scheme, wireless interference is been minimized or even eliminated in and between the nodes. Then MR can execute MC access, ingress and egress data transmission tasks concurrently.

The batman-adv is a layer 2 WMN routing protocols. It supports for independent networking, wireless roaming [10], and has been a relatively wide range of applications [11]. The mesh nodes use batman-adv routing protocol to realize independent network, forming a multi-hop backbone transmission network.

The multi-radio mesh node is designed as a layer 2 network devices, using the bridge to combine the four network module and forward data packets. As use of bridge module, it is easier to realize wire-speed data forwarding. To improve the basis bandwidth, with using the 802.11n standard for wireless transmission, the theoretical transfer speeds can be up to 300Mbps, the actual reach 100Mbps or more.

4 Experiment and Analysis

Based on the design scheme of multi-radio mesh node structure, this paper accomplished the wireless mesh router prototype system. In order to test the multi-hop WMN backbone link transmission performance, this paper established the eight-hop mesh network emulation test bed outdoor. In a main road on campus about 2,000 meters long, we have deployed nine MRs, the fifth node as the gateway through the cable is connected directly with the monitoring center computer room. Mesh nodes by using the batman-adv routing protocol to realize independent networking, based on the 802.11n wireless protocol. Each node used the directional antenna and the distance between every two nodes was flexibly adjusted according to the terrain.This paper mainly measured the transmission bandwidth and time delay for multi-hops network, and took an average of test results, the final results shown in Table 1.

As the above actual test results shown, the backbone's basis bandwidth of 802.11n-based multi-hop mesh network is higher, and it can be up to 165Mbps or more. After eight-hop wireless transmission, the bandwidth is still around 76Mbps, and the network transmission delay compared to other types of wireless transmission network is smaller. After eight-hop transmission, the network bandwidth dropped only about 50%, relative to traditional WMN 1/n of the downward trend, transmission performance has been improved tremendously.

Table 1. Transmission performance of multi-hop mesh backbone link

Name	Item	Bandwidth(Mbps)	Transmission delay (second)
One-hop	#1-#2	167	0.578
Two-hop	#1-#3	129.8	0.616
Three-hop	#1-#4	126.118	0.633
Four-hop	#1-#5	110.656	0.723
Five-hop	#1-#6	99.515	0.804
Six-hop	#1-#7	85.799	0.931
Seven-hop	#1-#8	82.443	0.972
Eight-hop	#1-#9	76.850	1.052

To improve the entire transmission performance of multi-hop network, the sending rate[12] would be further reduced., so we can consider the T3 standard as target speed and set MR's basis speed to 60 Mbps. Through re-testing these experiments, the test results shown in Figure 4.

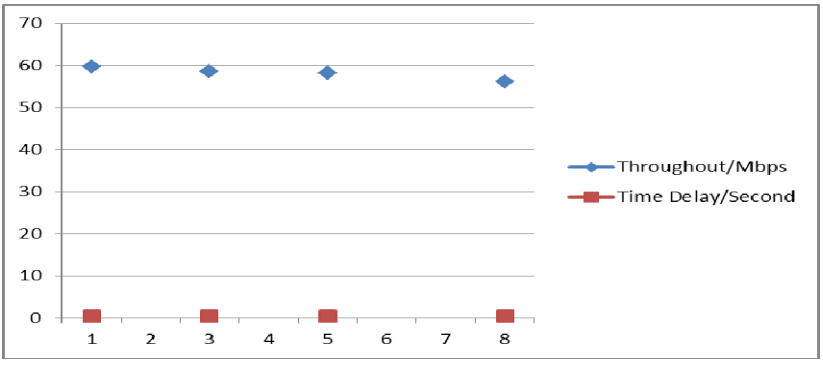


Fig. 4. Transmission performance at limited basis speed

In the limited 60Mbps speed, single-hop bandwidth was up to 59.322Mbps, after eight hops wireless transmission, the bandwidth was still has 56.79Mbps, and network transmission delay was nearly unchanged. After eight-hop transmission, the bandwidth of the simulation network only decreased by 4.3 percent, the results was consistent with the theoretical analysis and experimental results in Reference [12].

Above experimental results show that: the prototype system basically solved the multi-hop transmission problems of WMN backbone link.

5 Conclusion

This paper has studied the transmission performance of multi-hop WMN backbone link, analyzed the main factors which affect the transmission performance. It proposed 802.11n-based multi-radio node structure of multi-hop Mesh backbone network,

realized the prototype system of wireless mesh router, solved the two problems of multi-hop transmission. Next step, we will further refine the prototype system, and solve the problems: large-scale deployment and intelligent management for WMN.

Acknowledgment. This work is supported financially by the Ministry of Science and Technology of PRC through the project of WMN (61101139), Department of Science & Technology of Fujian Provincial by the project of AEED (No.2008J1002).

References

1. MORGAN STANLEY RESEARCH. The Mobile Internet Report. Morgan Stanley & Co. Incorporated, America (2009)
2. Jiang, X., Zou, F., Lin, Z., et al.: A Survey on the Internet Application on Passenger Trains. Journal of the China Railway Society 29(5), 103–110 (2007)
3. Federal Railroad Administration of US. Study of High-Speed Wireless Data Transmissions for Railroad Operation. Department Of Transportation, US (2007)
4. Lannoo, B., van Ooteghem, J., Pareit, D., et al.: Business model for broadband Internet on the train. Journal of the Institute of Telecommunications Professionals, Warsaw, Poland, 19– 27 (2007)
5. Zou, F., Jiang, X., Lin, Z., et al.: A Banyan-tree Topology Based Railway Wireless Mesh Network Architecture. Journal of the China Railway Society 32(2), 47–54 (2010)
6. Xu, H., Fang, X.-M., Xiang, Z.: Design of Train Broadband Access Prototype System Based on Wireless Mesh Networking. Journal of the China Railway Society 32(2), 41–46 (2010)
7. Song, W., Fang, X.-M.: Research and Development of Wireless Mesh Network. Journal of the China Railway Society 29(2), 96–103 (2007)
8. Zhu, Q., Jiang, X.-H., Zou, F.-M.: Experimental Research on the Influence Factors in the Multi-hop Multi-radio Mesh Network. Science Technology and Engineering 8(21), 5822–5826 (2008)
9. LAN MAN Standards Committee. IEEE Std 802.11nTM-2009 (Ver. 2009). IEEE, New York, America (2009)
10. B.A.T.M.A.N. advanced, `http://www.open-mesh.org/wiki/batman-adv`
11. Garroppo, R.G., Giordano, S., Tavanti, L.: Experimental evaluation of two open source solutions for wireless mesh routing at layer two. In: 5th IEEE International Symposium on Wireless Pervasive Computing (ISWPC 2010), Modena, Italy, pp. 232–237 (2010)
12. Li, W.-H., Lv, P., Chen, Y.-W., et al.: PSMR: Power control and scheduling scheme in multi-rate wireless mesh networks. Journal of Computer Applications 31(1), 208–211 (2011)

Application of Geographic Information System in Orienteering Sports

Xiaozhi He

P.E. Departmernt, Changsha University of Science and Technology,
410004 Changsha, P.R. China
hexiaozhi88@163.com

Abstract. This paper introduces the application of GIS in how to make Orienteering maps, in how to select the field of Orienteering and in how to design the route of Orienteering. It also analyzes the process of making the Orienteering map in detail and points out the shortages. At last it puts forward the frame and the development project of Orienteering map assistance design system based on Geographic Information System (shortened as GIS).

Keywords: Orienteering, Orienteering map, GIS.

1 Introduction

Orienteering is a sport in which a precise map and a compass are used to visit all the controls listed on the map, with the winner covering the controls in the shortest period of time. Orienteering, originating from Sweden, used to be a military sports activity. The word *Orienteering* was firstly used in 1886, which means crossing the unrecognized terrain with the help of a map and a compass. The first formal Orienteering match was held in a military camp in Stockholm(the capital of Sweden) and Oslo (the capital of Norway). It marks the formal birth of Orienteering as a sports event. The application of Geographic Information System (GIS) in Orienteering focuses on the three aspects of making the Orienteering map, selecting the field of Orienteering and designing the route of Orienteering[1].

2 Application of GIS in Orienteering Map-Making

2.1 Differences between Orienteering Maps and Ordinary Maps

The ordinary map is the premise for the fast spreading of Orienteering sports and a precise Orienteering map is a must for this sport. The Orienteering map can be divided into three types, namely, road map, field map and city map. The differences in Orienteering maps lie in the proportional scale, degree of details and degree of preciseness. The Orienteering map is actually a kind of special map, with features related to Orienteering sports. Firstly, it has a lower requirement of preciseness. A distance difference within 5% between two detailed parts can be satisfied. Secondly, man-made surface features are important for Orienteering sports, for, athletes need to

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 637–642.
springerlink.com

position by contrasting these features, and the controls of Orienteering are usually set near such surface features as artificial hills, statues, pavilions etc. Thirdly, Man-made surface features conveyed by Orienteering maps can only be divided into the five types of terrain features, rocks, river system, vegetation and man-made surface features and focus on the conveying of surface[2]. Therefore, there is a big difference between Orienteering maps and ordinary large scale digital maps (See Table 1).

The early Orienteering map is a kind of single-colored map or an ordinary large scale map, which can be easily made. However, reading an ordinary map requires certain professional knowledge, which leads to many athletes' not understanding it and failing in the competitions. So, ordinary maps are not quite suitable for matches. Compared with ordinary maps, orienteering maps emphasize on the description of the visibility and accessibility of the earth's surface, mostly with the terrain features of Northwestern Europe as the description subjects. The symbols and colors used in map making are also designed according to the habits of Europeans. In 1999, the Map Committee of IOF (International Orienteering Federation) modified *Map-making Standards of Orienteering Maps.* They noticed the terrain features in other places but still based on the symbols and colors of traditional Western Europe, that is ISOM 2000.

Table 1. Differences between Orienteering Maps and Large-scale Digital Maps

	Ordinary Maps	Orienteering Maps
Scale	1:500; 1:1000; 1:2000	1:5000; 1:10000
Map-making Standard	GB/T 7929-95	*Map-making Standards of Orienteering Maps*
Diagrammatic Presentation	over a hundred with different colors to symbolize the terrain surface	about 70 symbols, 7 colors to symbolize the terrain surface
Requirement and Features	precisely displays man-made surface features & terrain features	focusing on the description of the earth's surface visibility accessibility, regional surface terrain
Preciseness	1:500, 1:1000; 1:2000 map GB/T16819-1997	a lower requirement of preciseness with distance difference within 5%. The height difference between two neighbor -ing Man-made surface features should be correctly displayed.

2.2 Methods of Making Orienteering Maps

OCAD is a special software developed by Swedish and Norwegian orienteers in the late 1980s. It computerized the making of Orienteering maps. At present, most professional Orienteering maps are designed with OCAD. There are three methods of making Orienteering maps.

(1) Traditional hand-making. This method is seldom used now.
(2) Use a base map to make an Orienteering relief map, then use OCAD to digitalize it to be a special Orienteering map. This method is very complicated in map-making, which takes a longer map-making time.
(3) Apply GIS large scale terrain data in hand, introduce OCAD and edit into maps. This is a good application of GIS in Orienteering map making [3].

In 2007 and 2008, the Orienteering Team of CSUST (Changsha University of Science and Technology) completed the 1:5000 Orienteering map for the main campus and the new campus of CSUST respectively by applying Method (3). We used OCAD as the platform because it can provide us with standard ISOM 2000 diagrammatic presentations. The working procedure is as follows.

First, load a 1:500 digital map in AutoCAD.

Second, map-making synthesizing. Because the scale of an Orienteering map is usually 1:5000 or 1:10000, we need to synthesize the large scale digital map, usually in two ways: choice making and generalizing [4]. Choice making corresponds to the selecting of detailed parts and man-made surface features on the map. We should recognize its importance and its effects on the clarity of the map from the perspective of athletes. These two points seem to be contradictory, but beeping clarity does not necessarily mean abandoning a large number of details and surface features. It is necessary to confirm the minimum size of some detailed parts, which may be different on different maps, but the same standard of making choices should be applied in the same Orienteering map. Generalizing has a big influence on the clarity of a map, we can use methods like simplifying, displacing and exaggerating etc.

Third, stratify according to the main surface features.

Fourth, save the digital map as DXF 12.0 format.

Fifth, input the DXF 12.0 file into OCAD.

Sixth, make transforming files corresponding to AutoCAD image layer and OCAD diagrammatic presentation.

Seventh, use OCAD 6.0 to provide the function of doing wholesale transforming. Transform the ordinary map's surface features into OCAD according to layers. The diagrammatic presentations are to be transformed to ISOM 2000 standard ones.

Eighth, regional filling, partially modify and edit, then directly print the map out.

In the actual operation, we find some problems in OCAD map-making software.

(1) A weak editing function. Compared with the special map-making software Auto CAD, OCAD has a weak editing function, which is inconvenient in use. It has an even weaker editing function than Arcview and MapInfo.

(2) At present it can only accept DXF file, with a better accept of DXF 12.0, a higher version of DXF may lead to errors or a loss of surface features. Now we have a variety of GIS data format, such as SHP format and MID format. Although it can be first transformed to DXF format and then load into OCAD, some useful details which may be very important for athletes may be lost during the transforming process.

(3)It does not have any other GIS function besides map making. It cannot achieve the function of with assistance, designing the routes or fields for competitions [5].

In order to solve the problems above, new Orienteering map-making GIS software needs to be developed and designed.

3 Application of GIS in Orienteering Field Selection

The ordinary map is the premise for the fast spreading of Orienteering sports and a **3** Orienteering is a sport in which an athlete's abilities of map reading and using are tested, which requires that the competition fields should reflect the athletes' abilities of map reading and running. The fields selected should be able to test the athletes' comprehensive levels, i.e. the level of reading and using maps, selecting marching routes, running in the field, searching for controls and climbing over mountains. Therefore, as for the competition field, terrain of average rippling woods is preferred, with a height difference of 100 meters, terrain of rich detailed parts, obviously rippling terrain features, abundant man-made surface features and 70% of the forest covering area. If it is a mountain, we may choose its upper or lower part with a height difference of within 100 meters, an amount of the climbing height of 4% out of the whole race. With the virtual three-dimensional display to GIS, DEM analysis to mountains and the statistics of forest covering area, we can greatly improve the efficiency for the selection of competition fields.

4 Application of GIS in Orienteering Route Design

According to the present *Rules for Orienteering Competitions*, the design of competition routes should follow the following principles.

(1) The competition routes should fully reflect the fairness and justness of the games.
(2) The competition routes should be a comprehensive test for athletes' Orienteering skills, intelligence and physical fitness.
(3) If possible, the competition routes should be designed in the way that males and females use their own controls respectively.
(4) The start and finish can be set in the same place, or in different places.
(5) The distance between different controls should be arranged according to the actual situations of the competition field and the need of route design.
(6) Nursery gardens, sowing fields, fields with crops, railways, highways or areas marked with *No Entry* should be avoided.

When designing routes, the designing of controls, running distance and the amount of height climbing is important. They can be preliminarily conformed through GIS analysis and statistics function, which can greatly relieve the intensity in field work, more reasonably design the competition fields so as to guarantee the fairness of the competitions [6].

5 Development and Design of Orienteering Map Assistance Design System Based on GIS

5.1 Aim of Developing the System

Orienteering map assistance design system aims to improve the present backward situations of Orienteering map making and analysis, obtain a strong editing function, data compatibility and transforming abilities to achieve the assistance design of fields and routes.

5.2 Frame of the System

As is shown in Fig. 1, the frame of this system includes three parts. Three systems correspond to the three main applications of GIS in Orienteering.

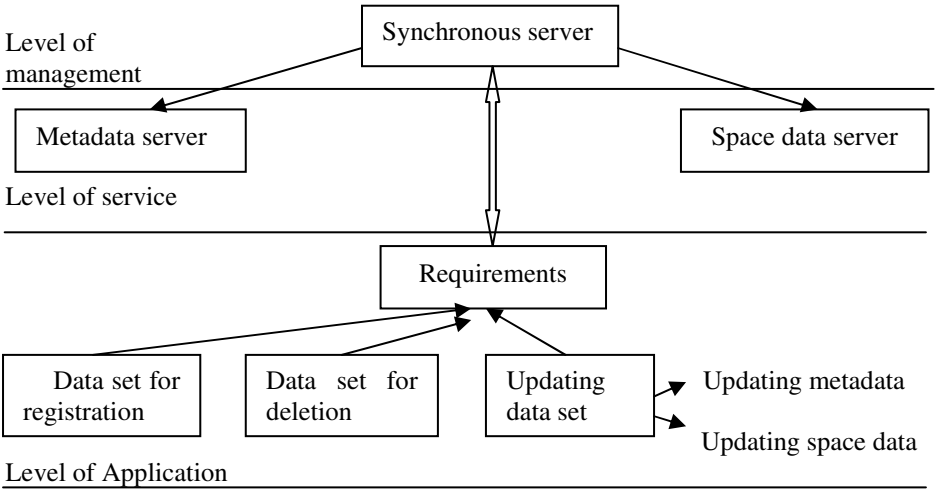


Fig. 1. Frame of Orienteering Map Assistance Design System

5.3 Development Projects for the System

At present, there are four development projects for small-scale Geographic Information System.

(1) Develop in a simple re-development way. In recent years, many excellent basic GIS platforms have sprung up, most of which offer the users macro-language for re-development and some of which offer Dynamic Link Library (shortened as DLL). The advantage of this project lies in that the development platform has a high starting point with strong functions and strong reliability [7]. The disadvantages are that it has a poor expandability and a rather large system, and also a great requirement of hardware.

(2)Use Visible Programming Language to develop from the bottom layer. This project is suitable for strong companies and individuals that have abundant financial resources because it is hard to develop. It also has a greater requirement for the developing team and a longer developing time. This system has such advantages as copyright, a flexible system, being easy to expand and upgrade.

(3) Apply OLE technique to develop compositively. The OLE automatic technique is to make an application program achieve programming, or enable other software to use various kinds of service offered by this program by programming. If we can offer a standard to provide programming ability, several types of software can achieve synchronous programming. COM offers a standard way called interface for the interaction of a certain software subject. Therefore, programming based on COM is natural, which is the OLE automation. This development method has both the

advantages and disadvantages of simply using the macro language offered by the platform for re-development. It has a better expandability and a more flexible developing process since it has utilized the OLE technique for its re-development with its visible advanced language.
(4)Develop with visible programming language and GIS components. Componential GIS develops with the computer software techniques, which can be considered as the developing trend of GIS system. Based on COM, Microsoft promoted ActiveX control technique, which has become the standard control for visible program design. The componential GIS of the new generation is mostly achieved by ActiveX control. What this kind of GIS offers is various kinds of standard ActiveX control for the purpose of completing the GIS system, which enables GIS developers to well know its properties, methods or events of different controls, utilize various visible developing languages such as VC, VB and these control systems to realize the GIS system. So, componential GIS has great advantages in seamless integration and flexibility. This method combines all the advantages of the three methods mentioned above, but with a disadvantage that a set of GIS components is in need of purchase.

Orienteering map assistance design system can be developed by adopting one of the four projects above. For the development of most small-scale GIS, visible programming languages and GIS components are preferred for its development. It stands for the developing trend of the present desk-top GIS. Project Four is the most favorable method for the development of the Orienteering map assistance design system. Languages and components used are to be determined by the developers' language proficiency and the components' properties [8].

6 Conclusion

The Orienteering map assistance design system based on GIS, introducing the advanced GIS method into Orienteering can not only improve the level of Orienteering map making in China, improve organizers' ability to select competition fields and design routes, but also promote the development of Orienteering sports in China.

References

1. Ottosson, T.: Cognition in Orienteering: Theoretical perspectives and Methods of Study in Scientific Journal of Orienteering, 66–72 (1996)
2. Anderson, G.: Orienteering, pp. 231–236. Military Press, Beijing (2002)
3. Ye, C.: Factors Affecting Orienteering Map Making. Shanghai P.E. Science and Technology, 25–26 (March 2003)
4. Hill, C.J., Moore, T.: Rapid Mapping with Post-Processed Garmin Data (DB/OL), http://www.research-tv.com/stories/technology/global/gringo.pdf
5. Dai, W., Zou, Z.: Problems on the Building of Small-scaled Integrated GIS. Surveying and Engineering, 18–21 (January 2002)
6. Zhao, Y., Song, K.: A Study of Orienteering Map Making.Map Making and Printing, 42–44 (April 2001)
7. Hejna, A.: A Mapper Makes a Map. Orieteering Today, 72–73 (January 2008)
8. Wu, X.: Design and Realization of GIS. Electronic Industry Press, Beijing (2002)

Formal Specification for Compiler Based Test Case Generation of Embedded Real-Time System

Yong Chen, Yanxiang He, Chao Xu, Wei Wu, and Jianbo Liu

School of Computer, WuHan University WuHan, China
{cyong,yxhe,xuch,wuwei,liujianbo}@whu.edu.cn

Abstract. Testing is an essential part of the development of embedded real-time system and it is necessity to generate the test case automatically. However, there is no good language to bridge the gap between the testing requirements and source codes. In this paper, we propose a compiler based automatically test case generation framework firstly, and then introduce a novel language call RCBL to specify the testing requirements that describe the relationship of source code and testing requirements for the embedded real-time system. From a simple but comprehensive testing requirement of an embedded system, we can see that the RCBL is powerful and flexible, and also can be learned and used easily.

Keywords: Testing Requirements Specification, RCBL, Automatically Test, Embedded System.

1 Introduction

Testing is a very important part of software development in embedded real-time system, and it spends plenty of time and people. It will be exciting if we can make it automatically. The first and important composition of automatic test is how to generate the test case automatically. The difficulty of automatic generating the test case is that we can't know the relationship between the requirements of testing and the final program codes. In this paper, we introduce a compiler based test case generation framework and a formal specification called RCBL (Requirements Codes Bridge Language) to bridge the gap between the testing requirements and the program codes.

The RCBL is oriented to the compiler and provide the testing requirements to the compiler along with the source code. It includes two parts specification to bridge the gap between the test requirements and program codes. One part is the entry functions that the testing requirement belonged to. The other is the map from the variables of code to the ports or signal of testing requirements. And in embedded real-time system, it has some special characters such as timing constraint, reliability, anti-interference and so on. But the important character is the timing constraint. To illustrate this

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 643–650.
springerlink.com

character of embedded real-time system, we add some timing information to the variables in the RCBL. We divide the variables into two types: var and TS. The var is for general variables and the TS is timing variables that conclude two timing parameters that to illustrate constraint timing factors. According the RCBL, we can link the testing requirement and program codes, and expect to get good test cases automatically.

The organization of this paper is as follows. Section 2 provides an overview of automatic test case generation that based on the compiler framework. Section 3 illustrates the RCBL specification in detail. And then we will provide a sample study case in Section 4. Section 5 compares our work with related work and concluding remarks are made in Section 6.

2 Compiler Based Automatically Test Case Generation Framework

As the good qualified test cases, they must cover the source codes sufficiently to evaluate the right of system's function and the minimal performance of the system such as timing performance in embedded real-time system. There are three mainly methods to evaluate the coverage: data coverage that depends on the coverage of defines or uses of each data, path coverage that depends on the coverage of each path of program and function coverage that is evaluated by the coverage of the system's functions. For getting the range of coverage, we need information both source codes and the requirement of system which can be accessed by the compiler and requirement model respectively.

According to the above description, we construct a compiler based automatically test case generation framework (see Fig 1). In this framework we construct source code and formal specification of test requirements from the requirement model firstly. It makes sure that the formal specification is corresponding to the source codes. Then we analysis the source code through the compiler and provide the path and data information to the scheme center. Scheme center synthesizes the test requirement to the data and path information and injected them to the compiler for generating the object codes that conclude the injected code of the test requirement information. At the same time, test cases generator gets the information from the scheme center and generates the test cases automatically. In the end, we use automatically generated test cases to verify the correct of the system.

The most important part of the framework is the formal specification of test requirements. It need reflect the requirement of test and must correspond to the source codes. And specially, it must easily describe and analysis by scheme center. So the specification must be simple but rich. We design a specification language called RCBL (Requirements Codes Bridge Language) to describe it that will be discussed in the next section.

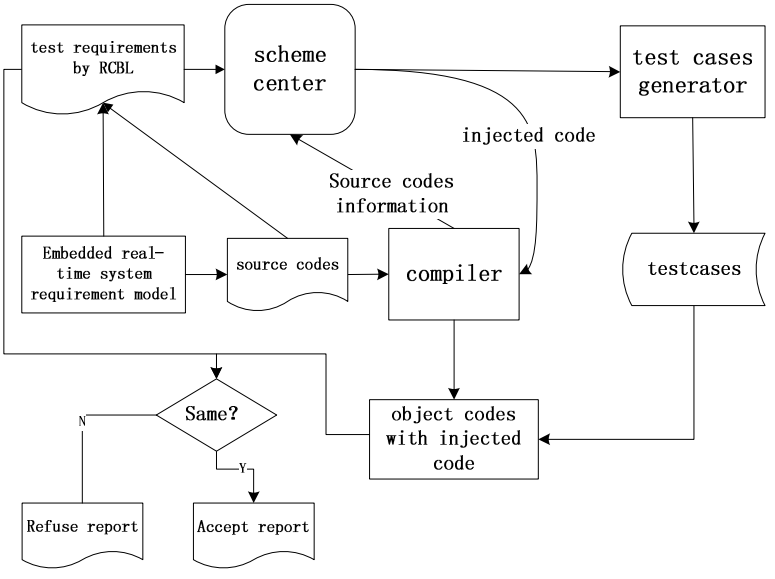


Fig. 1. Compiler based automatically test case generation framework

3 Requirements Codes Bridge Language (RCBL)

To the best of our knowledge, there is no formal language to specify the testing requirements for the embedded real-time system, especially to respect to the source codes. So we design the requirements codes bridge language (RCBL). In RCBL, the testing requirements are combined with source codes such as entry point to entry function, timing requirements to corresponding codes and so on. The syntax of RCBL is shown in Fig 2. The symbols that start with capital letter are nonterminal symbols and the symbols that start with lowercase or quotation marked are terminal symbols.

We divide the testing requirements into many modules that can be seen from the line 1 and 2 of Fig 2, and each module access to one testing requirements. In line 3, we can see that each module contains three parts: modID that identify the module solely, EntryFunc that illustrates the entry function of this module in source code and Functions. In embedded real-time system, it is usually running parallel and for a testing requirement, it may be mapping to more than one function of the source codes, so in RCBL, the EntryFunc can be include many funcIDs which identify the function of source code that connected by "&" (see in line 4). Functions describe the characters of this testing requirement. Because each testing requirement needs many characters to verify it, so the Functions contains many Fun that connected by ":". Each Fun represents one character of this requirement that include pre-modules, input data and output data. Of course, one module may be not require the pre-modules (it could be the first module), and so ModuleIDs can be assign "null" value (see line 7 in Fig 2). Input data and output data can be single variable or be the expression of variables. And for input data, it may be no needing any data because it may test the stability of

the system in no data inputting. Because it was for the embedded real-time system, the timing character is very important, and we need do some special modify for it. We use the symbol “∪” for two variables input or output at the same time and “∩” for two variables input or output in order. For the variables, it has two types: timing signal variables (TS) and combination signal variables. Considering the timing in embedded real-time system, we introduce the TS to describe the timing signal variable. It contains three parts: variable name, minimal delay time and maximal delay time. When generating the test cases, we can use the follower two time parameters to get the appropriate delay time to test the real-time characteristic of the embedded system.

```
1.  Start       ::= Modules
2.  Modules     ::= Module|Module;Modules
3.  Module      ::= modID EntryFunc “(“ Functions “)”
4.  EntryFunc   ::= funcID | funcID “&” EntryFunc
5.  Functions           ::= Fun| Fun ”:” Functions
6.  Fun         ::= ModuleIDs “\” Input “\” Output
7.  ModuleIDs::= modID | modID “&” ModuleIDs | null
8.  Input       ::= VariableID | “(“ ExpI ”)” | null
9.  Output      ::= VariableID | “(“ ExpO “)”
10. ExpI        ::= VariableID “∪” Input | VariableID “∩” Input
11. ExpO        ::= VariableID “∪” Output | VariableID “∩” Output
12. VariableID          ::= TS | var
13. TS                  ::= var “{“ Time “,” Time “}”
14. Time                ::= timeValue
```

Fig. 2. Syntax of RCBL

Through the RCBL, we can find the entry functions of each test requirement in source code, and we can use the white-box testing of these functions to test this test requirement. And we can also know that one testing requirement need which variables as input data and which variables as output data to verify it, then we can use data directed technics to generate such test cases. The syntax of RCBL is so simple that we can analysis it easily. In the next section, we will take an example to illustrate that how to use the RCBL in an embedded real-time system.

4 Case Study

For illustrating how to use our specification, we will take the following example. This example is come from a water heater to test the function of instant heat. The source codes of this function are shown in Fig 3 that modify from an IC company’s real system. In this source codes, the variable’s function is shown in Table 1.

```
  void main(){//the entry point of program
   while(true){
        ......
       if(!turnoff) //it is on the working status
          instant_heat();
......}......}}
void instant_heat (void){ // to instant heat water
  if(!setTemp_reachf){
    if(setTemp >curTemp)
      heat_on();
    else{
      setTemp_reachf = 1;heat_off();
    }
  else{
    if(setTemp <55) {
      if(setTemp - curTemp >=5)
        heat_on ();
      else
        goto label1;
     }
     else{
       if(setTemp - curTemp >=10)
        heat_on();
       else{
label1:   heatingf=0;stillstandf=1;
} } } }
static void  isr(void) interrupt {// update temperature
       ......
       curTemp = updateTemp();
       ......}
```

Fig. 3. Source codes of instant heat module of a water heater

Table 1. Variables to their descriptions

Variables	Descriptions
setTemp_reachf	Specify whether the set temperature is reached before
setTemp	Expected to gotten or hold temperature
curTemp	Current temperature
heatingf	Specify is heating on
stillstandf	Specify on the status of holding the temperature

The testing requirement is that:

1. We need heat the water until the temperature that set initially.
2. After heating the initial temperature, we need hold on this temperature by following scheme:

a) If the initial set temperature is less than 55 degree, it needs to heat the water to reach the initial temperature within 3 to 5 minutes if the current temperature is less than set temperature more than 5 degree.

b) If the initial set temperature is no less than 55 degree, it needs to heat the water to reach the initial temperature within 5 to 10 minutes if the current temperature is less than set temperature more than 10 degree.

From the testing requirement and source codes, we can construct the formal specification using RCBL as "1 isr&instant_heat(null\(setTemp∩curTemp{3,10})\ (curTemp U (heatingf U stillstandf))". It can be described as Fig 4 using the specification syntax tree.

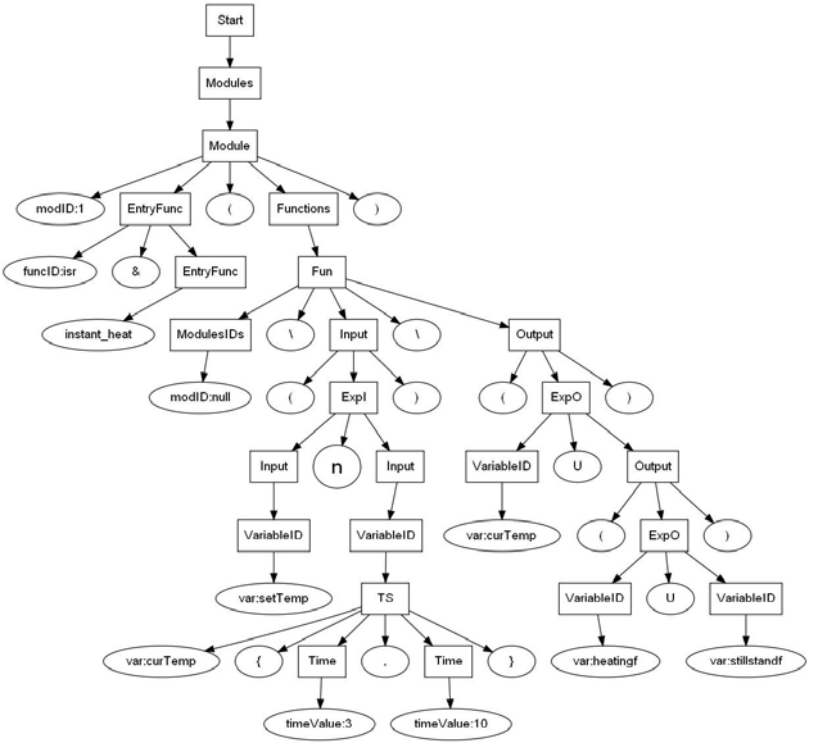


Fig. 4. The syntax tree of the testing requirement specification

From this formal specification, we can see that this testing requirement has two entry function: the "isr" function that update the current temperature in time and the "instant_heat" function that doing the corresponding action according to current temperature. The requirement has no pre-module because we suppose that the main

function is pre-module to all modules. Then the requirement needs current temperature and set temperature that are expressed by “curTemp” and “setTemp” of source codes respectively to check that current action is correct that can also expressed by “curTemp”, “heatingf” and “stillstandf” in the source code. For checking the time status, we can see “curTemp{3,10}” in the specification and know that the minimal heating time 3 minutes and the maximal heating time is 10 minutes. So when the compiler gets this specification, it can generate the test cases such as analysis the def-use path of input variable “curTemp” and “setTemp” and check out the output variables such as “curTemp”, “heatingf” and “stillstandf” in the right time that can be seen from “curTemp{3,10}”, then we can get five right time such as less than 3 minutes, 3 minutes, the time between 3 and 10 minutes, 10 minutes and more than 10 minutes.

From this example, we can see that this formal specification is rich and flexible, and we can get the testing requirement easily such as parser analysis of compiler that is the basic part of one compiler.

5 Related Work

In order to permit cross referencing within LNCS-Online, and eventually between different publishers and their online databases, LNCS will, from now on, be standardizing the format of the references. This new feature will increase the visibility of publications and facilitate academic research considerably. Please base your references on the examples below. References that don’t adhere to this style will be reformatted by Springer. You should therefore check your references thoroughly when you receive the final pdf of your paper. The reference section must be complete. You may not omit references. Instructions as to where to find a fuller version of the references are not permissible.

We only accept references written using the latin alphabet. If the title of the book you are referring to is in Russian or Chinese, then please write (in Russian) or (in Chinese) at the end of the transcript or translation of the title.

The following section shows a sample reference list with entries for journal articles [1], an LNCS chapter [2], a book [3], proceedings without editors [4] and [5], as well as a URL [6]. Please note that proceedings published in LNCS are not cited with their full titles, but with their acronyms!

6 Conclusion

In this paper, we introduce a compiler based automatically generation testing cases framework and RCBL to specify the testing requirement. It is expected to bridge the gap between the testing requirements and the program codes and to generate the current test cases that suitable to the testing requirements automatically. From the case study, we can see that RCBL is a powerful and simple specification language, and it can be parsed and learned easily.

References

1. Yuan-Hsin, T., et al.: A Novel Approach to Automatic Test Case Generation for Web Applications. In: Book A Novel Approach to Automatic Test Case Generation for Web Applications. Series A Novel Approach to Automatic Test Case Generation for Web Applications, pp. 399–404 (2010)
2. Martin, E., Tao, X.: Automated Test Generation for Access Control Policies via Change-Impact Analysis. In: Book Automated Test Generation for Access Control Policies via Change-Impact Analysis. Series Automated Test Generation for Access Control Policies via Change-Impact Analysis, p. 5 (2007)
3. Samarah, A., et al.: Automated Coverage Directed Test Generation Using a Cell-Based Genetic Algorithm. In: Book Automated Coverage Directed Test Generation Using a Cell-Based Genetic Algorithm. Series Automated Coverage Directed Test Generation Using a Cell-Based Genetic Algorithm, pp. 19–26 (2006)
4. Zhifang, L., et al.: Test automation on mobile device. In: Proc. Proceedings of the 5th Workshop on Automation of Software Test, pp. 1–7. ACM (2010)
5. Oh-Hyun, K., Sun-Myung, H.: Mobile GUI Testing Tool based on Image Flow. In: Book Mobile GUI Testing Tool based on Image Flow. Series Mobile GUI Testing Tool based on Image Flow, pp. 508–512 (2008)
6. Vemuri, R.: Testing Predictive Software in Mobile Devices. In: Book Testing Predictive Software in Mobile Devices. SeriesTesting Predictive Software in Mobile Devices, pp. 440–447 (2008)
7. Jiang, B., et al.: MobileTest: A Tool Supporting Automatic Black Box Test for Software on Smart Mobile Devices. In: Book MobileTest: A Tool Supporting Automatic Black Box Test for Software on Smart Mobile Devices. Series MobileTest: A Tool Supporting Automatic Black Box Test for Software on Smart Mobile Devices, p. 8 (2007)
8. Liu, Z., Gao, X.: SOA Based Mobile Device Test. In: Book SOA Based Mobile Device Test. Series SOA Based Mobile Device Test, pp. 641–644 (2009)
9. Zhi-fang, L., et al.: SOA based mobile application software test framework. In: Book SOA based mobile application software test framework Series SOA based mobile application software test framework, pp. 765–769 (2009)
10. Hackner, D.R., Memon, A.M.: Test Case Generator for GUITAR. In: Book Test Case Generator for GUITAR. Series Test Case Generator for GUITAR, pp. 959–960 (2008)
11. Kirchsteiger, C.M., et al.: Automatic Test Generation From Semi-formal Specifications for Functional Verification of System-on-Chip Designs. In: Book Automatic Test Generation From Semi-formal Specifications for Functional Verification of System-on-Chip Designs. Series Automatic Test Generation From Semi-formal Specifications for Functional Verification of System-on-Chip Designs, pp. 1–8 (2008)
12. Zhao, C., et al.: Automated Test Program Generation for an Industrial Optimizing Compiler. In: Book Automated Test Program Generation for an Industrial Optimizing Compiler. Series Automated Test Program Generation for an Industrial Optimizing Compiler, pp. 36–43 (2009)
13. Pacheco, C., et al.: Feedback-directed random test generation. In: Book Feedback-directed random test generation. Series Feedback-directed random test generation, pp. 75–84 (2007)
14. Burnim, J., et al.: WISE: Automated test generation for worst-case complexity. In: Book WISE: Automated test generation for worst-case complexity. Series WISE: Automated test generation for worst-case complexity, pp. 463–473 (2009)
15. Patrice, G., et al.: DART: Directed automated random testing. In: Book DART: Directed automated random testing. Series DART: Directed automated random testing 40 edn., pp. 213–223 (2005)

Fair Video Streaming over Wireless with Network Coding

Xiaoqiao He, Rui Wang, and Depei Qian

Computer Science Department, Beihang University,
100191 Beijing, China
xqhe@cse.buaa.edu.cn, {rui.wang,depei.qian}@jsi.buaa.edu.cn

Abstract. In this paper, we study fair video streaming over wireless network with network coding. Prior work demonstrated network coding can increase throughout over a broadcast medium, by mixing packets from different flows into single packet, thus increasing the information content per transmission and improve video quality. This paper aims to propose fair video streaming over wireless network which can improve video quality and balance video quality of each sink. Our key insight is (i) fairness and (ii) utilization of code packets. Simulation result shows that our scheme can improve and balance video quality.

Keywords: Video streaming, fairness, wireless, network coding.

1 Introduction

Our goal is to enhance the QoS of video streaming. Traditional studies focused on improving throughput with network coding. These studies neglected the fact that high throughput not always improve each sink's video quality while code packets cannot be decoded. In this paper, we propose novel technique which utilizes network coding and fair schedule scheme to improve video quality of each sink simultaneously.

Our work builds on recent work [1] [2] [3]. In particular, [1] [2] proposed wireless routers mix packets from different flows, to increase information content of each transmission and therefore throughput for applications, while [3] proposed scheduling scheme for transmission over wireless mesh network(WMN) to improve video quality.

2 Related Work

This paper introduces a novel technique for video streaming over wireless that combines network coding [4] and fair scheme to improve the video quality of each sink simultaneously. On the one hand, Our work build on [1], [2] to exploit the broadcast nature of the wireless medium and use network coding to pack several packets from different streams into single for transmission, thus increasing throughput. On the other hand, we focus to local throughput and fairness each sink achieves.

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 651–658.
springerlink.com

The recent work in [1], [2] applied these ideas from the network coding community in the context of WMN. [1] implemented a pseudo-broadcast mechanism together with opportunistic listening and coding layer between IP and MAC that is used to detect coding opportunities and pack packets from different flows into single transmission, thus increasing network throughput, while [2] proposed COPE that wireless routers mix packets from different flows to increase the information content of each broadcast transmission and therefore the throughput for data applications.

Our inspiration mainly comes from [3] [5] [6], these works propose video-aware opportunistic network coding schemes that take into account both aspects the decode-ability of network codes by several receivers and distortion values and playoff of video packets. In brief, they attempt to improve throughput as much as possible. We consider balancing throughput that improves the video quality for each sink.

3 Video Streaming Model over Wireless

3.1 Basic Framework

As Fig. 1 shown, we consider video streaming over WMN where mesh routers are able to forward packets to other routers and/or sinks. In this paper, we consider the scheduling by mesh router after gathering packets intended to different receivers and reached situation where network coding is available. We present fair packets schedule scheme to balance throughput each sink received and propose algorithms used at mesh router to maximize video quality and throughput with network coding. We assume mesh routers can perform simple network coding operations (bit-wise XOR) and combine packets from several incoming streams into a single outgoing packet.

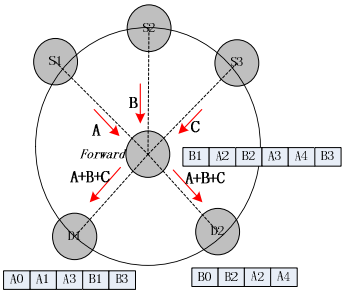


Fig. 1. Video Streaming flow model over wireless mesh network

It usually requires following steps when packet from source to target:

1) Raw packets send out from source; 2) Intermediate nodes/Mesh routers overhear all transmissions from upstream, whether they are intended for them or not, they decode it using overheard packets then perform simple network coding operations (bit-wise XOR) combining decoded packets from several incoming streams into single outgoing packet. Then broadcast coded packet to downstream neighborhoods; 3) Receiver overhears several encode packets and try its best to decode them. Then it inserts decoded packets into its play buffer if desired data and plays it in time.

3.2 Data Flow

We focus on that, when the transmitted flows are video streams, the fairness are considered. In this paper, we develop schemes and packet selection across different flows, to balance flows and improve video quality.

Code selection. Let's consider intermediate node receive packets from upstream and forward them to its downstream, each intermediate maintains *transmission queue* (T_x) with incoming video packets. At given transmission time slot several packets are selected from T_x and encoded for transmission. As [3] mentioned, we try to ensure one packet should be received and decoded by target and this packet is called *primary packet* while others are *candidate packets*. Our question is: *how to select and encode packets to improve video quality, while each client gets its matched video quality.*

Retransmissions. Once coded packet is transmitted, intermediate will wait for mean RTT until it receives ACK. During that period, all packets as part of coded packet will be moved to the *backup transmission queue* (T_b), packets in T_b will not be selected to be primary, but could be candidate. When receives ACK, it removes corresponding packet from T_b at once. If an RTT expires without ACK, the packet will return to T_x. Of course, packets in T_x or T_b will be remove when its playoff expires.

ACKs. Once packets are chosen, there are encoded and broadcasted to downstream. Because of wireless channel, some will receive successfully. When downstream receive code packets and can decode, it sends *acknowledgement* (ACK) back to upstream. Then downstream node stores packets (raw packet or decoded packets) into their *virtual buffer* (V_x). Packets in V_x play important role when decode step.

Requirements. We assume the following capabilities available:

1) We assume broadcast capability is available adopting research findings of [1], [2] in order to harvest the benefits of network coding.

2) Nodes no matter which role know the packets buffer (T_x or T_b or V_x) copy of all their neighbors in real time. This can be achieved by ACK feedback and guessing.

3) Nodes are capable of coding/decoding in real time for bit-wise XOR operation.

4 Coding Algorithms

Our main questions have to do with the construction and selection of network codes to achieve better video quality and fairness. The construction problem is concerned with finding primary and candidate packets that guarantee decode-ability by the target node. The code selection problem is concerned with selecting the best ones among the candidate so as to optimize video quality and sinks get its matched video quality. This chapter we will compare with traditional solution, then proposes our fairness solution.

4.1 Traditional Solution

In traditional transmission, intermediate adopts the best-effort service mechanism to forward the packets. It puts the packets into T_x by order when several video streams coming, while forwards packets from T_x in turn according to the principal of FIFO.

Obviously, this scheme is inefficient and results in low throughput and poor video quality. To solve it recent work in [5], [6] applied the network coding community in video streaming of wireless mesh networks. [5] proposed schemes NCV and NCVD to improve both video quality and throughput. Its key insight is that, when transmitted flows are video streams, network codes should be selected so as to maximize network throughput and video quality. They propose video-aware opportunistic network coding schemes that take into account both aspects: decode-ability of network codes by several receivers and distortion values and play-out deadlines of video packets.

4.2 NCF: Network Coding with Fair Policy

The scheme NCV and NCVD [5] proposed can improve video quality and throughput obviously, but we found that some sinks could achieve good video quality while some others could not achieve desired quality when wireless occur jitter. The fact makes us rethink how to combine network coding to improve video quality of each sinks fairly.

After analyzing closely, we found that NCV can improve throughput while ignore balancing video quality. In order to propose a better scheme let us follow the trail.

Flow statistics in Time Window. Our decisions depend on this statistics. At first, we define $Path < A, B >$ is path A to B, and *flow statistics* is a parameter over path while target of path keeps it. The *inflow* is defined to the total packets number from the upstream of path in the certain time window range. As the inflow, the *outflow* is defined to the total packets number from the upstream of path that have been forwarded in the certain time window range. Therefore the inflow and outflow are:

$$inflow(S,T) = \sum_{t} p < S,T > \tag{1}$$

$$outflow(S,T) = \sum_{t} p' < S,T > \tag{2}$$

$$F(S,T) = \frac{1 + outflow(S,T)}{1 + inflow(S,T)} \tag{3}$$

Where p<S,T> & p'<S,T> indicate packets number intermediate received and forwarded in fixed time *t* respectively. It is clearly that number statistics intermediate keeps is related to number of upstream nodes.

Primary Packet Selection. In order to choose the best primary packet, we first define a metric that evaluates fairness named forward ratio $F(S,T)$ as Eq.(3) shown.

After defining metric of fairness, we select packets in T_x of T from upstream S where F(S,T) is minimum. Then we can choose primary packet easily combining with FIFO and playoff. Primary Packet Selection algorithm is summarized in Alg.1.

Algorithm 1. Primary Packet Selection Algorithm

```
1.  Procedure SELECT()
2.      for( path : Ps ) do
3.              calculate F(Ps) with Eq.3
4.      end for
5.      select minimum F(Ps);
6.      candidates = {packets go through Ps};
7.      return first not playoff packet in candidates;
8.  end Procedure
```

Code Selection. After chosen the appropriate primary packet, improving video quality and throughput turns to be the key concern. More generally, we consider intermediate *I* with *N* downstream $\{n_1, n_2, ..., n_N\}$ and *K* packets in its T_x denoted $\{p_1, p_2, ..., p_K\}$ Choose p_α as the primary packet with target node n_m based on *Primary Packet Selection*. Then construct and transmit network code with some other packets. In order to guarantee decode-ability, candidate packets must be stored at n_m and among T_x or T_b at *I*. Assume *L* packets are overheard at node n_m and denoted $\{v_{m,1}, v_{m,2}, ..., v_{m,L}\}$ Therefore, the candidate packets among:

$$\Psi = \{p_1, p_2, ..., p_K\} \cap \{v_{m,1}, v_{m,2}, ..., v_{m,L}\} \tag{4}$$

And the network code should be:

$$C_m^k = p_\alpha \cup S_m^k, k = 1, 2, ..., 2^n \tag{5}$$

where *n* is the number element of Ψ, and S_m^k is the *k*th subset of Ψ.

In order to choose the best code, we define a metric that represent the contribution of each candidate to video quality improvement. Let $I_m^k(n_i)$ be the improvement in video quality at n_i when C_m^k is received.

$$I_m^k(n_i) = \alpha \sum_{l=1}^{L_k} (1 - r^i) d_l^i Q + (1 - \alpha) \rho F_k^N(n_i) \tag{6}$$

Each factor in this formula is defined as follows:

- α is indicator function expresses whether C_m^k can be decoded.
- L_k is the number of packets included in network code C_m^k.
- r^i is the loss probability due to channel error or latency.
- d_l^i is indicator function that express whether this packet is useful for n_i
- Q is the contribution of packet p_l to video quality improvement.
- $F_k^N(n_i)$ is the contribution of C_m^k which cannot be decoded.
- ρ is the correction value for $F_k^N(n_i)$.

Eq. (6) shows that any code packet could bring contribution. In other words, code packet is useful no matter whether can be decoded or not, different to [5].

The following discuss is nearly same with [5] description.

After defining the contribution of code packet to single node, the total video quality improvement to all downstream as following:

$$I_m^k = \sum_{i=1}^{N} I_m^k(n_i) \tag{7}$$

NCF choose the code that maximizes the total video quality improvement:

$$\max_k I_m^k \tag{8}$$

Alg.2 summarized NCF. It traverses all possible coding combination and selects the maximum contribution then encodes and forwards them.

Algorithm 2. The NCF Algorithm

```
1.  Procedure CODE()
2.      p = SELECT(); cP = sP - p; n = powernum(cP);
3.      for ( k=0; k<n; k++ ) do
4.              iC = subset(cP ,k) + p;
5.              calculate I of iC with Eq.(6)
6.              keep max I and iC.
7.      end for
8.      return iC;
9.  end Procedure
```

5 Performance Evaluation

In this section, we evaluate performance of proposed scheme in terms of video quality and network throughput under different condition upon a simply wireless network experiment platform. Simulation results show that fair scheme could improve throughout significantly while adjust and balance video quality of each client.

5.1 Simulation Setup

First, we describe the communication scenario, wireless channel models and the baseline algorithms for comparison.

1) Scenario Description: We focus on scenario which is two-hop network topology. Assume video streams are only data-stream hence no congestion. However, packets maybe lost due to error of wireless channel, and also experience a random delay. The maximum playoff delay for single hop set 100ms to allow for queuing, transmission and a few retransmissions. We also perform different playoff deadline from 50ms to 200ms. ACKs and retransmissions are used to combat errors and loss packets.

2) Wireless Channel Model: We consider packets transmitted in each link experience with fixed loss probability. We simulated a wide range of effective packet loss rates from 5% up to 20%. The effective loss rate simulates the use of retransmissions, FEC and other mechanisms that mask the error rate on the channel.

3) Algorithms under Comparison: We compare with NCF against two baseline algorithms: no Network Coding (noNC) and Network Coding for Video (NCV).

5.2 Simulation Result

In this section, we evaluate performance of NCF in terms of video quality and throughput under different conditions.

1) Throughput Improvements: Fig. 2 shows the total throughput added over all sinks, it shows that NCV and NCF achieved higher throughput compared to noNC, in addition, it is not much difference between NCV and NCF. The reason is that NCV and NCF improve throughput naturally by network coding. Furthermore, through correction of packets contribution, NCF could improve network throughput, while lose throughput adopting fair scheme. The comprehensive result indicates that the network throughput improvement is very little difference between NCV and NCF.

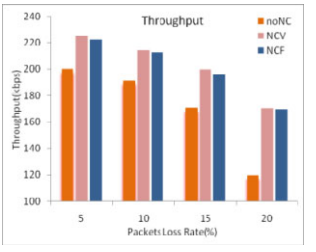


Fig. 2. Total throughput achieved by three schemes under different packet loss rate

2) Video Quality Improvements: Fig. 3 shows average PSNR achieved with different packet loss rate and different buffer delay. It is clearly shown the same trend between Fig.3 (a) and Fig.2 depict. The reason is that average PSNR is positively correlated but different with network throughput. This is well understood that not entire throughput can effective contribute to PSNR. It displays a marked difference between noNC and NCV/NCF. This is one of advantages NCV and NCF achieved.

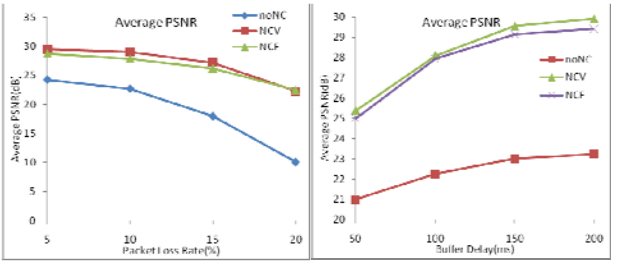


Fig. 3. Average PSNR achieved by setting different packet loss rates and different buffer delays

Fig. 3 (b) depicts average PSNR with different buffer delay. This figure shows that NCF/NCV improve video quality for entire range of delay values. The improvement is smaller for tight delay budget, because delay constraint limits retransmissions and lifetime of packets, thus decreasing selection opportunities. However, with tight delay constraints, NCV/NCF brings significant video quality improvement compare noNC.

3) Fairness Improvements: Fairness is the key point we focus on. The simulation is performed at packet loss rate of 10% with 150ms buffer delay, for comparison, same wireless channel trace is used and same event setting. We watched the PSNR of three nodes in 60 frames. During this period, we set once jitter manually. As Fig. 4 shown,

NCF and NCV can improve about 20% PSNR of video quality compared to noNC. But compared with noNC, NCV did not change fact PSNR wave range which closely affect video quality, under normal condition, the PSNR wave achieve average above 5-15dB with noNC and NCV. Chart (c) is PSNR wave NCF achieved. After adopting fairness algorithm, NCF improve video quality while decrease PSNC wave to below 5dB. Then all sinks can get fluent video streaming service in the same situation.

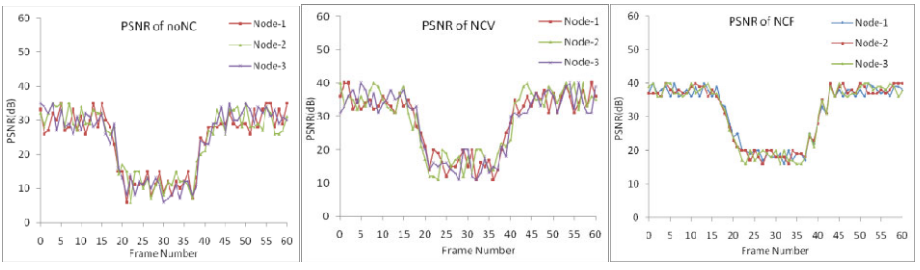


Fig. 4. Fairness achieved under noNC, NCV and NCF respectively

6 Conclusion

Network coding enables more efficient, scalable, and reliable wireless network. Each sink can achieve relatively fair bulk when combined network coding and fair schedule strategy. NCF we proposed brings better QoS of living stream service for any sinks.

Acknowledgments. This work was supported in part by a grant from NSFC under contracts 60673180 and 61073076, and by grants from National High-Tech Program of China (863 program) under contract 2009AA01A129, and by grant from Program of International Science & Technology Cooperation of China under contracts S2010GR0607 and 2009DFA12110.

References

1. Katti, S., Rahul, H., Hu, W., Katabi, D., Medard, M., Crowcroft, J.: XORs in the air: Practical network coding. In: Proc. of ACMSIGCOMM, Pisa, Italy, September 2006, vol. 36(4), pp. 243–254 (2006)
2. Katti, S., Katabi, D., Hu, W., Rahul, H., Medard, M.: The Importance of Being Opportunistic: Practical Network Coding For Wireless Environments. In: Proc. of Allerton Conference (2005)
3. Seferoglu, H., Markopoulou, A.: Opportunistic Network Coding for Video Streaming over Wireless. In: Proc. of the 16th Packet Video Workshop, Lausanne, Switzerland (November 2007)
4. Ahlswede, R., Cai, N., Li, S.R., Yeung, R.W.: Network information flow. IEEE Transactions on Information Theory (2000)
5. Seferoglu, H., Markopoulou, A.: Video-Aware Opportunistic Network Coding over Wireless Networks. IEEE Journal on Selected Areas in Communications 27(5), 713–728 (2009)
6. Yomo, H., Popovski, P.: Opportunistic Scheduling for Wireless Network Coding. In: Proc. IEEE International Conference on Communication, Glasgow, Scotland (June 2007)

The Functional Semantic for Home Service*

Moji Wei, Jianliang Xu, and Xinjun Song

Department of Computer Science, Ocean University of China, Qingdao, 266100, China
weimoji@126.com, xjl9898@gmail.com, songxinjun1226@gmail.com

Abstract. To retrieving home services semantically, this paper constructs function concept ontology for smart home. We classified function into five categories. Then the categories of function are refined through analyzing the way of realization. Finally a scenario of audible alarm of gas detective is given to present the architecture for service registry, retrieval and invocation based on the function concept ontology.

Keywords: Smart Home, Service, Function Concept Ontology.

1 Introduction

Smart home is different from the traditional home on its ability to perform function by integrating appropriate appliances automatically. To facilitate integrating services for households intelligently, ontologies are essential, as they act as the format explicit specification[1] for describing the services provided by networked devices, sensors, and other appliances. Thus the ontology-based semantic approaches could integrate services according to the specification automatically[2]. Currently the designed ontologies focus mainly on context of service [3, 4] or QoS [5, 6] related service properties. This paper focuses functional semantic of service and constructs function concept ontology for home service.

2 Function Concept Ontology

2.1 Categories of Service Function

The function of service is to generate effective information and then represent the information to consumer. So we abstract function by distinguishing the operations in the lifecycle of information. The information would refer to five operations: collection, processing, storage, transmission and representation. Accordingly the function could be divided into five categories: collect, process, store, transmit and represent. Figure 1 illuminates the organization of the categories of function.

*Support by: Natural Science Foundation of Shandong Province (No. ZR2010FM002).

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 659–665.
springerlink.com

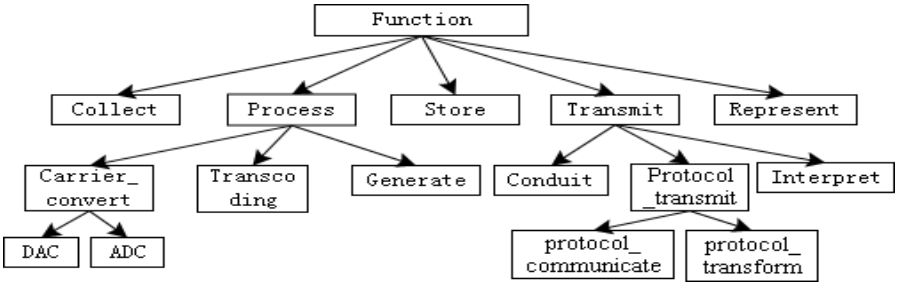


Fig. 1. Categories of Function

2.1.1 Process Function

In above five categories, process function is responsible for content processing. Content as the carrier of information includes three parts: carrier type, code format and the semantic. Consequently process function could be divided into three sub-categories according to the different facets of content. First Carrier_convert function converts specific analog signal to appointed digital signal (DAC function), or the opposite (ADC function). Seconde Transcoding function encodes the format of content into another format. Third Generate function transforms the semantic of content, viz. with Generate function the information of input and output is different.

2.1.2 Transmit Function

Transmit function provides channel for information transmitting between two services. According to the type of channel, transmit function could also be divided into three sub-categories: Conduit, Protocol_transmit and Interpret.

Conduit function transmits content without any operation, thus to realize conduit function dose not have to employ service. When two services are connected with conduit they could communicate with each other directly.

The service which is annotated by Protocol_transmit function transmits data with specific protocol. The Protocol_transmit function could be divided into two categories: Protocol_communicate function which is used for transmitting data for same protocol and Protocol_transform function used for different protocols.

Consulting the object that the service oriented, home service could be classified into two kinds. One is device-oriented service which provides content for device controlling. The content of this kind service is hard for users to understand. The other is user-oriented service which provides content for user consuming. When different kinds of services transmit data, they should search a service whose function is Interpret as medium.

2.2 Refinement of Function

The five categories of function describe all the functions that are referred in the lifecycle of information. However, the level of abstraction of above functions is still too high. To explicate the semantic of service exactly, the categories should be refined further more.

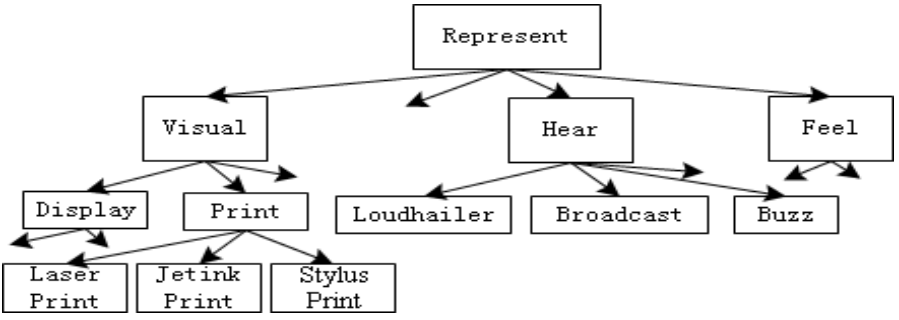


Fig. 2. Portion Sub-functions of Represent Function

This paper takes Represent function refinement for instance to present how to classify function with the guidance of ways. As figure 2 shown, by stimulating on different human senses, there would be many ways that could realize Represent function such as optical way, aurally way, tactile way, etc. Accordingly based on the different ways, the represent function could be classified into Visual function, Hear function, Feel function, etc. Then for Visual function, with different sources of light it could be divided into Display function and Print function. Obviously adopting different principle, Print function could be classified into LaserPrint, InkjetPrint, and StylusPrint.

3 Architecture for Ontology-Based Service Search

In this section a scenario of audible alarm of gas detective is used to show how to search service based on function concept ontology.

3.1 Scenario Description

The essential service that smart home should supply is alarm service. The service should inform the insecurity when dangerous affairs occur. In smart home, the gasDetectiveServcie() service which is provided by gas security device should detect the percentage of gas ceaselessly. When the poisonous gas exceeds secure threshold, the service would output bit 1 to denote the dangerous information.

Then the alarm service which is connected to the detective service may represent the information by aurally way, that is the service of which function is Hear should be invoked. Assuming that there are three devices would provide audio-play services. The first one is gasAlarmServcie() service which is provides by the bell that comes with gas security device. The second one is FMService() service which is run on the radio. The third one is playServcie() service which is hosted on Bluetooth loudspeaker. The three alarm services could be invoked to inform the insecurity automatically when the detective service output bit 1.

3.2 Service Annotation with Function Concept Ontology

Smart home usually employs a residential gateway to manage the home area network (HAN). When the device enters HAN, it would register the services to service registry. Hereafter the gateway would retrieve proper services by searching the service registry. To solve the ambiguous syntactic problems brought by current keyword-based approach, we enhance residential gateway by expanding service registry with function concept ontology.

3.2.1 Service Register

In the alarm scenario of gas detective, there are three devices: bell, radio and Bluetooth loudspeaker. When they attend into home area network, the devices should register services which are hosted on the devices to the service registry. In order to append addition semantic information to service, the service registry should be expanded as shown in figure 3. The Function instance item which stores the instance of the function provides the semantic information for service.

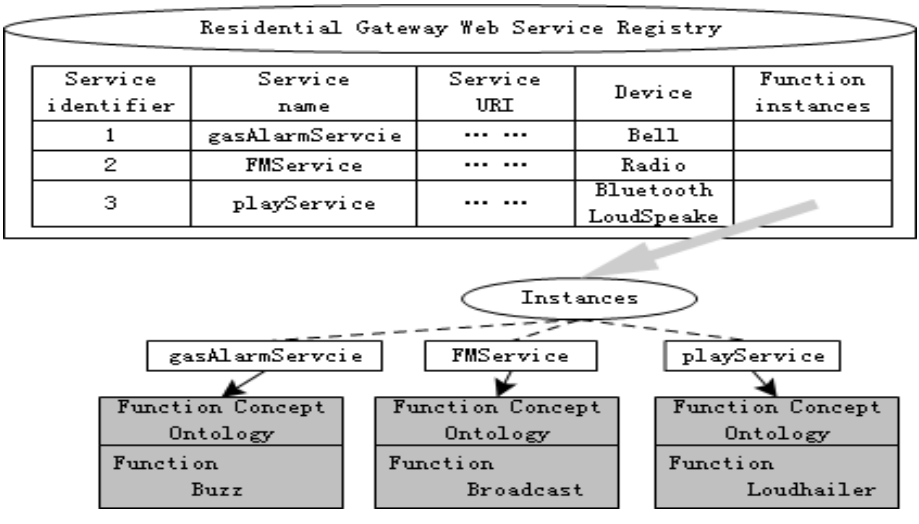


Fig. 3. Instances of Function Concept Ontology

In the scenario, the above three services realize same hear function with different ways. According to function concept ontology, the services could be annotated with lower degree of abstraction and more accurate functions. As shown in figure 3, the services are annotated by Buzz function, Broadcast function and Loudhailer function respectively. That is declaring instances for the function concepts and describing the services with the instances. The instances are declared as follow:

```
<Buzz        rdf:id=”Buzz_GasAlarmService ”/>
<Broadcast   rdf:id=”Broadcast_FMService ”/>
<Loudhailer rdf:id=”Loudhailer_PlayService”/>
```

Then the URI address of instance of function concept ontology would be stored into the corresponding Function instances item of service registry.

3.2.2 Service Retrieval

By appending service registry with function concept ontology, residential gateway could semantically search services by reasoning the instances of function which consumer requests. In the alarm scenario, to represent alarm information, the gateway could use service of which function is Hear, i.e. the gateway would reason instances of Hearing concept in function concept ontology. As figure 2 depicts, Buzz, Broadcast and Loudhailer are the sub-concepts of Hearing concept. According to individual reason of ontology, the instances of sub-concept are also instances of parent concept. Thus by reasoning instances of Hearing concept, gasAlarmService service, FMService service and playServcie service three services which are annotated by the sub-concept of Hearing concept would be retrieved.

3.2.3 Service Invocation

Once services which could represent alarm information are retrieved, the residential gateway should invoke the services automatically. The following would present how the gateway invokes above three audio-play related services.

(1) gasAlarmServcie

Both of gasAlarmServcie service provided by the bell and gasDetectiveServcie service are device-oriented services. When the input of gasAlarmServcie service is bit 1, it would control the bell to ring. Meanwhile the gasDetectiveServcie service would output bit 1, if the poisonous gas exceeds threshold. The contents that the two services operate are consistent, so the services could transmit data with conduit. Once the gasDetectiveServcie service detects danger, the gateway could invoke gasAlarmServcie service directly, as shown in figure 4(a).

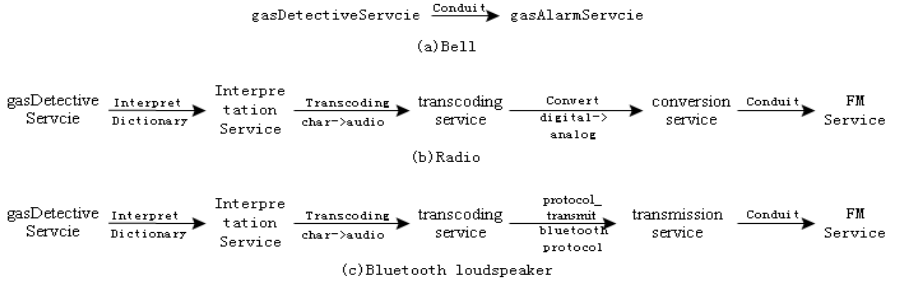


Fig. 4. Service Chain

(2) FMService

Unlike gasDetectiveServcie the FMServcie of radio is user-oriented service. It means that FMService could not process the content given by gasDetectiveServcie. Then two services should be connected with an interpretation service as medium, namely the gateway should search a service of which function is Interpret. The interpretation

service would interpret the output bit "1" of gasDetectiveServcie to text "gas leak" as the input of FMService.

Moreover FMService requests frequency modulation signal as its input, while the output of interpretation service is digital signal. The carrier types of the contents are still inconsistent. The gateway has to search a service whose function is Carrier_convert function for unifying carrier type. By reasoning with function concept ontology, several services which are annotated with Carrier_convert function would be retrieved. Among them there may be a service calls digital signal of which format is WAV as input and outputs frequency modulation signal. Then the service could be connected to FMService. Although both outputs of interpretation service and conversion service are digital signal, the formats of outputs are still inconsistent. One is audio format data, the other is character format data. Accordingly a transcoding service is needed to harmonize format of the content. It means that the gateway has to search a service of which function is Transcoding. If there exists a service transcodes Unicode text to WAV audio format, then through interpretation service, conversion service and transcoding service, the output of gasDetectiveServcie service and input of FMService service are finally compatible, and the gasDetectiveServcie service and FMService service could be connected with assistant of above three services. The service chain is shown as figure 4(b).

(3) playService

Similar with FMService service of radio, the playService service hosted on Bluetooth loudspearker is user-oriented service as well, so an interpretation service should be connected to gasDetectiveServcie service. Additionally the format of interpretation service and playService service are inconsistent, thus a transcoding service is also wanted to harmonize format of the content.

Furthermore Bluetooth loudspeaker has special demand for communication, it transmits data complying with the Bluetooth protocol. Thereby a service of which function is Protocol_transmit should be employed to transmit data between playService and transcoding service. Figure 4(c) shows the service chain for playService and gasDetectiveServcie connection.

4 Conclusion

This paper constructs function concept ontology to annotate the semantic of home service. Then we take alarm scenario of gas detective as an example to represent how to register, select and invoke a service based on the ontology.

Besides function, content and device also have close relationship with service. Consequently offering semantic information for content and device could assist service filtering. One step further, and part of our future work, is to construct content and device ontology by abstracting effect elements of service searching.

Acknowledgment. This research is supported by National Natural Science Foundation of China (No.40806040) and Natural Science Foundation of Shandong Province (No.ZR2010FM002).

References

1. Studer, R., Benjamins, V.R., Fensel, D.: Knowledge Engineering: Principles and Methods. Data and Knowledge Engineering 25(1-2), 161–197 (1998)
2. Roussaki, I., Papaioannou, I., Tsesmetzis, D., Kantorovitch, J., Kalaoja, J., Poortinga, R.: Ontology Based Service Modelling for Composability in Smart Home Environments. Constructing Ambient Intelligence 11, 411–420 (2008)
3. Kim, E., Choi, J.: An Ontology-Based Context Model in a Smart Home. In: Workshop on Ubiquitous Web Systems and Intelligence (UWSI 2006), pp. 11–20 (2006)
4. Gu, T., et al.: An Ontology-based Context Model in Intelligent Environments. In: Communication Networks and Distributed Systems Modeling and Simulation Conference, pp. 270–275 (2004)
5. Maximilien, W.M., Singh, M.P.: A framework and ontology for dynamic web services selection. IEEE Internet Computing 8(5), 84–93 (2004)
6. Dobson, G., Lock, R., Sommerville, I.: Qosont: A qos ontology for service-centric systems. In: EUROMICRO-SEAA, pp. 80–87. IEEE Computer Society (2005)

Research on DICOM Compliant Medical Image Interactive Visualization on Android Platform

S.H. Zheng, J. Chen, C.Y. Xu, and L. Yu

College of Physics and Information Engineering, Fuzhou University, Fuzhou, Fujian, China

Abstract. Medical image diagnosis plays a key role in modern medical science, and the medical image is the core of Digital Imaging and Communications in Medicine (DICOM) standard. This work presented an implementation of interactive visualization of DICOM 3.0 compliant medical image on the Google's Android platform which was the most popular mobile operating system at present. The application enabled an Android terminal (version 2.2) with multi-touch screen and Wi-Fi function to visualize the medical images (supporting DICOM 3.0 file format and JPEG2000 coding). And then, several DICOM compliant images of different modalities (MR, CT) and different file sizes had been used to evaluate the transmission time in two network types: WLAN and 3G network. This research indicated an Android terminal (might be a mobile phone or tablet PC) with the application was suitable to be an auxiliary facilities for medical staff, and it would improve medical efficiency and reduce the medical delay.

Keywords: DICOM, Android, Medical Image, JPEG2000.

1 Introduction

Digital Imaging and Communications in Medicine (DICOM) is widely used as a transfer protocol as well as a storage format for medical applications. DICOM not only contains the medical image data itself, but also stores extensive information about the patient, the medical examination process, and the imaging modality setup, etc. DICOM plays a key role in the establishment of Picture Archiving and Communication Systems (PACS), which might provide the electronic platform for radiology images interfacing with other medical automation systems such as Hospital Information System (HIS), Electronic Medical Record (EMR), and Radiology Information System (RIS)[1].

It is well known that the digital medical image represents tremendous amounts of data, from tens of MB for a MRI (Magnetic Resonance Imaging) image to hundreds for a multi-resolution spiral CT-scan [2]. Therefore, Medical image transfer requires higher bandwidth. So, it is a great challenge to the mobile medical application with the emphasis on real-time performance and higher resolutions.

Various types of mobile devices (e.g., Tablet PCs, Personal Digital Assistants (PDAs), etc.) supported applications for retrieving and examining patient data [3-7]. But for the reasons of limited wireless bandwidth, image size and the device cost, those applications were difficult to be widely used by the medical personnel. Wavelet-based

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 667–673.
springerlink.com

compression solution with ROI coding support for DICOM images was presented, but it was not on Android platform [8].

Android was a open source software stack for mobile devices that included an operating system, middleware and key applications, and it was announced in November 2007 by Google[9]. Since 2010, Android was listed as the best-selling smartphone platform worldwide. With the feature: Wi-Fi, 3G, Capacitive Multi-touch Screen, etc. Android 2.2 (or higher version) would be suitable for mobile healthcare terminal. Likewise, it was necessary to reduce the size of medical images.

As a consequence, DICOM format and JPEG2000 coding should be supported by Android mobile application, while the JPEG2000 standard has been implemented to support lossy and lossless compression, progressing coding and Region of Interest (ROI) coding [10]. The progressive coding allows the user to decode large image files at different resolution levels. The ROI coding enables scalable compression, retrieval, and decompression of medical images for advanced image examination of the specific areas within the image.

2 Experiment

2.1 DICOM File Format

The current release, DICOM Version 3.0 (since 1993), supported many different diagnostic imaging modalities, including CR (Computed Radiography), MR (Magnetic Resonance), CT (Computed Tomography), Nuclear Medicine, Ultrasound, X-ray Angiography, and so on.

The DICOM specification is extremely complex, and it defines the protocols for data interchange and communications as well as the specifics of the data format. The DICOM standard models real-world processes involving medical image data via a set of interdependent information entities. Each entity contains the data covering a certain aspect of the actual process (e.g. image acquisition, printing) and the real-life entities involved (e.g. patient, image modality) [11].

The DICOM File Format is described in Part 10 of the DICOM Standard: "Media Storage and File Format for Data Interchange". A DICOM file consists of a file header followed by a File Meta Information Data Set and a Data Set representing a single SOP (Service Object Pair) Instance.

The File Header is made up of a 128-byte File Preamble followed by a 4-byte prefix. The File Meta Information follows the File Header and has the structure of a Data Set with several Data Elements. All Data Elements have group number 0x0002 and contain specific information about the DICOM file. The Transfer Syntax of the File Meta Information Data set is always Explicit Little Endian. The Transfer Syntax of the Data Set following the File Meta Information is encoded as a Data Element within the File Meta Information [11].

The Data Set is stored in the portion of the file following the file header. A data set represents a single SOP instance of a particular SOP class. The Data Set contains public Data Elements identified by a tag with an even group number and it may also contain private Data Elements identified by a tag with an odd group number. Data Elements are the individual units of information within a DICOM file. Data Elements may be nested.

This means the value of a Data Element may contain a number of Data Sets. Data Elements that contain nested Data Sets are called Sequences and a single Data Set that is nested within a Sequence is called a Sequence Item [12]. Fig.1. shows the structure of the DICOM File Format. Fig. 2 represents a real DICOM compliant image file.

File header		File Meta Informatic	Data set	
Preamble	Prefix	(0002 0000)	Data element	...

Fig. 1. DICOM File Format structure

DICOM file information		
Group No.	Definition	Value
Null(File Header)	Preamble, first 128 bytes	00H, general unused
	Prefix, 4 bytes	characters 'D', 'I', 'C', 'M'
0002 0000	File Meta Elements Group Length	132
0002 0001	File Meta Info Version	256
0002 0010	Transfer Syntax UID	1.2.840.10008.1.2.4.91 (JPEG2000)
0008 0000	Identifying Group Length	152
0008 0020	Study Date	20090902
0008 0060	Modality	MR
0010 0010	Patient Name	Chen lihui
0010 0040	Patient Sex	F
…	…	…
0028 0004	Photometric Interpretation	MONOCHROME2
0028 0008	Number of frames	1
7FE0 0000	Pixel Data group length	186,716
7FE0 0010	Pixel Data	186,704

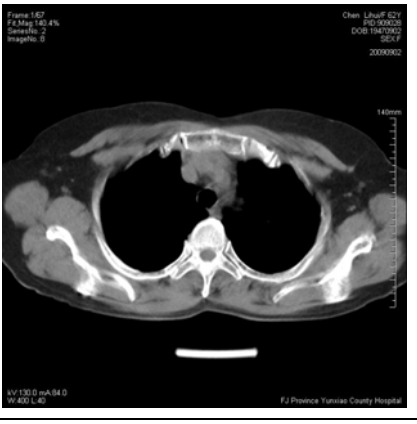


Fig. 2. DICOM-compliant image file format representation

Earlier DICOM3.0 standard only supported JPEG compression, RLE (Run-length Encoding) compression and original image data (non-compression). And the latest DICOM3.0 standard already supported JPEG2000 compression. Comparing the JPEG2000 and JPEG algorithms, it was observed that JPEG2000 algorithm was more successful than JPEG for all compression rates [13].

In DICOM file format, JPEG2000 transfer Syntax UID as follows, defined by group number: 0002 0010.

* 1.2.840.10008.1.2.4.90 (only support JPEG2000 non-compression)
* 1.2.840.10008.1.2.4.91 (support both JPEG2000 compression and non-compression)

2.2 Methods

Android platform was built on the Linux kernel version 2.6, and Android applications were written in Java. Java classes were compiled into *Dalvik* executables and run on Dalvik, a specialized virtual machine designed specifically for Android and optimized for battery powered mobile devices with limited memory and CPU.

JPEG2000 encoding and decoding procedures required more memory and higher computing speed. It was necessary to optimize the system design and the encoding / decoding algorithm, such as, using byte stream transmission, reducing data shift times to optimize wavelet transform algorithm. The JPEG2000 encoding procedure which run in the DICOM sever was as following Fig.3.

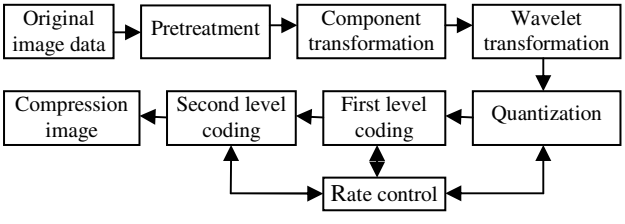


Fig. 3. JPEG2000 encoding procedure

And then, the main work in the client (DICOM terminal) was to decode the DICOM file with the JPEG2000 compression. The DICOM file format analysis procedure was as follows.

1) Acquire DICOM format file from remote DICOM server through Wi-Fi or WLAN by using *URL*, *HttpURLConnection*, *InputStream*, and *ByteArrayOutputStream* class in Package java.io, and put DICOM byte stream into a byte array (byte[] data).

2) Read the transfer Syntax UID to ascertain the information: byte sequence, whether compression or not, and so on.

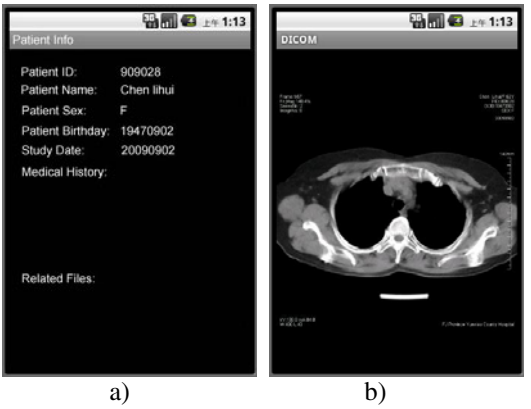
3) Obtain file elementary information (e.g. patient info, Photometric Interpretation, Number of frames, Pixel Data group length, etc.) until find the pixel data Group No..

4) Read the Group No. (7FE0 0010) to acquire the pixel data, and store them into byte arrays (byte[] frameData) corresponding to each of frames.

5) According to the elementary information (e.g. compression type, image modality, frame count, bit distribute, tec.) has been obtained to decode the files and store them in new byte arrays (byte[] decData).

6) Extract the DICOM grayscale value and map them into the scope 0-255 according to Window Width and Window Center. It should extract the color pixel values while the DICOM image was color, and at last combine them to the view data.

DICOM image data had been decoded should be stored as a general image format such as JPEG or BMP which Android platform supported. In this work, BMP format was adopted to display the image. The method *getPixels* in android.grap- hics.Bitmap package could load a separate pixel to an integer array (gray range 0-255). And the method *setPixels* could load this integer array to an image which created by the method *createBitmap*. Fig.4. showed the patient information and the visualization of a DICOM compliant image on an Android device.



a) b)

Fig. 4. Screenshots of the application: a) a patient information; b) Display of JPEG2000 decoding of a MRI

Android 2.2 had native support for multi-touch which was convenient for interaction visualization of DICOM image, e.g. zoom, rotation, mirror, pseudo-color, anti-color, and measure of distance, area, angle, and so on. A GesturesBuilder program preinstalled in Android could create character gestures. Some methods, such as *getAction, getPointerCount, getX, getY, getPressure*, etc. belonging to the class *MotionEvent*, could acquire touch action and its features, and change the image through the interface *setOnTouchListener*.

In the same way, we could append Wi-Fi function in the application if the WLAN was available. Wi-Fi network card could be operated through modifying the configuration file *AndroidMenifest.xml* by inserting the permission related network. And an object of the class *WifiManager* in the package android.net.wifi needed to be created to manage the Wi-Fi operation states. The permission sentences were as follows.

```
<uses-permission android:name = “android.permission.CHANGE_NETWORK_STATE” />
 <uses-permission android:name = “android.permission.CHANGE_WIFI_STATE” />
  <uses-permission android:name = “android.permission.ACCESS_NETWORK_STATE” />
  <uses-permission android:name = “android.permission.ACCESS_WIFI_STATE” />
```

2.3 Application Evaluation in Practice

In order to prove the application's usability, it should make some initial experiments to evaluate the application's performance. In this work, HTC Legend G6 mobile phone running Android OS version 2.2 had been used to present the application. And several DICOM compliant images of different modalities (MR, CT) and different file sizes had been used. Table 1 showed the transmission times of DICOM compliant medical images in different network types: WLAN and 3G network. It indicated that the performance of both WLAN and 3G networks was suitable be an auxiliary facilities for medical staff.

Table 1. Transmission time of DICOM compliant medical images in different network types

Image type(encoding)	File size(MB)	Time(secs)	
		WLAN	3G
MR(JPEG)	0.808	3.618	8.270
MR(JPEG2000 Lossy)	0.096	1.280	1.835
MR(JPEG2000 Lossless)	0.186	2.245	4.860
CT(JPEG)	0.608	2.834	7.626
CT(JPEG2000 Lossy)	0.102	1.560	2.032
CT(JPEG2000 Lossless)	0.204	2.818	3.612

3 Conclusions

Mobile browser of medical information resources (medical image data and relevant electronic medical record) plays a key role in modern medical science. In this paper, an implementation of interactive visualization of DICOM compliant medical image was presented on the Android platform which was the most popular mobile operating system at present. The application enabled the Android terminal with multi-touch function to visualize the medical image (supporting DICOM 3.0 file format and JPEG2000 coding) via two different wireless networks: WLAN and 3G network. This research would improve medical efficiency and reduced the medical delay.

Future work might include improving security by implementing advanced user authentication techniques on the mobile device (e.g., through voice recognition) and deploying the platform in real healthcare environment for evaluating the system in terms of user acceptability and performance.

Acknowledgments. The authors thank the other members of the Laboratory of Medical Image for their support and advice, and thank the radiologists of Fujian Province People's Hospital for DICOM images.

References

1. Pianykh, O.S.: Digital Imaging and Communications in Medicine (DICOM): A Practical Introduction and Survival Guide, pp. 3–6. Springer, Heidelberg (2008)
2. Ramakrishnana, B., Sriraam, N.: Transmission of DICOM Images with Effective Low Bandwidth Utilization. Digital Signal Processing 16, 825–831 (2006)
3. Andrade, R., von Wangenheim, A., Bortoluzzi, M.K.: Using Mobile Wireless Devices for Interactive Visualization and Analysis of DICOM Data. In: Proceedings of the 16th IEEE Symposium on Computer-Based Medical Systems (CBMS 2003), pp. 1082–1086 (2003)
4. Andrade, R., von Wangenheim, A., Bortoluzzi, M.K.: Wireless and PDA: A Novel Strategy to Access DICOM-Compliant Medical Data on Mobile Devices. International Journal of Medical Informatics 71(2), 157–163 (2003)
5. Lin, Y., Jan, I., Ko, P., Chen, Y.: A Wireless PDA-based Physiological Monitoring System for Patient Transport. IEEE Trans. Inf. Technol. Biomed. 8(4), 439–447 (2004)
6. Mendonça, E.A., Chen, E.S., Stetson, P.D., et al.: Approach to Mobile Information and Communication for Health Care. International Journal of Medical Informatics 73, 631–638 (2004)
7. Bojovic, M., Bojic, D.: Mobile PDR: A Mobile Medical Information System Featuring Update Via Internet. IEEE Trans. Inf. Technol. Biomed 9(1), 1–3 (2005)
8. Maglogiannis, I., Doukas, C., Kormentzas, G., et al.: Wavelet-Based Compression With ROI Coding Support for Mobile Access to DICOM Images Over Heterogeneous Radio Networks. IEEE Transactions on Information Technology in Biomedicine 13(4), 458–466 (2009)
9. Zhang, J.: Android vs iPhone (April 2010)
10. http://www.jpeg.org/jpeg (2000)
11. Güld, M.O., Kohnen, M., Keysers, D.: Quality of DICOM Header Information for Image Categorization. In: Proc. of SPIE, pp. 1068–1073 (2002)
12. http://en.wikipedia.org/wiki/DICOM
13. Gulkesen, K.H., Akman, A., Yuce, Y.K.: Evaluation of JPEG and JPEG2000 Compression Algorithms for Dermatological Images. J. Eur. Acad. Dermatol Venereol 24(8), 893–896 (2010)

Using a Single-Layer Neural Network to Generate a FIR Filter That Improves Digital Images Using a Convolution Operation

Jakub Peksinski and Grzegorz Mikolajczak

West Pomeranian University of Technology,
Faculty of Electrical Engineering
26 Kwietnia 10, 71-126 Szczecin, Poland
jakub.peksinski@zut.edu.pl, grzegorz.mikolajczak@zut.edu.pl

Abstract. The paper features the method of application of a neural network for improving the quality of the digital images generated by means of devices for backup and processing of data into a digital form which construction is based on the Charge Coupled Device (CCD) structure. In order to introduce the problem, the digital images were generated by means of two scanners (including a high class and a low class scanner) and afterwards the images were subject to an objective and a subjective evaluation. An objective evaluation was performed using two quality criteria, i.e. MSE (Mean Square Error) and NMSE (Normalized Mean Square Error). A FIR (Finite Impulse Response) filter applied for filtration of a low quality image was obtained as the result of the neural network learning process. The image so generated as the result of filtration features a superior quality in comparison to the original.

Keywords: Image processing, neural network, FIR filter.

1 Introduction

Dynamic development of the image acquisition equipment, which bases on CCD matrice, that has occurred in recent years was the cause of the abundance of those devices in household usage. However, this equipment has many disadvantages caused mainly by the quality of the CCD matrice and electronics used to build it. Professional equipment that has high quality parameters and truly mapps an image is still expensive and is owned mainly by professional companies that deal with digital image processing. The aim of this article is to present the possibilities of improvement the quality of images acquired by low class devices, so that their quality is comparable to images acquired by high class devices.

A digital image is a more or less accurate mapping of an analog image [1]. The better the quality of the equipment located in the image acquisition path, the more accurate the mapping. Using a device to convert an analog image to a digital form has a considerable impact on the quality of the mapping. Nowadays, the devices featuring construction based on the CCD structure are commonly used for conversion. Various interferences and errors affecting the quality of the obtained digital image occur

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 675–681.
springerlink.com

during acquisition of data by means of such devices. Following factors are responsible for the quality of images acquired by those devices:

- Number of pixels (resolution),
- Quantum efficiency (sensitivity),
- Noises in matrice,
- Quality of the A/C converter.

The occurrence of any error always causes a decline in quality of the obtained images. The basic method of improving the quality of an image is the linear filtration of an image [2]. Alternatively, any device of very favorable parameters can be used but it usually implies a high device cost. It can be easily found that the quality of the obtained digital images depends strongly on the quality of a device applied for image acquisition.

The purpose of the paper is proving that it is possible to obtain an image generated by means of a low class scanner of the quality similar to the quality of a digital image generated by means of a high class scanner. Firstly, an image is subject to the linear filtration. The paper features application of a neural network for generation of the impulse response as a FIR filter. Design and optimization of FIR filters using a neural network [3] features a huge potential in the field of digital image processing. The results of the tests proved that the mask of a filter used for correction of the quality of an image can be obtained as the result of the neural network learning process.

2 Comparison of Images Obtained from Two Different Scanners

A practical experiment was performed in order to show a difference in the quality of digital images generated by means of a low class scanner and images generated by means of a high class scanner. For the comparison of images, the following criteria:

- *Mean Square Error* (MSE) according to formula no. 1 [4]:

$$MSE = \frac{1}{MN}\sum_{x=1}^{M}\sum_{y=1}^{N}\left[f_{in}(x,y) - f_{out}(x,y)\right]^2 \tag{1}$$

- *Normalized Mean Square Error (NMSE)* according to formula no. 2 [4]:

$$NMSE = \frac{\sum_{x=1}^{M}\sum_{y=1}^{N}\left[f_{in}(x,y) - f_{out}(x,y)\right]^2}{\sum_{x=1}^{M}\sum_{y=1}^{N}\left[f_{in}(x,y)\right]^2} \tag{2}$$

Where: f_{in} - the image generated by means of a high class scanner;
f_{out} - the image generated by means of a low class scanner.

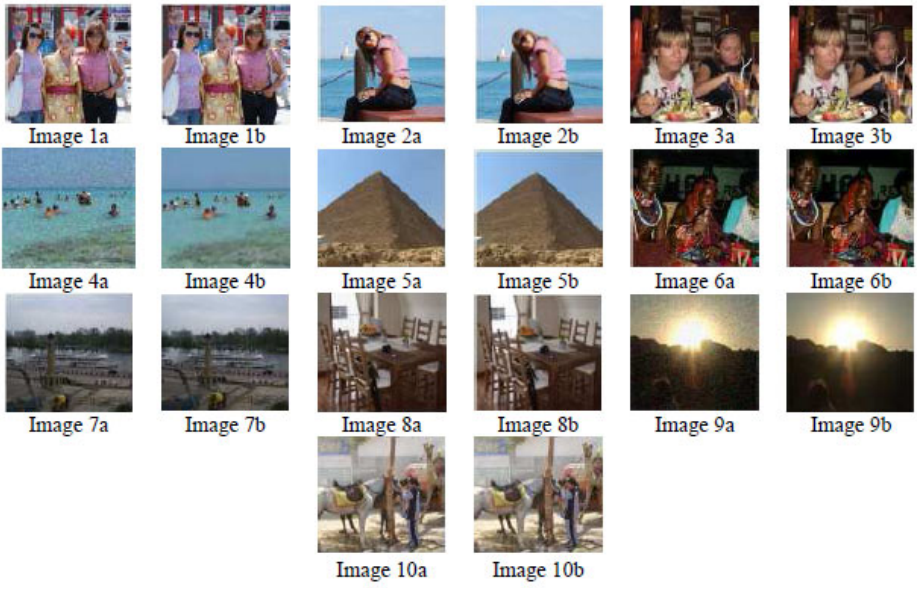


Fig. 1. The images used in the test. *a* – image from high class scanner, *b* – image from low class scanner

The results of comparison of the images are presented in table no. 1.

Table 1. The values of criteria MSE and NMSE

Image pair	MSE	NMSE
Image 1a and 1b	52,381	0,771
Image 2a and 2b	57,293	0,765
Image 3a and 3b	57,511	0,791
Image 4a and 4b	50,022	0,712
Image 5a and 5b	54,135	0,783
Image 6a and 6b	56,772	0,737
Image 7a and 7b	58,635	0,754
Image 8a and 8b	54,494	0,777
Image 9a and 9b	55,765	0,798
Image 10a and 10b	49,987	0,734

The analysis of the results shown in table1 proves that slight differences between the images can be noticed. This fact is caused by the errors specified in section 1 of the paper. It means that two digital images of different quality were obtained.

The images generated by means of a low class scanner can be fitted to the quality of the images generated by means of a high class scanner by applying the digital image processing. The digital image filtration is one of the most often used methods of improving the quality of the digital images. The operation guarantees that some undesirable objects such as interferences or noises can be eliminated from an image.

However, it is necessary to possess the knowledge related to the image interferences in order to perform the operation efficiently.

3 Generation of a FIR Filter by Means of a Neural Network

The authors applied a neural network for generation of the impulse response as a FIR filter according to the following assumptions:

1. A filter shall be a 5x5 mask
2. A filter shall execute the discrete convolution according to formula no. 3

$$f'_{out}(x,y) = \sum_{i,j \in k} f_{out}(x-i, y-j) \cdot w(i,j) \tag{3}$$

Where: $f_{out}(x,y)$ - low class image before filtering;
$f'_{out}(x,y)$ - image after filtering; $w(i,j)$ - filter mask.

Thanks to the above specified assumptions, a filter can be efficiently applied for improving the quality of the images using common graphical software such as *PhotoShop* or *Paint Shop Pro*.

A neural network used for generation of a FIR filter and applied by the authors is a unidirectional network operating in the mode of supervised learning. This configuration implies that the input images in such networks are processed by neural network from the input layer to the output. It means that the network outputs at a given moment *t* are input-dependent only at the same moment. Therefore, the connection weights serve as parameters.

To obtain a digital filter with *5x5* mask the authors of the article have used one direction neural network working in mode with teacher – figure no 2. This network has 25 inputs ($x_0, x_1, ..x_{24}$), into which in the course of learning, we give image acquired by low class scanner (y1). In the course of learning, in which the actualization of wages was based on recurrent presentation of the images acquired by low class scanner, the change in wage rate of the neural network occurred ($w_0, w_1, .., w_{24}$) according to formula no 4 [4].

$$w_i(j+1) = w_i(j) + n \cdot \delta_\mu \cdot m_i^\mu \tag{4}$$

Where: j – step number, n – learning parameter, $w_i(j)$ - wage, δ_μ - error described by formula no 5, y – result value,

$$\delta_\mu = y1 - y = \begin{cases} -1 \; for \; y1 = 0 \; and \; y = 1 \\ 0 \; for \; y1 = y \\ 1 \; for \; y1 = 1 \; and \; y = 0 \end{cases} \tag{5}$$

The whole learning process continues according to the following steps:

- Input of an image of low quality,
- Determining an error according to formula no 5,

- Determining new wages according to formula no 4,
- Repeating steps no 1 to 3 as long as the error becomes satisfactorily small.

The result of the operation of the neural network is the filter of *5x5* mask (25 wages), which in the course of realization of discrete tangle operation described by formula no 3, improve the quality of digital images acquired by low class scanner in such a way, that they will correspond with images from the model scanner.

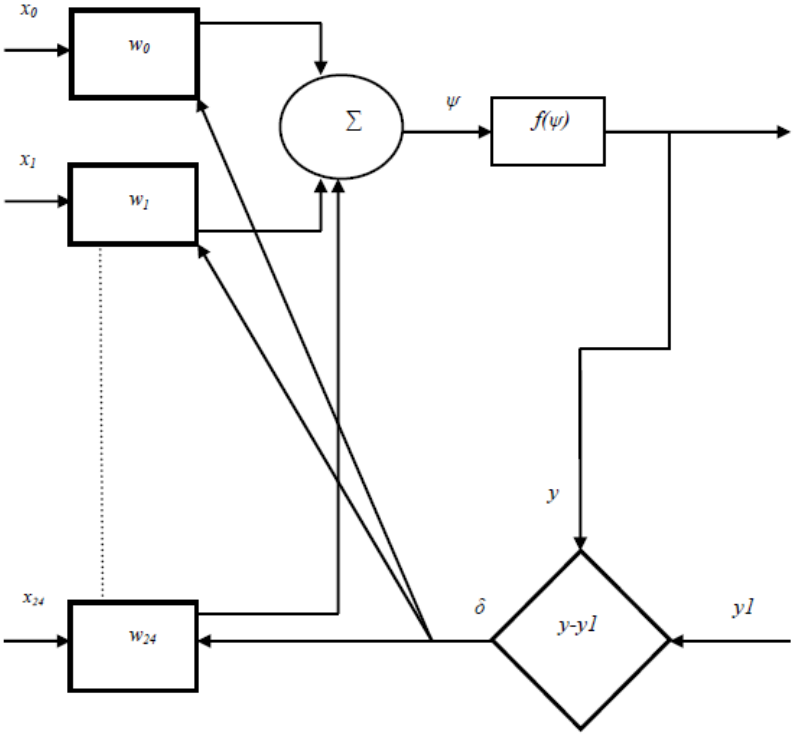


Fig. 2. Neural network configuration

During the network learning process were obtained for 10 image pairs ten filter masks, some of which were presented in formula:

$$w = \begin{bmatrix} -0.0082 & -0.0098 & 0.0231 & -0.0073 & -0.0036 \\ -0.0110 & 0.0094 & 0.1588 & -0.0047 & 0.0008 \\ 0.0270 & 0.1084 & 0.4435 & 0.1087 & 0.0244 \\ 0.0012 & -0.0058 & 0.1558 & 0.0098 & -0.0103 \\ -0.0064 & -0.0059 & 0.0230 & -0.0074 & -0.0076 \end{bmatrix} \quad (6)$$

The amplitude characteristics of the generated filters are presented in figure no 3 During the analysis of weight coefficients of the filters featured in formula no. 8 and the amplitude characteristics shown in figure no. 3, it can be found that the filters generated during the neural network learning process are low-pass. Therefore, most interferences of an image generated from a low class scanner can be classified as noises.

In order to check the efficiency of operation of the obtained filters, each of ten images generated from a low class scanner was subject to the filtration process according to formula no. 3. A new image so generated was compared with a reference image by means of the criteria according to formulas 1, 2. The results are shown in table 2.

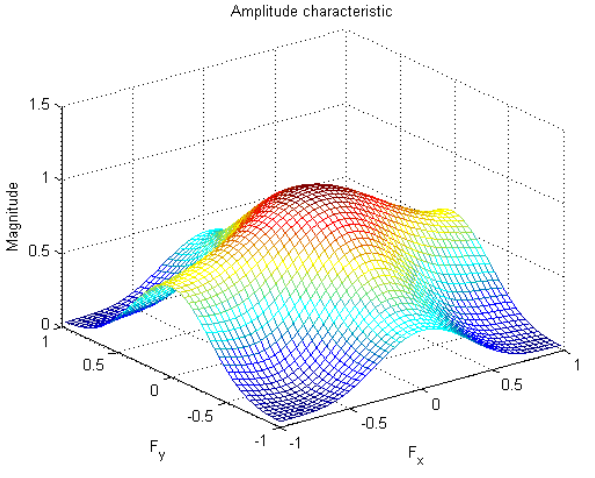


Fig. 3. The amplitude characteristics of the generated filters w, presented in formula no 6

Table 2. The values of criteria for the images after filtering by filter *w* (6)

	MSE before filtering	MSE after filtering	NMSE before filtering	NMSE after filtering
Image 1a	52,381	23,311	0,771	0,998
Image 2a	57,293	30,764	0,765	0,982
Image 3a	57,511	32,349	0,791	0,924
Image 4a	50,022	21,576	0,712	0,943
Image 5a	54,135	26,023	0,783	0,997
Image 6a	56,772	29,351	0,737	0,964
Image 7a	58,635	34,976	0,754	0,972
Image 8a	54,494	27,009	0,777	0,998
Image 9a	55,765	29,731	0,798	0,951
Image 10a	49,987	20,097	0,734	0,978

The results from table 2 prove explicitly the efficiency of the filters generated by means of a neural network. The efficiency is related to the objective evaluations performed by means of *MSE* and *NMSE* criterion.

4 Final Conclusions

Analyzing the results presented in table no 2, it can be stated that the digital filter of 5x5 mask obtained in the course of learning process of the neural network fulfils its task. After applying tangle operation, images acquired by low class scanner have become comparable in quality with model images. Surely, there is no 100 percent improvement, because the filter mask is relatively small. However, the advantage of it is that through realization of the discrete tangle operation, it can be used in popular programmes for image procession such as CorelDraw or Photoshop.

Of course, the learning process of the neural network is long but it occurs only once and filter obtained in such a way can be applied in signal processors, so that the improvement of photos can be implemented practically in real time.

References

1. Pratt, W.K.: Digital Image Processing, PIKS Inside. Willey (2001)
2. Bovik, A.C.: Handbook of Image and Video Processing, 2nd edn. Department Of Electrical And Computer Engineering, The University Of Texas AT Austin, TEXAS. Elsevier Academic Press (2005)
3. Gupta, P.K., Kanhirodan, R.: Design of a FIR Filter for Image Restoration using Principal Component Neural Networks. In: IEEE International Conference on Industrial Technology, ICIT 2006, December 15-17, pp. 1177–1182 (2006)
4. Wang, Z., Bovik, A.C.: Mean Squared Error: Love it or Leave it? IEEE Signal Processing Magazine (2009)
5. Wang, W., Bovik, A.C.: A Universal Image Quality Index. IEEE Signal Processing Letters XX(Y) (2002)
6. Korbicz, J., Obuchowicz, A., Uciński, D.: Sztuczne sieci neuronowe. Akademicka Oficyna Wydawnicza, Warszawa 1994;
7. Grainger, E.M., Cuprey, K.N.: An Optical Merit Function (SQF) Which Correlates With Subjective Image Judgments. Fotografic Science and Engineering 16(3) (1992)
8. Wyszecki, G.: Color appearance. In: Boff, K.R., Kaufman, L., Thomas, J.P. (eds.) Handbook of Perception an Human Performance (1986)
9. Ackenhusen, J.G.: Real-Time Signal Processing: Desing and Implementation of Signal Processing Systems. PTR Prentice Hall, Upper Saddle River (1999)
10. Bellanger, M.: Digital Processing of Signal. Theory and Practice. Wiley, Chichester (1989)
11. Thimm, G., Fiesler, M.: Neural network initialization, in Form Neural to Artificial Neural Computation. In: Mira, J., Malaga, pp. 533–542 (1995)
12. Sheikh, H.R., Bovik, A.C.: Image information and visual quality. IEEE Trans. Image Processing 12(11), 1338–1351 (2003)
13. Papas, T.N., Safranek, R.J., Chen, J.: Perceptual criteria for image quality evolution., Handbook of Image and Video Processing, 2nd edn. (May 2005)
14. Cammbell, C.: Neural Network Theory. University Press, Bristol (1992)

Improved Particle Swarm Optimization Algorithm and Its Application to Global Optimization for Complex Function

Jing Zhang and Ze Zhang

College of Electronic Information and Engineering, Inner Mongolia University, Hohhot, China, 010021
{zj8705,zhangzeimu}@163.com

Abstract. Particle Swarm Optimization (PSO) is a kind of group intelligent heuristic global evolution algorithm, which was proposed recently. It is used to solve complex optimization problems which are non-linear and non-differentiable, as well as combinatorial optimization problems. With advantages such as simple principle, it has fast convergence rate and it is liable to be implemented. Firstly, this paper proposes an improved PSO to search the global optimum of function after researching principle of the original one, as well as online and offline evaluation criterion of this improved algorithm. Secondly, an optimal algorithm is proposed to solve the problems with restrictions based on the combination of PSO and penalty function. Finally, analyzing the testing function and inertia weight, the paper affirms effectiveness of this improved algorithm through simulations. It is revealed that this improved algorithm is liable to search the global optimum of function and realize the global optimization.

Keywords: Particle Swarm Optimization, Weight, Performance Evaluation, Penalty Function.

1 Introduction

Particle Swarm Optimization (PSO) proposed by Kennedy Eberhart in 1995[1][2] is a kind of heuristic group intelligent evolution computing technology, which searches the optimum by means of iteration. Over other technologies based on group optimization, PSO has advantages those less complex operations and less parameter. So it develops rapidly and is wildly applied to fields of artificial intelligence, image processing, pattern recognition and multi-objection optimization [9], etc.

This paper researches principle and performance of PSO with emphasis on application of it in function optimization. Associating with performance evaluation criterion of GA, it proposes online and offline evaluation criterion of PSO. Besides, this paper proposes an optimization algorithm which is used to solve the problems with restrictions based on the combination of PSO and penalty function, and affirms effectiveness of this algorithm through simulations.

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 683–690.
springerlink.com

2 Principle of PSO

In PSO, each particle has position and speed, the position of it represents a potential solution of the problem and a random solution is formed during initializing. The potential solution of each optimization problem is compared to a bird that is searching space, and it is also called 'particle'. In the whole searching process, adaptive values of all the particles depend on the value of the chosen optimization function, and each particle has the following information: position of particle; current speed of particle; the optimal position of particle by far (P_{best}). This optimal position is called individual extreme value. The optimal position found by the whole particle swarm is viewed as the shared flying experience of the whole particle swarm, it is represented to be G_{best}, which is the global extreme point.

Suppose that M particles exist in D-dimensional searching space, position of the i^{th} particle is $X_i=(x_{i1},x_{i2},...x_{iD})$, its speed is $V_i=(v_{i1},v_{i2},...,v_{iD})$, i = 1, 2,…, M. Assume that the optimal position searched by the i^{th} particle is $P_i=(p_{i1}, p_{i2}, ..., p_{iD})$, which is known as P_{best}, the optimal position searched by the whole particle swarm is $P_g=(p_{g1}, p_{g2}, ..., p_{gD})$, which is also known as G_{best}. The position and speed of particle in PSO are respectively shown as Equation (1) and Equation (2), which were creatively proposed by Knenedy and Eberhrtn.

$$V_{id}(t+1) = V_{id}(t) + c_1 r_1 (p_{id} - x_{id}(t)) + c_2 r_2 (p_{gd} - x_{id}(t)) \tag{1}$$

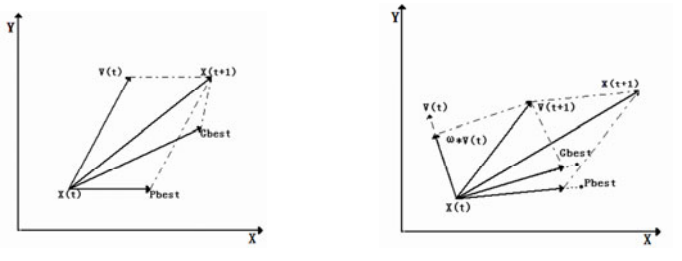
$$X_{id}(t+1) = x_{id}(t) + V_{id}(t+1) \tag{2}$$

Where, i= 1, 2, ..., M; d=1, 2,..., D; r_1 and r_2 are independent with each other, both of them distribute evenly in [0, 1]. Parameters c_1 and c_2 are acceleration constants (learning rate).

In order to improve optimization of PSO more effectively, Shi Y.H and Eberhart [4] introduced inertia weight ω [5][8] into memory part, then Equation (1) is subsequently modified into Equation (3) shown as follow.

$$V_{id}(t+1) = \omega V_{id}(t) + c_1 r_1 (p_{id} - x_{id}(t)) + c_2 r_2 (p_{gd} - x_{id}(t)) \tag{3}$$

Where, ω is negative, it represents effect of the former speed on the current speed. Ability of PSO in global and local optimization is adjusted according to the variance of ω. The larger ω is, the greater the effect is, the more robust the global optimizing ability is, and the weaker the local optimizing ability is, the opposite is the other way around.



a. Vectors of original PSO b. Vectors of standard PSO

Fig. 1. Vectors of PSO

3 Steps of PSO

Steps of PSO are shown as follows:

I. Initialize a group of particles (size of group is N), including position and speed at random. Searching space is n-dimensional. Then a matrix generating at random is defined as follow.

$$\begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n} & v_{11} & v_{12} & \cdots & v_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} & v_{21} & v_{22} & \cdots & v_{2n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ x_{N1} & x_{N2} & \cdots & x_{Nn} & v_{N1} & v_{N2} & \cdots & v_{Nn} \end{bmatrix}$$

Where, $\{x_{ij} | i=1, 2, ..., N, j=1, 2, ..., n\}$ denotes that position of the i^{th} particle is j, v_{ij} is its speed, both of them distribute evenly in $[R_{1j}, R_{2j}]$.

II. Calculate adaption of each particle (objective function).

III. Calculate the optimal position each particle finds, which is shown as Equation (4).

$$P_{best_i}(t) = (x_{i1}, x_{i2}, \ldots, x_{in}) \tag{4}$$

Particles in this position have the optimal adaption; this optimal position is defined as follow.

$$P_{best_i}(t+1) = \begin{cases} P_{best_i}(t) & f\left(x_1(t+1), x_2(t+1), \ldots, x_n(t+1) < P_{best_i}(t)\right) \\ x_i(t+1) & f\left(x_1(t+1), x_2(t+1), \ldots, x_n(t+1) \geq P_{best_i}(t)\right) \end{cases} \tag{5}$$

Calculate the optimal position the whole group finds; which is the optimal position in globe.

$$G_{best_i}(t) = \max\left\{f(P_{best_1}(t)), f(P_{best_2}(t)), \ldots, f(P_{best_3}(t))\right\} \tag{6}$$

IV. Modify speed and position of each particle according to Equations (7).

$$\begin{cases} v_{ij}(t+1) = \omega \cdot v_{ij} + c_1 \cdot r_1 \cdot \left(P_{best_i}(t) - x_{ij}(t)\right) + c_2 \cdot r_2 \cdot \left(G_{best_i}(t) - x_{ij}(t)\right) \\ x_{ij}(t+1) = x_{ij}(t) + v_{ij}(t+1) \end{cases} \tag{7}$$

V. If condition for iteration terminating is not satisfied, the algorithm will go to step II. If it is, PSO finishes its work.

Generally, the condition for iteration terminating is chosen to be either the maximal iterating times ‘run’ or that the optimal position found by far is smaller than the minimal threshold predefined.

4 Inertia Weight

As far as the three components in the speed calculation, relationship between inertia speed component $v_{inertia}$ and adaption is weak.In order to ensure convergence at the end of iteration, weight of $v_{inertia}$ must be gradually weakened. Inertia factor [7][8] ω will monotonously decrease with the raising of iteration.

Shi Y.H and Eberhart proposed a formula to make inertia weight decrease linearly; it does well in global optimization in earlier phase and well in local optimization in later phase. However, speed of convergence is slow.

Assume d_{ig} is distance between the i^{th} particle and the global optimal particle P_g, $d_{\max}$ is the maximal distance and $d_{\min}$ is the minimal distance.When d_{ig} varies from $d_{\min}$ to $d_{\max}$, its inertia weight ω increases in a monotonic way with the distance; the formula is shown as follow.

$$w_i = w_{\min} + (d_{ig} - d_{\min})(\bar{w}_{\max} - w_{\min})(d_{\max} - d_{\min}) \tag{8}$$

Then we will get Equation (9) with iteration process.

$$\bar{w} = w_{\max} - \frac{(w_i - w_{\min})t}{T_{\max}} \tag{9}$$

And weight of each particle is calculated by Equation (10), which will be available on the basis of Equation (8) and Equation (9).

$$\bar{w} = w_{\max} - \frac{(d_{ig} - d_{\min})(w_{\max} - w_{\min})t}{(d_{\max} - d_{\min})T_{\max}} \tag{10}$$

5 Performance Evaluation Criterion of PSO

Two measurements used to analyze GA[3] quantitatively in this paper can be introduced into PSO. Dynamic characteristics are tested by online performance, and convergent characteristics are tested by offline performance [6].

I. Online Performance Evaluation Criterion.

Theorem 1: Assume $X_e(s)$ is online performance of strategy s under environment e; $f_e(t)$ is an objective function or average adaption function under environment e at time t or t^{th} generation. Then $X_e(s)$ is defined as Equation (11).

$$X_e(s) = \frac{1}{T}\sum_{t=1}^{T} f_e(t) \tag{11}$$

It is revealed that if online performance is able to be represented by average adaption, online performance will be available by means of calculating the ratio of sum of average adaption of all generations and the number of generation. Parameter $f_e(t)$ in Equation (11) is the average adaption of all generations.

II. Offline Performance Evaluation Criterion

Theorem 2: Assume $X_e(s)$ is offline performance of strategy s under environment e. Then $X_e(s)$ can be defined as Equation (12).

$$X_e^*(s) = \frac{1}{T}\sum_{t=1}^{T} f_e^*(t) \tag{12}$$

Where, $f^*_e(t)=best\{ f^*_e(1), f^*_e(2), f^*_e(3), \ldots, f^*_e(t)\}$. Equation (12) reveals that offline performance is cumulative average of optimal performance at some specified time.

In choosing evaluation criterion of genetic algorithm, we must take the effect brought by random factors on searching results of PSO into consideration. In this paper, however, we can directly use online performance and offline performance to test the dynamic and convergence characteristics respectively without considering those issues.

6 Penalty Function

The basic principle of penalty function is fabricating some kind of punish function according to characteristics of constraints, and introducing punish function into objective function, transforming the solution for problems without any constraints into those with constraints. Traditional penalty function method is generally divided into external and internal penalty function method. And the former is the key researched by this paper.

General form of external penalty function is shown as Equation (13).

$$B(x) = f(x) + [\sum r_i G_i + \sum c_j H_j] \qquad G_i = \max[0, g_i(x)]^a \qquad H_j = |h_j(x)|^b \tag{13}$$

Where, $B(x)$ is a new objective function in optimization process; respectively, G_i and H_j are functions of $G_i(x)$ and $H_j(x)$, which are constraints; r_i and c_j are constants, known as penalty factors; generally, both a and b are chosen to be 1 or 2. In improved PSO, the adaption function is modified into Equation (14).

$$adaption = f(X) + f'(X) \qquad f'(X) = e^{g_j(X)} \tag{14}$$

7 Common Testing Function

These are some common testing functions in PSO: Griewank, Rastrigrin, Sphere, Schwefel, Ackly, Rosen-brock and so on. Three of them are listed in Table 1 in order to compare effectiveness of PSO with different testing function.

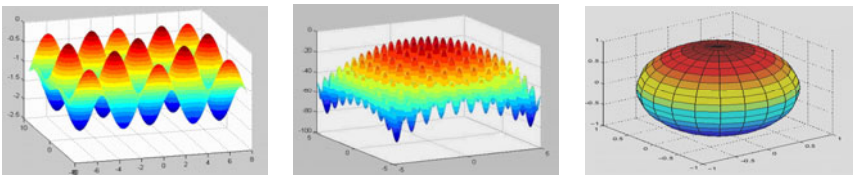
Table 1. Testing function table

Function name	Expression	Dimension
Griewank	$f(x) = \frac{1}{4000}\sum_{i=1}^{n} x_i^2 - \prod_{i=1}^{n} \cos\frac{x_i}{\sqrt{i}} + 1$	30
Rastrigrin	$f(x) = \sum_{i=1}^{n} [X_i^2 - 10\cos(2\pi x_i) + 10]$	30
Sphere	$f(x) = \sum_{i=1}^{n} X_i^2$	30

Although these functions are rich in local optimization, it is still difficult to search optimization. Griewank function is a multi-crossing concept function which interacts significantly. Its global minimum is 0 at the point x=0 and this function reaches local minimum when Equation (15) is satisfied.

$$x_i \approx \pm k\pi\sqrt{i} \quad (i = 1, 2, \ldots, m; k = 1, 2, \ldots, n) \tag{15}$$

Rastrigrin function is a typical multi-modal function which is non-linear and has a wide searching space, many local minimum points and tall obstructions. It is generally viewed as a complex multi-modal problem which is difficult for GA to deal with sphere function reaches its minimum 0 at the point (x_1, x_2, ..., x_n)= (0,0,..., 0). It has only one extreme point and is easy to be solved, so this paper pays more attention to the solving speed of this problem.



a. Griewank b.Rastrigrin c. Sphere

Fig. 2. Three testing function graphs

8 Parameter Coding

In MATLAB, we use real codes to represent position and speed of particles. Format of coding for particle parameter is shown as Fig.3. Where, ParticleSize and SwarmSize denote dimension of the parameter and the size of swarm, respectively.

$X_1, X_2, X_3, \ldots, X_{ParticleSize}$	$V_1, V_2, V_3, \ldots, V_{ParticleSize}$	$F(X)$

Fig. 3. Format of coding for particle parameter

Position, speed and current adaption of each particle are recorded in Table 2. *W*, *V* and *F* are used to denote location, speed and the current adaption, respectively. Assume that the number of particles is *N*, and dimension of the each particle is *D*.

Table 2. ParSwarm table

$W_{1,1}$	$W_{1,2}$	$W_{1,D}$	$V_{1,1}$	$V_{1,2}$	$V_{1,D-1}$	$V_{1,D}$	F_1	1st particle
$W_{2,1}$	$W_{2,2}$	$W_{2,D}$	$V_{2,1}$	$V_{2,2}$	$V_{2,D-1}$	$V_{2,D}$	F_2	2nd particle
…	…	…	…	…	…	…	…	……
$W_{N,1}$	$W_{N,2}$	$W_{N,D}$	$V_{N,1}$	$V_{N,2}$	$V_{N,D-1}$	$V_{N,D}$	F_N	N^{th} particle

The historic optimal solution of each particle (particle historical optimal adaption) and the global optimal solution searched by all of the particles are recorded in Table 3. W_g and W_i denote the global and local optimal solution, respectively.

Table 3. OptSwarm table

$W_{j,1}$	$W_{j,2}$	… $W_{j,D-1}$	$W_{j,D}$	historic optimal solution of the 1st particle
$W_{k,1}$	$W_{k,2}$	… $W_{k,D-1}$	$W_{k,D}$	historic optimal solution of the 2nd particle
…	…	… …	…	…
$W_{l,1}$	$W_{l,2}$	… $W_{l,D-1}$	$W_{l,D}$	historic optimal solution of the N-1th particle
$W_{m,1}$	$W_{m,2}$	… $W_{m,D-1}$	$W_{m,D}$	historic optimal solution of the Nth particle
$W_{g,1}$	$W_{g,2}$	… $W_{g,D-1}$	$W_{g,D}$	historic optimal solution of the global particles

9 Simulations

Initialize particle swarm and draw location of initialized swarm in testing function graph, as well as the results after iterations. We could simplify this process by means of the example shown as follow. Find the maximum of function y=1-cos (3*x)* exp(-x) in [0, 4]. Results of each step of this algorithm are shown as follows.

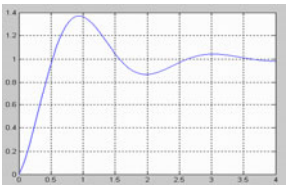


Fig. 4. Objective function

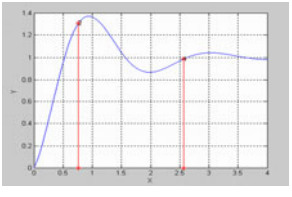


Fig. 5. Initialization

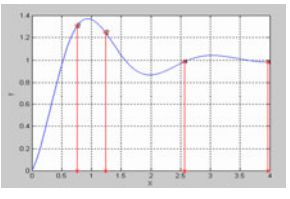


Fig. 6. First modification

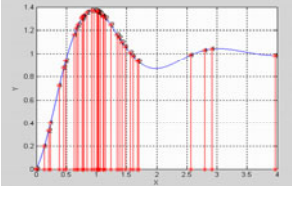


Fig. 7. 21th modification

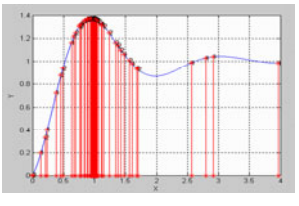


Fig. 8. Final modification (30 iterations)

The function eventually achieves its maximum y = 1.3706 when x=0.9350-0.9450 after 30 iterations. We propose PSO to search the maximum of multi-modal function (Equation (16)) in MATLAB based on characteristics of the particle swarm and get a satisfied result, which converge to the optimal solution 0.1275, with the maximal of the function is 19.8949.

$$f(x) = 10 + \frac{\sin(1/x)}{(x-0.16)^2 + 0.1}, \qquad x \in (0,1) \tag{16}$$

We use online and offline performance mentioned above to evaluate performance of this algorithm, Figure 9 is reveals the characteristic curve of Equation (16) in interval (0,1), Figure 10 and Figure 11 represent online and offline performance curve obtained by means of the algorithm refereed in this paper, respectively, the number of

evolutionary generation is 100. Assume the number of swarm is 500, the number of evolutionary generation is 100, learning factor is 2, and weighted coefficient varies from 0.1 to 0.9.

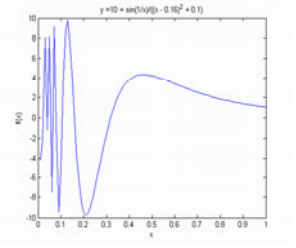


Fig. 9. Characteristic of (18)

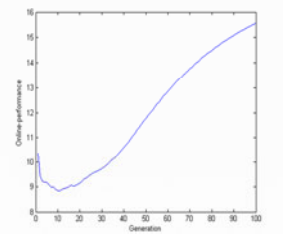


Fig. 10. Online performance

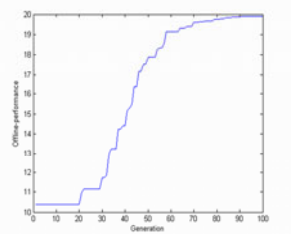


Fig. 11. Offline performance

10 Conclusions

This paper deeply researches PSO and proposes an improved PSO to solve problems those have restrictions based on testing function, penalty function and performance evaluation criterion of this algorithm. And then this improved PSO is applied to global optimization for functions, result of simulation reveals that the optimal solution in globe is correct and this algorithm is effective when it is applied to global optimization for functions.

References

1. Kennedy, J., Eberhart, R.C.: Particle Swarm Optimization. In: Proc. IEEE Int'l Conference on Neural Networks, IV, pp. 1942–1948. IEEE Service Center, Piscataway (1948)
2. Eberhart, R.C., Kennedy, J.: A New Optimizer Using Particle Swarm Theory. In: Proc. Sixth International Symposium on Micro Machine and Human Science, pp. 39–43. IEEE Service Center, Nagoya, Japan (1995)
3. Kristinsson, K., Dumont, G.A.: System Identification and Control using Genetic Algorithms. IEEE Trans., Systems Man and Cybernetics 22(5), 1033–1046 (1992)
4. Shi, Y.H., Eberhart, R.C.A.: Modified Particle Swarm Optimization. In: IEEE World Congress on Computational Intelligence, pp. 69–73 (1998)
5. Wang, W.-B., Lin, C., Zheng, Y.-K.: Experiment and Analysis of Parameters in PSO. Journal of Xi Hua University (Natural Science Edition) 27(1), 76–80 (2008)
6. Chen, J.-H., Wang, G., Wang, D.-Y.: Study on Performance of Particles Swarm Optimization and Its Application. Journal of Zhong Yuan University of Technology 18(4), 1 (2007)
7. Wang, Q.-F., Wang, Z.-J., Wang, S.-T.: A Modified Particle Swarm Optimizer Using Dynamic Inertia Weight. China Mechanical Engineering 16(11), 945–948 (2005)
8. Chen, G.-M., Jia, J.-Y., Han, Q.: Study on the Strategy of Decreasing Inertia Weight in Particle Swarm Optimization Algorithm. Journal of Xi'an JiaoTong University 40(1), 53–56 (2006)
9. Zhang, L.-B., Zhou, C.-G., Ma, M., Liu, X.-H.: Solutions of Multi-Objective Optimization Problems Based on Particle Swarm Optimization. College of Computer Science and Technology of Jilin University 41(7), 1286–1291 (2004)

A Distributed Estimation Algorithm in Binary Sensor Network for Tracking Moving Target

Zixi Jia[1], Mo Chen[1,2], and Chengdong Wu[1]

[1] College of Information Science and Engineering
[2] Software College
Northeastern University, 110819
Shenyang, China
{jiazixi,chenmo,wuchengdong}@ise.neu.edu.cn

Abstract. At the present time, some filter algorithms are generally employed for the research of tracking target in wireless sensor network. However, due to the difficulty of collecting accurate sensor data and the huge computational burden, using particle filter for tracking in wireless sensor network is not suitable for practice. In this paper, a distributed estimation algorithm is proposed, and the work is concentrated on how to obtain the moving trajectory effectively in binary sensor network. For rational use of network resources, a new routing mechanism is designed based on Gabriel Graph. Moreover, we present the velocity estimation, which can efficiently estimate the speed of the moving target. And according to the relative position between nodes, two distributed tracking algorithms are proposed for different situations. As a result, the performance evaluation demonstrates that the proposed algorithm has a good performance on accuracy of tracking and effective energy saving.

Keywords: distributed estimation, binary sensor network, tracking.

1 Introduction

Wireless sensor network (WSN) is a distributed self-organizing network, where the nodes, as the minimum physical element in the network, monitor the natural environment, such as temperature, sound, vibration and stress, through the sensors equipped in nodes. With the development of communication, information and electronic technologies, wireless sensor network is becoming a significant way of attaining and processing a variety of informations. One of the most important applications of wireless sensor network is tracking target of interest, and the application has been used into several fields, such as environmental protection, health care, traffic management, agricultural production, intelligent building and etc. For instance, when fire occurs in a intelligent building, the trapped persons can use a handheld mobile terminals to access into WSN for determining the location of fire and figuring out an optimal escape path[1].

In research of tracking in WSN, Kalman filter (KF) and Particle Filter (PF) are widely used. Kalman filter is a traditional filter method for smoothing the measurements, but it only can be applied to linear system and Gaussian distribution.

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 691–698.
springerlink.com

To overcome this limitation, particle filter is presented by the researchers. Despite PF is fit for non-linear system and non-Gaussian distribution, it triggers the huge computational load of WSN. Although in recent years many researchers have improved it, such as auxiliary particle filter and unscented particle filter, the computational load has not reached a satisfactory level. So far, the complicated computational process is always an important factor restricting the development of WSN. In addition, due to complicated computational process the centralized data processing is always used in research of WSN, however the centralized approach is not an efficient way to a large network with lots of nodes. It is know that the power consumed by communicating is much greater than that by processing, therefore, for transmitting the information collected from all the nodes to sink, the whole network will consume a large amount of power. The above problems are the obstacles to tracking object in WSN power-savingly and efficiently[2,3].

In this paper, for finding a fundamental solution to the high-energy and low efficiency problems, we propose a distributed estimation algorithm in binary sensor network. This algorithm applies a new routing mechanism, which allows the node to wake up its neighbors and send messages to the neighbors temperately, thus the mechanism can effectively control the energy consumption of WSN. Additional, the binary sensor nodes are used for detecting the moving target in this paper, and to the binary sensor network we propose a velocity estimation method and the distributed tracking process . The velocity estimation is a method which can estimate the speed of target by the locations of nodes and the moving time interval between some special areas. In distributed tracking process, we analyze the two kinds of cases in detail, and propose a approach to determine the position which the moving target most likely appears at in the sensing region. This work can effectively solve the above problems of traditional filter algorithm, and achieve efficiently tracking of moving target.

The outline of the paper is as follows. Section 2 presents the distributed estimation algorithm, which includes four subsections, they are assumptions, the routing mechanism, velocity estimation of moving target and the distributed tracking process. In section 3, the performance evaluation is given. Conclusions and ongoing research issues are highlighted in the last Section 4.

2 Distributed Estimation Algorithm

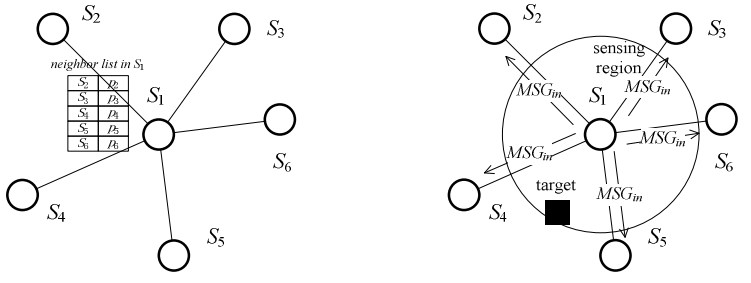
2.1 Assumptions

- We assume that the moving target in the monitoring area tries to keep the forward direction, i.e., the moving target cannot turn sharp bend, but can only turn smoothly.
- The moving target is assumed to keep uniform motion, i.e. moving at constant speed.
- In this paper, we apply binary sensor network to detect and track moving target. The binary sensor network is composed of binary nodes, which can judge whether the target is in their sensing regions[4].
- The nodes are deployed uniformly in monitoring area, and all the nodes keep time synchronization.

2.2 The Routing Mechanism

We consider that the whole sensor network can be planarized as a undirected graph, in which each vertex corresponds to a sensor node and the edges represent the minimum communication routings. In the paper, we adopt Gabriel Graph (GG) to planarize the sensor network[5].

On the basis of Gabriel Graph, we design a new routing mechanism. In the mechanism, unlike LEACH, the nodes don't need to be divided into several clusters, but they need to maintain a neighbor list which records the *ID* and position of all the neighbor nodes, as shown in Fig. 1(a). Due to using binary network, when moving target enters the sensing region of S_1, as shown in Fig. 1(b), S_1 will detect the target and immediately send a message MSG_{in} to the neighbor nodes recorded in the neighbor list. The format of MSG_{in} is {*ID*, *seq*, *P*, v_e, t_{in}}, where *ID* is the identity of S_1, *seq* is the set of the nodes detecting the target in successive sequence, *P* is the position of S_1, v_e is estimated velocity and t_{in} is the accessing time to the sensing region of S_1. In the same way, when the target leaves the sensing region, S_1 will send another message MSG_{out} to all the neighbor nodes. the format of MSG_{out} is {*ID*, *seq*, *P*, v_e, t_{out}}, where t_{out} refers to the exit time from the sensing region.



(a) Neighbor list (b) Sending MSG_{in}

Fig. 1. The routing mechanism

The proposed routing mechanism can guarantee the tracking result in maximum extent, meanwhile, try to save the power as much as possible. Before the target enters the monitoring area of wireless sensor network, all the nodes are in alternate *sleep* situations. Once the target is detected by some nodes, network will keep them into *wake* situations and send MSG_{in}s to their neighbors. In this processing, any node receiving MSG_{in} should also be in *wake*. Because the MSG_{in} is transmitted to all the neighbor nodes by the sender, the probability of miss-detecting target can effectively decline. If one *wake* node doesn't detect target and doesn't receive any MSG_{in} from other node, it will switch to *sleep* in a while.

2.3 Velocity Estimation of Moving Target

In this distributed estimation algorithm, the estimated velocity of moving target is a quite significant parameter, which can influence the accuracy of target-tracking directly. In the

above assumptions, the moving target keeps uniform motion, thus estimating velocity can be simplified to find out the constant speed. Obviously, the speed cannot be obtained intuitively because of the lack of accurate information about the moving distance and time, even though the target maintains uniform motion. Therefore, we propose an approximate method to estimate the constant speed, in which the approximate estimate of the speed will be updated when a node detects the target out of its monitoring area. The process of approximate method is shown in Fig. 2.

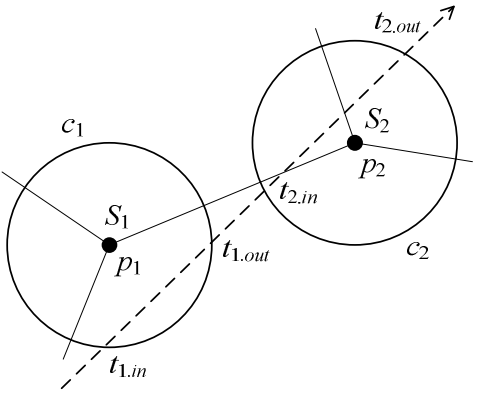


Fig. 2. The approximate estimation method

In Fig. 2, there are two sensor nodes S_1 and S_2, whose positions are p_1 and p_2, and the monitoring areas of S_1 and S_2 are c_1 and c_2 respectively. The dotted line is the trajectory of moving target, and the arrow points to the moving direction. Thus the target enters c_1 firstly, where $t_{1.in}$ is the access time and $t_{1.out}$ is the exit time, and then enters c_2, where $t_{2.in}$ is the access time and $t_{2.out}$ is the exit time, obviously $t_{1.in} < t_{1.out} < t_{2.in} < t_{2.out}$. When target enters c_1, S_1 will send a message $MSG_{in}\{ID, P, v_e, t_{in}\}$ to all its neighbors. When target exits from c_1, S_1 will send another message $MSG_{out}\{ID, P, v_e, t_{out}\}$ to all its neighbors. As a neighbor of S_1, S_2 will receive the two messages MSG_{in} and MSG_{out}. When target moves out of c_2, v_e will be updated in S_2, and the updating function is

$$v_e = \frac{\|p_1 - p_2\|}{\frac{t_{2.out} + t_{1.in} - t_{2.in} - t_{1.out}}{2}} \tag{1}$$

Because the estimated velocity v_e is highly susceptible to the trajectory of moving target, we employ $\sum_{i=1}^{n} v_e^i / n$ as v_e^n, where n is the number of updatings.

2.4 The Distributed Tracking Process

The subsection discusses the distributed tracking process. In Markov Chain, let $p(X_{n+1}|X_n)$ denote the conditional PDF of transiting the state from time n to n+1, so

that the value of X_n can be regarded as priori knowledge of X_{n+1}. In this paper, however, we regard X_{n+1} as the priori knowledge of X_n, where X is a random position variable whose value space is the whole sensing region, and the conditional PDF $p(X_n|X_{n+1})$ expresses that the position of moving target at next time is the determinative factor in judging where the target is at present.

In actual calculation, however, $p(X_n|X_{n+1})$ hardly can be attained directly, thus we apply a distributed estimation algorithm to determinate one target's possible position in sensing region. In Fig. 1, if S_5, S_1 and S_3 detect the target successively, intuitively, the probability of the target appearing in right side of the sensing region of S_1 will be much larger than any other possibilities. Due to the random deployment of the nodes, we will discuss the two cases, which are "the intersecting of two sensing regions" and "the separateness of two sensing regions".

The intersecting of two sensing regions. This case is shown in Fig. 3. There are two nodes S_a and S_b, and their sensing regions are c_a and c_b which have an intersection.

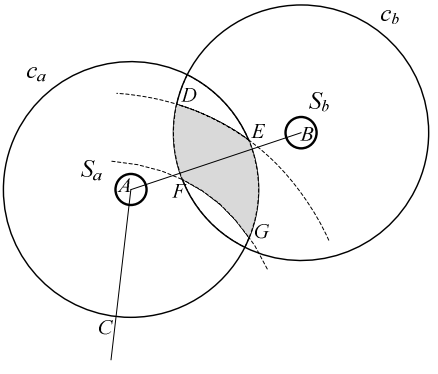


Fig. 3. The intersecting of two sensing regions

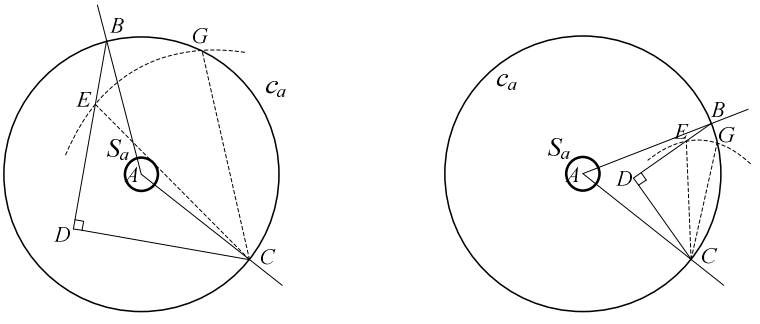
As the above statements, when the target enters c_a, S_a will transmit $MSG_{in\_S_a}$ to S_b. If the target enters into the intersection of c_a and c_b, S_a and S_b can detect it simultaneously, and S_b will transmit $MSG_{in\_S_b}$ back to S_a. In S_a, the $t_{in\_S_a}$ and $t_{in\_S_b}$ are obtained by the transiting information, and $t_{in_a \to in_b} = t_{in\_S_b} - t_{in\_S_a}$ is the time from accessing into S_a to accessing into S_b. Therefore, we can estimate the moving distance $d_{in_a \to in_b} = v_{e\_S_a} \cdot t_{in_a \to in_b}$ in the interval between $t_{in\_S_a}$ and $t_{in\_S_b}$, where $v_{e\_S_a}$ is the estimated velocity discussed in the above subsection. If C is assumed to be the point where the target enters S_a, a circle can be drew with C is the center and $d_{in_a \to in_b}$ is the radius, and two points F and G can be attained by the circle and the intersection of c_a and c_b.

Similarly, when the target exits from c_a, we can gain $t_{in_a \to out_a} = t_{out\_S_b} - t_{in\_S_a}$ which is the time from accessing into S_a to exiting from S_a, thus the distance

$d_{in_a \to out_a} = v_{e\_S_a} \cdot t_{in_a \to out_a}$ also can be obtained. With C as the center and $d_{in_a \to out_a}$ as the radius, we can get the two points D and E.

There is no denying that the probability of the target appearing in shade shown in Fig. 3 is larger than any other parts in the intersection. For improving the result, we consider the centroid of the four points D, E, F and G as the estimated result $E\left(p\left(X_{S_a} \mid X_{S_b}\right)\right)$. The result indicates the position which the target most likely appears at.

The separateness of two sensing regions. This case is shown in Fig. 4. There is a node S_a, and its sensing region is c_a which has no intersection with any other sensing region. Note that nodes S_b and S_c are omitted in the figure, and their Gabriel edges with S_a are AB and AC. According to the angle of two Gabriel edges AB and AC, this case is divided into two situations shown in Fig. 4(a) and Fig. 4 (b) respectively.



(a) The first situation (b) The second situation

Fig. 4. The separateness of two sensing regions

In Fig. 4(a), $\angle BAC$ is an obtuse angle, and the probability of appearing in sector BAC is relatively greater. Under the above assumption that the moving target in the monitoring area tries to keep the forward direction, we definite DB and DC which are at right angles as the limit boundary, which can guarantee that the target wouldn't cruise into the left side of BDC.

When the target moves through c_a and enters the sensing region of S_b, S_b will transmit $MSG_{in\_S_b}$ back to S_a. By using $MSG_{in\_S_b}$, $\angle BAC = \left|\arctan\frac{y_c - y_a}{x_c - x_a} - \arctan\frac{y_b - y_a}{x_b - x_a}\right|$ can be attained in S_a, where (x_c, y_c) can be obtained from MSG_{out_Sc} before the target enters c_a. After that, it is easy to determinate where the point D is with the help of $\angle BAC$. As the above case, we can get $t_{in_a \to out_a} = t_{out\_S_b} - t_{in\_S_a}$, which is the interval between accessing into S_a and exiting from S_a. Therefore, the moving distance $d_{in_a \to out_a} = v_{e\_S_a} \cdot t_{in_a \to out_a}$ can be estimated. If C is assumed to be the point where the

target enters S_a, a circle can be drew with C is the center and $d_{in_a \to out_a}$ is the radius, and point E which is the intersection of the circle and the limit boundary and point G which is the intersection of the circle and c_a can be attained. Finally, we consider the centroid of the sector ECG as the most likely point.

In second situation, $\angle BAC$ is an sharp angle shown in Fig. 4(b). According to the assumption, it is likely that the situation happens, but the likelihood of occurrence is small. As the first situation, by the similar way, we can find point D which is inside sector BAC due to the sharp angle $\angle BAC$. Here, the moving distance $d_{in_a \to out_a}$ is still equal to $v_{e\_S_a} \cdot t_{in_a \to out_a}$, where the interval $t_{in_a \to out_a}$ should be very short. Therefore, we can determinate the two point E and G by employing moving distance $d_{in_a \to out_a}$. As a result, we still utilize the centroid of the sector ECG to estimate $E\left(p\left(X_{S_a} \mid X_{S_b}\right)\right)$.

3 Simulation Results

There are 100 sensor nodes deployed randomly and uniformly in the 100×100m surveillance area. In the simulation energy model, the initial battery of each node is 30J, the transmitting power is 700mW and the receiving power is 360mW. The communication range is 30m and the sensing region is 15m. The velocity of moving target is set to be 2m/s.

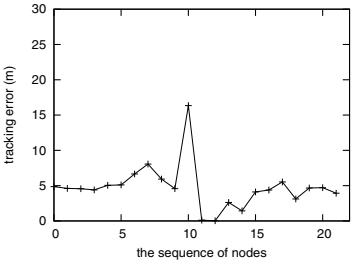
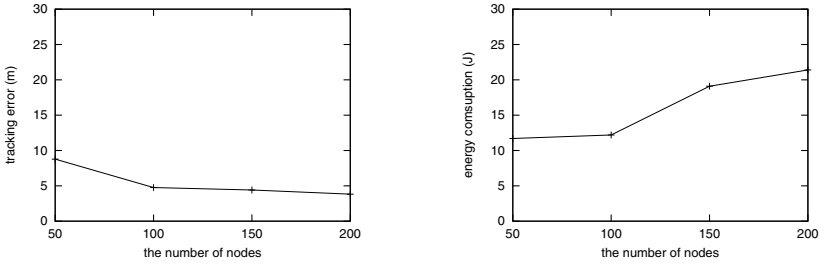


Fig. 5. The tracking error

Fig. 5 shows the tracking error as the evaluated result of the proposed algorithm. In the simulation process there are 22 nodes detecting the target, and 22 most likely points are attained, which are arranged by the order of appearing sequence. According to the settings, the range of the error is 0 to 30m, and the average of tracking errors is 4.76m. Although the error of No. 10 node is 16.34m, this error do not causes the accumulation of error, which demonstrates that the proposed algorithm is robust.

Fig. 6 shows the comparisons of tracking errors and energy consumptions respectively. In Fig. 6(a), when the number of nodes increases, the tracking error decreases. The main reason is that the more nodes, the higher likelihood of intersection of sensing regions, and it is obvious that the tracking accuracy of case 1 is higher than that of case 2. In Fig. 6(b), the energy consumption shows the upward trend with the increasing of nodes, which is caused by too much communication load.



(a) the comparison of tracking errors (b) the comparison of energy consumptions

Fig. 6. The comparison with different nodes

4 Conclusion

In this paper, we propose a distributed estimation algorithm to track the moving target by binary sensor network. For the accurate tracking, we present a new routing mechanism, velocity estimation method and distributed tracking process. As a result, we can effectively tracking the mobile target by using the mechanism or methods proposed here, and the tracking accuracy can satisfy general applying requirement. However, to the target following the complicated trajectory, our algorithm shows the limitation of the tracking capability. Therefore, we will focus on the complicated-trajectory tracking and multi-target tracking in the future in order to extend the application range of the algorithm.

Acknowledgments. This research was supported by the National Natural Science Foundation of China (60874103) and the Fundamental Research Funds for the Central Universities (N090304003).

References

1. Jia, Z., Wu, C., Zhang, Y., et al.: Based on Adaptive Variable Rate Particle Filter Tracking Research in Wireless Sensor Networks. In: WiCOM 2009, pp. 1–4 (2009)
2. Djurić, P., Vemula, M., Bugallo, M.: Target Tracking by Particle Filtering in Binary Sensor Networks. IEEE Transactions on Signal Processing 56(6), 2229–2238 (2008)
3. Maskell, S., Gordon, N.: A Tutorial on Particle Filters for On-line Nonlinear/Non-Gaussian Bayesian Tracking. Signal Processing 50(2), 174–188 (2002)
4. Yang, Z., Liu, Y.: Quality of Trilateration: Confidence-Based Iterative Localization. IEEE Transactions on Parallel and Distributed Systems 21(5), 631–640 (2010)
5. Tsai, H., Chu, C., Chen, T.: Mobile object tracking in wireless sensor networks. Computer Communications 30, 1811–1825 (2007)

A Method of Image Segmentation Based on Adaptive Bidirectional Balloon Force Model

Jinyao Yang[1], Yugui Xu[2], Boling Zheng[2], Haihong Shen[2], and Minfen Shen[2]

[1] Shantou Institute of Ultrasonic instruments Co., Ltd., Shantou, China
[2] College of Engineering, Shantou University, China
mfshen@stu.edu.cn

Abstract. A novel image segmentation method is proposed using geometric active contour model based on an adaptive bidirectional balloon force (ABBF). Since active contour model was used extensively in image segmentation, great progress has been achieved. The evolved active contour model combined with balloon force could effectively detect object boundaries, however, this model require that we must put the initial active contour inside or outside the boundary completely, also it is easy to cause leaking through weak edges. In order to solve these problems, we propose a new ABBF-snake model based on magetostatic field. According to our experiment, it is obvious that our method make great improvement in convergence ability and broken edges, furthermore, compare with GVF and VCF, our method achieved more satisfactory results in segmentation of complex shape objects and real images.

Keywords: Segmentation, active contour, adaptive bidirectional balloon force.

1 Introduction

Image segmentation refers to partition a digital image into segments of homogeneous properties, it is the basis of image registration and image analysis, and now is widely used in medical image recognition, manufacturing production, and other advanced applications [1], [2]. Since Kass et al. proposed the active contour model [3] (or snake) to segment images, two general types of active contour models have appeared in literatures: parametric active contours and geometric active contours. Although the classical active contour model [3] can implement image segmentation effectively, some disadvantages are still exist. For example, only when the initial contour is placed inside or outside the object boundary completely, can the model converge to the edges accurately [4]. For this reason, there are many improved methods, such as GVF, VCF in [5], [6] respectively while Cohen proposed a balloon-snake model to detect the object boundary.

The balloon-snake model [7] is based on balloon force, which needs to combine with the Gaussian potential force. Though this model made some improvements, it still had a serious problem, that is unidirectional movement property —unidirectional inflating or unidirectional shrinking. This problem limits the application of the balloon model severely. Responding to the problem, many researchers proposed different improved

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 699–706.
springerlink.com

methods [8], [9]. But essentially, they didn't improve the performance of the complex shape object segmentation.

In this paper, we propose a novel ABBF-snake model to overcome these defects. Our paper is structured as follows. In section 2, the balloon-snake model will be reviewed. Then, section 3 describes our method in detail, including the magnetostatic field, which acts as the criterion to control the direction of balloon force, and the ABBF-snake model. In the following section 4, the performance of the ABBF-snake model in synthesis images and real images is shown, compare with some traditional snake model. Section 5 presents our conclusion.

2 Balloon-Snake Model

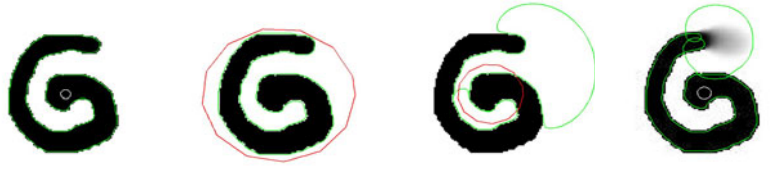
In 1991, Cohen proposed the famous balloon force [7], the force is defined as

$$F_B(X) = k_B \vec{N}(X). \tag{1}$$

Where $\vec{N}$ is the unit inward normal of the evolving contour X, and k_B denotes a constant. When k_B is negative, the balloon force inflates and pushes the active contour outside as if we introduce air inside; when k_B is positive, the balloon force shrinks and pulls the active contour inside. Combining the balloon force with the Gaussian potential force, we get the balloon-snake model as following

$$\frac{\partial X}{\partial t} = \frac{\partial}{\partial s}\left(\alpha \frac{\partial X}{\partial s}\right) - \frac{\partial^2}{\partial s^2}\left(\beta \frac{\partial^2 X}{\partial s^2}\right) - \nabla E_{ext} + F_B. \tag{2}$$

The balloon-snake model's convergence ability is shown below. When the initial contour (see the white circle) was placed inside the object boundary (see Fig.1 (a)), and let k_B be negative, the model could detect the edges accurately (see the green curve). When the initial contour was placed outside the object boundary, and let k_B be positive, the result also was successful (see Fig.1 (b)), however, when the snake contour was placed across the object boundary initially (see Fig.1 (c)), the converge result failed. And according to Fig.1 (d), if the object boundary was blurred partially, using the same parameters as Fig.1 (a)'s, it's clear that the snake contour didn't stop in the blurred part because of lack of Gaussian potential force.



(a) (b) (c) (d)

Fig. 1. Comparison of balloon model results with different initial active contours

3 Proposed Model

To overcome the defects of the above balloon-snake model, ABBF-snake based on the magnetostatic field due to the object boundary was proposed. First of all, we introduce the magnetic field, which partitions an image into the regions of targets and backgrounds, decides the magnitude and the direction of balloon force at each point in the snake contour. Then, next, based on the magnetic field, the adaptive bidirectional balloon force model is presented.

3.1 The Magnetostatic Field

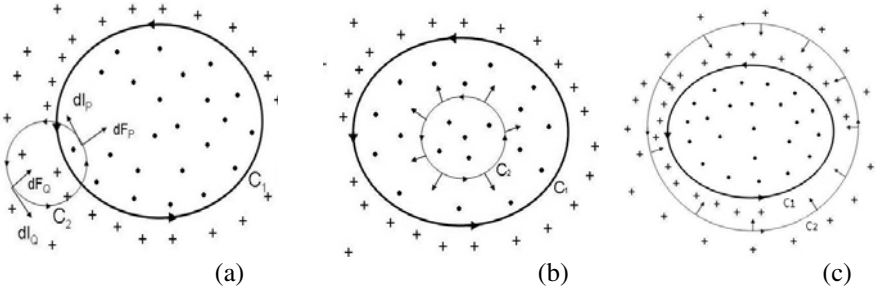
Consider that loops C_1 and C_2 have the same currents I_1 and the position of C_1 is fixed. According to the principle of the electromagnetic field, C_1 will introduce a magnetic field and the magnetic flux density at each point in the field can be computed from the Bi ô t-Savart Law

$$B(a) = \frac{\mu_0}{4\pi} I_1 \int_{C_1} dI_b \times \frac{\vec{R}_{ba}}{R_{ba}^2} \cdot \tag{3}$$

Where $B(a)$ is the magnetic flux density at point a, μ_0 is the permeability constant, dI_b is infinitesimal current segment at point b in C_1, R_{ba} denotes the distance between point a and b, $\vec{R}_{ba}$ is the unit distance vector, and $\times$ denotes the cross product. Based on Ampere's law, the magetostatic force exerted upon each point due to C_1 is

$$F(x) = I_2 dI_x \times B(x) . \tag{4}$$

Where $F(x)$ is the magnetic field force at point x, I_2 is the current in C_2, dI_x is infinitesimal current segment at point x in C_2. Obviously, $F(x)$ is always perpendicular to the current flow direction dI_x.



(a) (b) (c)

Fig. 2. The force on C_2 where C_2 is cross, inside, and outside C_1

Fig.2 shows the magnetic field forces exerted upon C_2 due to C_1 with different initial positions. Where the "+" denote the magnetic field going perpendicularly into the image plane, and the "." denote the opposite direction. It's clear that the initial contour drove by the magnetic field forces will inflate in the regions of targets and shrink in the regions of backgrounds adaptively.

3.2 The ABBF-Snake Model

According to the concept of the external field [10], we define a modified magnetic field in the image plane. From (1), we know the magnetic flux density $B(a)$ is always perpendicular to I_b and vector $\vec{R}_{ba}$, so $B(a)$ only has a nonzero component that is perpendicular to the image plane. The three-dimensional coordinates of point a in B and current element I_b are set to (x_a, y_a, z_a), (x_b, y_b, z_b), respectively. Let $\vec{R}_{ba}$ to be (x_{ab}, y_{ab}, z_{ab}). Using the cross multiplication, the perpendicular component at point a in the magnetic field can be obtained

$$z_a = x_a \bullet y_{ab} - y_a \bullet x_{ab} . \tag{5}$$

If we plot Z_a of one object, we can see that the direction of the magnetic flux density inside the object boundary is negative, and the direction outside the object boundary is positive. Besides, the magnitude of B at a point near the edges is large, and decrease rapidly with increasing distance between the point and the edges. For this reason, we modify k_B to be k_{DB}

$$k_{DB}(x) = \lambda_B(x) \bullet \min(R_{xl}) . \tag{6}$$

Where R_{xl} denotes the distance between point x and the object boundary, and

$$\lambda_B(x) = \begin{cases} 1 & if \ \ B(x) > \Pi \\ 0 & else \\ -1 & if \ \ B(x) < -\Pi \end{cases} . \tag{7}$$

Here Π is the threshold value. Using k_{DB}, (1) can rewritten as

$$F_{DB}(x) = k_{DB}(x)\vec{N}(x) . \tag{8}$$

From the above equation, we can see that when the snake contour is far away the object boundary, k_{DB} is large, that is, the balloon force is strong. With the model evolving, the contour gradually approaches the edges, k_{DB} decreases. This means with the contour evolving to the edge, the balloon force reduces gradually, finally the snake contour stops on the object boundary.

Given the balloon force $F_{DB}(x)$, the evolving active contour under this force is

$$\frac{\partial C}{\partial t} = (F_{DB}(x) \bullet \vec{N})\vec{N}. \tag{9}$$

According to the level-set method, the level-set representation of (9) is

$$\frac{\partial \phi}{\partial t} = \alpha g(x)\nabla \bullet \left(\frac{\nabla \phi}{|\nabla \phi|}\right)|\nabla \phi| - (1-\alpha)F_{DB}(x) \bullet \nabla \phi. \tag{10}$$

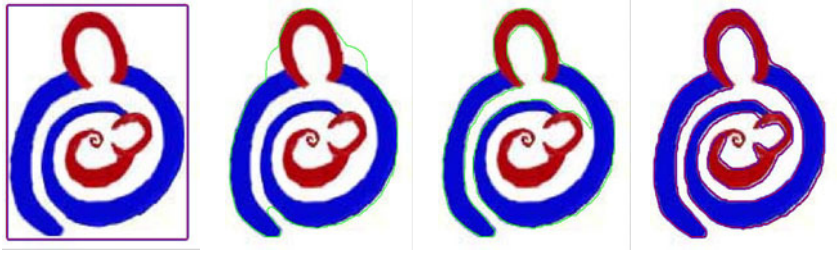
The numerical implementation of (10) can be found in [11].

4 Experimental Results and Analysis

Next, we perform the proposed ABBF-snake model in some synthesis images and real images, moreover, compare with the traditional GVF and VCF models. The experiment environment is Matlab2008a, AMD 2800+, and Memory 512MB.

4.1 Convergence

Fig. 3 shows the evolution of the ABBF-snake model. The red rectangle in Fig. 3(a) is the initial contour, and the final segmentation result is shown in Fig. 3(d). Fig. 3(b-c) gives the evolution of the snake contour (see the green curve) after 10 and 40 iterations. It's obvious that our proposed model has good convergence ability.

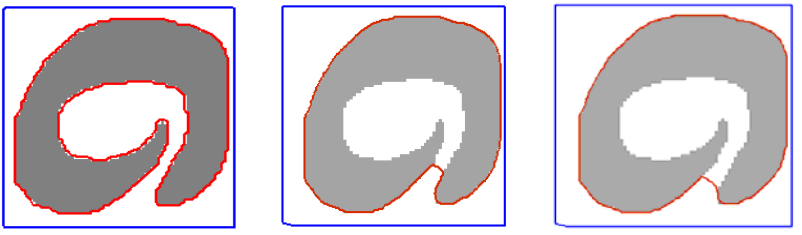


(a) (b) (c) (d)

Fig. 3. The evolutionary of active contour using ABBF-snake model

4.2 The Segmentation of Concave Object

Next, we show the concave object boundary detection results by our proposed method, GVF, and VCF respectively. In Fig.4, blue rectangles represent the initial contours, and the red curves denote the final segmentation results. It is obvious that GVF and VCF display a disappointing result. However, our proposed model succeeded in evolving through the narrow structures and reaching into concavity.

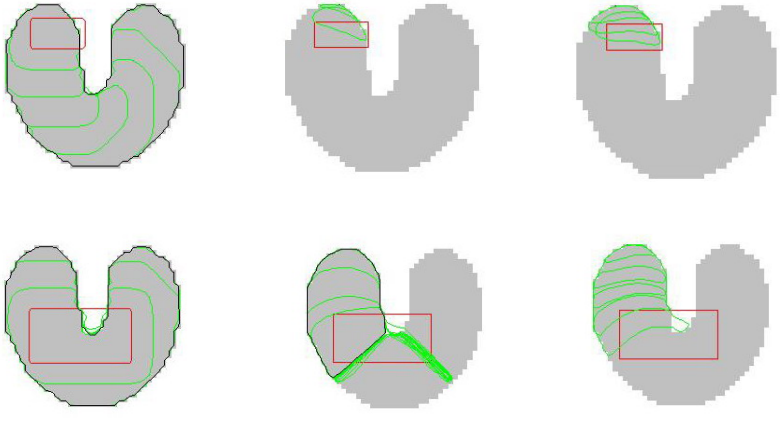


(a) (b) (c)

Fig. 4. (a-c) show the results of the proposed method, GVF and VCF respectively

4.3 Comparison of Different Original Positions

Then, we will show the segmentation results of the three models at arbitrary initial positions. In Figure 5, the initial positions (see the red rectangles) are set on the top-left (the first row) and inside the boundary (the second row). Green curves are the evolution of the snake contour. From left to right, they are the proposed method, GVF, and VCF respectively. Obviously, our proposed method can effectively overcome the limit of initial positions of the traditional snake models.



(a) (b) (c)

Fig. 5. Column (a-c) show the results of the proposed method, GVF and VCF respectively

4.4 The Recovery of Broken Boundary

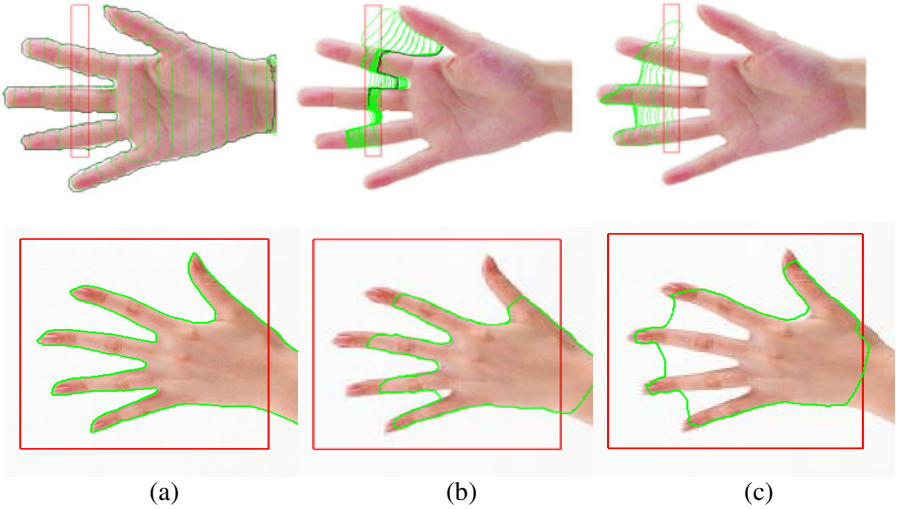
In Fig. 6(a), we blurred an object boundary. The edge can be obtained as shown in Fig. 6(b), clearly, the blurred boundary is unable to be detected. But after processed by our method, the distribution of λ_B is shown in Fig. 6(c), we can see the broken boundary was recovered well.

(a) (b) (c)

Fig. 6. (a) the blurred object, (b) the object boundary, (c) the distribution of λ_B

4.5 The Real Image Segmentation

In this section, we use our proposed model, GVF and VCF to perform some real images. As shown in Fig. 7, we set the two different initial contours (see the red rectangles), and the green curves denote the evolution of the active contours. From column (a-c), we can clearly see that our proposed method can extract the hand boundary accurately, while GVF and VCF both get wrong results. This experimental result further verifies the effectiveness of our proposed method.



(a) (b) (c)

ABBF GVF VCF

Fig. 7. Segmentations of real images using ABBF-snake, GVF and VCF

5 Conclusion

In order to overcome defects of traditional balloon force model, we propose a new ABBF-snake model in this paper. First of all, the magnetostatic field method was used to determine the magnitude and the direction of the balloon force, and then we deduced the bidirectional balloon force. Thus, combining with this force and snake, the

ABBF-snake model was then obtained. Experimental results show that our proposed method has good convergence ability, at the same time, our method can recover the broken boundary well that stop leaking through weak edges. Moreover, compare with the traditional snake methods, the proposed method shows significant improvement in concavity reaching and different initial positions.

References

1. Gu, D.-B., Sun, J.-X.: EM image segmentation algorithm based on an inhomogeneous hidden MRF model. In: IEE Proceedings-Vision, Image and Signal Processing, April 8, vol. 152(2), pp. 184–190 (2005)
2. Bhanu, B., Peng, J.: Adaptive integrated image segmentation and object recognition. Applications and Reviews. IEEE Transactions on. Systems, Man, and Cybernetics, Part C 30(4), 427–441 (2000)
3. Kass, M., Witkin, A., Terzopoulous, D.: Snakes: Active Contour Models. In: Proceedings of the 1st International Conference on Computer Vision, pp. 259–268. IEEE Computer Society Press, London (1987)
4. Radeva, P., Serrat, J., Marti, E.: A snake for model-based segmentation. In: Proceedings, Fifth International Conference on Computer Vision, June 20-23, pp. 816–821 (1995)
5. Xu, C., Prince, J.L.: Gradient vector flow:a new external forces for snakes. In: Proceeding of IEEE International Conference on CVPR, vol. 6, pp. 66–71 (1997)
6. Bing, L., Acton, S.T.: Active Contour External Force Using Vector Field Convolution for Image Segmentation. IEEE Transactions on Image Processing 16(8), 2096–2106 (2007)
7. Cohen, L.D.: Note: on active contour models and balloons. CVGIP: Image Understanding 53(2), 211–218 (1991)
8. Paragios, N., Mellina-Gottardo, O., Ramesh, V.: Gradient Vector Flow Fast Geometric Active Contours. IEEE Transactions on Pattern Analysis and Machine Intelligence 26(3), 402–407 (2004)
9. Wu, H.-H., Liu, J.-C., Chui, C.: A wavelet-frame based image force model for active contouring algorithms. IEEE Transactions on Image Processing 9(11), 1983–1988 (2000)
10. Xie, X., Mirmehdi, M.: MAC: Magnetostatic Active Contour Model. IEEE Transactions on Pattern Analysis and Machine Intelligence 30(4), 632–646 (2008)
11. Osher, S., Sethian, J.A.: Fronts propagating with curvature dependent speed: algorithms based on Hamilton-Jacobi formulations. Journal of Computational Physics 79(1), 12–49 (1988)

Variable Step-Size Algorithm on Feedback and Feed-Forward Coefficients for Modified Decision Feedback Equalizer Turbo Code DS/CDMA System

T. Udomsripaiboon[1] and C. Benjangkaprasert[2]

[1] Faculty of Information and Communication Technology,
University of Phayao, 56000 Thailand
thana.ud@up.ac.th
[2] Faculty of Engineering, King Mongkut's Institute of Technology Ladkrabang,
Bangkok 10520, Thailand,
kbchawal@kmitl.ac.th

Abstract. In this paper presents a variable step-size algorithm for modified decision feedback equalizer structure with turbo coded for direct-sequence code division multiple access (DS/CDMA) digital communication systems. The main objective of the proposed equalizer is to minimize the bit error rate (BER) of the data due to the disturbances of noise and intersymbol interference (ISI) phenomenon on the channel. By the proposed technique, the performances of the proposed adaptive equalizer cooperate with the attractive Soft Output Viterbi Algorithm (SOVA) algorithm of turbo decoder for the DS/CDMA system. It provides a good statistical in non-stationary channel condition scenario. The simulation results are shown the superior bit error rate performances of the proposed adaptive equalizer over other equalizers.

Keywords: DS/CDMA, FIR, DFE, variable step-size algorithm, turbo code.

1 Introduction

Direct-sequence code division multiple access has been adopted by industry as a key technology for the third generation wireless communication systems [1]. Due to time variation and frequency selectivity of wireless multipath channels, the presence of time variant fading, intersymbol interference (ISI), and multiple access interference (MAI) are the main limitation to the performance of DS/CDMA systems. The methods to eliminate this problem include code waveform design, power control, adaptive equalizer, forward error control codes and multiuser detection [2]. Combination of these techniques would be more effective due to difference advantages of each technique [3]. In [4], it is shown that the turbo code provided significant improvements in the quality of data transmission over a noisy channel. And for the turbo decoding algorithms, the iterative Soft Output Viterbi Algorithm (SOVA) algorithm is an attractive algorithm of turbo decoding process [9,13]. Various types of adaptive equalizers that applications for mitigate the effect of MAI and ISI have been discussed in the literatures [3, 5-8, 12].

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 707–714.
springerlink.com

This paper proposes a variable step-size algorithm for the adaptive equalizer scheme and using the SOVA algorithm in the turbo decoding of the DS/CDMA communication system.

The paper is organized as follows. In section 2, the system model is described. In section 3, presents the structures of the decision feedback equalizer and the modified decision feedback equalizer. Section 4, the adaptive least mean square algorithm and the proposed variable step-size algorithm for update coefficients of the adaptive equalizer receiver is described. In section 5, the turbo decoder is described. Numerical results for the suppression of ISI and MAI for DS/CDMA communication system are presented in section 6, and finally, section 7 is conclusion.

2 System Model

We consider the forward link of a DS/CDMA system with *K* users. We assume that the signals of the users arrive at the receiver synchronously and that the spreading codes are known at the receiver. The modulation scheme is linear. d_i^k is the i^{th} symbol of duration T_s, transmitted by user *k*.

The spreading code $c_k(n)$ of length L_c assigned to user *k* can be written as follows:

$$c_k(n) = \sum_{i=0}^{L_c-1} c_k^i x(n - iT_c) \tag{1}$$

herein, we let

$$x(n) = \begin{Bmatrix} 1 & if & n \in [0, T_c] \\ 0 & otherwise & \end{Bmatrix}$$

where T_c is the chip duration $(T_c = T_s / L_c)$ and $c_k^i \in \{-1, 1\}$.

The spread signal is transmitted by user *k* given by:

$$s_k(n) = \sum_{i=0}^{\infty} d_i^k c_k(n - iT_s) \tag{2}$$

The spread signal is transmitted over the frequency and time selective channel, where $h_k(n)$ is given as follows [5]:

$$h_k(n) = \sum_{i=0}^{W-1} a_i^k \delta(n - iT_c) \tag{3}$$

whose impulse response is given by:

$$a_i^k = \left\{ \begin{array}{ll} 0.5\left\{1+\cos\left(\frac{2\pi}{A}(i-2)\right)\right\}, & i=1,2,3 \\ 0 & ,otherwise \end{array} \right\} \tag{4}$$

where the factor A is introduced to allow scaling to customize of the simulated ISI.

The path gain a_i^k is a complex Gaussian random process with zero mean. Let us now introduce a guard interval before the transmission of information symbol in order to cancel ISI. T_s must therefore be replaced by τ in Equation (2). The minimum value of τ is: $\tau = T_s + (W-1)T_c$.

The received signal is expressed as follows:

$$r(n) = \sum_{i=0}^{\infty}\sum_{k=0}^{K-1} d_i^k c_k(n-i\tau) * h_k(n) + n(n) \tag{5}$$

where $n(n)$ is a complex additive white Gaussian noise with zeros mean and a variance of σ^2.

3 Modified Decision Feedback Equalizer (MDFE) Structure

The structure of the conventional decision feedback equalizer (DFE) [11], consists of feed-forward filter and decision feedback filter. The DFE that has feed-forward filter of *M*-1 taps and feedback filter of *N*-1 taps.

We assume that the input vector, R_n^k, and filter coefficient vector, $W_{w,v}$ with time index *n* is given by:

$$R_n^k = [r^k(n)\, r^k(n-1)...r^k(n-(M-1)) \quad y_{df}^k(n-1)...y_{df}^k(n(N-1))]^T \tag{6}$$

$$W_{w,v} = [w_0 \; w_1...w_{M-1} \quad v_1...v_{N-1}]^T \tag{7}$$

where $r^k(n)$ is the input signal and $y_{df}^k(n)$ is the decision value of the equalizer output.

Then, the output before decision of the DFE can be expressed as:

$$y_{d1}^k(n) = \sum_{i=0}^{M-1} w_i r^k(n-i) + \sum_{i=0}^{N-1} v_i y_{df}^k(n-(i+1)) \tag{8}$$

$$= W_{w,v}^H R_n^k \tag{9}$$

where the *H* stands for Hermitian transpose, $y_{d1}^k(n)$ is the output before decision of the DFE can be expressed as:

$$y_{d1}^{k}(n) = y_{ff}^{k}(n) + y_{fb}^{k}(n) \tag{10}$$

The error between the desired signal $d_k(n)$ and the feed forward filter output $y_{ff}^{k}(n)$ is defined as:

$$e_{ff}(n), e_{fb}(n) = d^{k}(n) - y_{ff}^{k}(n) \tag{11}$$

In next subsection, the modified decision feedback equalizer (MDFE) structure is presented [12].

The output of the MDFE $y_{ff2}^{k}(n)$ of the feed-forward part is given by:

$$y_{ff2}^{k}(n) = \sum_{i=0}^{M-1} w_i r^{k}(n-i) \tag{12}$$

The output of the MDFE $y_{fb2}^{k}(n)$ of the feedback part is given by:

$$y_{fb2}^{k}(n) = \sum_{i=1}^{N-1} v_i y_{df2}^{k}(n-(i+1)) \tag{13}$$

where $y_{df2}^{k}(n)$ is the decision output of $y_{ff2}^{k}(n)$ can be expressed as:

$$y_{df2}^{k}(n) = sign[y_{ff2}^{k}(n)] \tag{14}$$

Then, the output $y_{O}^{k}(n)$ of the MDFE can be expressed as:

$$y_{o}^{k}(n) = y_{ff2}^{k}(n) + y_{fb2}^{k}(n) \tag{15}$$

The error between the desired signal $d_k(n)$ and the filter output $y_{ff2}^{k}(n)$ of MDFE is given as:

$$e_{ff2}(n), e_{fb2}(n) = d^{k}(n) - y_{ff2}^{k}(n) \tag{16}$$

4 Variable Step-Size Algorithm

The weights of feed-forward part of the equalizers using the classical least mean square (LMS) algorithm are updated according to :

$$w(n+1) = w(n) + \mu_{ff} e_{ff,ff2}(n) r^{k}(n) \tag{17}$$

where μ_{ff} is a fixed step-size parameter of feed-forward filter, $e_{ff,ff2}$ is e_{ff} or e_{ff2}.

The weights of feedback part are updated as:

$$v(n+1) = v(n) + \mu_{fb} e_{fb,fb2}(n) y_{df,df2}^{k}(n) \tag{18}$$

where μ_{fb} is a fixed step-size parameter of feedback filter, $e_{fb,fb2}$ is e_{fb} or e_{fb2} and $y_{df,df2}^{k}$ is y_{df}^{k} or y_{df2}^{k}.

Then in order to improve the performance of the system. The weights of feed-forward part and feedback part of the equalizer are using the variable step-size (VS) algorithm.

The weights of feed-forward part are updated according to:

$$w(n+1) = w(n) + \mu_{ff}(n) e_{ff,ff2}(n) r^{k}(n) \tag{19}$$

where $\mu_{ff}(n)$ is a variable step-size parameter of feed-forward filter.

Also, the weights of feedback are updated as follows:

$$v(n+1) = v(n) + \mu_{fb}(n) e_{fb,fb2}(n) y^{k}_{df,df2}(n) \tag{20}$$

where $\mu_{fb}(n)$ is a variable step-size parameter of feedback filter.

The objective of using variable step-size algorithm Equations (19) and (20) are that to ensure large values of $\mu_{ff}(n)$ and $\mu_{fb}(n)$ when the algorithm is far from the optimum and the values of $\mu_{ff}(n)$ and $\mu_{fb}(n)$ decreasing as the updated value approach the optimum. That is described as follows:

$$\varphi(n+1) = \alpha\varphi(n) + \sigma(1 - BER) \tag{21}$$

Thus, the variable step-size parameter for update the equalizer coefficients are given by:

$$\mu_{ff}(n+1) = \gamma_1 \mu_{ff}(n) + \beta_1 \varphi^2(n) \tag{22}$$

$$\mu_{fb}(n+1) = \gamma_2 \mu_{fb}(n) + \beta_2 \varphi^2(n) \tag{23}$$

where α, σ, γ_1, γ_2, β_1 and β_2 are positive constants, *BER* is bit error rate of the system.

5 Turbo Decoder

In this paper SOVA solutions have to be considered for the turbo decoder of the system. Therefore, we will briefly describe the attractive SOVA decoder in the next sub-section.

The SOVA decoder [9,13] estimates the soft output information for each transmitted binary symbol in the log-likelihood function $L(x'_n)_{SOVA}$, as follows:

$$L(x'_n)_{SOVA} = \log \frac{P_r(s_n = 1 \mid y_1^{\tau})}{P_r(s_n = 0 \mid y_1^{\tau})} \approx \log \frac{e^{-\mu_{\tau,\min}}}{e^{-\mu_{n,s}}} \approx \mu_{n,s} - \mu_{\tau,\min} \tag{24}$$

where P_r is a *posteriori* probability of the transmitted symbol, s_n is the spread signal of the transmitted symbol, y_1^{τ} is the received sequence, $\mu_{\tau,\min}$ is the minimum path

metric as the maximum likelihood (ML) , $\mu_{n,s}$ is the minimum path metric of the paths with complementary symbol to the ML symbol at time *n*. If the ML symbol at time *n* is 1, then its complementary symbol is 0. Therefore, we can write [9]:

$$P_r(s_n = 1 \mid y_1^\tau) \approx e^{-\mu_{\tau,\min}} \tag{25}$$

$$P_r(s_n = 0 \mid y_1^\tau) \approx e^{-\mu_{n,s}} \tag{26}$$

See [9,13] for more details on turbo codes and turbo decoding algorithm.

6 Numerical Results

The BER performance of the proposed variable step-size algorithm for the adaptive equalizer structure with turbo code for the DS/CDMA system receiver is presented in this section. The simulation results are obtained from the communication system over both ISI and MAI channels.

The main system parameters are defined as follows: length of transmission information 15 *kbits* (15 x 1024), support 4 users, filter length of feed-forward part M = 11, feedback part N = 10, power of each user = 8 dB, and with 10 independents computer runs. And for turbo code parameters: channel code rate R = ½ code generator metric:

$$G = \left[1, \quad \frac{1 + D^2}{1 + D + D^2} \right] \tag{27}$$

with constraint length = 3, and the number of iteration for turbo decoder based on SOVA decoding = 1.

Table 1 shows the algorithms of feed-forward part and feedback part that is used to adjust the coefficients of the equalizers. In this paper there have 3 types of equalizers: first is FIR equalizer, second is DFE and the last one is MDFE.

Table 1. Type of Algorithms Using on the Equalizers

	Eq. (17)	Eq. (18)	Eq. (19)	Eq. (20)
FIR	✓			
DFE, MDFE	✓	✓		
VS-DFE$_1$,VS-MDFE$_1$		✓	✓	
VS-DFE$_2$,VS-MDFE$_2$			✓	✓

Fig. 1 shows the convergence process of feed-forward and feedback part with adaptive algorithm for the equalizers with the *Eb/No* = 15 dB. From Fig. 1, it is seen that the variation of the coefficients of the equalizer using the variable step-size algorithm (VS) is

stabilize than the variation of the coefficients of the equalizer using the least mean square algorithm (LMS).

$w(n)$, $v(n)$

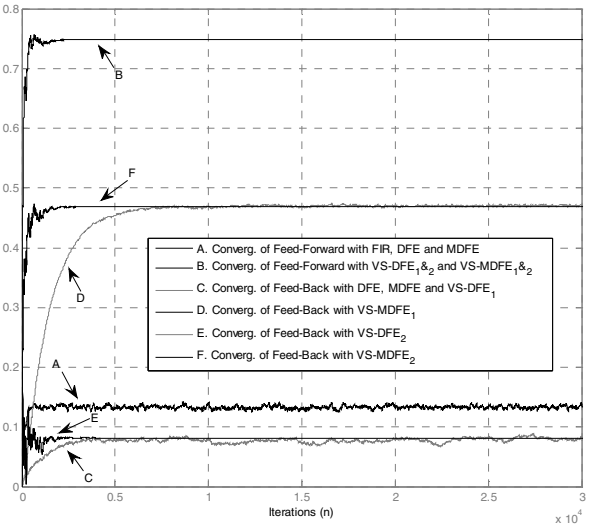
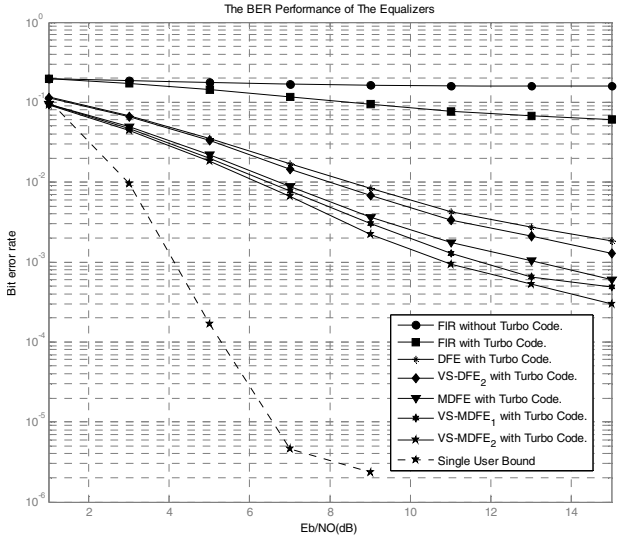


Fig. 1. The convergence process of feedback and feed-forward coefficients of the equalizers with *Eb/No* = 15 dB



BER

Fig. 2. BER performance of the adaptive equalizers

Fig. 2 shows the average BER performance of the perfect power control versus *Eb/No* for various values of the receiver using the number of iteration for turbo decoder = 1. From Fig. 2, it is evident that the VS-MDFE$_2$ adaptive equalizer provides a lowest BER performance for every *Eb/No* values.

7 Conclusion

In this paper, we proposed a powerful adaptive equalizer that using the variable step-size algorithm on feedback and feed-forward coefficients for modified decision feedback equalizer (VS-MDFE$_2$) to control the adaptation process of the equalizer. The proposed adaptive equalizer is worked cooperative with the attractive SOVA algorithm for the turbo decoding process of the communication system. Our simulation results are shown that the BER performances of the proposed adaptive equalizer for the DS/CDMA communication systems are greatly outperformed than other equalizers.

References

1. Mandyam, G., Lai, J.: Third-Generation CDMA Systems for Enhanced Data Services. Academic Press (2002)
2. Haykin, S., Moher, M.: Modern Wireless Communications. Prentice-Hall (2005)
3. Haykin, S.: Communication Systems. John Wiley & Sons, Inc. (1994)
4. Berrou, C., Glavieux, A., Thitimasjshima, P.: Near Shannon Limit Error-Correcting Codings and Decoding: Turbo Code. In: Proc. IEEE ICOC, pp. 1064–1070 (1993)
5. Haykin, S.: Adaptive Filter Theory. Prentice-Hall (2002)
6. Satorius, E.H., Alexander, S.T.: Channel Equalization Using Adaptive Lattice Algorithms. In: Proc. IEEE Trans. on Communication, pp. 899–905 (1979)
7. Sklar, B.: Digital Communications Fundamentals and Applications. Prentice-Hall (2004)
8. Kafle, P.L., Makelainen, K., Rajatheva, R.M.A.P.: Performance of Parallel Concatenated Interleaved Codes in Correlated Multipath Fading Channels in a DS-CDMA System. In: Proc. IEEE Trans. on Communication, pp. 1826–1830 (1999)
9. Divsala, D., Pollara, F.: On the Design of Turbo Codes. The Telecommunication and Data Acquisition Progress Report, Jet Propulsion Laboratory, pp.99–121 (1995)
10. Jung, P., Nashan, M.: Performance Evaluation of Turbo Codes for Short Frame Transmission Systems. Electronics Letters 30(2), 111–113 (1994)
11. Austin, M.: Decision-Feedback Equalization for Digital Communication Over Dispersive Channels. M. I. T. RES. Lab. Election., Tech. Rep. 461 (1967)
12. Udomsripaiboon, T., Sanpote, N., Sooksomsatarn, K., Benjangkaprasert, C.: Modified Decision Feedback Adaptive Equalizer for Turbo Coded DS-CDMA Systems Using SOVA Decoding. In: Proc. ECTI-CON 2007, pp. 892–895 (2007)
13. Benedetto, S., Divsala, D., Montorsi, G., Pollara, F.: Soft-Output Decoding Algorithms in Iterative Decoding of Turbo Codes. The Telecommunication and Data Acquisition Progress Report, Jet Propulsion Laboratory, pp. 63–87 (1996)

Research of a Ubiquitous-Computing Enabled Service Pattern in Converged Network

Xiao-Yan Yu[1,2], Rong-Heng Lin[1,2], and Peng-Cheng Huang[1,2]

[1] Institute of Network Technology, Beijing University of Posts and Telecommunications
[2] State Key Lab of Networking and Switching Technology,
Beijing University of Posts and Telecommunications
Beijing, China, 100876, China
{yuxy,rhlin,gnuhpc}@bupt.edu.cn

Abstract. The service triggering mechanism of current IMS not only leads to heavy load of the S-CSCF entities, but also limits the provision of feasible service which makes use of context information. Thus, a ubiquitous-computing enabled service triggering and executing pattern in converged network is introduced. Via analyzing service pattern of current IMS architecture, a novel mechanism that introduces P2P like service triggering based on terminal side coordination is introduced, which results in the conversion from central service control to the negotiation service triggering pattern based on context information among terminals. In this new service pattern, context aware triggering is supported. Moreover, service can be executed in a distributed way which means service logic can be composed of terminal and server parts. A detail scenario is also discussed in this paper.

Keywords: Ubiquitous Computing, P2P, IMS, Service Trigger, Context-awareness.

1 Introduction

The rapid development of the Internet has become an enormous impact on telecommunication networks. Internet technology provides users with more powerful services comparing with the services providing mechanisms in traditional telecommunication network; On the other hand, the traditional telecommunications services face the challenges of the newly rich and personalized Internet business model. Therefore, the integration of telecommunications and the Internet has become a popular trend and gradually deepened. This trend is both of the network's own development needs, but the personalized, integrated business needs as well.

IMS was introduced by 3GPP in R5 stage. It is a kind of integrated network architecture to provide multimedia services to mobile users in an all-IP network. And it defines a unified control network and services trigger mechanism. More importantly, it achieves the multi-services integration. However, the service triggering mechanism of the IMS would result in load pressure and network response latency, and moreover couldn't adapt to the trends of providing context-aware services.

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 715–722.
springerlink.com

This paper presents a mechanism for mobile terminal to support ubiquitous computing. A detailed description of the trigger and migration of this pattern is given. In this pattern, the mobile terminal, making use of the context-aware P2P middleware for context-awareness and communications capabilities, achieves the service trigger matching based on the context of users through the consultation mechanism between the terminals. And the advantages of the mobile communications and computing power are taken to complete the part of the service logic migration among terminals or between terminals and application server (AS).

This paper is organized as follows: section 2 introduces related research on mobile P2P middleware and context-aware middleware; section 3 describes the IMS service trigger and execution pattern, and then presents a P2P-based converged network service model; section 4 elaborates the basic framework to implement this pattern; section 5 illustrates the effectiveness of the mechanism by the analysis of specific business scenario; section 6 concludes the paper.

2 Background and Related Works

2.1 Context-Aware Middleware

The important area of ubiquitous computing research is context-aware computing. By context-aware system, the various factors in surroundings of a user can be calculated and analyzed, and then make the appropriate behaviors [1]. What's more, the middleware can offer a unified interface to the underlying capabilities. With an increasingly number of intelligent mobile terminals owning various sensors and richer user interaction means, the middleware has been of great significance in mobile applications. The common sources of context information collected by mobile terminal are light, temperature, video, voice, body movement, location information, etc. The middleware obtains the information through a variety of sensors, cameras and GPS positioning device for data collection. And then the context data is analyzed, classified, as well as present for the transmission and the provision of a unified data interface. In paper [2], a framework of location-based services (LBS) system in IMS was introduced. It was through the software of the terminal side to implement the basic service of LBS. However, there was no more in-depth analysis of how to support a variety of service in general. The paper [3] presented a P2P-based context-aware mobile middleware called PnPAP. It has the ability to support multi-protocols for P2P communication and context information management for service providers. In the terminal side, it provides a distributed service environment to achieve the free exchange of contextual information between the terminals. But PnPAP didn't specify the trigger and execution process of services in the IMS in detail.

2.2 Mobile P2P Middleware

In order to support the direct communication between mobile terminals for ubiquitous computing, mobile P2P network is needed. This is an autonomous wireless communication network to accomplish terminals interaction, collaboration, and data sharing in a dynamic, distributed and self-organization way. For some applications such as ubiquitous computing, the paper [4] described the mobile P2P application in voice

communication, text messaging, traffic information, disaster emergency information dissemination, video photo sharing, local services and other fields in detail. The paper [5] also mentioned the P2P in content sharing and collaborative computing has commercial value.

The middleware of mobile P2P is to meet the requirement of the P2P communication and business scenario with supporting mechanism for heterogeneous mobile data transmission and sensor technology while providing specific application software with uniform and convenient interface. The paper [6] listed and compared a number of mobile P2P middleware. Although most of them have a similar structure, in terms of communication and resource discovery, the scenarios are of essential difference from each other.

3 Analysis of IMS Service Trigger and Execution

In the IMS architecture, CSCFs (Call Session Control Function) are the core network entities, which contain P-CSCF, I-CSCF, S-CSCF. P-CSCF is the access entry of terminals IMS system; I-CSCF provides service nodes and routing queries to the users within the domains, and makes topology hiding between different IMS domains; S-CSCF is the core control of services trigger. It is responsible for the registration and authentication of a terminal, and it also provides the basic session routing functions to the end users. In addition, it triggers AS's value-added services when service rules are met. In addition, the user's subscription information is stored in Home Subscriber Server (HSS).

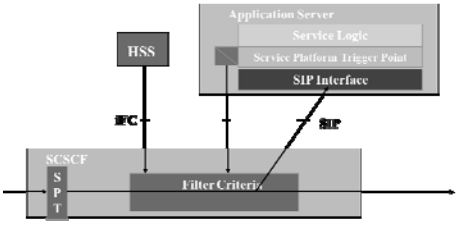


Fig. 1. IMS Service Trigger Mechanism

In 3GPP standard, the iFC includes the server address to be connected, priority, service trigger point, and other default options, as part of the user services subscription attributes stored in the HSS. These information will be downloaded to the S-CSCF when the user registration. Figure 1 shows that when the user request arrives at S-CSCF, S-CSCF analyzes iFCs of the user's services subscription data [7]. In accordance with the matching priority order, if the trigger conditions are met, the specified AS will be triggered, and then the corresponding AS will complete the business logic processing.

Through the above analysis of the traditional IMS trigger and execution mechanism, it can be summarized that there are some inadequacies listed below:

- **The iFC Matching Way Cannot Adapt to the Development of Context-Aware Service**

As the iFC only provides basic information about services itself, and does not involve the triggered user's context information, triggering the services through the iFC has to be only in strict accordance with the user subscription information and parameters to complete. This mechanism cannot give fully consideration to the user's context environment, terminal capabilities and services order status. Thus, it has no way to achieve a more comprehensive analysis to accomplish context-aware services triggering.

- **The Trigger and Execution of Service Logic Cannot be Distributed**

That the unified placement of service logic in the AS allows the network structure and functions clear, but this kind of centralized architecture is not flexible for the ubiquitous computing services which require the coordination among terminals in the light of the context information.

- **Load Pressure Is Caused by the Centralized Service Mechanism**

The process of triggering a communication services needs to go though from the initial user registration during which the filtering rules will be downloaded to S-CSCF, to matching iFC when the request comes. These processes are concentrated in the S-CSCF network entity, so that an enormous load pressure has been created to the network communications, storage and processing power of S-CSCF. As a result, the delay or even failure of services match and trigger are caused.

4 Service Pattern Based on P2P Middleware

4.1 Pattern Architecture

For the several issues mentioned above, a service trigger model with P2P middleware based is present.

Built on the original IMS architecture, this converged network model adds a function entity called Service Profile Description Server (SPDS) to provide the information of service subscriptions and abstraction. This provides the unified description of the service logic process. And with the P2P ability of middleware in intelligent mobile terminal, it makes it become a reality that the triggering decision and part of fundamental service logic can be executed in the terminals.

In this mode, the newly added entity SPDS in the side of network is to support context-aware services trigger and dynamic migration for pervasive computing. The service information and the registered user who has subscripted the service are bind. After that, the services logic information will be downloaded to the user terminal.

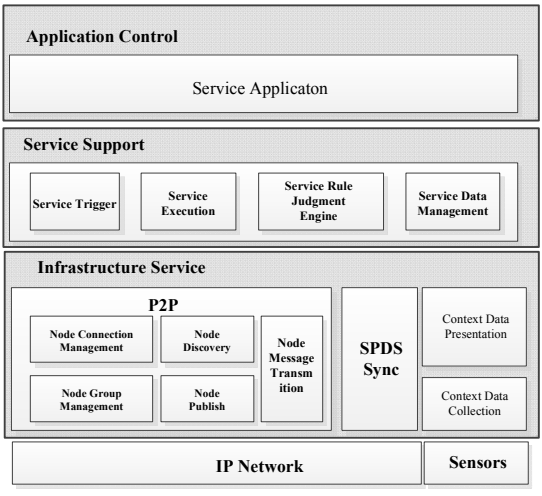


Fig. 2. Terminal Middleware Architecture

In the mobile terminal side, in order to build a P2P overlay network to support direct communication between the terminals, the P2P middleware is designed on top of the operating system (Figure 2). The middleware is divided into three layers: the lowest layer is the infrastructure services layer to construct P2P network based on all-IP network and to provide context awareness capabilities. The P2P module contains messaging exchanging function and node connection management. The context-aware module offers the data acquisition, representation and management of the sensors in terminal. In addition, in this layer, there is communication interface to synchronize with the SPDS. The middle layer is the service logic support layer, which relies on low-level ability to complete service trigger decision making, trigger action and the execution migration of part of the service logic. Besides, it manages the service data synchronization of the SPDS. The top layer is the application control layer, by which the terminal application can be designed and implemented for specific business scenarios using the ability of two lower layers.

4.2 Services Trigger and Execution

Pervasive computing requires the service trigger and executing migration could be done according to the terminal capabilities and scenarios [8]. Therefore, the terminal must control its own subscription service's features (such as service-related subscriber, basic service logic running on the terminal, trigger conditions and migration conditions, etc.). In this pattern, firstly via HSS, SPDS get the user's subscription data, and then synchronize with AS to obtain service information which is needed for the coordination trigger and migration between terminals. Afterwards, it maps the service information and registered users, and then tries to send them to the user's terminal.

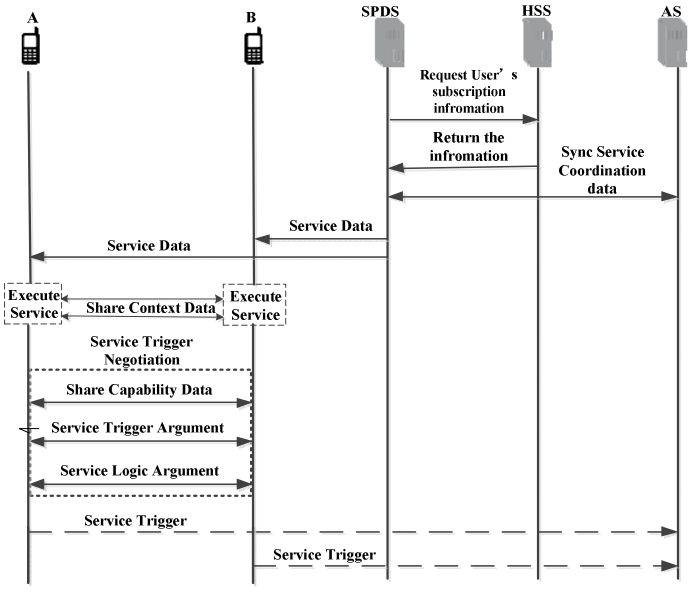


Fig. 3. A Ubiquitous-Computing Enabled Service Trigger Mechanism

After accessing to service information, the middleware use these data to store and analyze for the services trigger and execution. Firstly, terminals will share data with each other according to the context and service logic. When the trigger criteria are met, the terminal will negotiate services triggering. By sharing terminal capacity and context information, one or more terminals will finish the negotiation process of the trigger decision. The mobile middleware can also complete dynamic migration consultation of the services logic between the terminal and the AS.

According to the results of negotiation, one or more terminals will proceed with service trigger directly to the AS. When encountering network or power problems, the middleware can dynamically migrate part of the service logic to AS in order to continue executing.

5 Smart Logistics Service Scenario

The analysis of the workflow-based smart logistics service is showed as below to demonstrate the specific application of the model.

As shown in Figure 4, at first, the SPDS gets the subscription information of users from the HSS. For example, it found that users A, B, C are logistics workflow nodes (the case of three nodes is discussed here due to the limited space of paper). Then it will synchronize data with the AS to access the service logic of A, B, C node. Then these service logic data will be pushed to the A, B, and C's intelligent terminal. After the terminal receives the information and analyzes the logic, it is found that both of the node A and C are to monitor the goods delivery process on node B, while node B is responsible for triggering three services after the goods delivered from node A and

C arrives: one is to send SMS to a preset number, the second is to trigger the 3rd-party video call service to connect node A and node C terminal for the interpersonal confirmation, and the three is to send notification to control platform that the process of the goods delivery among node A, B, and C have been finished. For the communication among Node A, B, C during service execution, the middleware on the terminal could provide P2P message ability to support directly data and notification exchange among the nodes. After the completion of the service logic, the terminals will go through the service triggering negotiation. In this case because it contains video conferencing, the contents of the service negotiation involves the decision of initiator of the video, audio and video codecs, terminal capacity and network capacity. After negotiation, the corresponding service is triggered in each node respectively.

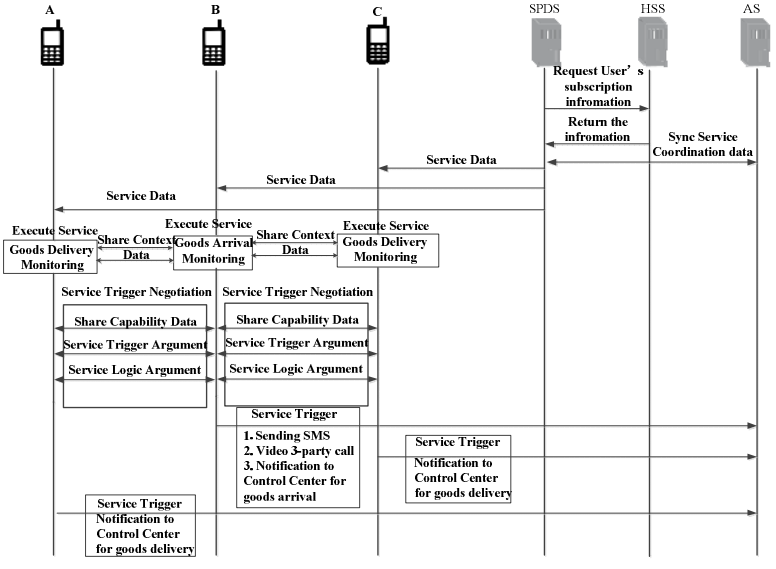


Fig. 4. Service Process of Smart Logistics Platform

It can be seen that due to the introduction of P2P, the communication between the nodes and service negotiation mechanism takes the context information of each terminal into account, making triggering more flexible. What's more, part of the service logic could be migrated to the terminal. This makes the central server has no need to poll service logic checks, reducing the load on the central server and improving the efficiency of the service logic.

6 Conclusion

In this pattern, mobile P2P middleware is introduced to the smartphone operating system in the terminal side, and in the network side, one function entity is added on top of the existing IMS architecture for the description of the service information.

It completes the semi-distributed, dynamic service trigger and execution migration. Thus it is able to provide users with a flexible and convenient means to support pervasive computing telecommunication service.

By analyzing the model and the specific service processes, it can been seen clearly that this model not only makes the original S-CSCF function entity greatly reduce the load pressure, and thanks to the introduction of user context and consultation mechanism between users' terminal, the distributed service trigger and the execution of the basic service logic is more diverse and flexible. In a word, all of these are a giant support for mobile service provisioning in a pervasive computing way.

Acknowledgement. This paper is sponsored by the China Core-High-Base project 2009ZX01039-001-001, 863 project 2008AA01A317, 973 project 2009CB320406, and innovation team 60821001. And this research is also supported by the ZTE University project.

References

1. Lee, S.G., Chang, J.: Survey and Trend Analysis of Context-Aware Systems. 2011 Information-An International Interdisciplinary Journal 14(2), 527–548 (2011)
2. Zündt, M., Deo, G., Naumann, M., et al.: Realizing Peer-to-Peer Location-Based Services in Mobile Networks. In: Proceedings of the 2nd Workshop on Positioning, Navigation and Communication (WPNC 2005) & 1st Ultra-Wideband expert Talk (UET 2005), pp. 175–182 (2005)
3. Ylianttila, M., Harjula, E., Koskela, T., et al.: Mobile Plug-and-Play Architecture for Collaborative Hybrid Peer-to-Peer Applications. In: 2008 IEEE Congress on Services Part II (services-2 2008), pp. 81–87 (2008)
4. Motta, R., Pasquale, J.: Wireless p2p: problem or opportunity? In: AP2PS 2010 The Second International Conference on Advances in P2P Systems, pp. P32–P37 (2010)
5. Hummel, T., Stramme, O., La Salle, R.M.: Earning a living among peers - the quest for viable P2P revenue models. In: 36th Annual Hawaii International Conference on System Sciences, pp. 219–224 (2003)
6. Niu, X.-Z., She, K., Lu, G., et al.: Overview of Mobile Peer-to-Peer Computing. Application Research of Computers 24(5), 269–272, 277 (2007)
7. Lai, J.: Research on IMS Service Provision and Invoking Mechanism. Information and Communications Technologies 3(2), 11–16 (2009)
8. Chen, M.-E., Su, F., Song, L.-F., et al.: Study of Adaptive Service Provision Platform Based on Middleware. Application Research of Computers 23(7), 147–149 (2006)

A Service Model Based on Stochastic Network Calculus in Wireless Mesh Networks

Gaocai Wang, Nao Wang, and Mingxing Lai

School of Computer and Electronic Information,
Guangxi University, Nanning, 530004, China P.R.
wanggaocai@yahoo.com.cn

Abstract. Stochastic service model is abstract expression of service capability of network node from probability meaning in stochastic network calculus. It is determined by channel capacity and scheduling strategy of one node. In this paper, we firstly propose a stochastic service curve for individual traffic flows based on improved WFQ (Weight Fair Queuing) scheduling in wireless mesh networks. So, our service model has greater fairness for arrival flows. To simplify performance evaluation of the whole network, we present an end-to-end stochastic service curves for aggregation flows. It can encapsulate and shield QoS requirements of different type arrival flows. The numerical results show that the stochastic service model is effective based on real-time and non-real-time stochastic service curve of arrival flows.

Keywords: Stochastic network calculus, wireless mesh networks, service model, improved WFQ scheduling.

1 Introduction

In recent years, wireless broadband networks are being increasingly deployed in a multi-hop wireless mesh networks configuration by the development of standards such as IEEE 802.11 and 802.16[11]. Experimental and commercial deployment of multi-hop wireless mesh networks are already in use widely. For example, many universities and cities have developed and deployed mesh networks. Those deployed mesh networks can provide commercial Internet access to residential homes and corporate offices. Therefore, wireless mesh networks are believed to be a highly promising technology and will play an increasingly important role in future generation wireless mobile networks.

With the increasing deployment of wireless mesh networks, researchers gradually concern about service provisioning and performance analysis of wireless mesh networks, such as delay and backlog. The related results play an important role in QoS (Quality of Service) guarantee, admission control and network programming. At present, there exist two methods for performance analysis in QoS guarantee, such as Deterministic Network Calculus (DNC or NC) and Stochastic Network Calculus (SNC). DNC means that one must compute and deduce a certain bound for service performance and then provides QoS guarantee transmission with probability 1 based on it, and SNC is based on tail distribution of service capacity and provides a probabilistic QoS guarantee. Usually, arrival

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 723–730.
springerlink.com

traffic is a stochastic process and has burstiness in wireless mesh networks, so the arrival traffic should be fluctuant. In DNC, we must ensure a certain bound in terms of probability 1, and that means a worst case QoS guarantee. So, a drawback of the worst-case view of traffic in DNC is that it does not reap the benefits of statistical multiplexing, which can result in an overestimation of the actual resource requirements and a low utilization of network resources. In SNC, by allowing even a small fraction of traffic to violate its traffic description or performance guarantees, one can achieve significant resource savings, and provides reasonable performance evaluation results.

The rest of the paper is organized as follows: Section 2 outlines related research work in the literature; Section 3 presents our main results, in which including a stochastic service curve based on improved WFQ scheduling for individual flows and an end-to-end stochastic service curve for aggregation flows; Section 4 provides our numerical result with a detailed stochastic analysis. Section 5 concludes the paper with directions for future research.

2 Related Work

Stochastic network calculus has already widely applied to modeling and evaluation of QoS due to overcome some disadvantages of traditional theories and methods for performance analysis in quantity and quality. It is extended edition of NC in probabilistic meaning. The basic model of NC is arrival curve and service curve. Arrival curve and service curve will be extended to probability meaning, they are stochastic arrival curve and stochastic service curve (SSC).

In performance analysis based on SNC, researchers mainly focus on the bounds of service performance, traffic model, service model and their corresponding characteristic. For example, EBB model [1] is the foremost proposed traffic model, its base function $\alpha(\tau) = \rho\tau + \sigma$ is linearly increasing, and the violation probability function $f(\sigma) = Me^{\theta\sigma}$ is an exponential function. SBB model [2] does not limit violation probability function, and only requires its arbitrary integral to be convergent. LEE model [3] extended the base function to general function. LEE model will transform into SBB if $\varepsilon = f(\sigma)$ and $G(\tau) = \rho\tau + f^{-1}(\varepsilon)$. SBB model can be used to Gaussian self-similar process, such as Brownian motion. In fact, most traffic is non-Gaussian process, such as α-stable is self-similar process. In addition, its distribution of burstiness is exponential function, and so does not satisfy with integral convergence. SBB extended to gSBB [4]. Incoming of gSBB includes non-Gauss self-similar process, and gSBB only limits the violation probability function *f* to a non-negative decreasing function. So gSBB relaxes *f* function than EBB, SBB and LEE. GEE [3] is a specific case for gSBB, but have to obtain the bound of busy period. In addition, m.b.c arrival curve based on maximum virtual backlog not only has some advantages similar to gSBB, but also can be applied to deduce stochastic service model with better performance solution.

Service model includes output characterization, concatenation property, superposition property, service guarantees, and per-flow service. Stochastic service curve characteristic a lower bound of service capacity of network node in probability meaning, so it different from service curve of NC. wSSC [7] is extended edition of NC in

probability meaning. EBF [5] and ESC [6] are special example of wSCC. Combined with wSCC and gSBB, one can obtain the random bounds of backlog and delay of single node and multi-nodes networks. But wSCC doesn't have output characterization, concatenation property and per-flow service. In the development of SNC, one type SNC based on moment generating function plays a very important role, but it isn't a part of above stochastic arrival curve and stochastic service curve.

3 Stochastic Service Model for Wireless Mesh Networks

According to the different requirements of QoS, the arrival flows are classified into four types by the IEEE 802.16 standard in the wireless mesh networks. The four types are: unsolicited grant service (UGS), real-time polling service (RTPS), non-real-time polling service (NRTPS) and best effort (BE). In the four services type, UGS and RTPS are related to the real-time periodic arrival flows. The arrival flows with fixed-rate can be supported by UGS, such as VoIP arrival flows, and the arrival flows with variable rate can be supported by RTPS, such as MPEG arrival flows. In contrast, NRTPS and BE are related to the non-real-time arrival flows. The arrival traffics with variable rate can be supported by NRTPS, such as FTP Server. The arrival flows without requirements of the rate and delay can be supported by BE, such as E-mail and short messaging services, etc. The priority of four services type is respectively: UGS has the highest priority, then RTPS and NRTPS. BE has the lowest priority. For convenience of analysis, the arrival flows are classified into the real-time and non-real-time flows in wireless mesh network.

3.1 A Stochastic Service Curve Based on Improved WFQ Scheduling for Individual Flows

IEEE 802.16 established a suite of detailed rules about channel interactive mechanism and different service types for wireless networks, and so solved the problem that there is no service quality guarantee in traditional wireless networks. On the other hand, it doesn't set up a framework about QoS scheduling algorithm. At present, the Weighted Fair Queuing (WFQ) is a widely used scheduling strategy in wireless networks. The WFQ assigns a weight for every arrival flow, which is related to the priority of flows and decides the amount of transmitted bits to cache. So, the scheduling strategy can guarantee sharing bandwidth for every flow.

Usually, a flow will pass through multi-router nodes in order to reach destination in a wireless mesh network, so, for any relay-node, it not only sends itself traffic, but also forwards other traffic that comes from other nodes. Real-time traffic has priority in our scheduling strategy. In addition, we assume as follows:

(1) Arrival traffic is a stochastic process $A(t)$ with fractal Brownian motion in time interval $(0,t)$. Average rate is m, variance $\sigma^2 > 0$, and a self-similar parameter $H \in [0.5,1]$;

(2) All flows can start execution on the processor as long as they are released; and the buffer of node is infinite;

(3) Network link is undamaged, so the arrival flows always are transmitted successfully. As long as the arrival traffic is released, it will be served. The node service rate denotes by R;

(4) The arrival flows will be assigned different weights in terms of priority. Real-time arrival traffics are prior to non-real-time arrival traffics.

Fig.1 shows a network model that the arrival traffic flows pass through a series of mesh nodes and then depart network. For the flow i, it is served by a mesh node, we have Theorem 1 in terms of our scheduling strategy.

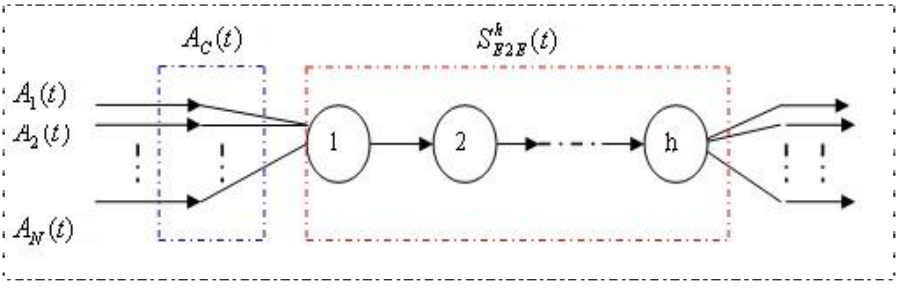


Fig. 1. A network model with h nodes and N arrival flows

Theorem 1: For given flow i with arrival process $A_i(t)$ and departure process $D_i(t)$, we have the following stochastic service curve

$$Pr\{D_i(t) \geq A_i \otimes S_i(t)\} \geq 1 - \varepsilon_d$$

$$S_i(t) = \frac{\phi_i}{\sum_{i=1}^{N} \phi_i}(Rt - T)^+, \quad T = \begin{cases} 0 & \text{for real-time flows} \\ G_{i-1}(t) & \text{for non-real-time flows} \end{cases}$$

Where, ε_d is violate probability, ϕ_i is weight which is assigned to each flow by network system, R is service rate, T is service delay of non-real-time arrival flows, and $G_i(t)$ is stochastic arrival curve of flow i.

Above stochastic service curve of individual flow is based on improved WFQ algorithm, and so can ensure the fairness for each arrival flow. By means of min-plus algebra, we can easily calculate an end-to-end stochastic service curve of each arrival flow. However, it is quite complicated to evaluate QoS (such as delay and backlog) in terms of stochastic service curve of individual flow. Therefore, we will obtain an end-to-end stochastic service curve of aggregation flows base on SNC in the following subsection.

3.2 An End-to-End Stochastic Service Curve for Aggregation Flows

For given aggregation arrival process $A_c(t)$ with N input flows, $A_c(t)$ encapsulates different type traffic flows, so it can occupy the whole bandwidth. We have the following stochastic service curve $S_c(t)$ of aggregation flows.

Theorem 2: For given aggregation arrival process $A_c(t)$ and departure process $D_c(t)$ with N input flows, we have the following stochastic service curve

$$Pr\{D_C(t) \geq A_C \otimes S_C(t)\} \geq 1-\varepsilon_d$$

$$S_C(t) = Rt$$

We can easily prove the stochastic service curve in wireless mesh network by the min-plus algebra and end-to-end stochastic service curve. Even if each router nodes service rate of the network links is different, we can also calculate end-to-end stochastic service curve by using convolution formula in whole network.

$$S_{e2e}(t) = \min\{R_1, R_2, \cdots, R_h\}\, t$$

Proof: Suppose the network consists of h router nodes, and the router node j has a service rate R_j, the node severs aggregation process $A_c(t)$ with stochastic service curve $S_j(t) = R_j t$. So, we have

(1) When $h=1$, $S^1_{e2e}(t) = S_1(t) = R_1 t$;

(2) When $h=2$, $S^2_{e2e}(t) = S_1 \otimes S_2(t)$

$$= \inf_{0\leq s\leq t}\{S_1(s) + S_2(t-s)\}$$

$$= \inf_{0\leq s\leq t}\{R_1 s + R_2(t-s)\}$$

$$= \inf_{0\leq s\leq t}\{R_1 s - R_2 s + R_2 t\}$$

For the above formula, if $R_1 \leq R_2$ $s=0$, the right side of the equality get the minimum $S^2_{e2e}(t) = R_1 t$. If $R_1 \geq R_2$, $s=t$, the right side of the equality get the minimum $S^2_{e2e}(t) = R_2 t$. So, we have $S^2_{e2e}(t) = \min\{R_1, R_2\} t$. The proposition is right.

(3) Suppose $h=k$, the proposition is right $S^k_{e2e}(t) = \min\{R_1, R_2, \cdots, R_k\} t$. We need to prove that, when $h=k+1$, the proposition is right.

$$S^{k+1}_{e2e}(t) = S_1 \otimes S_2 \otimes \cdots \otimes S_k \otimes S_{k+1}(t)$$

$$= \inf_{0\leq s\leq t}\{S^k_{E2E}(s) + S_{k+1}(t-s)\}$$

$$= \inf_{0\leq s\leq t}\{\min\{R_1, R_2, \cdots, R_k\} s + R_{k+1}(t-s)\}$$

$$= \inf_{0\leq s\leq t}\{\min\{R_1, R_2, \cdots, R_k\} s - R_{k+1} s + R_{k+1} t\}$$

$$= \min\{R_1, R_2, \cdots, R_k, R_{k+1}\}\, t$$

The proposition is proved.

In general, for given aggregation arrival process $A_c(t)$ with N traffic flows, whatever the service rate of router nodes is alterable, let $R_{\min}$ denote the minimum rate in the network link, then the end-to-end stochastic service curve is $S_{e2e}(t) = R_{\min} t$ in a wireless mesh network.

4 Numerical Results

In this section, we apply Matlab tool to study the differences between the real-time stochastic service curve and non-real-time stochastic service curve of individual flow when the weight ϕ_i is changed. Firstly, we have the following for the real-time stochastic service curve and non-real-time stochastic service curve of individual flow, respectively.

$$S_i^{real-time}(t) = \frac{\phi_i}{\sum_{i=1}^{N}\phi_i} R\,t \qquad S_i^{non-real-time}(t) = \frac{\phi_i}{\sum_{i=1}^{N}\phi_i}(Rt - T)^+$$

Here, ϕ_i is the weight of each flow which is assigned by network system, $\phi_i \in [0,1]$, T is the service delay of non-real-time arrival flow.

Fig.2 shows that the stochastic service curve $S_i^{real-time}(t)$ of real-time arrival flows changes when time t increase under difference weight ϕ_i.

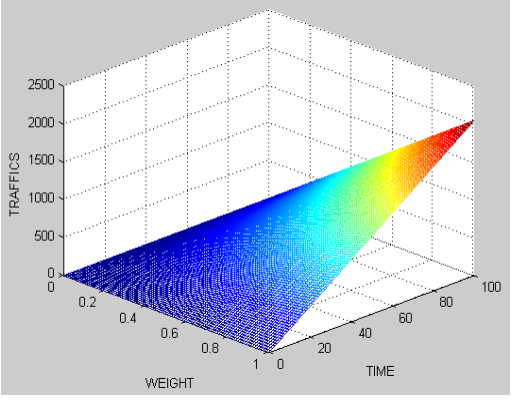


Fig. 2. The service capacity for real-time flows under different time t and weight ϕ_i

From Fig.2, we observe that there is no delay for the providing service by network node to the real-time flows because the real-time arrival flows are prior to non-real-time arrival flows.

Similarly, Fig.3 shows that the network provides stochastic service curve $S_i^{non-real-time}(t)$ of non-real-time arrival flows changes when time t increase under difference weight ϕ_i. To simplify, let stochastic service curve delay T=10 ms, that is to say, the services delay T of each non-real-time arrival flow is 10ms in a network system.

From Fig.3, we can observe that the provided service by network node for the non-real-time arrival flows is similar to the real-time arrival traffics. But there is certain delay; the value of delay is relevant to the node scheduling strategy and the traffic arrival sequence.

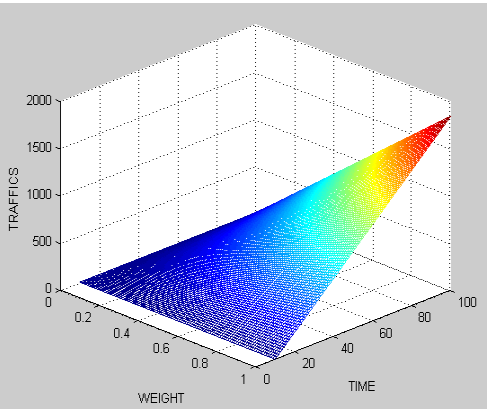


Fig. 3. The service capacity for non-real-time flows under different time t and weight ϕ_i

5 Conclusions

In wireless mesh networks, service performance is a very important topic for researchers and users. In this paper, we firstly propose a stochastic service curve for individual traffic flows based on improved WFQ scheduling in wireless mesh networks. So, our service model has greater fairness for arrival flows. To simplify performance evaluation of the whole network, we present an end-to-end stochastic service curves for aggregation flows. It can encapsulate and shield QoS requirements of different type arrival flows. The numerical results show that the stochastic service model based on real-time and non-real-time stochastic service curve of arrival flows is effective and efficient.

Acknowledgement. This research is supported in part by the National Natural Science Foundation of China under Grant Nos. 60763013, 61063045, in part by the Natural Science Foundation of Guangxi Province under Grant No. 2010GXNSFC013013. Guangxi Science Research & Technologies Development Program: 11107006-1

References

1. Yaron, Q., Sidi, M.: Performance and Stability of Communication Networks Via Robust Exponential Bounds. IEEE/ACM Trans. Networking (June 1993)
2. Starobinski, D., Sidi, M.: Stochastically Bounded Burstiness for Communication Networks. Trans. IEEE Inform. Theory 46, 206–212 (2000)
3. ZZBoorstyn, R.R., Burchard, A., Liebeherr, J., Oottamakorn, C.: Statistical service Assurances for Traffic Scheduling Algorithms. IEEE Journal on Selected Areas in Communications 18(12), 2651–2664 (2000)
4. Yin, Q., Jiang, Y., Jiang, S., Kong, P.Y.: Analysis on Generalized Stochastically Bounded Bursty Traffic for Communication Networks. In: Proc. IEEE LCN (2002)

5. Lee, K.: Performance Bounds in Communication Networks with Variable-Rate Links. In: Proc. ACM SIGCOMM 1995, pp. 126–136 (1995)
6. Burchard, A., Liebeherr, J., Patek, S.D.: A Calculus for End-to-End Statistical Service Guarantees. Technical Report. CS-2001-19, University of Virginia, Computer Science Department (May 2002)
7. Cruz, R.L.: Quality of Service Management in Integrated Services Networks. In: Proc. the 1st Semi-Annual Research Review, CWC, USCD (1996)
8. Norros, I.: A Storage Model with Self-Similar Input. Queueing Systems 16, 387–396 (1994)
9. Mayor, C., Silvster, J.: Time Scale Analysis of an ATM Queueing System with Long-Range Dependent Traffic. In: IEEE INFOCOM 1997 (1997)
10. Liu, T., Liao, W.: Location-Dependent Throughput and Delay in Wireless Mesh Networks. IEEE Transactions on Vehicular Technology 57(2), 1188–1198 (2008)
11. Wang, X., Liu, J., Cheng, G.: Analysis of Bottleneck Delay and Throughput in Wireless Mesh Networks. In: Proceedings of the 3rd IEEE International Conference on Mobile Ad-hoc and Sensor Systems, Vancouver, Canada, pp. 765–770 (2006)
12. Cruz, R.L.: Quality of Service Management in Integrated Services Networks. In: Proc. the 1st Semi-Annual Research Review, CWC, USCD (1996)
13. Cheng, Y., Zhuang, W., Ling, X.: FBM Model Based Network-Wide Performance Analysis with Service Differentiation. In: Proceedings of the 4th International ICST Conference on Heterogeneous Networking for Quality, Reliability, Security and Robustness, ICST Qshine 2007 (2007)
14. Cheng, H.T., Zhuang, W.: QoS Driven Node Cooperative Resource Allocation for Wireless Mesh Networks with Service Differentiation. In: Proceedings of IEEE Global Communications Conference (Globecom 2009), Miami, USA (2009)
15. Qi, H.: Research on Network Calculus based QoS Performance for Wireless Mesh Networks. Central South University (2010)
16. Lai, M.X.: Performance Evaluation of Wireless Mesh Networks with Self-Similar Traffic. B.S. dissertation, Guangxi University (2011)

The Recent Condition and Relevant Algorithms of the Study on Sensor Network Deployment

Wu Yi

School of Computer Science and Technology,
Heilongjiang University,
150080, Harbin, China

Abstract. Sensor network deployment directly influence the management of network resources and distribution of sensor network data. this paper will introduce the recent condition and relevant algorithms of the study on sensor network deployment, the international situation and the future trend of development, etc.

Keywords: Sensor network, deployment algorithm, sensing field.

1 Introduction

Integrating the sensor technology, embedded computing technology, distributed information processing technology and communication technology, sensor network can help to conduct monitoring, perceiving and processing the information of the monitored object in the network distribution area. Therefore, the network system can be widely used in national defense military, environmental monitoring, traffic management, anti-terrorist and disaster relief, etc. Sensor network is a revolution of information perception and acquisition. As a new field of study, it puts forward a lot of challenging research to scientific and technical workers[1]. In the sensing field, it works as a network node. Not only does it collect and pocesses sensor data, but also sends data and information through routers.

Recently, the algorithms of the study circle studying the sensor deployment adopts artificial deployment and self deployment. In artificial deployment, the sensor nodes can disperse themselves in the target area through the airplane tossing or artillery firing. When the sensor is placed to the initial positon, the sensor nodes are fixed in that positon and do not have their own mobility. In the self deployment, the sensor nodes are placed to the original position by various means in that they are either centrally placed to some specific position or to the target area at random and uniform. The sensor nodes themselves are movable, and they can be moved to another place in accordance with need so as to conveniently achieve the sensor nodes configuration program. Now, I will introduce some kinds of algorithms that the study circle uses in studing the sensor deployment.

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 731–736.
springerlink.com

2 The Relevant Algorithms of the Study on Sensor Network Deployment

2.1 Algorithm of Randomly Distributed Deployment

The probloms that needs to be solved in the process of adopting the random deployment are as follows: on one hand, in order to achieve the covering requirment, we need the layout the number of sensors in the region. On the other hand, if the adding sensors need to be deployed again in the perceptive region, then what is the minimum cost in per deployment. When scaning covering, the technique of random research is more effective than many other ones.

T. Clouqueur in Duke university proposed to describe the region[2] of perception using two-dismensional grid, and adopting the relevant algorithms can calculate the possibility that regional tagets are perceived by the layouted sensors. We can use the integrative method to calculate the possibility of perceptive information in net nodes. This kind of method can also be used in the random deployment detecting the regional obstacles covering, which can provide the path-detecing possibility through the lowest detecting one.

Through the consumption of layouted sensors, we can calculate the deployment cost of the entire perceptive region. According to experiment, simulating the distributed position of deployed sensor nodes can meet the need of deployment of perceptive region. This kind of deployment is also in line with the need of deployment in the barrier region.

2.2 Algorithms of Non-random Deployment

If the fixed sensor nodes layouted by man in perceptive region is safe or communication between the moving sensors is effective, then in the smaller perceptive region, we can adopt direct deployment algorithm to achieve the layout of sensor network nodes; if layouting the moving sensor nodes in the perceptive region, and date communication between nodes、nodes location and network construction in protecing the covering region qualifies the possibility, then we can study the related algorithms to achieve deployment according to these characteristics, which promotes the study field in deployment of sensor nodes.

Dhillon of the Duke universtiy proposed one directed deployment algorithm [3].When ensuring that every grid achieves the covering taget in layouting sensor nodes, we can turn to arranging nodes for fault-tolerant sensing area. The study shows that comparing with algorithm of random deployment, that of direct deployment can decrease quantity of the needing sensors. Dhillon also proposed that one sensing area in full-covering using in sensors resouce management has optimized structure [4],and also proposed to use grid to describe sensing area,and establish detecting model of inaccurate possibility to deploy the sensing model which contains obstacle. The taget is to calculat the arranged minimum quantity of sensors that covering area needs. At the same time, carrying out the experiment and getting the result.

Chakrabarty of Duke university proposed the idea [5]of coding theory of distributing the position of sensor nodes in the sensor network. He thought that the maximum quantity is in accordance with the requiring sensors, and adopts the method that in

sensors area,using interger linear programming to decide the layouting sensors position. Althoug the target of deployment is strict regional covering, the basic idea has formed the position of sensors based on the sensing target. It needs a high-density deploying sening area, and uses more quantity of sensors. Complicated calculation limits the further development of this kind of method.

Boyoon Jung of Southern Calfornia proposed that it is potential sensor deployment to use remote mobile robots as sensor nodes, and adopts various controlled movable devices owning the moving and controling behavior as the carrier of sensors, which can ensure the specific position of sensors moving. The problom that using the movable or non-movable sensors tracks the moving target in the sensing area. We can adopt two testing standards based on regional mul-targets tracking. One is to use topoloty and distributed density as standard corse deployment. The other is to use taget- tracking controller to ensure the maximum number of target-tracking in sensing area. According to sensing ,requiring deployment can deploy the movable and non-movable nodes to area with a contain percentage[6].

Himanshu Gupta of NewYork university introduced the given region-covering and self-orgonizing algorithms which are used to decrease the resouce cosumption of network nodes[7].The author proposed a kind of technique differing from coverage metrics which is used to reduce the communication cost.Their goal is to deal with and choose the layouted sensors and the core problom is the optimization of sensors.

The art gallery problem (AGP)proposed by Dutch scholar Berg M in reference[8] is a classic one of sensor network deployment, studied by many scholars. This problom regards sensing area as art gallery. Arranging video in art gallery is regarded as arranging sensors in the sensing area. Every point in art gallery at least is within the visual range of certain video. The study shows that we can find to use minimum videosto ensure any point in atr gallery within the visual range of certain video.AGP can solve a class of problems of using visual sensors to complete sensor network regional coverage. Differing from the sensor of random possible perception, this problem will ensure the distance supervised by video in art gallery is effective and ensure the supervised quality by video within the effective distance. Such kind of problems show that it can sovlve sensing problem of two-dismensional spacial information, but will not adapt to the three-dismensional one.

Seapahn Meguerdichian, form Southern California, Los Angeles Seapahn Meguerdichian university, put forward to look for the maximum and minimum coverage problem in the sensing area from deterministic, statistical,the worst coverage and the best coverage etc. several aspects. Combining with computational geometry and graphic theoretic technique ,he chose VORONOI map to describe sensing area and cut it into polygonal regions. Using the polynominal time to achieve the design of the worst coverage and average coverage algorithms.

2.3 Autonomous

When the quantity of the sensor nodes is large, communication capacity is limited and cannot maintain covered sensing area network system, the autonomous is very effective.

Howard A, from Southern Calfornia, Los Angeles, described effectvie deploying algorithms of adopting movable sensor nodes in sensor network[9]. This stuty is very

meaningful. It adopts distributed nodes collecting and constituting the sensor network and every node has sensible, computational, communicative and movable ablility. It not only published to prove the autonomous deployment of sensor network, but also used the stimulated robots to be mobile sensor device. According to the relevant information of arranged precursor nodes, we can add the deploying sensor nodes to the unknown condition. This deploying plan is open to the entire process and limits the regional coverage problem. Reference[9] proposes that using distrubution and measurement to solve the sensing area coverage, which firstly proposes the deployment algorithm in such aspect.

Zou Y of the Duke university in the reference[10] proposed the sensor deploying process described by the "visual force" method. When the sensor nodes are layouted randomly the sening area, according to the distance between sensing area and nodes, we can define the visual force relation between repulsion and attractiveness existing between nodes. According to the effection of visual force existing between regional nodes, we can move the nodes and layout them to the optional position. In the algorithms, it neither consider the effection nor the deploying tact of obstacle

Maxim A. Batalin of Southern California university in reference[11] also proposed to use multi-dimensional moblie sensors to achieve two algorithms of antonomous deployment ,adopted the natural deploying idea of from the patial coverage to the global coverage and analyzed the interactive factors between nodes of mobile sensors. Experiment shows that comparing with artificial deployment, antonomous deployment can optimize 5%-7%

3 The Progress and Future Developing Trend of the Study on the Sensor Network Deployment

Sensor network nodes deployment needs the covering level of measuring nodes in the sensing area and the effectiveness of objectively envaluing grid coverage is a direction of research. K.Chakrabarty of Duke university put forward to adopt different levels of granularity to use network barrier coverage of sensor of direction grid measurement, high-density grid can measure higher testing accuracy and give an grid model of apropriate-granularity descriptive sensing area in accordance with the sensing requirement of regional information and regional natural factors. Guoliang Xing of Washington university proposed VORONOI map can also describe sensing area[12], can compare grid of direction with VORONOI map and studies the effectiveness of granularity of grid model.

Xiaorui Wang of Washington university ect. proposed the central coverage and connecting deployment in the sensor internet. According to applying different covering need, they put forward the agreement of covering deployment and analyzed the relevant relation between coverage and connection. The futher research can point at the requirement of distributed test and target tracking, discussing the sele-adating covering tecnique of re-deployment and studying the covering algorithm that adapts to the nodes connection.

Chih-fan Hsin of Michigan university and Li(Erran)Li of Lucent research institute proposed effective deployment, miantianed sensor nodes of covering area and solve the effective problem of sensor network resouce[13].When the distrubuted density of

sensor nodes is quite high, the effective redundant sensor nodes、alternatively waking and sleeping nodes ensure the effectiveness of sensor network resouce. The further work is how to arrange network nodes better、rationally build the routing path, effectively balance distribution of network nodes resouce and at the same time design the effective routing algorithm of resouce, which has a key effection on fully extending the life cycle of sensor network and ersuring the effective use of resouce and which is the very valuable studing direction of sensor deployment.

Using the heterogenous sensor nodes to deploy sensor network,it is a developing trend to adjust the structure of nodes in time adapting to the requirement of sensing information in the target region according to the need. In the fixed sensing region, we can layout certain number of simialr class of sensor nodes and constitute the frame-testing stucture of sensor network; At the same time, layout certain number of hetergenous sensor nodes again and move the position in terms of the need of sensing target test. Adopting relevant self-adapting testing model keeps the dynamic adjustment of network structure and ensures the various stages in the limited cycle of sensor network to stay in a optional state.

Improving the scalability of sensor network. First, patially deploy the every part of sensing area,then connect every part to form the whole deployment of sening region. It can also extend the testing range of sensing region in accordance with collecting need. Requiring algorithm of deployment to achieve sensor nodes layouting to form the distributive strtucture of Distributed**.**

4 Conclusion

In China,Harbin institute of tecnolougy and Heilongjiang university、Tsinghua university and PLA university of science and other institutes also have studied on the related tecnique of sensor netword. Howerer, the papers on the deployment of sensor network is less. Yao Xianghua of Xi'an transport university put forward algorithm[14] of sensor deployment based on constraint mechanism. The author of this paper proposed the possibility-testing deploying algorithm etc. of sensor network. This indicates that the demestic research in this area needs further study.

References

1. Li, J.Z., Li, J.B., Shi, S.F.: Concepts issues and advance of sensor networks and data management of sensor networks. Journal of Software 14, 1717–1727 (2003)
2. Clouqueur, T., Phipatanasuphorn, V., Ramanathan, P., Saluja, K.K.: Sensor Deployment for Target Detection. In: Proc. International Workshop on Wireless Sensor Networks and Applications, pp. 42–48 (September 2002)
3. Dhillon, S.S., Chakrabarty, K.: A Fault-Tolerant Approach to Sensor Deployment in Distributed Sensor Networks. Dept. of EE & CE, Duke University, Durham, NC (2002)
4. Dhillon, S.S., Chakarabarty, K.: Sensor Placement for Grid Coverage under Imprecise Detections. Dept. of EE & CE, Duke University, Durham, NC (2002)
5. Chakrabarty, K., Iyengar, S.S., Qi, H., Cho, E.: Grid coverage for surveillance and target location in distributed sensor networks. IEEE Transactions on Computers 51 (2002)

6. Jung, B., Sukhatme, G.S.: A Region-Based Approach for Cooperative Multi-Target Tracking in a Structured Environment. In: 2002 IEEE/RSJ International Conference on Intelligent Robots and Systems (2002)
7. Gupta, H., Das, W.R., Gu, Q.: Connected Sensor Cover: Self-Organization of Sensor Networks for Efficient Query Execution. In: Mobihoc 2003, Annapolis, MA (2003)
8. Berg, M., Kreveld, M., Overmans, M., Schwarzkopf, O.: Computational Geometry: Algorithms and Applications. Springer (2000)
9. Howard, A.: Mobile Sensor Network Deployment using Potential Fields: A Distributed,Scalable Solution to the Area Coverage Problem. In: Proceedings of the 6th International.Symposium on Distributed Autonomous Robotics Systems (DARS 2002), Fukuoka, JA (2002)
10. Zou, Y., Chakrabarty, K.: Sensor Deployment and Target Localization Based on Virtual Forces. Dept. of EE & CE, Duke University, Durham, NC (2003)
11. Batalin, M.A.: Spreading Out: A Local Approach to Multi-Robot Coverage. In: Proceedings of the 6th International Symposium on Distributed Autonomous Robotics Systems, Fukuoka, JA (2002)
12. Xing, G., Lu, C.: Greedy Geographic Routing is Good Enough in Sensing Covered Networks, Department of Computer Science and Engineering, Washington University (2003)
13. Hsin, C.-F.: Network Coverage Using Low Duty-Cycled Sensor: Random & Coordinated Sleep Algorithms. In: IPSN 2004, April 26-27 (2004)
14. Sh, Y., Qy, Y.: Research of Algoriththm for Placement Program Base on Constraint Propagation. Computer Scencce 34, 176–178 (2007)

Mobile Book Management Application Using Camera of Smart Phone

MyoungBeom Chung and IlJu Ko

Department of Media, Soongsil University,
Seoul 156-743, Korea
{nzin,andy}@ssu.ac.kr

Abstract. This thesis suggests an application that will allow users to manage their books easily by using an open API based book information available from smart phones or on-line. A camera of a smart phone can recognize bar-codes and QR-codes on books automatically and help users retrieve more detailed information through the open API based on thus recognized information. The suggested application will help users save the provided content of a book in Scan List or My List, and create a memo for the page users read last, and write a simple comment when they finished the book. In other words, the suggested application will provide useful functions for users to manage books effectively.

Keywords: SmartPhone, Book Management, Bar-Code, QR-Code.

1 Introduction

The development of the mobile environment as well as the widespread of the internet not just have helped people gain an easy access to more plentiful and deep information, but also has popularized smart phones, which are otherwise called small computers in hands, as the mobile terminals have transcended the simple role of getting a call connected. In case of the cameras of smart phones, they were used mainly for taking and store pictures. However, they are now responsible for providing diverse types of information, such as augmented reality (AR: Augmented Game) games, AR navigation systems, Bar-code and QR-code (Quick Response) [1].

When it comes to the current augmented reality function that uses the camera of a smart phone, it has been applied for the technologies that present desired images or animation on a certain position through the recognition of a marker on that position, but also applied for the navigation function that uses the GPS information to mix the position values of users and find specific locations. The animation technology based on the marker recognition presents animation by using OpenGL ES on the camera overlay and is mainly used for games [2], [3]. The navigation function that combines the GPS function uses the heading and location data of Location Manager available from smart phones. iPhone ARkit provides open sources for the function to be used more easily [4].

The bar-code and QR-code recognition technologies, well-known due to QR-Dic and QROOQROO, will allow the data stored within the codes to be recognized by the camera of a smart phone, the recognized data to be sent to its server and the detailed

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 737–744.
springerlink.com

illustration relevant to the code to be provided through web pages. Such research regarding code recognition technologies has been carried out very much based on the existing computer platforms as well as the mobile platform [5], [6], while they are now available in libraries in Zxing and Zbar to allow them to be used easily on the iPones's iOS platform [7], [8]. The QR-code and bar-code recognition technologies have been highlighted as social issues. The QR code was used for the advertisement of Avante car in 2010. There have been vigorous marketing activities using QR codes through various portals (Naver QR code. Daum code).

Accordingly, this thesis suggests a mobile book management application that allows users to manage books using the bar-code and QR-code recognition technologies. In terms of its functions, the suggested application is similar to Books, which is an application provided by Apple to help individuals manage their books on a Mac PC and a Mac Book. However, Books can be used only on a computer or a notebook, recognize the bar-code only and has a problem that its users are required to update the database of books themselves periodically. This suggested application can be used on iPhone that can be carried easily by users and can recognize the bar-code as well as the QR-code. As this suggested application uses the Naver's open API and the Google's API for books, there is no need for users to update the database of books. Accordingly, the suggested application can recognize the bar-code or the QR-code of books, regardless of the locations of users, whether they are at libraries, bookstores or home.

This thesis includes following sections. The Chapter 2 introduces the bar-code and QR-code and describes Zxing and Zbar which are the libraries for recognizing codes. The Chapter 3 explains the configuration of the suggested application and the overall structure of the system. The Chapter 4 presents the results of the execution of the application developed based on the design of the Chapter 3. The Chapter is the conclusion.

2 Related Research

2.1 Bar-Code and QR-Code

The bar-code and the QR-code are the methods for managing information to process large quantities of data in a prompt and accurate manner. Starting with the research for "Auto calculation in supermarkets", conducted in 1923 by Wallace Flint in America, the 1-dimensional bar-code has been applied widely in industries such as production and distribution of products, offices, hospitals, post offices, communications and aerospace. The bar-code displays information in a patterned arrangement of black bars and white ones (spaces) of various widths as shown in Fig. 1(a). However, due to its weak points in terms of accuracy of data, sizes of symbols, volume of information, and displayed information, the 2nd-dimensional code is created to make up for such shortcomings as shown in Fig. 1(b). The 2nd dimensional code consists of codified second-dimensional symbols along the X and Y axes in a square of certain size.

(a) (b)

Fig. 1. One and two-dimension bar code: (a) one-dimension bar code and (b) two-dimension code

The two-dimensional code is divided into a stacked bar code and a matrix code depending on how to configure data. A representative stacked code is PDF417, which consists of some individually recognizable characters that are put together and configured in a row horizontally, just like the 1-dimensional bar-code [9]. The QR code is a representative matrix code, which consists of square-shaped black elements with the same width that are arranged in a mosaic pattern to form data [10]. The QR-code is a sign with a two-dimensional structure developed by Denso Wave in Japan. It can contain and display 4,295 numerical characters for the maximum and 1,817 Chinese as well as Korean characters for the maximum. In addition to this, there are other two-dimensional codes, such as the data code and the maxi code. The two-dimensional bar-code makes it possible to register large data, helpful containing numerical data, diverse languages and graphic data.

2.2 Zxing and ZBar

Zxing is otherwise called zebra crossing. It is a Java based open source that can recognize 1D/2D bar-codes. The main purpose of Zxing is to analyze photos on a mobile phone itself and decode bar-codes without communicating with a server. The code types that are currently supported include UPC-A, UPC-E, EAN-8, EAN-13, Code 39, Code 93, Code 128, QR Code, ITF, Codabar, RSS-14, Data Matrix, PDF 417 and Aztec. As it is created based on the Java language, it is realized in a way that all the above codes can be recognized on the android mobile OS. The Java language can be converted into the C/C++ language, allowing it to be used on iOS, which is the OS for iPhone. However, only the QR-code can be supported and recognized. In addition, as it is required to take a picture and then analyze the photo accordingly, there is some inconvenience of taking a picture in multiple times if the recognition process keeps failing.If it is needed to be used it for other applications, necessary files need to be copied and corrected for a specific purpose.

Zbar is also an open source that can recognize the bar-code from diverse sources. It is based on C++, which is different from Zxing. It can recognize EAN-13, UPC-A, UPC-E, EAN-8, Code 128, Code 39, 2/5 Code and QR-Code. As it is based on C++, it can be used on Linux/Unix, Windows and iPhone, but does not run on the android mobile OS. However, it has an advantageous point that it makes it possible for a user to take a real time scanning through a mobile camera and analyze the scanned picture,

allowing the bar-code to be tracked and recognized automatically, without the need for the user to take a picture manually. Accordingly, for this paper, Zbar, which is more optimized for iOS, is used to recognize the bar-code and QR-code.

3 Design and Configuration of the Application

This chapter will describe the overall structure of the system of the suggested application and also about how to run the application. The application will search through books in the sequence as shown in Fig. 2.

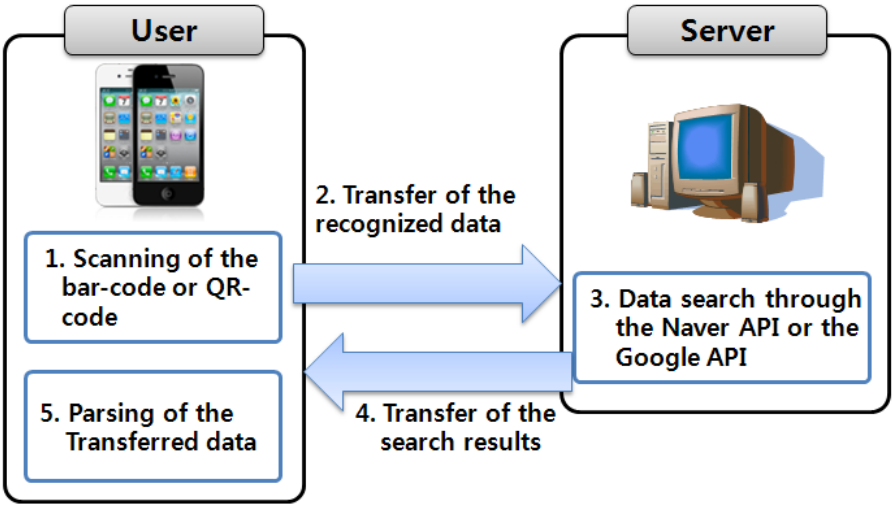


Fig. 2. Process of book search application

A user who wants to store the information of a book needs to scan the bar-code or the QR-code of the book using the iPhone's camera, and then the automatically recognized data will make a request to the Naver API and the Google API for searching a book in order to get more detailed information transferred. The reason why the Naver API and the Google API need to be used at the same time is because most Korean books can be searched through the Naver API while some foreign books can not be searched through it, so the Google API needs to be used for supplementing the search the foreign books that would not be searched otherwise. The way of making a request the search is different between the Naver API and the Google API as shown in Table 1. The Naver API and the Google API deliver key values, which are assigned to users. The number of requests made corresponding to the key values will be counted. In case of Naver, 25,000 requests can be made for one key for one day for the maximum, while, in case of Google, there is no limit to the number of requests to be made as Google runs a lab-based operation.

Table 1. Book search example using Naver and Google API

API	Example
Naver	http://openapi.naver.com/search?key=user_key&query=9780135705995&target=book&d_isbn=9780135705995
Google	https://www.googleapis.com/books/v1/volumes?q=9780135705995+isbn:9780135705995&key=user_key

As it is required for the data recognized from the bar-code and QR-code to be transferred as the search word, the value of the query to be searched needs to be entered after the query. In case of more detail queries, the desired result can be gained when d_isbn, the additional query statement, is prepared when the searching is carried out using the ISBN for books. In case of Google, the form will be made shorter in a way that the query language is prepared after 'q' and the ISBN, a transfer factor, needs to be used additionally as a search word to progresses the searching process. In other words, the query request as described in Table 1 shows that the book requested through both Naver and Google is a book whose ISBN is 978-01-3570-599-5. When it comes to the search results, Naver transferred a result in the form of XML (Extensible Markup Language) and Google transferred a result in the form of JSON (JavaScript Object Notation). The result acquired through the search request will be sent back to iPhone. Then, in order to process the transferred data, the Naver API uses the NSXML Parser and the Google API uses the JSON Parser. The data processed by using each Parser will be displayed on the screen showing the book cover, the title, the author, the publisher, the published date and the detailed web site link. Then, the user can store the results in Scan List or My List to manage books. At this time, it would be necessary to show the information in My List by author, so it is required to process data for listing books by author. The relevant configuration of the book and author information is shown in Fig. 3.

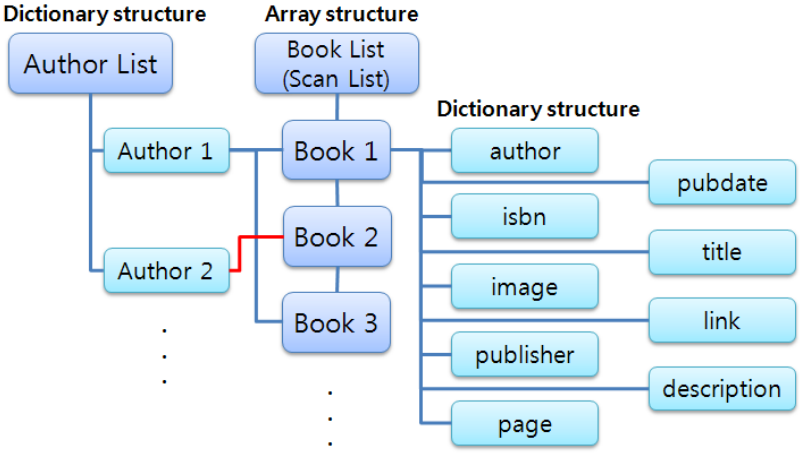
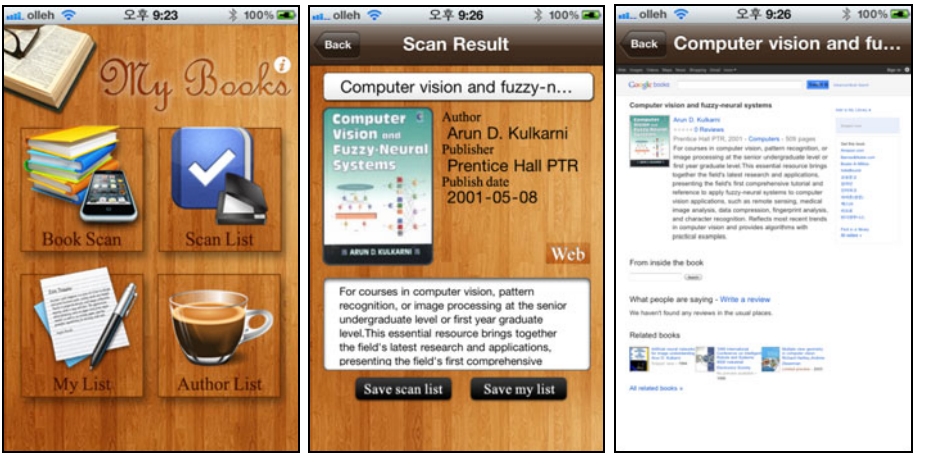


Fig. 3. Data process structure of book search result

The Book List and the Scan List shown in Fig. 3 are to be used for showing the information of Scan List and My List. The detail information shows the data in the dictionary structure shown at the right side of a book. When it is needed to generate the list by author, the individual books connected to the author will be retrieved through Author List and used. Then, the initial data of the page node below each book are 0 and the node will be used when a user enters the number of read pages later.

4 Application Development Results

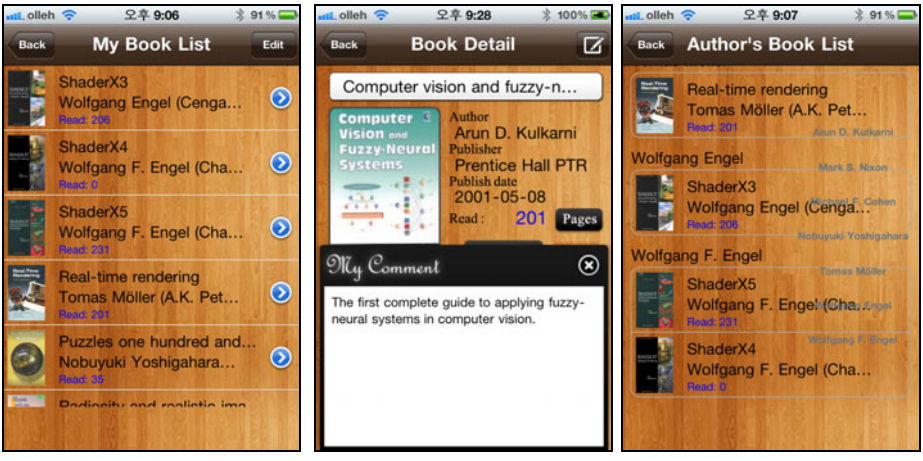
This chapter shows the result of the application developed based on the design of Chapter 3. Fig. 4(a) shows the first page of the suggested application. In order to enter the book information, the user needs to select the Book Scan menu and scan the bar-code and QR-code. When the Book Scan menu is selected, the application will turn on the embedded camera using the ImagePicker function of iPhone, allowing the ZBar library to run and the real-time code scanning to take place. Fig. 4(b) shows the result of the scanning of the book "Computer Vision and Fuzzy-Neural Systems" As the summary of the book can be acquired form the Naver API, the application will express the information using the result acquired from the Google API. When the character "web" on the center right is touched, you will see the process of linking to the web site with detail information or Google as shown in Fig. 4(c).



(a) (b) (c)

Fig. 4. Scene of book management application: (a) Main scene, (b) Search result scene, and (c) Web view scene

The user can store the search result in Scan List or My List. In Scan List, you just store the books that are scanned but not purchased and you would like to review at library or a bookstore for a moment. In My List, you sort out purchased books and write your thoughts and comments. Fig. 5(a) shows the list of the books stored in My List.



(a) (b) (c)

Fig. 5. Scene of book management application: (a) My list view, (b) Book detail view, and (c) My list view scene according to author group

Books in My List and Scan List will be listed in a way that the most recently registered books are to be placed on the top of the list. If it is needed to delete registered books or change the order of registration, the Edit button on the top right on the figure needs to be touched to delete or change the order of books. Fig. 5(b) shows the books stored in My List and also the comments left by the user. The icon on the top right in Fig. 5(c), looking like a Write menu, is for you to leave some comments. It will not be displayed when the Book Detail is selected in Scan List. When the Pages button on the center right of the figure is touched, the number of pages of books that are read up to now can be recorded, and the number of the record will be displayed in numeric information next to the button. The books stored in My List can be displayed by author. The menu for it is Author List shown in Fig. 4(a). When a user touches it, the list as well as the index can be shown. As explained up to now, you can now know the suggested application is useful in a way that the application can help you sort out books and store the information regarding at which page you stopped reading a book. Yon can also notice that Books, an application provided by Apple, is embodied in a small mobile terminal as it is. Accordingly, the suggested application can be applied usefully for the users who might have difficulties in managing books or have to take a look at books for the moment at libraries or bookstores, or want to store the titles of books they love to buy.

5 Conclusion

The application, suggested in this paper, not only can help a user see detail information of a book by using the camera of iPhone to recognize the bar-code or the QR-code, but also can sort out one's own books, thoughts after reading and comments. It will also make it possible for you to scan a book simply at a library or at

a bookstore to review a brief introduction of the scanned book and also to store the information regarding at which page you stopped reading it. This application realizes what has to be processed using expensive computers and cameras that operate on a network wholly into a mobile terminal. It also has developed the existing technology by adding convenience.

As a subject for our next search, we will develop a service supplemented with the SNS (Social Networking Service) function by preparing a server to help people communicate with each other.

Acknowledgments. The work was supported by the Soongsil University Research Fund.

References

1. JungHyun, P.: Mobile Marketing of Smart Phone Period. Marketing 45, 33–38 (2011)
2. Oda, O., Feiner, S.: Rolling and shooting: two augmented reality games. In: 28th International Conference Extended Abstracts on Human in Computing Systems, pp. 3041–3044 (2010)
3. Tan, C.T., Soh, D.: Augmented Reality Games: A Review. Gameon-Arabia, Eurosis (2010)
4. iPhone ARKit Project, `http://code.google.com/p/iphonearkit`
5. Muniz, R., Junco, L., Otero, A.: A Robust Software Barcode Reader Using the Hough Transform. In: International Conference on Information Intelligence and Systems, pp. 313–319 (1999)
6. Ohbuchi, E., Hanaizumi, H., Hock, L.A.: Barcode readers using the camera device in mobile phones. In: International Conference on Cyberworlds, pp. 260–265 (2004)
7. Multi-format 1D/2D barcode image processing library, `http://code.google.com/p/zxing`
8. ZBar bar code reader, `http://zbar.sourceforge.net/`
9. ISO/IEC 15420:2000: Automatic identification and data capture techniques. Bar code symbology specification. EAN/UPC. Information technology (2000)
10. ISO/IEC 18004:2000: Automatic identification and data capture techniques. Bar code symbology. QR code. Information technology (2000)

Modeling Nonfunctional Requirements in Software Product Line

Sun Lianshan and Wang Jinyu

College of Electronical and Information Engineering
Shannxi University of Science and Technology,
XI'AN, Shannxi Province, China
{sunlianshan,wjyokok}@gmail.com

Abstract. Several approaches of modeling nonfunctional requirements (NFRs) of single product have emerged and little attention has been paid to model NFRs of software product line (SPL). This paper first argues the necessity and challenges of modeling and managing NFRs in SPL, and then presents an approach built on an existing SPL requirements modeling approach--feature modeling, to explicitly model NFRs-related assets in feature model. The approach is designed to be extensible to model NFRs of different types, and to effectively manage complexity of feature model with NFRs.

Keywords: Non-Functional Requirements, Feature Model, Software Product Line, Domain Modeling.

1 Introduction

Software requirements include functional requirements (FRs) and non-functional requirements (NFRs). FRs tell what a software system should do while NFRs tell how FRs should be implemented [1]. It is increasingly accepted that explicitly modeling NFRs is a prerequisite to build high-quality software system. To this end, several NFRs modeling approaches have been developed, such as quality attribute scenario [2] and NFR framework [1]. However, they were suitable for developing a single product, and did not consider how to model, manage and reuse NFRs in a software product line (SPL), which consists of multiple software products covering similar functional domain [3].

SPL engineering aims at building a solid basis to be re-used in developing new software products in application engineering, and generally uses feature modeling technology [4] to capture the commonality and variability of a SPL in a domain feature model (DFM) [3]. Feature modeling initially emphasizes on FRs of a SPL and provides little support for NFRs [4]. It is difficult for developers to model and manage complex NFRs in a feature model, and to easily integrate application-specific NFRs into application feature model (AFM) tailored from DFM [5].

This paper explores the necessity and challenges of modeling and managing NFRs in SPL, and presents an approach extending feature modeling approach to explicitly model and manage NFRs in SPL. This research is the first step to our long term goal of enabling the derivation of new product satisfying NFRs in SPL by tailoring DFM.

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 745–753.
springerlink.com

2 Necessity and Challenges

The necessity of modeling NFRs for a single product had been widely recognized [1] and we argue that it is even more important to model NFRs in SPL due to the increasing adoption of SPL methodologies and the increasing requirements for high quality software in various domain.

SPL methodologies increasingly become a routine of software developers. Nowadays, plenty of software systems have been applied in almost all area of business and life. Development of a new software system is often undertaken with multiple similar software systems as available references. Developers are experiencing a transition from "develop from scratch" to "develop from reusable assets", and are increasingly required to deliver a product with high customizability and reusability.

To build a software system with high quality from SPL assets, a prerequisite is, as the case of developing single product, to explicitly model NFRs in SPL model. DFM without NFRs is not enough to shape the SPL software architecture with good quality and consequently new product shaped by AFM tailored from DFM may not embody expected quality attributes.

Besides the fact that NFRs are subjective, relative, interactive and tending to have a broad impact on the system as a whole [1], SPL engineering places additional challenges on modeling NFRs. We identified three challenges (C1~C3) of developing NFRs modeling approach as follows.

C1 NFRs should be treated as first class citizens in SPL model due to the fact that developers may derive a new product with functionalities same to those of other products while with different NFRs. Assets related to NFRs cannot be reused in the disciplined way of reusing assets related to FRs until explicitly capturing NFRs in SPL model.

C2 NFRs in a SPL model will be more complex than NFRs in requirements model of a single product. Products in a SPL may operate in different context and may be required to satisfy NFRs of different types and levels.

C3 Different SPL may compile different knowledge base for eliciting, refining and analyzing NFRs. The approach of modeling NFRs should be flexible enough to effectively use different domain knowledge.

3 Modeling NFRs in SPL

3.1 The Base Feature Modeling Approach

In SPL, feature model captures commonality and variability of requirements in SPL and can be tailored to produce the specification of application product [6]. Mei et al. presented a concrete form of feature model [4], as shown in Fig. 1. It partitions a feature model into four parts: functional section describing FRs in layers of service, function, and behavior characteristics (BC), use cases section describing how the target system are used to perform services, quality section describing NFRs of SPL, and constraints describing various relationships among different parts of feature model.

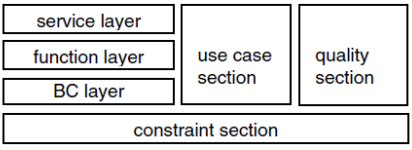


Fig. 1. A concrete form of feature model [4]

The concrete from of feature model in Fig. 1 identified NFRs as an integral part of feature model. Corresponding to the concrete form in Fig. 1, Mei et al further elaborate the constructs of feature model in [7] by present a meta-model of feature model, as shown in Fig. 2, which consists of three basic concepts: feature, refinement, and constraint, deriving from Classifier and Relationship in UML Core Package.

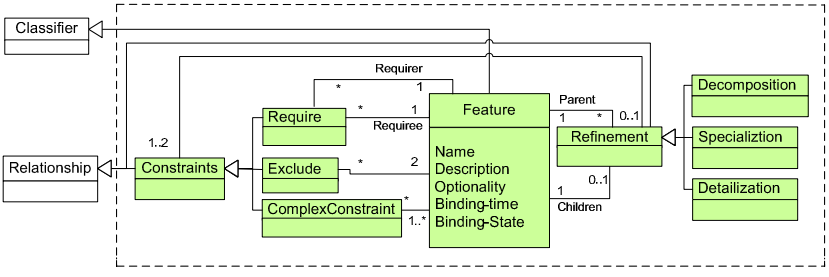


Fig. 2. A meta-model of feature model [7]

Feature is characterized by five attributes: name, description, optionality, binding-time, and binding-state. *Name* is a short description of feature. *Description* is a detailed specification of feature and is often specified in natural language to make it easy for various stakeholders to understand what a feature is about. For FRs, that works well because that a feature's name and description can clearly describe the meaning of a FR. *Refinement* is for structural dependencies among features. A feature can be decomposed into its constituent features, specialized as features incorporating further details, or characterized into its attribute features. Refining FRs often results in a SPL model enabling fine-grained customization in deriving new product and forms a hierarchy structures. Hierarchy structures is effective in describing complex systems, since it is easier to understand a complex system from general to specific and from high levels to low levels. *Constraint* is for capturing dependencies between the binding-state of features. For two features A and B, A require B means A should be work together with B at business level; A exclude B means A is not permitted to work together with B at business level.

3.2 A Running Example

We use an e-shop product line as an example to illustrate our approach due to the plethora of various e-shops deployed on the Internet. An e-shop is a web-based application selling goods to customers. Customers can order goods on e-shop. Employees of e-shop are responsible for processing the orders, monitoring e-shop operation and configuring operation policies.

Fig. 3 is a feature model segment of e-shop SPL. Feature *ordering* means that customer can create, modify, and send a purchase order to e-shop, and pay for the order, and that the system can set default value and compute price of the order. Optionally, the sent order or modified order can be automatically saved. Feature *Configuring* states the functionalities of analyzing orders, defining operation policies of the online store, managing the way of payment and the responsibilities of employees. Feature *OrderingNFRs* and *ConfiguringNFRs* are NFRs constraining the feature *Ordering* and *Configuring*.

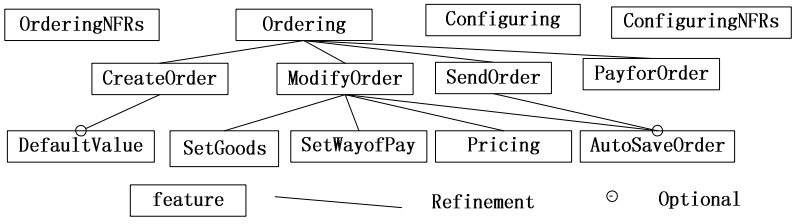


Fig. 3. Feature model segment of E-shop SPL

3.3 Deficiencies of Feature Model in Modeling NFRs

We further reveal the deficiencies of feature model in modeling and managing NFRs.First, modeling NFRs as ordinary features often makes it difficult to understand, manage, and customize complex NFRs during the derivation of new product from SPL, and the evolution of SPL. In fact, NFRs are actually constraints on FRs. The description of NFR features often interweave with descriptions of other features modeling either NFRs or FRs. For example, the feature OrderingNFRs in Fig. 1 may state that "registered customers need to process order via a usable interface and to send order and pay in a secure and reliable way." In real project, developers often need to scrutinize the whole complex feature model to ensure that all NFRs and FRs affected by NFRs are completely identified and correctly processed. When the feature model evolves, the whole process has to restart again.

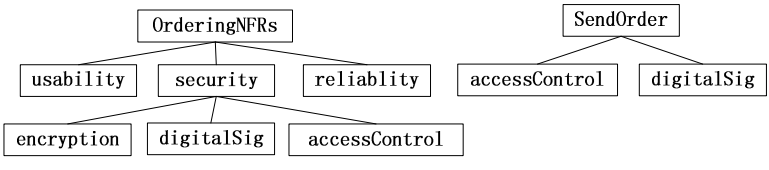


Fig. 4. Feature model segment refined from Fig. 1

Second, NFRs can also be refined. For example, as shown in Fig. 4, feature *OrderingNFRs* in last paragraph can be refined as feature *usability* (usable ordering interface), *security* (registered customer can process orders, and send and pay order in a secure way), and *reliability* (send and pay order in a reliable way), and *security* can be in turn refined as *encryption*, *digitalSig*, and *accessControl*. Obviously, we can formulate a hierarchy structure in Fig. 4 to organize all NFRs constraining a specific FR. Each hierarchy structure for NFRs may cover various types of NFRs and tactics

implementing the NFRs. That will result in a very complex feature model which is very difficult to manage and use.

In addition, without explicitly differentiating NFRs and FRs, simultaneous refinement along with NFRs and FRs may introduce ambiguity, especially when feature model is created and maintained by multiple persons. For example, *accessControl* in Fig. 4 may be refined from a functional feature *SendOrder* to denote that *SendOrder* can only be accessed by legitimate users (registered customers) or from a *security* feature constraining ordering to denote that access control is a means of achieving securely ordering. In this case, tailoring *accessControl* feature in DFM may introduce ambiguity in AFM.

Furthermore, some dependencies among NFRs and FRs cannot be model by refinement relationships. 1) Each NFR constrains specific FRs. For example, both *security* and *reliability* constrain functions of "send and pay order". 2) Some NFRs can be implemented as FRs, such as access control for security. 3) NFRs may conflict or cooperate with each other. For example, replication mechanisms implementing reliability often impose performance overhead; it is necessary to encrypt data first, then pass data via reliable channel, and finally decrypt data.

Third, there are no special requirements on modeling constraints among binding-state of NFR features. Sometimes, NFRs may *require* other NFRs when they need to be implemented simultaneously and work together to deliver better users experiences. For example, employing access control to protect sensitive information *requires* ensuring the confidentiality of meta-data of access control, such as user identity and password. On the other hand, NFRs rarely *exclude* other NFRs. For example, although reliability may conflict with performance, both of them are still expected and need to be well-balanced in target system.

In summary, the feature model in Fig. 2 does not provide facilities for differentiating FRs and NFRs and modeling relationships among them, as well as facilities for controlling complexity of feature model.

3.4 Extending Feature Model to Model NFRs

To model NFRs in SPL, we extended the meta-model of feature model in Fig. 2 as shown in Fig. 5.

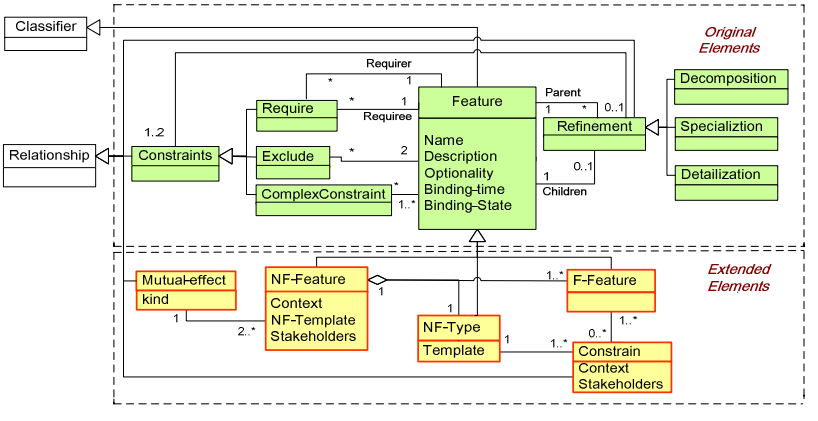


Fig. 5. An extended meta-model of feature model

First, we define two concepts *NF-Feature* and *F-Feature,* as the stereotypes the concept *feature* to explicitly model NFRs and FRs respectively. For example, feature *OrderingNFRs* in Fig. 3 should be modeled as a *NF-Feature* and *SendOrder* is a *F-Feature*. In this way, feature model can capture the expertise of NFRs experts on types of software requirements. That is helpful for mitigating the possible ambiguities in feature model.

Second, developers as well as other stakeholders often talk about NFRs using these well-known terms, such as security, reliability, usability and so on. These terms are actually used as umbrella to cover common semantics of a group of similar NFRs and common means to achieve these NFRs, such as access control for security, layered architectural style for maintainability and reusability. These terms are often called NFRs types. They are independent from specific application and are actually knowledge reusable in developing different applications. NFRs of a specific application are compositions of these types with concrete FRs in specific context [1][8]. Obviously, using NFR types to refer to concrete NFRs specific to application will confuse the semantics of NFRs in real development.

With these insights, we define a concept *NF-type* as the stereotype of the concept *feature* to model types of NFRs, and explicitly define a new relationship *constrain* among *NF-type* and *F-Feature* to enable the developers to explicitly composing them together as *NF-Feature*. To be noticed, NF-types can be refined into other NF-types but never *require* or *exclude* with each other.

For the newly defined concepts, we also define a series attributes besides those attributes inherited from the concept *feature*. *NF-Type* has an attribute *NF-Template,* which includes actually a set of attributes and differs from one *NF-type* to another. For example, *NF-Template* of *security* may include subjects and certificates that are allowed to access specific FRs, keys and algorithms used in encryption or digital signature while *NF-Template* of performance may includes the allowed interval of response time of specific FRs, or throughput of a specific business transaction. To achieve flexibility, we do not define a general template for all NF-types but leave the selection and customization of the template to developers. *Constrain* among NF-type and *F-Feature* has an attribute *Context*, which describes the context in which software systems are required to satisfy given NFRs. For example, e-shop may run in a context (normal) in which accesses to the e-shop retain in a relatively low and stable level, or in a context (peak) in which accesses to the e-shop retain in a high level. Another attribute of *Constrain* relationship is *Stakeholders*, which refers to stakeholders who may be affected by the degree to which the composed NFRs are satisfied, such as users, developers.

When composing *NF-Type* with *F-Feature* via *Constrain* relationships, the attributes *NF-template* of a *NF-Type*, *stakeholders* and *Context* of *Constrain* will be automatically integrated into the resulted *NF-Feature*. Each *NF-Type* will associate with one or more *F-Features* via *Constrain* relationship, and each *NF-Feature* consists of one *NF-Type* and one or more constrained *F-Features*.

At last, we also introduce a new relationship *Mutual-effect* among NF-Features to denote the possible relationships among NFRs. The attribute *kind* of *Mutual-effect* denotes whether the *NF-Feature*s do harm to each or not and has two possible value as *confliction* or *cooperation*.

The extended feature model can be employed to explicitly model NFRs and FRs, as well as the relationships among them. However, the complexity of feature model is still not mitigated and complexity management facilities are required to make the approach practicable.

3.5 Facilities and Techniques to Control Complexity

We introduce five views to separate concerns on complex feature model, *FR view, NF-type view, composition view, NFR view*, and *correlation view*. Multiple views in combination of automated facilities can effectively control the complexity of feature model.

FR view consists of functions originated from business requirements, such as *ordering* and *configuring* in e-shop. FR view forms a basis for interaction between developers and business users of the software systems. Business users are free from abstruse NFRs and their implementing tactics.

NF-type view consists of NF-types that encapsulate expertise knowledge of NFRs, including well-known words delivering common understanding of NFRs, such as security, performance, and reliability and so on, as well as metrics and tactics to achieve these NFRs. This view actually captures reusable knowledge across products or even product line.

Composition view captures associations among hierarchy structures of functional features and hierarchy structures of NF-types. Developers can elicit NFRs by checking whether each functional feature should be constrained by NF-types or not in some contexts, and specify the value of attributes of NF-template of NF-type, such as subjects of access control and response time. By associating a NF-type to a functional feature, subordinate features of them are implicitly associated together. In this way, developers are relieved from defining voluminous NFRs one by one.

NFR view is a view that is automatically generated from composition view and needs to be customized by developers when necessary. Each association among NF-types and a functional feature in composition view is transformed into a hierarchy of NFRs which are combination of NF-types, functional features, and contexts. By this hierarchy, developers can customize fine-grained NFRs.

Correlation view is a view that is automatically generated from composition view and NFR view and needs to be customized by developers when necessary. It captures interactions among NFRs attached to one or more FRs. Correlation view is the foundation of trade-off between different architectural alternatives.

4 Discussion and Related Work

Existing work focused on the structure and constraints aspects of feature model as well as the validation of feature model while pay little attention on how to model and manage NFRs in feature model[3][4][6][7]. One exception is the work of Peng et al.[5]. Peng's work implicitly assumed that NFRs of SPL have been modeled as NFR graph built in NFR framework [1] and focused on how to identify nonfunctional variability from NFR graphs. However, Peng's work did not clearly differentiate and model the relationships among features and goals representing functional features, and soft-goals representing NFRs. That may lead to confusion when models evolved.

In contrast, we do not assume the adoption of any specific NFR modeling approaches for single products, but focus on building a general framework that is extensible enough to incorporate various NFRs in feature model. We model NFRs in only primitives of feature modeling, avoiding the unnecessary efforts of matching concepts from different requirements modeling approaches.

Our previous work [9] has explored the possibility of extending feature model to model access control requirements. In this paper, we propose a more general solution for modeling various NFRs in feature model.

5 Conclusion

This paper analyzes the necessity and challenges of modeling NFRs in SPL and presents a feasible solution. First, we explicitly differentiate NFRs and FRs and enable the refinement and customization of NFR features. Second, our approach provides multiple views and several automated facilities to control complexity of feature model with NFRs. Finally, we use hierarchy of NF-type to model knowledge common to NFRs in different context and applications. Developers can even to define a new NF-type to codify their knowledge in a specific domain. In this way, our approach is flexible enough to effectively use and codify different domain knowledge. In summary, our approach treated NFRs as first class citizen in SPL and can be used to enable the reuse of NFRs related assets in a more disciplined way than ever. Our future work is to extend our existing feature modeling tool to implement our approach and to apply our approach in more real cases.

Acknowledgments. This work is sponsored by the Research Foundation of Shannxi University of Science and Technology of China No. BJ09-13 and the Research Foundation on Opening Issues of Key Laboratory of High Confidence Software Technologies of Ministry of Education of China No. HCST201004.

References

1. Chung, L., Nixon, B.A., Yu, E., Mylopoulos, J.: Non-Functional Requirements in Software Engineering. Kluwer Academic (2000)
2. Barbacci, M.R., Ellison, R., Lattanze, A.J., Stafford, J.A., Weinstock, C.B., Wood, W.G.: Quality Attribute Workshops (QAWs) Third Edition. Technical report CMU/SEI-2003-TR-016, Software Engineering Institute Carnegie Mellon University (2003)
3. Pohl, K., Böckle, G., van der Linden, F.J.: Software Product Line Engineering: Foundations, Principles and Techniques. Springer, New York (2005)
4. Hong, M., Wei, Z., Fang, G.: A Feature Oriented Approach to Modeling and Reusing Requirements of Software Product Lines. In: 27th Annual International Computer Software and Applications Conference, Compsac, p. 250 (2003)
5. Peng, X., Lee, S.-W., Zhao, W.-Y.: Feature-Oriented Nonfunctional Requirement Analysis for Software Product Line. Journal of Computer Science and Technology 24(2), 319–338 (2009)
6. Kang, K., Cohen, S., Hess, J., Nowak, W., Peterson, S.: Feature-Oriented Domain Analysis (FODA) Feasibility Study. Technical Report CMU/SEI-90-TR-21, Pittsburgh, PA, Software Engineering Institute, Carnegie Mellon University (November 1990)

7. Zhang, W., Zhao, H., Mei, H.: A Propositional Logic-Based Method for Verification of Feature Models. In: Proc. of 6th International Conference on Formal Engineering Methods, ICFEM 2004, Seattle, WA, USA (2004)
8. Dörr, J., Kerkow, D., von Knethen, A., Paech, B.: Eliciting Efficiency Requirements with Use Cases. In: Proceedings of the International Workshop on Requirements Engineering: Foundations of Software Quality (2003)
9. Sun, L., Huang, G.: Modeling Access Control Requirements in Feature Model. In: Proceedings of 16th Asia-Pacific Software Engineering Conference, pp. 241–248. IEEE Comput. Soc. (2009)

A Case Study on High Trustworthy Environment Construction for Smart Home Applications

Xinyan Gao[1,⋆], Xianlu Luo[2], Ning Zhou[3], and Dakui Li[1]

[1] School of Software of Dalian University of Technology, Dalian City, Liaoning, China
[2] Department of Computer, Neusoft Vocational College of Information Technology, Nanhai, Guangdong, China
[3] School of Computer and Information Technology, Beijing Jiaotong University, Beijing, China
Tyler.gxy@hotmail.com, luoxianlu@21cn.com

Abstract. Nowadays, trusted computing plays an important role in improving information security and building trusted environment for network related applications. In this paper, an effective implementation of smart home based on existing trusted hardware components and software stacks is studied. We investigate the digital certificate generation and its use, key hierarchy and its establishment, software integrity measurement, firmware upgrading and default recovery processes, and related key steps. We present a practical establishment for a high trustworthy environment over smart home applications by using TPM. The case study aims to act as a useful supplement and practical solution reference for trusted system deployment.

Keywords: Trustworthy Environment, TPM, Smart Home, Trusted Computing.

1 Introduction

With the rapid advances in wireless communication, embedded system, sensors network, and other technologies, the concept of **Smart Home** [1] proposed in the end of last century is becoming much closer than ever to reality. The housing-based ubiquitous computing [2,7] plays an important role in the planning of future housing-based models of living environment. More and more companies, labs, institutions and research groups are working on this domain.

In such ubiquitous computing environment, security has become an important problem and should be paid more attention in practice. Thus, more attention will mainly be paid on smart home networks security currently, in particular,

⋆ The project is supported by "the Fundamental Research Funds for the Central Universities" (1600-852014) and partly supported by the National Natural Science Foundation of China under Grant No. 60873118 and 60973147 and partly supported by Grants (HCIC201102) of Guangxi Key Laboratory of Hybrid Computational and IC Design Analysis Open Fund.

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 755–763.
springerlink.com

the mechanics based on hardware for trust establishment and its application scenarios.

As well known, **Trusted Computing** [4,11] is an emerging paradigm that addresses information security concerns in a wide variety of computing systems. Trusted computing standards are driven by the computing and communications industries through the TCG (Trusted Computing Group). Trusted Computing uses a hardware device, a separate trusted coprocessor named Trusted Platform Module (TPM, for short), to provide encryption key generation, encryption key storage, tamper proof, secure reporting, and other services. Additionally, TPM also provides tamper-resistance techniques to prevent a wide range of physical and hardware-based attacks.

In the smart home project, we proposed and implemented a trustworthy environment solution based on TPM hardware coprocessor which acts as a trusted root for the whole system that contains it.

In this paper, we will mainly discuss the system architecture of project, key security solution techniques based on TPM hardware and related processes of manufacturing, authentication, upgrading and other auxiliary work flow.

2 System Architecture and Security Requirements

In this section, we will firstly illustrate the high level overview of our smart home project architecture. Then, we will discuss the actual security requirements.

2.1 System Architecture

In smart home, the system consists of five major components: (1)an external network, (2)a gateway, (3)services, (4)a local network, (5)devices or household appliances.

A sensor network based on ZWave[6] or Zigbee[5] connects all key electrical appliances and intelligent devices, and allows them to be remotely controlled, monitored or accessed. Additionally, a data/communications network connects all computation devices and cooperates with the previous sensor network seamlessly.

As shown in Fig.1, the SOC processor is based on ARM (Advanced RISC Machines) core with the 32KB data/instruction cache. Peripheral interfaces and devices mainly include the following components:

- one USB 2.0 high speed controller interface with two ports
  - port 1# connected with 3G module (HSUPA, High-Speed Uplink Packet Access)
  - port 2# connected with a high speed USB 2.0 hub
- three UART interfaces connected with
  - "ZWave" sensor network server controller
  - "Zigbee" controller
  - console

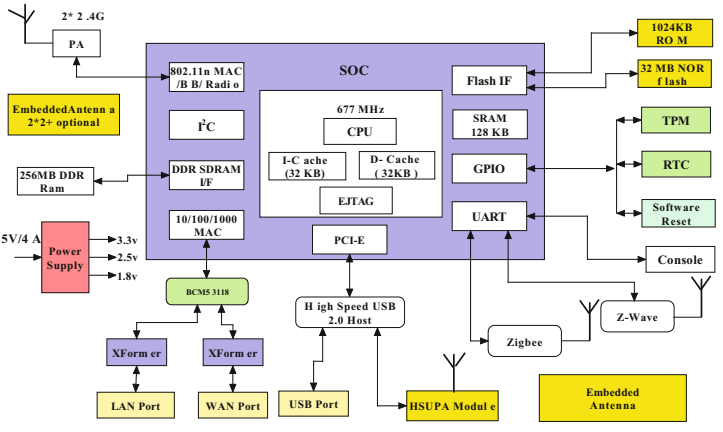


Fig. 1. Smart Home Platform System Architecture

- two flash memory interfaces connected with
  - NOR flash memory with 32MB capacity
  - readonly flash memory for CRTM(Core Root of Trust for Measurement)
- I^2C bus interface connected with TPM chip
- eight ports ethernet switch named "BCM53118" for Local Area Network
- one ethernet port for Wide Area Network
- one 802.11n MAC for 2.4GHz radio frequency device

In this project, a TPM chip (AT97SC3204C, from "Atmel" company) is integrated and used as secure key device to protect platform and applications. CRTM is saved in a special flash room that will not be changed once written. BOOTLoader is stored in a partition of normal NOR flash memory and can be upgraded when needed.

Our project implementation is done on an embedded Linux distribution running 2.6.29 kernel. Note that, Linux kernel and application layer components reside on external storage(NAND flash memory on SD card). We modify Linux kernel in order to measure and verify the integrity of binary executables and libraries as soon as they are loaded.

2.2 Security Requirements

In this section, we will discuss the concrete security requirements under smart home application environment.

Clearly, from the system description mentioned previously, smart home is a comprehensive highly integrated area which consists of a large variety of content sources (sensors), multiple information carriers (wired and wireless media) and communication standards and can provide a lot of humanized services, such as smart information management, personal application, home entertainment, remote medical, remote communication, remote control, and etc.

One of the smart home application scenarios is where users are interacting with their home banking or making transaction via web service, it is essential to make sure whether the service has been tampered with or not.

Another scenario, in no distant future, the **Internet of Things** (IOT, for short) will be used throughout the country. On one hand, the IOT is the extended network based on the Internet and smart home applications may benefit greatly from trusted computing; On the other hand, the flaw and imperfect software of critical technologies may cause serious secure problems. Evidently, any of these issues should deserve more focused attention.

Therefore, in this paper, we will focus on the issues regarding the construction of trustworthy environment for smart home applications and emphatically discuss how we implement such an environment and other related processes.

3 Trustworthy Environment Construction

In this section, we will discuss the construction process for trustworthy environment. Five key aspects are involved, which are digital certificate generation and its use, key hierarchy and its establishment, software integrity measurement, software upgrading and default recovery.

3.1 Using Digital Certificate

In this section, we will discuss how digital certificate (DC, for short) is used for smart home applications.

A digital certificate is the combination of a statement and a digital signature of the statement that can establish credentials when making electronics transactions over internet. The well known digital certificate is the $X.509$ [9] public-key digital certificate adopted in the project.

The signer of the digital signature is normally a trusted certificate authority (CA, for short). The X.509 public-key digital certificate has been widely used in public-key infrastructure (PKI, for short) to provide authentication on the users public key contained in the certificate. The user will be authenticated if he can prove that he has the knowledge of the private key corresponding to the public key specified in the $X.509$ public-key digital certificate. There are three entities in a digital certificate application, which are certificate authority, owner of a DC, and verifier respectively.

The issuing process of DC is described as follows: the user creates a key pair firstly, then he creates CSR (Certificate Sign Request) with public-key and own information, and sends it to CA. After receives the CSR, CA will check it and create the corresponding certificate. In the end, CA will send the produced certificate to the user.

The following is an actual information flow between a publisher's server and a user's computer in using digital certificates, as shown in Fig.2.

(1) The client attempts to access a controlled resource from a publisher, such as a database or digital library, usually through a Web interface.

(2) The publisher's server then asks the client to provide a certificate.
(3) The client presents a certificate, and the publisher's server verifies that the certificate is issued by a recognized certificate authority, asserts that the holder is a member of a licensed institution, and has not been revoked.
(4) The publisher extracts a URL from the certificate, which provides the means to retrieve from the campus or library additional information (attributes) needed for authorization decisions.
(5) The publisher then connects to the specified attribute server using the prescribed secure protocol and presents its own X.509 certificate to establish the secure connection. The attribute server verifies whether the publisher's certificate is valid and uses the publisher's identity to determine the access permissions by using the information from the directory service.
(6) The attribute server executes the related query. The query result is presumed to be a list of attribute name-value pairs, including the service type or access authorized for the individual. A list of the results is returned to the publisher.
(7) The publisher checks the value(s) of the "Service Class" attribute. If at least one value is valid for the publisher and service requested, the user is granted access.

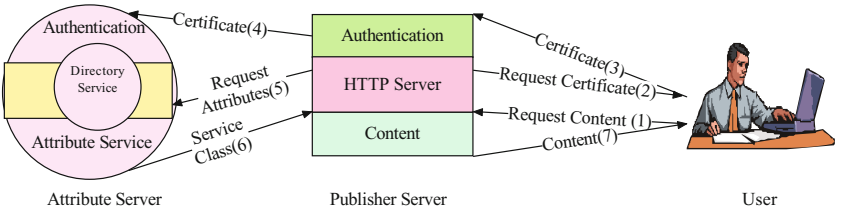


Fig. 2. Authentication Process

3.2 Key Hierarchy and Its Establishment

In order to implement secure protection functionality and perform other trust based operations, some important information generated during system running should be protected via keys by using TPM.

In TCG, seven key types are defined, which are *Signing keys*, *Storage keys*, *Identity keys*, *Endorsement keys*, *Bind keys*, *Authentication keys*, and *Legacy keys* respectively. Since all keys are not independent with each other, in this project, we organized all the keys by their relationships (parent key and leaf key) to form a treelike logical architecture, as shown in Fig.3.

From this figure, we have:

(1) The root key-SRK (Storage Root Key) must be created firstly when manufacturing via TPM as all keys' root; other keys can be created following it.

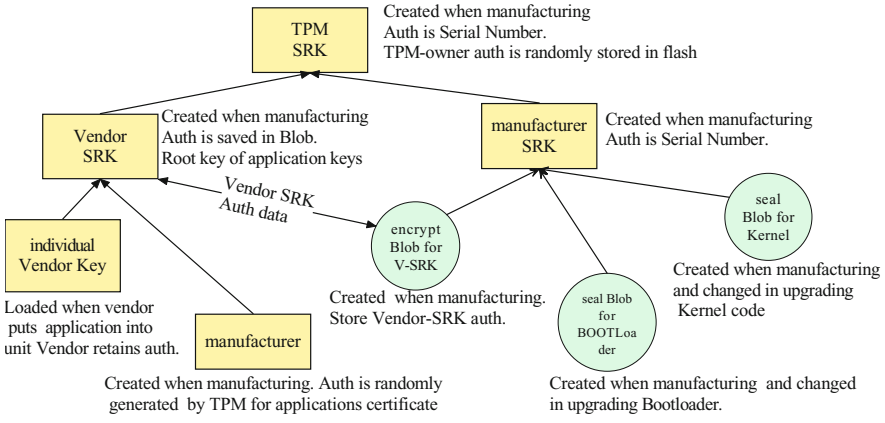


Fig. 3. Key Architecture

(2) Manufacturer SRK is also created when manufacturing as a son key of SRK, which will take charge of protecting platform information created during manufacturing, such as a Seal-Blob for bootloader.
(3) Vendor SRK, a son key of SRK is also created when manufacturing and it creates all keys for applications.
(4) Certificate key is generated via TPM when manufacturing which is under V-SRK. Certificate key's information can be directly stored in flash. All these keys can be used through TSS (TCG Software Stack)[10].

3.3 Measurement

Integrity measurement[12,8] is a process to obtain metrics of platform characteristics that affect the platform integrity (trustworthiness) and put the metrics digests into PCRs of TPM.

In our project, the CRTM is developed from Redboot[3], which is customized on demanded. TCG software stack has been ported to both Redboot and CTRM packages to implement the measurement functionalities.

For any file, regardless of its type, either executable code or sensitive data, we can compute a $SHA1$ hash over the complete contents of the file and get the 160bit hash value that unambiguously identifies the file's contents. A measurement code generates measurement events including two classes of data:

1) Measured value - a representation of embedded data or program code segment;

2) Measurement digest - a hash1 (SHA-1) of those values.

Algebraically, a PCR updating is described as the following:

$$PCR[n]_{new} = SHA-1(PCR[n]_{old}||SHA-1(code)). \quad (1)$$

The content to be measured is scanned by the measurement code to generate a message digest. The two data elements are stored separately. The measurement

digest is stored in the PCR of TPM, while measured values may be stored virtually anywhere(in this project, it is saved in a flash partition). Other important information, such as system profile, configuration data, will be saved in different partitions.

From above discussion, we have the following process pseudo-code description of the integrity measurement process.

Process.01: "Integrity Measurement Process"
BEGIN

```
   /*    Step 1: System is powered on and starts up*/
01    Powered On;
02    System Initialization(CPU, Cache, DDR, GPIO, ...);
03    TPM starting up and TPM driver Initializing;
04    TPM_DisableForceClear();
05    TPM_Measure(Bootloader, SN) with PCR0; /* SN: Serial Number */
06    TPM_Unseal( PCR0 );
07    if(success! = true){
08          return false;
09      }
   /*    Step 2: Measure Bootloader ; */
10    Bootloader Starting Up and TPM drivers Initializing;
11    TPM_Measure( Kernel ) with PCR1;
12    TPM_Unseal( PCR1 );
13    if(success! = true){
14          return false;
15      }
   /*    Step 3: Measure Linux kernel and other modules; */
16    Linux Kernel Starting Up and TPM drivers Initializing;
17    Mount Linux File Systems
18    while(MeasureList! = Empty){
19        file = getlist(MeasureList);
20        TPM_Measure( file ) with PCRn;
21        TPM_Unseal( PCRn );
22        if(success! = true){
23            return false;
24        }
25    }
26    return true;  /*   to execute applications; */
```

END;

The starting point of measurement is the CRTM. CRTM can perform tasks such as, initialize hardware, start-up TPM, measure BOOTLoader code and serial number, extend the digests to PCR0, and unseal a Seal-Blob in flash, and etc. After hardware is powered on and initialized, TPM will start up firstly and disable "forceclear" command. Note, CRTM is the root of the chain of transitive trust and absolutely trusted, which need not be measured at all. Once CRTM was programmed when manufacturing, it can not be changed or upgraded any more. Since CRTM should be protected by hardware, it is saved in secure boot ROM (a readonly flash memory chip) in this project.

3.4 Software Upgrading

Software upgrading means that common user can update their system with latest version firmware from the vendor by general interfaces. As we all know, to ensure the trustworthy property holds, upgrading process should be performed in a trustworthy environment and the upgraded system should still keep trustworthy.

Therefore, before upgrading, a checking process will be carried out firstly. Then, when upgrading, a new seal-blob for the new firmware will be generated and overwritten to the old one. In our project, in order to support rolling back, we implemented a set of loop queues for software upgrading. The data structure of the queue is defined as the following.

```
class UpgradingQueue{
    uint8 type; /* bootloader, kernel, ... */
    int version; /* current version number */
    Blob blob*; PCR pcr*; uint8 blobValid;
    TimeStamp time; }
```

The upgrading operation sequence can be illustrated as the following: (1) Start up the system and enter upgrading interface; (2) Ask authentication to the vendor to access the new firmware; (3) Download the new firmware from vendor and save it as a file to local file system; (4) Reset TPM-chip and clear the related PCR value; (5) Hash the new firmware file content into the PCR; (6) Create a new blob locked to the new PCR value and overwrite the old one or insert it to upgrading queue; (7) Reboot the system and finish upgrading.

3.5 Factory Default Recovery

There are a lot of scenarios where the user need to recover the smart home system to factory default status. Additionally, the user need to roll back to the previous configurations when upgrading fails.

To this end, we have saved the backup information for BOOTLoader and kernel codes when manufacturing. When doing recovery, we just overwrite the firmware space and corresponding Seal-Blob from backup space to TPM data space directly. For rolling back, user will firstly select the previous version to be recovered by $UpgradingQueue :: version$, then a backup process for this version will be performed.

4 Conclusion

In this paper, we discussed how the trustworthy environment is constructed and presented a practical implementation case of the TPM based secure system over smart home project. In our design, the measurement system is highly extensible and the secure mechanism can perform the measurement of BOOTLoader, kernel, kernel modules and sensitive data in addition to executed files. We have applied our trust measurement architecture to smart home applications and shown how our system detects undesirable invocations and measures the system

integrity. This case ultimately achieved a high trust guarantees. Hence, we can conclude that the TPM based implementation will evidently be efficient and trustworthy with recent devices so that other groups will be greatly motivated to build secure systems by applying this idea.

We are now developing the integrity measurement program for high level applications based on our current solution. Part of future work is to explore the encryption and decryption functionality of the trusted chip for high-level applications. Additionally, TCG-based trusted computing system development in sensor network environment and related experimental results analysis will be performed in the near future.

Acknowledgments. The project is supported by "the Fundamental Research Funds for the Central Universities" (1600-852014) and partly supported by the National Natural Science Foundation of China under Grant No. 60873118 and 60973147 and partly supported by Grants (HCIC201102) of Guangxi Key Laboratory of Hybrid Computational and IC Design Analysis Open Fund. The authors would like to thank their colleagues for participating in the research. They also appreciate the anonymous reviewers for their helpful comments.

References

1. Valtchev, D., Frankov, R.: Service gateway architecture for a smart home. IEEE Communications Magazine 40(4), 126–132 (2002)
2. Yamazaki, T.: Ubiquitous Home. International Journal of Smart Home 1(1), 17–22 (2007)
3. http://www.cygwin.com/redboot/
4. Felten, E.W.: Understanding Trusted Computing. IEEE Security & Privacy 1(3), 60–66 (2003)
5. ZigBee Alliance, ZigBee Specification (2006)
6. Zensys, SDS10242, Software Design Specification, Z-Wave Device Class Specification
7. Kuntze, N., Rieke, R., Diederich, G., Sethmann, R., Sohr, K., Mustafa, T., Detken, K.-O.: Secure mobile business information processing. In: 2010 IEEE/IFIP 8th International Conference on Embedded and Ubiquitous Computing (EUC), pp. 672–678 (December 2010)
8. Dietrich, K., Pirker, M., Vejda, T., Toegl, R., Winkler, T., Lipp, P.: A practical approach for establishing trust relationships between remote platforms using trusted computing. In: Barthe, G., Fournet, C. (eds.) TGC 2007. LNCS, vol. 4912, pp. 156–168. Springer, Heidelberg (2008)
9. Network Working Group, Internet X.509 public key infrastructure certificate and crl profile, RFC: 2459 (January 1999)
10. Trusted Computing Group. TCG Software Stack Specification (TSS) Version 1.2, Level 1, Errata A (March 2007)
11. Trusted Computing Group, Trusted Platform Module Main Specification, Part 1: Design Princi ples, Part 2: TPM Structures, Part 3: Commands, Version 1.2, Revision 62 (October 2003), http://www.trustedcomputinggroup.org
12. Sailer, R., Zhang, X., Jaeger, T., van Doorn, L.: Design and Implementation of a TCG-Based Integrity Measurement Architecture. In: Thirteenth Usenix Security Symposium, pp. 223–238 (August 2004)

A Modbus Protocol Stack Compatible with RTU/TCP Frames and Embedded Application

Jian Kuang, Guibao Wang, and Jiali Bian

Beijing Key laboratory of Intelligent Telecommunications Software and Multimedia,
Beijing University of Posts and Telecommunications, Beijing
{jkuang,wangguibao,jlbian}@bupt.edu.cn

Abstract. Modbus communications are carried out by different frame formats on different physical links. As a result, to use Modbus on two distinct links, it's common to design two sets of protocol stack software. In this article a Modbus protocol stack compatible with RTU and TCP frames is proposed, which can be adopted on resource-limited micro controllers to greatly reduce software size, as well as speeding the system performance. Detailed layered model, as well as function descriptions and flow charts for each layer are given. In the end, we will examine an application case based on ST Semiconductor's STM32 micro controoler. The result shows it achieves the design expectation.

Keywords: Modbus, RTU, TCP, STM32.

1 Introduction

Modbus is an application layer controller communication protocol, which is mainly used in master-slave mode device monitoring and management[1]. On a physical bus there can be at most 255 slave hosts, each listening on the bus for data frames from the master host and determining whether to accept the frame or not via the address code field in the frame. The design has made it not possible for very high data rate but very suitable for remote monitoring. Designed to be work on serial lines in the early days, Modus has now been adopted to also work with TCP/IP protocol, or Modbus over TCP/IP. However, because of the differences between the two frame formats caused by design background (data rate, error control, etc), it is hard for an application to benefit from the two at the same time. It will be of practical use for microcontrollers with limited resource to have a protocol stack compatible with both frame formats.

2 Differences between Modbus RTU and Modbus over TCP/IP

Before addressing the problem, we have to begin with a brief introduction to these two frame formats and comparison, leading us to a reasonable solution.

Further more, there're two transmit modes for Modbus over Serial Line, namely RTU mode and ASCII mode. And because the former is more widely used in practice, we will concentrate on RTU mode in this paper.

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 765–770.
springerlink.com

2.1 Modbus RTU Frame Format

The Modbus RTU format is illustrated in the bellow figure[1]:

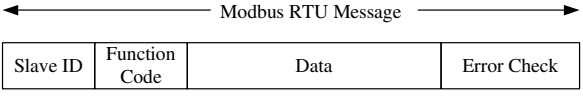


Fig. 1. Modbus RTU frame format

The meanings for each field are:

Slave ID: 1 byte, slave polling address. Slave host having this ID will accept messages from master host; Each slave on the bus has an unique ID ranging from 0 to 255 (0 is the broadcast address, 1 - 247 valid addresses, 248 - 255 reserved addresses)

Function Code: 1 byte. This field is used by master to tell slaves what operation to perform, ranging from 1 to 255

Data: N * 1 bytes, the actual data length and content is determined by Function Code

Error Check: 2 bytes, frame CRC checksum sequence, used by message receiver to verify if has any error

2.2 Modbus over TCP Frame Format

Modbus over TCP frame format is illustrated in the figure bellow[3]:

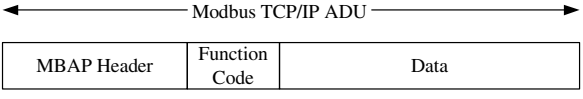


Fig. 2. Modbus over TCP frame format

The meanings for each field are:

MBAP Header: Modbus Application Protocol Header, 7 bytes long, used to mark a Modbus frame over TCP/IP protocol stack

Function Code: 1 byte, used by master to tell the slave what operation to perform, ranging from 1 to 255

Data: N * 1 bytes, the actual data length and content is determined by Function Code

2.2.1 MBAP Header

MBAP header is used to mark a Modbus frame on a TCP/IP network, which has the following structure:

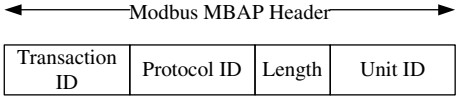


Fig. 3. Modbus MBAP header

The meanings for each field are:

Transaction ID: 2 bytes, Modbus request/response transaction ID

Protocol ID: 2 bytes, 0 = Modbus protocol

Length: 2 bytes, byte count of the rest of the frame not including the length field itself

Unit ID: 1 byte, the identification code for a remote slave on the bus (might be a serial line or TCP/IP network)

2.3 Major Differences between the Two Frame Formats

From the above brief introduction we can tell that in most part the two formats are the same, except:

1) The address in the RTU format is a 1-byte Address field, while in Modbus over TCP it is a 7-bytes MBAP Header

2) There's a 2-bytes frame check sequence in RTU frames

Inspired by the layered protocol model in TCP/IP, a new layer is introduced to address the differences, then pass the common parts upward for further processing.

3 A Modbus Protocol Stack Model Compatible with Both Frame Formats

The new model is illustrated as below:

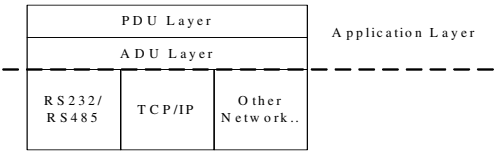


Fig. 4. A Modbus protocol stack model compatible with Modbus RTU and Modbus over TCP

In this model, the first layer resembles functions of layers under the application layer in the OSI model, while layer 2 and layer 3 are both in application layer. Layer 2 is called ADU layer, while layer 3 the PDU layer. Modbus frames from both serial lines and TCP/IP networks are passed upward from layer 1 to the ADU layer. The ADU layer is responsible for extracting the header information from the frame source, possibly making some change to the frame, then pass it to the PDU layer. In the other direction, application software data are passed to ADU layer after some kind of encapsulation by the PDU layer. ADU layer then prepends the frame with an appropriate ADU header according to the message destination in the PDU. For frames destined to serial lines, ADU is also responsible for calculating the FCS. Finally the response frame is passed to respective transmit media.

3.1 ADU Layer

The ADU layer is responsible for extracting the PDU from Modbus frames came up from different transmit media then pass it to the PDU layer. In the downward direction, ADU packs the PDU with appropriate header information according to message destination. For messages destined to serial lines, ADU will also calculate the FCS and append to the end of the frame. At last the packed frame is passed down for transmitting.

We are now checking the work ADU will do in each direction in detail.

In the upward direction:

1) If the frame is coming from TCP/IP network, jump to step 4)

2) Checks if the Address field equal to local address. If not, the frame is dropped, the processing is finished

3) Do checksum on the entire frame, if success, pass the buffer beginning from the Address field, and subtracting the last two FCS bytes, to the PDU layer; else the frame is dropped, the processing is finished

4) Check for the MBAP header:

a. Copy and save the Transaction ID field

b. If the Protocol ID is not zero, the frame is dropped and the processing is finished

c. Check if the Unit ID is equal to local address. If not equal, the frame is dropped and the processing is finished

d. Pass the buffer beginning from the 1 byte after the Unit ID to the PDU layer

The upward processing path is illustrated as follow:

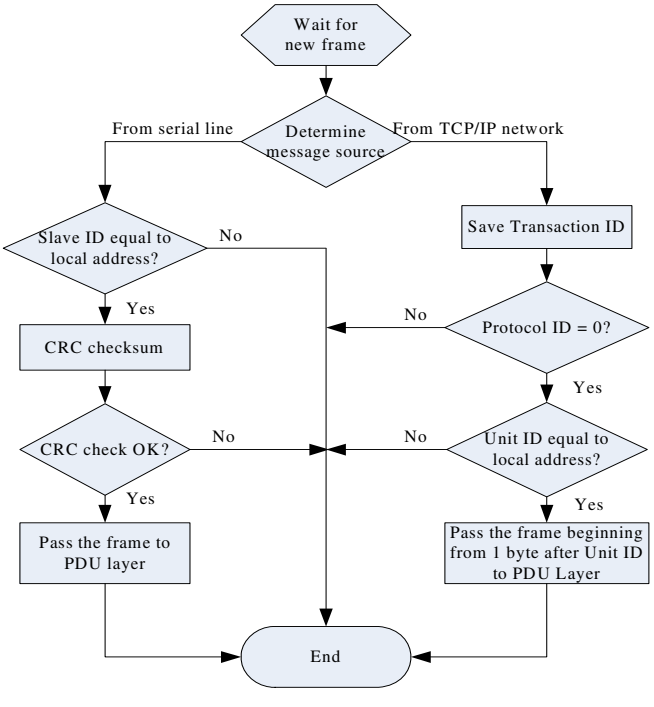


Fig. 5. ADU layer upward processing dataflow

In the downward direction:

1) If the message is destined to TCP/IP network, jump to step 3)
2) Prepend the PDU with a 1-byte local address, calculate the FCS for the entire frame and append to the end of the PDU, then jump to step 4)
3) Prepend the PDU with a 7-bytes MBAP Header
a. Transaction ID is a previously stored value
b. Protocol ID is 0
c. Length is the length of the PDU, plus 1 for the Unit ID
d. Unit ID is the local address
4) Pass to lower layer transmit media for transmission

This processing is simple so we are not giving a flow chart here.

3.2 PDU Layer

The major function of PDU layer consists of parsing the PDU according to the Modbus protocol, performing the requested function, possibly calling other application layer routines such as reading and writing coils, etc. It is also responsible for encapsulating the response frame and passes it down to the ADU layer.

4 An Application Based on STM32 Microcontollers

In a solar dish power system, dishes were mounted on motors to track the sun, so that solar energy utilization can be maximized. An ARM family microcontroller, STM32F107 from ST Semiconductors, was adopted to communicate with both remote monitoring host and the motor controller, both via Modbus protocol - with different underlying physical media: communication with remote host was carried on TCP/IP for long distance communication, while the motor controller only has serial line interface. STM32 is an ARM Cortex-M3 microcontroller, with 5 UARTs and one Ethernet controller, provided at a low price. On the other hand, the processor is only equipped with 64K on-chip memory and there's no way to expand it's external memory, making the software design a great challenge[4][5][6]. By adopting the stack model detailed in this paper, code size was reduced notably and finally fixed into the 64K memory.

5 Conclusion

In practice, much optimization can be taken by observing the protocol behavior. For example, in the ADU layer, very large buffer can be allocated during initialization time, so that after receiving a Modbus frame, a pointer into the large buffer can be directly passed to PDU. In this way, the PDU layer could process the frame in place, without needing to worry about memory overwrite. By doing this, one memory copy is saved, and ADU layer doesn't have to remember the address field for every newly incoming frame.

Both in theory and in practice, the Modbus protocol stack model compatible with Modbus RTU frames and Modbus over TCP/IP frames has been proved to be useful.

References

1. The Modbus Organization, Modbus Application Protocol Specification v1.1b, http://www.modbus.org/docs/Modbus_Application_Protocol_V1_1b.pdf
2. The Modbus Organization, Modbus over Serial Line Specification and Implementation Guide v1.02, http://www.modbus.org/docs/Modbus_over_serial_line_V1_02.pdf
3. The Modbus Organization, Modbus Messaging On TCP/IP Implementation Guide V1.0b, http://www.modbus.org/docs/Modbus_Messaging_Implementation_Guide_V1_0b.pdf
4. STMicroelectronics, STM32F10xxx Cortex-M3 programming manual, V3.0, http://www.st.com/internet/com/TECHNICAL_RESOURCES/TECHNICAL_LITERATURE/PROGRAMMING_MANUAL/CD00228163.pdf
5. STMicroelectronics, Connectivity line, ARM-based 32-bit MCU with 64/256 KB Flash, USBOTG, Ethernet, 10 timers, 2 CANs, 2 ADCs, 14 communication interfaces, http://www.st.com/internet/com/TECHNICAL_RESOURCES/TECHNICAL_LITERATURE/DATASHEET/CD00220364.pdf

Intelligent Education Services Based on Context Mediator in Pervasive Computing

Taijong Kim[1], Mincheol Kim[1], and DoHyeun Kim[2,*]

[1] Dept. of Management Information Systems,
Jeju National University, South Korea
jong4333@hanmail.net, mck1292@jejunu.ac.kr
[2] Dept. of Computer Engineering,
Jeju National University, Jeju, Republic of Korea
kimdh@jejunu.ac.kr

Abstract. In the pervasive computing environment, humans and devices or machines with computing ability become interoperable. In order to implement intelligent services in this environment, hardware technologies such as sensing and security are needed. However, the development of technology that enables devices to take contexts into account is also necessary. This paper proposes architecture of intelligent education services using the context mediator. This context mediator recognizes the situation of the user and provides the user with personalized service suitable for the user. This study is able to support students to more efficiently study on their own by using context mediator.

Keywords: Intelligent education, Customized pervasive services, Context awareness.

1 Introduction

A pervasive computing environment, one in which all objects are connected by means of networks, is seamlessly connected with the semantic web. The semantic web refers to the next generation intelligent web [1]. The semantic web is able to be easily understood by computers, meaning that these devices will be able to read and understand information from the web. In addition, these semantic web-enabled devices will be able to process and create new information which has been handled by humans.

In addition, every agent running on various clients can understand and process the information contained in the Semantic Web. Under such conditions, ubiquitous services are quite dynamic rather than stationary. A high level of interoperability should be offered to coordinate intricate relations between numerous objects, agents and devices. Also, resources and information objects should be processed flexibly according to the contexts. Therefore, what is needed are both agents that are aware of the context in which computing is carried out and semantic web technology through which the state of each device can be expressed semantically.

* Corresponding author.

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 771–777.
springerlink.com

Nowadays, context awareness technology is being actively researched throughout the world; however most of the research is focused on expressing context rather than on recognizing it. Providing services which users need has been considered a middleware issue and has received less attention. Some of the existing research of context-aware services includes the context broker system, which has been presented by many researchers as a framework where devices can interact with humans through context-aware computing in the ubiquitous environment. Other research in the field includes the 'Context Toolkit' [2], the 'context-awareness architecture' [3], and the 'Smart Room' [4] which are popular studies which suggest frameworks in which Pervasive Computing devices can interact with humans. There has been much subsequent research focusing on virtual realization based on the above frameworks.

This study illustrates the role that the context mediator plays in a Pervasive Computing environment. One of the goals of this study is to enable students to more efficiently study on their own by using customized learning services through a scenario. The scenario is concerned with providing students with good test items. Good Items are items for which item analysis has been completed. Item analysis includes item discrimination and item difficulty and is an extremely useful set of procedures available to students and teaching professionals. Item analysis enables students to improve the quality of their study time and raise their test scores. Therefore if students can solve good test items, they can increase their capabilities as well.

This study is an activity that is ultimately performed by the individual. The student needs to have a large amount of time for drilling new material repeatedly and regularly. The Context Mediator is able to support the student in this regard. The Context Mediator gathers and judges simple context information which each device acquires, delivers messages appropriate for each object. However, the information collected through sensors is insufficient for intelligent service applications, because it is lacking in context information. To cover this deficiency, ontology which can infer contexts based on information gathered through sensors is needed. In addition, different kinds of devices must be able to communicate and share information with each other through the network. This requires an expressive language for inter-device communication. Ontology is one such language making inter-device communication possible, independent of technologies, information structures, and applications of individual systems. To illustrate the important role of ontology, this chapter will cover two scenarios using OWL ontology. In addition, we will show how the context mediator infers rules within each scenario.

2 Architecture of Intelligent Education Services

This section proposes the architecture of customized u-Services for learning that use the semantic web and ontology technology. The proposed architecture is represented in Figure 1. Components of the architecture are introduced as follows.

Pervasive computing refers to the environment where all objects are connected via networks, on-line or off-line. Different kinds of agents communicate with each other through the network. Semantic web technology is based on machines able to process the semantics of data and agents which can understand and process such information. Semantic web technology enables agents to communicate effectively amongst themselves and share information with others.

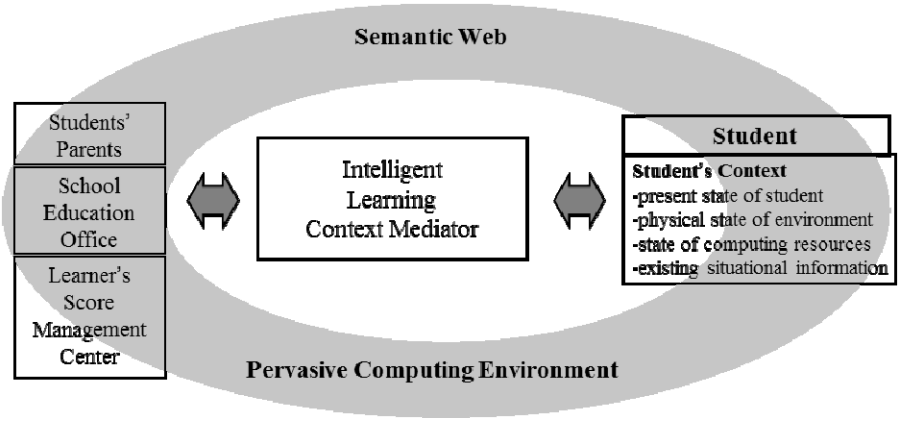


Fig. 1. Architecture for intelligent education services

In terms of a student's context, a student's location, actions, work, and emotional status can be referred to as objects. Students themselves, personal object information rules, and changes in information can be defined as contexts. It is not easy, however, to define context well. Context should include more information necessary for computing such as the student's identity, location, past achievements, and surroundings including the current time, season, and temperature. The process of getting context information from the student's surroundings becomes the core part of the context-awareness [5].

The ontology located on the RDF stratum in the semantic web defines vocabularies as the relationships between terms and inference rules. Ontology is a formal theory suitable for implementing the semantic web, and is a new technology that attempts to achieve effective retrieval, integration, and reuse of web resources [6]. Ontology is necessary for context-awareness in a Pervasive Computing environment. The reason is as follows. Information is necessary to provide personalized services in a Pervasive Computing environment and such information can be gained through sensors implanted in physical objects. However, the information collected through sensors is not sufficient for providing intelligent services, because it lacks context. To address this shortcoming, ontology that enables agents to infer context based on information gathered through sensors is needed. In order to provide services that users want in a Pervasive Computing environment, an exact understanding of user context information is necessary. Different kinds of devices communicate with each other through the network and share information, necessitating an expressive language for inter-device communication. Ontology is such a language that can facilitate communication between different devices, independent of technology, information structure, and applications of individual systems. The shared ontology of devices based on agents is interoperable between devices. In addition, it ensures high levels of spontaneity, flexibility, and promptness. Devices become interoperable in any context because they are able to share context knowledge, even when unforeseen situations suddenly arise. Without needing to define legacy parameters in advance, ontology helps independently-existing agents share the context knowledge necessary for interoperability. Also, ontology enables dispersed agents to share context knowledge as well as inferences on contexts.

3 A Framework of Intelligent Education Services

The Intelligent Learning Context Mediator (I-LCM) is a system in which each device can put together pieces of information on contexts, judge them, and then send proper messages accordingly. The intelligent learning context mediator (I-LCM) consists of the LCAA(Learning Context Acquisition Agent), LRA (Learning Reasoning Agent), LKDB (Learning Knowledge Database), and LTA (Learning Transfer Agent). Brief explanations for each agent are as follows. The role of the LCAA is to put together pieces of information captured by sensors from objects. The LKDB is a space to store items, facts of context, and rules underlying the operation of the I-LCM system.

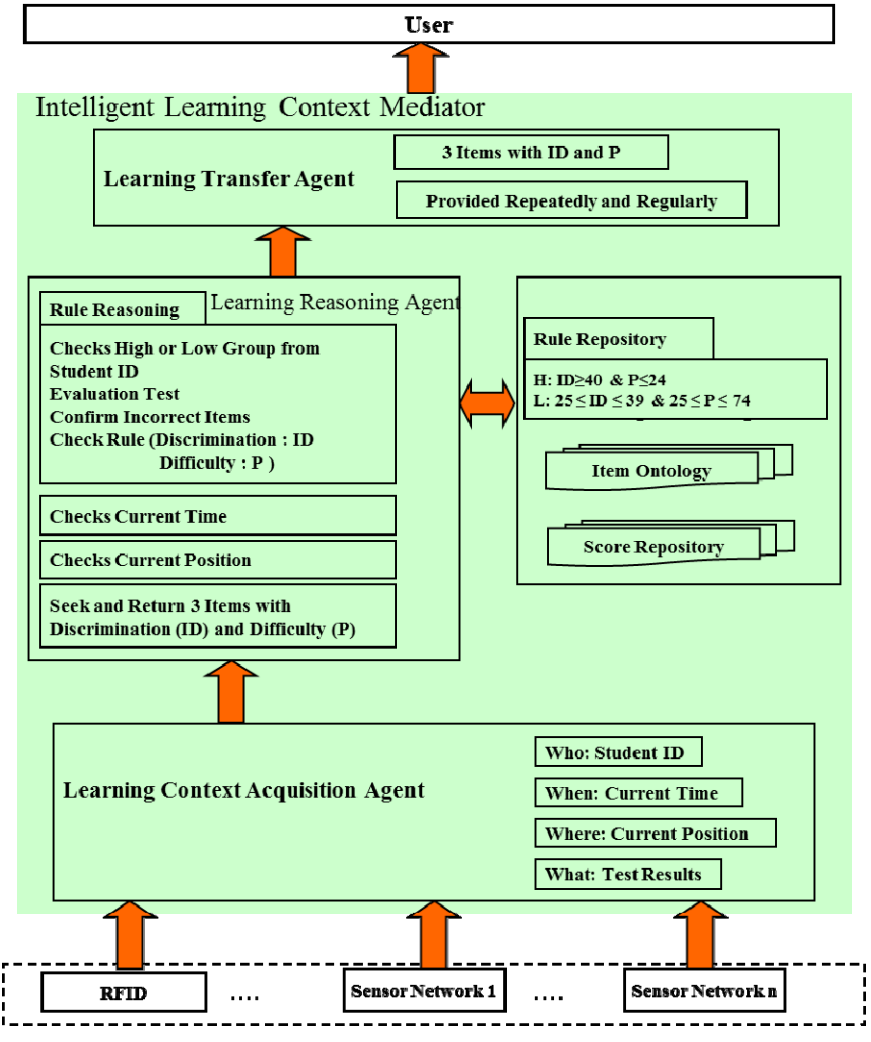


Fig. 2. Context awareness framework for intelligent education services

The LRA compares new information collected from the LCAA with old information such as facts and rules stored in the LKDB. And the LRA infers rules from facts. LRA is a core function of the mediator. The LRA infers context knowledge from acquired context information and identifies informational inconsistencies within the Learning knowledge Database.There are mainly two necessary functions of the inference. First, the LRA needs ontology to infer information from various contexts, enabling the LRA to comprehend the situation based on collected information and must also induce situation-appropriate actions or rules. Second, the LRA must also be able to draw conclusions from pieces of sometimes inconsistent information. The LTA sends messages which the LRA infers to people and devices. Figure 2 shows the context awareness framework for intelligent education services. And figure 3 shows the intelligent learning context mediator for a student.

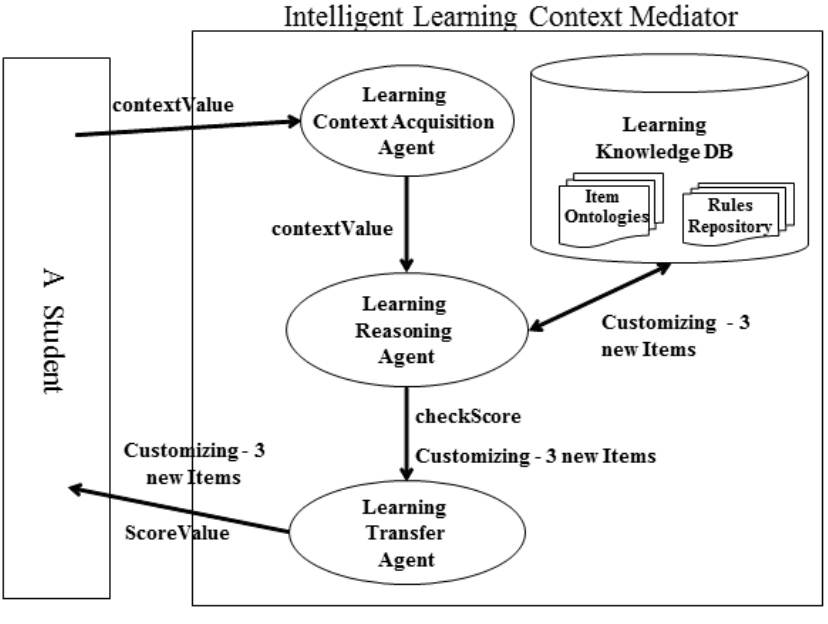


Fig. 3. Intelligent learning context mediator for a student

The I-LCM is able to automatically inform a Student of 3 new Items similar to the incorrectly answered items. The scenario is as follows: A student finished the Test and presses the completion button on their PDA. The student belongs to the high group and is currently located in the park. Figure 4 show you the sequence diagram of intelligent education services for a student using the context awareness framework.

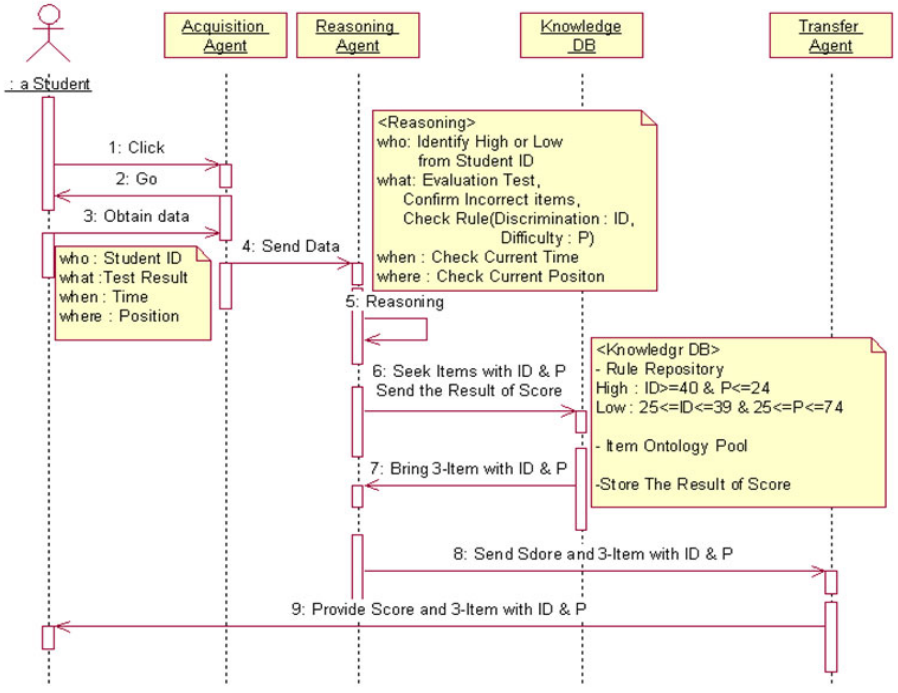


Fig. 4. Sequence diagram of intelligent education services for a student

4 Conclusions

In the pervasive computing environment, humans and devices or machines with computing ability become interoperable. In order to implement intelligent services in this environment, hardware technologies such as sensing and security are needed. However, more importantly, the development of technology that enables devices to take contexts into account is also necessary.

Context awareness must be able to infer unknown facts from known facts. Unknown facts include users' future actions or machine-based information collected from devices. We also showed how ontological technologies could be integrated, in an interoperable manner, with a Ubiquitous learning environment.

This study presents a framework for intelligent education services in which the context mediator recognizes the situation of the user and provides the user with personalized service suitable for the user. The context mediator supports as the intermediation agent for the system while the semantic web and ontology technologies were used for awareness of learner's contexts. This study was able to build an ontological sample using Protégé work.

Acknowledgments. This work was supported by the Industrial Strategic Technology Development Program funded by the Ministry of Knowledge Economy(MKE, Korea). [10038653, Development of Semantic based Open USN Service Platform]. This research was supported by the project of regional technology innovation of the Ministry of Knowledge Economy (MKE, Korea).

References

1. Berners-Lee, T., Hendler, J., Lassila, O.: The Semantic Web. Scientific American, 284 (2001)
2. Salber, D., Dey, A.K., Abowd, G.D.: The Context toolkit: Aiding the development of context-enabled applications. In: Proceedings of ACM SIGCHI Conference on Human Factors in Computing Systems (1999)
3. Schilit, B., Adams, N., Want, R.: Context-aware computing applications. In: Proceedings of the IEEE Workshop on Mobile Computing Systems and Applications (1994)
4. Chen, H., Finin, T., Joshi, A.: A Context Broker for Building Smart Meeting Rooms. In: Proceedings of the Knowledge Representation and Ontology for Autonomous Systems Symposium, 2004 AAAI Spring Symposium, Stanford, CA (2004)
5. Dey, A.K.: Context-Awareness Computing: he Cyber Desk Project. In: Proc. of the AAAI 1998 Spring Symposium on Intelligent Environments (AAAI Technical Report SS-98-02) (1998)
6. Hendler, J.: Agents and the Semantic Web. IEEE Intelligent Systems, 16 (2001)

On-Line Error Detection and Off-Line Test Design in Polynomial Basis Multiplier over GF(2^m) Using Irreducible Trinomials

Chi Hsiang Chang[1], Chiu-Ching Tuan[1], Wen-Tzeng Huang[2], and Che Wun Chiou[3]

[1] Institute of Computer and Communication Engineering,
National Taipei University of Technology,
1, Sec. 3, Chung-hsiao E. Rd.,Taipei City, 10608, Taiwan, R.O.C.
t6419004@ntut.org.tw, cctuan@ntut.edu.tw
[2] Department of Computer Science and Information Engineering, Minghsin University,
1, Xinxing Rd., Xinfeng Hsinchu 30401, Taiwan, R.O.C.
wthuang@must.edu.tw
[3] Department of Computer Science and Information Engineering, Ching Yun University,
229, Jianxing Road, Zhongli City, Taoyuan County. 320, Taiwan, R.O.C.
cwchiou@cyu.edu.tw

Abstract. The concurrent error detection capability can give countermeasure to recent developed fault-based cryptanalysis. The design-for-testability is one of evaluated indexes to detect the faulty element of VLSI chips for manufacturability and maintainability issues. Thus, design of multipliers in GF(2m) with both concurrent error detection and design-for-testability is an important issue for elliptic curve cryptosystem. In this study, a novel self-checking alternating logic (SCAL) multiplier in GF(2m) is presented for achieving both on-line test and off-line test purposes. The proposed polynomial basis multiplier using irreducible trinomials requires only about 33% extra space complexity of existing multipliers. As our best knowledge, the proposed polynomial basis multiplier is the first polynomial basis multiplier which can provide both on-line error detection and off-line test capabilities.

Keywords: Elliptic curve cryptosystem, self-checking alternating logic, concurrent Error Detection, finite field multiplication.

1 Introduction

The finite field arithmetic has been applied and employed in the coding theory [1], cryptography [2], and digital signal processing [3,4]. The finite field arithmetic includes addition, multiplication, division, and inversion operations. The multiplication operation among these operations is the most important arithmetic operation for cryptosystem applications, such as Elliptic Curve Cryptosystem (ECC). Unfortunately, cryptosystems on smart phones are highly vulnerable to side-channel attacks such as fault-based attacks.

ECC strongly depends on finite field arithmetic operations, especially $GF(2^m)$. Hence, efficient implementation of multiplication is fundamental in cryptographic

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 779–786.
springerlink.com

applications. The National Institute of Standards and Technology (NIST) has suggested five binary finite fields for ECC implementation, there are 163, 233, 283, 409, and 571, where 233 and 409 are two of trinomials.

Fault-based cryptanalysis is a new attack for cryptosystems. Boneh et al. [6] had proved that the fault-based cryptanalysis is efficient for RSA cryptosystem in 1996. Kelsey et al. [7] used differential fault analysis to easily recover the key of symmetric Data Encryption Standard (DES), and only requires about two hundred ciphers. Biham and Shamir and Boneh and DeMillo [5] also showed that the fault-based cryptanalysis can effectively attack public-key cryptosystem. Therefore, the simplest method to avoid attacks from the fault-based cryptanalysis is to check the correctness of ciphers before output the ciphers. For this reason, technique to detect errors in cryptosystems is an important research issue in recent years.

In 1998, Fenn et al. [8] firstly proposed an on-line error detection for bit-serial PB multipliers in $GF(2^m)$, and it used parity prediction method. For modern cryptographic applications, the field size m can be very large and it can even range from 160 to 2048 bits. For example, the time overhead in [9] takes about 20% or more in average. Chiou [10] used the REcomputing with Shifted Operands (RESO) method to provide a concurrent error detection method for polynomial basis multipliers using all-one polynomials. Lee et al. [11] extended the Chiou's results [10] to the PB multiplier generated by a general irreducible polynomial. This study will present a self-checking alternating logic (SCAL) bit-parallel PB multiplier using irreducible trinomials. Our SCAL approach can detect both permanent and transient faults and has the fault-secure property which any occurred fault in fault model can be detected by at least one alternating input pair. But, some occurred single faults are not excited and then not detected by existing error detection approaches. Such a SCAL PB multiplier has the features of remaining regular structure and taking only a little time overhead just like the existing time redundancy methods in [10-12].

2 Preliminaries

In this section, we will review basic concepts with PB multiplication over $GF(2^m)$ and Self-Checking Alternating Logic (SCAL).

(a) Polynomial Basis

Let P be an irreducible primitive polynomial over $GF(2)$ with a degree m and is shown in Eq. (1).

$$P = p_m x^m + p_{m-1}x^{m-1} + p_{m-2}x^{m-2} + \cdots + p_1 x^1 + p_0 = \sum_{i=0}^{m} p_i x^i \tag{1}$$

where $p_i \in GF(2)$ and $p_0 = p_m = 1$. Then, let a set $\{1, x, x^2, \ldots, x^{m-2}, x^{m-1}\}$ be a polynomial basis of $GF(2^m)$ generated by P [2]. Thus, we can represent any three elements A,B, and C in $GF(2^m)$ defined by P

Now, let C be the product of A and B and can be represented by Eq. (2).

$$\begin{aligned} & C \\ & = AB \bmod P \\ & = \left(a_0x^0 + a_1x^1 + a_2x^2 + \ldots + a_{m-1}x^{m-1}\right)B \bmod P \\ & = \begin{pmatrix} a_0x^0B \bmod P + a_1x^1B \bmod P + a_2x^2B \bmod P + \ldots \\ + a_{m-1}x^{m-1}B \bmod P \end{pmatrix} \\ & = c_0x^0 + c_1x^1 + c_2x^2 + \ldots + c_{m-1}x^{m-1}, \\ & \text{where } c_i \in \{0,1\}, \text{ for } 0 \le i \le m-1. \end{aligned} \tag{2}$$

(b) Self-checking alternating logic

A circuit whose output is encoded in an error-detecting code is called a self-checking circuit. Alternating logic design technique can achieve the fault detection capability; it is one of the time redundancy techniques. Yamamoto et al. [13] pointed out that any combinational circuits can be made self-dual with only one extra input if they are not self-dual. The alternating logic technique can be easily extended to the multipliers and other arithmetic units if functional modules are self-dual. Reynolds and Metze [14] showed that an arbitrary non self-dual function F with m variables could become a self-dual function F^* with $m+1$ variables by setting

$$\text{(i) } F^*\left(X,\overline{x_{m+1}}\right) = F(X), \text{ and}$$

$$\text{(ii) } F^*\left(X,x_{m+1}\right) = \overline{F(\overline{X})}.$$

When x_{m+1} is the logical value "0", F^* realizes the original function F. When x_{m+1} is "1", F^* performs the dual function $\overline{F}$ of F. Let X^* represent a group of $m+1$ variables (X, x_{m+1}). The function $F^*(X^*)$ is a self-checking alternating logic circuit if it satisfies the following condition when both alternating inputs X^* and $\overline{X^*}$ are applied: $\forall f, \exists X^* \ni \overline{F_f^*(X^*)} \neq F_f^*\left(\overline{X^*}\right)$, where f represents a stuck-at fault in F^*, and F_f^* denotes the function F^* with an existing fault f. In [14,15], they provided the design method for having the alternating logic with self-checking capability. Thus, the self-checking alternating logic strategy can be extended to arithmetic units in $GF(2^m)$ such as multipliers and dividers with hardware redundancy.

3 The Proposed Bit-Parallel Self-Checking PB Multiplier Using Irreducible Trinomials

Based on the above properties, the polynomial basis multiplication over $GF(2^m)$, $C=AB \bmod P$ is represented as follows:

$$\begin{aligned} & C = AB \bmod P \\ & = c_0 + c_1x + c_2x + \ldots + c_{m-1}x^{m-1}, \\ & \text{where } c_i \in GF(2) \text{ for } 0 \le i \le m-1. \end{aligned}$$

Based on Horner's rule, C then can be computed as follows:

$$\begin{aligned} C &= AB \\ &= \left(a_0 + a_1 x + a_2 x^2 + \ldots + a_{m-1} x^{m-1}\right) B \\ &= a_0 B + a_1 xB + a_2 x^2 B + \ldots + a_{m-1} x^{m-1} B \end{aligned} \tag{3}$$

Because finite field operation needs the mod P operation, it can be simplified by using the following equation, Eq. (4).

$$x^m = p_0 + p_1 x + p_2 x^2 + \ldots + p_{m-1} x^{m-1} \tag{4}$$

Let xB mod P can be shown as the following Eq. (5).

$$\begin{aligned} xB &= b_0 x + b_1 x^2 + b_2 x^3 + \cdots + b_{m-2} x^{m-1} + b_{m-1} x^m \\ &= b_0 x + b_1 x^2 + b_2 x^3 + \cdots + b_{m-2} x^{m-1} + b_{m-1}(p_0 + p_1 x + p_2 x^2 + \cdots + p_{m-2} x^{m-2} + p_{m-1} x^{m-1}) \\ &= b_{m-1} p_0 + (b_{m-1} p_1 + b_0) x + (b_{m-1} p_2 + b_1) x^2 + \cdots + (b_{m-1} p_{m-1} + b_{m-2}) x^{m-1}. \end{aligned} \tag{5}$$

Hence, let $B^i = x^i B = b_0^i + b_1^i x + b_2^i x^2 + \cdots + b_{m-1}^i x^{m-1}$, where $\beta_j^i \in \{0,1\}$ for $0 \le j \le m-1$.

Single stuck-at fault model is assumed in this study. In this paper, the problem of concurrent error detection in bit-parallel PB multiplier will be coped through the self-checking design with alternating logic. In other words, self-checking alternating logic design of bit-parallel PB multiplier is necessary. The proposed SCAL bit-parallel PB multiplier includes self-dual AND and three-input XOR gates. Since the three-input XOR gate is a self-dual gate, thus only a self-dual AND circuit needs to be designed. Table 1 lists the truth table of the self-dual AND circuit. Its simplified equation is $z = tb + ab + ta = \overline{(\overline{tb})(\overline{ab})(\overline{ta})}$.

Table 1. Truth table of self-dual AND circuit

t	a	b	z
0	0	0	0
0	0	1	0
0	1	0	0
0	1	1	1
1	0	0	0
1	0	1	1
1	1	0	1
1	1	1	1

The design example of the proposed PB multiplier generated by trinomials is discussed here. An irreducible polynomial consisting of three non-zero terms, such as is $P = x^m + x^n + 1 \, (m > n > 0)$ called a trinomial of degree m. If the proposed PB multiplier is generated by such an irreducible trinomial $P = x^m + x^n + 1$, Eq. (4) can be rewritten as shown as Eq. (8).

$$\begin{aligned} B^i &= xB^{i-1} \\ &= b_{m-1}^{i-1} + b_0^{i-1} x + \ldots + b_{n-2}^{i-1} x^{n-1} + \left(b_{m-1}^{i-1} + b_{n-1}^{i-1}\right) x^n + b_n^{i-1} x^{n+1} + \ldots + b_{m-2}^{i-1} x^{m-1}. \end{aligned} \tag{6}$$

Based on Eqs. (1), (4), and (6), the output result C of the proposed PB multiplier in $GF(2^4)$ generated by $P = x^4+x^3+1$.

The proposed SCAL bit-parallel PB multiplier with $P = x^4+x^3+1$ is shown in Fig. 2. Two steps must be executed in sequential order. The first step performs the original multiplication function: C=A×B. The second step executes the complemented multiplication function: $\overline{C} = \overline{A} \times \overline{B}$. The results of both steps are compared. A mismatch indicates an existed error. The detailed execution status is depicted in Fig. 2. The self-checking algorithm on the proposed PB multiplier is described as follows:

Algorithm-MulSCAL:

/* To execute and compare both C=A×B and $\overline{C} = \overline{A} \times \overline{B}$ */

Step 1: Performing C=A×B.

Step 2: Performing $\overline{C} = \overline{A} \times \overline{B}$.

Step 3. Comparing the result C of Step 1 and the complemented value of the result of Step 2. An error=1 is signaled if at least a mismatch occurs.

Theorem 1: Algorithm-MulSCAL can detect any single stuck-at fault occurred in any cell *K* of the proposed SCAL PB multiplier in Fig.2.

Proof:

The cell *K* is a SCAL circuits. Woodard [15] showed that any single stuck-at fault can be detected in such SCAL circuit. At least one output of the cell *K* will be affected and at least one mismatch will also be found for both final results *C* and $\overline{C}$, Therefore, any single stuck-at fault in any cell *K* of the proposed PB multiplier in Figure 2 can be detected. □

Theorem 2: Any single stuck-at fault occurred in three-input XOR gate of the proposed SCAL PB multiplier in Fig. 2 can be detected by Algorithm-MulSCAL.

Proof:

Since the three-input XOR gates are with self-dual features. If a single stuck-at fault occurs in one of the three-input XOR gates, then the output of such three-input XOR gate will be affected and at least one mismatch will also be found for both final results *C* and $\overline{C}$, Therefore, any single stuck-at fault in the three-input XOR gates of the proposed PB multiplier in Fig. 2 can be detected. □

Theorem 3: Any single stuck-at fault occurred in the proposed SCAL PB multiplier in Fig. 2 can be detected by Algorithm-MulSCAL.

Proof:

Based on Theorems 1 and 2, at least one output will be affected if any single stuck-at fault occurred in the SCAL PB multiplier in Fig.2 and thus at least one mismatch will

also be found by comparing both final results C and $\overline{C}$, Therefore, any single stuck-at fault in the proposed PB multiplier can be detected. □

Any single stuck-at fault occurred in the proposed SCAL PB multiplier is detectable by our proposed SCAL PB multiplier. Moreover, all unidirectional faults are also detectable by our algorithm. Because unidirectional faults will cause normal logical value 0 to faulty logical value 1 or normal logical value 1 to faulty logical value 0 on output, but not both for all inputs. Thus, SCAL circuits can detect all unidirectional faults in nature.

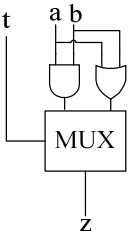


Fig. 1. The circuit K for self-dual AND gate

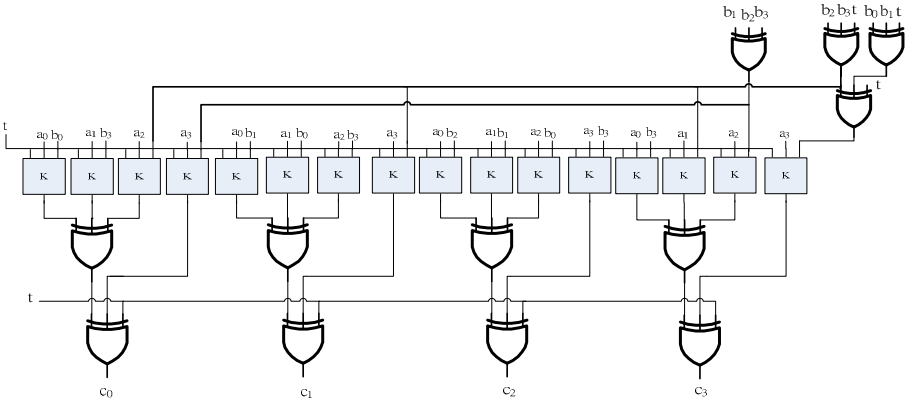


Fig. 2. An example of the proposed SCAL bit-parallel PB multiplier with m=4

Table 2 lists the space complexities of bit-parallel PB multipliers with and without self-checking capability. The proposed SCAL PB multiplier requires 33% space overhead as compared to the original PB multiplier. While comparing with other existing PB multipliers with error detection capability, our proposed SCAL PB multiplier requires less space overhead than existing PB multipliers in [16] with error detection capability. It is noted that our proposed SCAL PB multiplier remains regular structure. Comparison results are listed in Table 2. The CMOS VLSI technology [17] is used to evaluate the space complexity.

Table 2. Space complexities of bit-parallel PB multipliers

	The traditional bit-parallel PB multiplier	Siavash Bayat-Sarmadi Multiplier in [16]	The proposed SCAL PB multiplier using irreducible trinomials in Fig.2
On-line error detection	No	Yes	Yes
Off-line error detection	No	No	Yes
2-input AND gate	m^2	m^2	m^2
2-input OR gate	0	0	m^2
2-to-1 Multiplexer	0	0	m^2
2-input XOR gate	m^2-m	m^2+4m+12	0
3-input XOR gate	0	0	$\leq(\lceil m/2 \rceil)m$
Latch	0	$2m^2$+m	0
Transistor count	$6m^2$+12(m^2-m)	26 m^2+56m+48	18 m^2+12($\lceil m/2 \rceil$)m
m=233	974406	1424610	1304334
Time complexity	$T_A + (2 + \lceil \log_2(m-1) \rceil)T_x$	$T_A + (5 + \lceil \log_2(m^2 - m) \rceil)T_x$	$T_A + (3 + \lceil \log_3 m \rceil)T_{3x} + T_{mux}$

Note: T_A= Two input AND gate execute time.
T_x= Two input XOR gate execute time.
T_{3x}=Three- input XOR gate execute time.
T_{mux}=Two to one Multiplexer execute time.

4 Conclusions

The self-checking alternating logic bit-parallel PB multiplier over $GF(2^m)$ using irreducible trinomials is presented. It just modifies the logic gate structure to have self-dual and self-checking properties. The proposed SCAL PB multiplier requires less space overhead than the existing similar PB multipliers. Furthermore, the proposed SCAL PB multiplier has fault-secure property which all single stuck-at faults occurred in such multiplier are detectable by at least one input pair while existing similar studies have not such property. Both permanent and transient faults are considered in our proposed SCAL PB multiplier. All single stuck-at faults and all unidirectional faults in the proposed multiplier are detectable.

References

1. Macwilliams, F.J., Sloane, N.J.A.: The theory of error-correcting codes. North-Holland, Amsterdam (1977)
2. Lidl, R., Niederreiter, H.: Introduction to finite fields and their applications. Cambridge University Press, NewYork (1994)
3. Blahut, R.E.: Fast algorithms for digital signal processing. Addison-Wesley, Reading (1985)
4. Reed, I.S., Truong, T.K.: The use of finite fields to compute convolutions. IEEE Trans. Inf. Theory IT-21(2), 208–213 (1975)
5. Biham, E., Shamir, A.: Differential fault analysis of secret key cryptosystems. In: Kaliski Jr., B.S. (ed.) CRYPTO 1997. LNCS, vol. 1294, pp. 513–525. Springer, Heidelberg (1997)
6. Boneh, D., DeMillo, R.A., Lipton, R.J.: On the importance of checking cryptographic protocols for faults. In: Fumy, W. (ed.) EUROCRYPT 1997. LNCS, vol. 1233, pp. 37–51. Springer, Heidelberg (1997)
7. Kelsey, J., Schneier, B., Wagner, D., Hall, C.: Side channel cryptanalysis of product ciphers. In: Quisquater, J.-J., Deswarte, Y., Meadows, C., Gollmann, D. (eds.) ESORICS 1998. LNCS, vol. 1485, pp. 97–110. Springer, Heidelberg (1998)
8. Fenn, S., Gossel, M., Benaissa, M., Taylor, D.: On-line error detection for bit-serial multipliers in GF(2m). Journal of Electronic Testing: Theory and Applications 13, 29–40 (1998)
9. Bayat-Sarmadi, S., Hasan, M.A.: On concurrent detection of errors in polynomial basis multiplication. IEEE Trans. VLSI systems 15(4), 413–426 (2007)
10. Chiou, C.W.: Concurrent error detection in array multipliers for GF(2m) fields. IEE Electronics Letters 38(14), 688–689 (2002)
11. Lee, C.Y., Chiou, C.W., Lin, J.M.: Concurrent Error Detection in a Polynomial Basis Multiplier over GF(2m). Journal of Electronic Testing: Theory and Applications 22(2), 143–150 (2006)
12. Chiou, C.W., Lee, C.Y., Deng, A.W., Lin, J.M.: Concurrent Error Detection In Montgomery Multiplication Over GF(2m). IEICE Trans. on Fundamentals of Electronics, Communications and Computer Science E89-A(2), 566–574 (2006)
13. Yamamoto, H., Watanabe, T., Urano, Y.: Alternating logic and its application to fault detection. In: Proc. 1970 IEEE International Computing Group Conference, Washington, D.C., pp. 220–228 (June 1970)
14. Reynolds, D.A., Metze, G.: Fault detection capabilities of alternating logic. IEEE Trans. Computers 12(c-27), 1093–1098 (1978)
15. Woodard, S.E.: Design of digital systems using self-checking alternating logic. Ph.D. Thesis, University of Illinois at Urbana-Champaign, U.S.A (1977)
16. Siavash, B.-S., Hasan, M.A.: Concurrent Error Detection in Finite-Field Arithmetic Operations Using Pipelined and Systolic Architectures. IEEE Transactions on computers, 58(11) (November 2009)
17. Baker, R.J.: CMOS-circuit, design, layout, and simulation, 2nd edn. IEEE Press (2004)

Design of the Low Cost Scalar Multiplier-on-Chip for 163bits Elliptic Curve Cryptosystem

Chi Hsiang Chang[1], Chiu-Ching Tuan[1], Guo-Ming Sung[2], Wen-Tzeng Huang[3], and Xi-Yan Wu[2]

[1] Institute of Computer and Communication Engineering, National Taipei University of Technology, 1, Sec. 3, Chung-hsiao E. Rd., Taipei, 10608, Taiwan, R.O.C.
t6419004@ntut.org.tw, cctuan@ntut.edu.tw
[2] Department of Electrical Engineering, National Taipei University of Technology, 1, Sec. 3, Chung-hsiao E. Rd., Taipei, 10608, Taiwan, R.O.C.
{gmsung,t8318096}@ntut.edu.tw
[3] Department of Computer Science and Information Engineering, Minghsin University, 1, Xinxing Rd., Xinfeng Hsinchu 30401, Taiwan, R.O.C.
wthuang@must.edu.tw

Abstract. As mobile communication flourishes, great importance has been attached to the security of communication data. The keys of the Elliptic Curve Cryptosystem (ECC) are shorter, which is characteristic, so necessary storage components will be significantly reduced, and the square measure of circuits will also be reduced. The cryptosystems are thus very suitable for mobile devices whose demand for low space complexity. In the ECC, scalar multiplication represents the core operation of a system. In recent years, the circuit architecture of triple processor cores or the above was addressed in domestic and international literature. A parallel processing concept is mainly used in this type of architecture to accelerate circuit operation. In this thesis, equation calculation and circuit design were employed to integrate the pipeline architecture and the parallel processing architecture and further propose an elliptic curve scalar multiplier for dual processor cores. In this paper, our proposed architecture Compared with the literature in recent years, this thesis not only provides a high-speed computation, but also effectively reduce about 28% in hardware area.

Keywords: Elliptic curve cryptosystem, Montgomery scalar multiplication, Finite fields, Cryptography.

1 Introduction

As network communication develops rapidly in recently years, Internet has already become an indispensable part in many people's life. The platform for people to use internet has shifted gradually from PCs to mobile devices, such as smart phones and tablet PCs. Furthermore, the communication medium for Internet has also progressed from traditional wired networks to wireless networks. The most significant feature of this type of Internet accessing pattern, in which portable devices and wireless network technology are integrated, lies in mobility, that is, users can use Internet services

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 787–792.
springerlink.com

anytime anywhere, and the pattern has become the new Internet mainstream. However, the information transmitting and receiving processes of this type of mobile Internet accessing may contain a great deal of secret personal data. When those important data are transmitted through public channels, such as wireless Internet, how should they be tackled with safely to avoid the stealing of people with certain intentions, which may further result in serious loss? Relevant issues also cause people to value Internet information security.

In 1976, Diffie & Hellman brought up the concept of public-key cryptosystem 2 which was applied to solve the problems of traditional cryptosystems in communication. A public-key cryptosystem is also called an asymmetric cryptosystem, and famous public-key cryptosystems include RSA (Rivest, Shamir, and Adleman) and the Elliptic Curve Cryptosystem (ECC). This type of cryptosystems needs two different keys, respectively a public key and a private key, which should correspond to each other. When a transmitting end hopes to transmit information to a receiving end, the transmitting end should at first use the public key to encrypt the information. After the receiving end receives the encrypted information, it should apply the corresponding private key to decryption in order to restore correct information.

The ECC brought up by Koblitz3 and Miller4 belongs to public-key cryptosystems. Based on Elliptic Curve Discrete Logarithm Problem (ECDLP), the greatest advantage of the ECC lies in that it uses a shorter key than the keys of other cryptosystems to achieve an identical security standard, as indicated by Table 15. The key of the ECC contains 160 bits, and the key of RSA contains 1024 bits, but the two cryptosystems have an equal degree of security. Hence, under the same degree of security, the space that ECC needs to save the key is only 1/6 of the space that RSA needs. The platforms of mobile communication systems usually have higher demand for square measure, so the EDD is very suitable for portable communication devices.

Table 1. The comparison between key lengths under the same security level

ECC Key Size (Bits)	RSA (RIVEST、SHAMIR、ADLEMAN) Key Size (Bits)	Key Size Ratio
163	1024	1:6
256	3072	1:12
384	7680	1:20
512	15360	1:30

In the ECC, scalar multiplication is responsible for the core operation of a system, and the scalar multiplier occupies most of the square measure of the entire system. In recent years, many experts researched scalar multiplication. In 2008, Bijan 15 established the first single-processor-core scalar multiplier using pipeline architecture. In this architecture was cut into four grades of pipelines, and the algorithm in the scalar multiplication was re-sequenced. In 2008, Kim 16 brought up a multiple-processor-core scalar multiplier, in which the Gaussian Normal basis was applied to the finite field mathematic basis. In 2010, both Lai 17 and Zhang 18 used three processor cores as the architectures of scalar multipliers. In these two architectures, the parallel processing concept was used to accelerate circuit operations.

To design a more effective and low-cost ECC, we employ pipeline and parallel processing architectures, to proposed the new elliptic curve scalar multiplier with a dual-processor-core architecture.

2 Preliminary

Scalar Multiplication Algorithm

Two inputs are necessary in traditional algorithms, respectively a reference point, *P*, and a scalar, *K*, and the output is *K* times of Point *P*. Traditionally, *KP* is calculated by *Doubling* Point *P* and then continuously *Adding* Point *P* for *K-2* times, as indicated by the following equation.

$$KP = P + P + \underbrace{P + \cdots + P}_{K-2}$$

For example, if *K*=18, and the calculation result is saved in Reg, the process for calculating *KP* is as follows:

Step 1: Reg ← $P + P = 2P$ (*Double*)
Step 2: Reg ← Reg + $P = 3P$ (*Add*)
Step 3: Reg ← Reg + $P = 4P$ (*Add*)
Step 4: Reg ← Reg + $P = 5P$ (*Add*)
Step 5: Reg ← Reg + $P = 6P$ (*Add*)

•
•
•

Step 16: Reg ← Reg + $P = 17P$ (*Add*)
Step 17: Reg ← Reg + $P = 18P$ (*Add*)

The algorithm used in this thesis was the Montgomery Scalar Multiplication [14], as shown in Figure 1. This algorithm can be divided into three parts. The first part includes Step, 1, 2, and 3 in Figure 1, in which the input value is initialized. The second part displays the scalar multiplication loop, namely Step 4 in Figure 1. In the

Input: An integer $K > 0$ and a point $P=(x,y) \in E$
Output: $Q=KP$

1. if $K=0$ or $x=0$ then output $(0,0)$ and stop .
2. Set $K \leftarrow (k_{n-1}k_{n-2}\cdots k_1k_0)_2$.
3. Set $X_1 \leftarrow x, \quad Z_1 \leftarrow 1, \quad X_2 \leftarrow x^4 + b, \quad Z_2 \leftarrow x^2$.
4. for i from $n-2$ downto 0 do
   if $k_i = 1$ then
   $Z_1 \leftarrow (X_1 \cdot Z_2 + X_2 \cdot Z_1)^2, \qquad X_2 \leftarrow X_2^4 + b \cdot Z_2^4$
   $X_1 \leftarrow x \cdot Z_1 + (X_1 \cdot Z_2) \cdot (X_2 \cdot Z_1), \qquad Z_2 \leftarrow Z_2^2 \cdot X_2^2$
   else
   $X_1 \leftarrow X_1^4 + b \cdot Z_1^4, \qquad Z_2 \leftarrow (X_1 \cdot Z_2 + X_2 \cdot Z_1)^2$
   $Z_1 \leftarrow Z_1^2 \cdot X_1^2, \qquad X_2 \leftarrow x \cdot Z_1 + (X_1 \cdot Z_2) \cdot (X_2 \cdot Z_1)$
5. $x_1 \leftarrow \frac{X_1}{Z_1}, x_2 \leftarrow \frac{X_2}{Z_2}, y_1 \leftarrow \{(x_1 + x)[(x_1 + x)(x_2 + x) + x^2 + y]/x\} + y$
6. return($Q = (x_1, y_1)$)

Fig. 1. The Montgomery Scalar Multiplication

operation of scalar multiplication, the scalar multiplication loop will be operated for 162 times in total. The third part includes Step 5 and 6 in Figure 1, in which the coordinates are transferred.

3 System Design and Results

The high-speed scalar multipliers can be divided into two categories, namely, single-processor-core architectures [16-17] and multiple-processor-core parallel processing architectures [18-20]. The single-processor-core architectures save circuit space, but it is difficult for them to improve the system performance. The multiple-processor-core architectures have high system performance, but the circuit space and cost will increase, which is disadvantageous to portable platforms.

The scalar multiplier in this thesis was designed by means of the Montgomery Scalar Multiplication introduced in Figure 1. It is obvious in Figure 1 that the overall scalar multiplication included Add and Double, the formulas of which are as indicated by Equation (1) and (2) as follows:

$$Add = \begin{cases} Z_2 \leftarrow (X_1 \cdot Z_2 + X_2 \cdot Z_1)^2 \\ X_2 \leftarrow x \cdot Z_1 + (X_1 \cdot Z_2) \cdot (X_2 \cdot Z_1) \end{cases} \quad (1)$$

$$Double = \begin{cases} X_1 \leftarrow X_2^4 + b \cdot Z_2^4 \\ Z_1 \leftarrow Z_2^2 \cdot X_2^2 \end{cases} \quad (2)$$

Through different operations and sequencing, different clock schedules will be obtained from the aforementioned formulas. If the quantity of processor cores are increased or decreased, the results will be different. Therefore, seeking for optimal clock scheduling is critical in this paper. The clock scheduling of the scalar multiplier with a dual-processor-core architecture, which was brought up in this thesis, will be introduced in the following part.

Dual-processor-core Clock Scheduling

If a single processor core is used as the architecture of a system, the system performance will be limited due to insufficient circuit resources although the hardware cost will be reduced. Nevertheless, if triple processor cores are applied to the design of system architecture, incredible hardware cost will be necessary although the system operation speed will be increased. Consequently, a scalar multiplier with a dual-processor-core architecture was designed in this thesis. It has not only the feature of low cost but also the capability for high-speed calculation. When dual processor cores are used to fulfill a scalar multiplier, the processor cores should have an addition, square, and multiplier circuit. In Figure 2, ⊛ represents the multiplication, ⊕ represents the addition, and ^2 represents the square.

In this architecture, the feature of pipeline finite field multiplier was used to accelerate operation. Since the circuit architecture was of two processor cores, there should be at least two multiplications, one addition, and two square operation. In this architecture, it needs 9 clock cycles to complete a loop operation of scalar multiplication. It needs 8 clock cycles to complete a loop operation of scalar multiplication. Moreover, it needs totally 162 loop operations of scalar multiplication to complete a loop operation of scalar multiplication, so each scalar multiplication needs 162×8=1296 clock cycles. The result is identical to that of the triple processor

core architecture[18], but this architecture needs only two processor cores the compared results as Table 2, so two advantages, namely high-speed operation and low cost, are achieved.

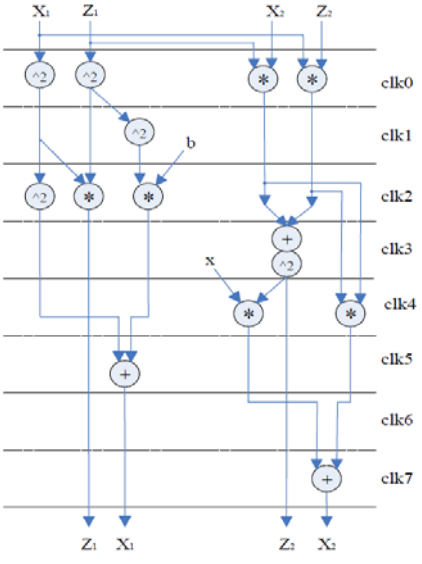


Fig. 2. The sequence diagram of proposed dual processor cores

Table 2. Comparisons Scalar Multiplier for 163bits Elliptic Curve Cryptosystem

	Ansari and Hasan [15]	Y. Zhang et. al [18]	Present here
Number of ALU	1	3	2
Multiplication	163	163	163
Addition	1	2	1
Square	1	2	2
Power of four	0	1	0
Clock	4050	1296	1296

4 Conclusion

Cryptosystems are usually fulfilled by means of VLSI since the operation speed of hardware circuits are faster than that of software programs. Moreover, scalar multiplication operation is very important in the decryption operation of the ECC. If the square measure of a scalar multiplier can be reduced, the opportunity for the ECC to be applied to mobile Internet devices will be greatly increased. Consequently, a high-speed, low-cost dual-processor-core elliptic curve scalar multiplier was brought up in this thesis.

Comparing with similar types of multipliers which are currently available, the dual-processor-core elliptic curve scalar multiplier brought up in this paper saves 28% of circuit cost and, meanwhile, provides high-speed scalar multiplication operation.

References

1. AES home page, http://www.nist.gov.aes/
2. Diffie, W., Hellman, M.E.: New Directions in Cryptography. IEEE Transaction on Information Theory 22(6), 644–654 (1976)
3. Koblitz, N.: Elliptic Curve Cryptosystems. Mathematics of Computation 48(177), 203–209 (1987)
4. Miller, V.S.: Use of elliptic curves in cryptography. In: Williams, H.C. (ed.) CRYPTO 1985. LNCS, vol. 218, pp. 417–426. Springer, Heidelberg (1986)
5. Huang, J., Li, H., Sweany, P.: An FPGA Implementation of Elliptic Curve Cryptography for Future Secure Web Transaction. In: International Conference on Parallel and Distributed Computing Systems, pp. 296–301 (September 2007)
6. Seberry, J., Pieprzyk, J.: Cryptography: An Introduction to Computer Security. Prentice-Hall (1986)
7. Hankerson, D., Menezes, A., Vanstone, S.: Guide to Elliptic Curves Cryptography. Springer-Verlag Hall/CRC (2004)
8. IEEE 1363,: Standard Specifications for Publickey Cryptography. IEEE Press (2000)
9. NIST: Recommended elliptic curves for federal government use (May 1999), http://csrc.nist.gov/encryption
10. Itoh, T., Tsujii, S.: A Fast Algorithm for Computing Multiplicative Inverses in GF(2m) Using Normal Bases. Information and Computation 78(3), 171–177 (1988)
11. Montgomery, P.L.: Speeding the pollard and elliptic curve methods of factorization. Mathematics of Computation 48, 243–264 (1987)
12. López, J., Dahab, R.: Fast multiplication on elliptic curves over $GF(2_m)$ without precomputation. In: Koç, Ç.K., Paar, C. (eds.) CHES 1999. LNCS, vol. 1717, pp. 316–327. Springer, Heidelberg (1999)
13. Sutikno, S., Surya, A., Effendi, R.: An Implementation of ElGamal Elliptic Curves Cryptosystems. In: IEEE Asia-Pacific Conference on Circuits and Systems, pp. 483–486 (1998)
14. Ansari, B., Hasan, M.A.: High performance architecture of elliptic curve scalar multiplication. Tech. Report CACR 2006-01 (2006)
15. Ansari, B., Hasan, M.A.: High-Performance Architecture of Elliptic Curve Scalar Multiplication. IEEE Transaction on Computers 57(11), 1443–1452 (2008)
16. Choi, H.M., Hong, C.P., Kim, C.H.: High performance elliptic curve cryptographic processor over GF(2163). In: Proc. IEEE Int. Symposium on Electronic Design, Test, and Applications (DELTA), pp. 290–295 (January 2008)
17. Lai, J.Y., Hung, T.Y., Yang, K.H.: High-performance architecture for Elliptic Curve Cryptography over binary field. In: International Symposium on Circuits and Systems (ISCAS), pp. 3933–3936 (August 2010)
18. Zhang, Y., Chen, D., Choi, Y., Chen, L., Ko, S.B.: A high performance pseudo-multi-core ECC processor over GF(2163). In: International Symposium on Circuits and Systems (ISCAS), pp. 701–704 (August 2010)
19. Wu, H.: Bit-parallel finite field multiplier and squarer using polynomial basis. IEEE Transaction on Computers 51(7), 750–758 (2002)

Research on Interference Rejection of TT&C in Formation Flying of Small Satellite

Pan Xi[1] and Chen Ming[2]

[1] Beijing Institute of Technology, Beijing, 100081
[2] Beijing Research Institute of Telemetry, Beijing, 100076

Abstract. Formation flying of small satellite has become research hotspot in space field today. Tracking Telemetry and Command (TT&C) is the key technology to ensure the whole formation to cooperate and self-running normally. Therefore, the TT&C system is becoming the main hitting target in the electronic warfare. This paper is to improve anti-interference ability of formation flying of small satellite, the TT&C controlling system characteristics are discussed and interference forms are provided. The methods of interference rejection are mainly studied.

Keywords: Formation Flying of Small Satellite, TT&C, Anti-interference.

1 Introduction

As an important application of small satellite, formation flying of small satellite is generally considered as the novel application of small satellite which can simulate a large or distributed satellite. Small satellite formation has the advantages of low-cost and high performance that has a vital performance in the area of targets detection and tactical communication and thus it will be chief attacked in electronic warfare. In this paper, we discuss the TT&C system characteristics and interference forms, and study the methods using for interference rejection.

There are many application of satellite formation presently which are different in the purpose (height measurement, deep space detection), formation characteristics (number, size and distance of satellites) and orbit altitude (low orbit, middle orbit and high orbit). Such as EO-1、DS3 (Deep Space 3)、ST3 (Space Technology), TechSat21, ORION, ION-F (Ionospheric Observation Nanosatellite Formation), LISA (Laser Interferometer Space Antenna), Gemini (GPS-based Orbit Estimation and Laser Metrology for Inter-satellite Navigation)、TanDEM-X、GRACE (Gravity Recovery and Climate Experiment),etc.

2 Characteristics of TT&C System in Formation Flying of Small Satellite

The main work for TT&C system in small satellite formation is to determine the orbit of each satellite, acquire the datum of subsystem and operating parameters of satellites and correct the attitude of satellites. TT&C in small satellite formation takes the following characteristics:

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 793–798.
springerlink.com

1) Monitoring and controlling multi satellites, spread spectrum has more superiority;
2) High autonomy and adaptive networking capability;
3) Multi-band communication and manifold modulation methods;
4) Communication safety between satellites is the vital issue.

3 Interference Analysis of TT&C System

TT&C system of formation flying of small satellite (FFOSS) consists of two types of link: one is between the satellites, another is between satellite and ground. When man-made interference comes forth, the lethal mistakes will occur when satellite communicates with each other in the link and the formation flying will be disarranged. Therefore, the reliable TT&C communication method between satellites is very important for self-running formation flying of small satellite, that the information should be insured to communicate accurately, orderly and non-repeatedly. In the link between satellite and ground, TT&C system on the ground measure the velocity, angle and distance for formation flying of small satellite. Thus, the research of anti-interference technology applied to TT&C is significant and will be key technology in the field of formation flying of small satellite in the future.

3.1 Interference Mode

Involuntary and hostility interference always appear to the two type links, which contains narrowband interference and broadband interference. The narrowband modes have single frequency and multi-frequency wave, digital narrowband interference, AR random interference, and barrage jamming, swept and pulse jamming are the broadband interference. When these interferences attack in time, frequency and space, together, TT&C between small satellites becomes destroyed badly.

3.2 Attack Methods

Man-made interferences entering from antenna into receiver have three sorts: the first is the low power interference. When it is added to the receiving signal the BER (bit of error rate) will deteriorate as the affection to random component when demodulating. Secondly, the interference power become more to lead the receiver lost lock then acquire repeatedly, communicate discontinuously which lead to serious BER that can't be corrected by error correct encode; Thirdly, when the interference is pulse, whose instantaneous power is more intense than signal, AGC of RF front will be saturation and broken.

4 Methods Analysis of Anti-interference

The anti-interference tolerance J/S of communication link system can be denoted as follow[4]:

$$J/S = (\psi) \times AGC \times J_{anti-Jam} \times (W/R) \times \left[(E_b/N_o)_{th} \right]^{-1} \quad (1)$$

Where, (ψ) stands the antenna spatial identification coefficient which can attain 0~30dB processing gain generally. $J_{anti-Jam}$ represents the adaptive interference restrain technique before de-spread that using corresponding algorithm for different type of interference can enhance the anti-interference ability of spread spectrum system. (W/R) is the tolerance brought by spread spectrum system. $(E_b/N_o)_{th}$ is the receiver's demodulation threshold, which can attain 0~10dB gain by error correct encode technology.

4.1 Anti-interference Technology for Antenna

Antenna is the first defense for the communication in ECM and the most commonly used method for anti-interference, which can be designed to restrain many kinds of interferences from different direction, even the one with the same frequency of signal. Such antennas include multi-beam antenna, auto-adaptive force zero antenna and intelligent antenna. Multi-beam antenna can control direction and beam form to enhance anti-interference ability. The principle of adaptive zeroing antenna is to auto-weight for each antenna cell which can bring higher gain at signal direction and zeroing at interference direction by auto-controlling and optimizing the antenna pattern according to the difference between both sides in frequency, amplitude and direction. Intelligent antenna is becoming research focus in military TT&C field, which is an antenna array consisting of multi-beam antenna. Antenna array can provide many single beams that can overlay desire communications area. Every single beam can make system work at high-point by adjusting antenna direction and zero according as a rule.

4.2 Smart AGC

Smart AGC is a new anti-interference method which is an adaptive envelop deflexed limiter circuit, having more 0~20dB gain than general hard limiter. This technology is very effective to restrain pulse interference which use the difference between weak signal and strong interference to pick-up envelop of strong interference then to auto adjust amplifier to depress strong interference and magnify weak signal. With this processing, we can effectively enhance SNR, restrain interference and improve system performance by weakening strong pulse interference, continues wave interference barrage interference.

4.3 Spread Spectrum Communication with Higher Tolerance

Spread spectrum communication mode contains DS-SS (direct sequence spread spectrum), FH-SS (frequency hop spread spectrum) and mix spread spectrum which can reduce interference by spreading interference spectrum. Spread spectrum has been widely used in satellite communications, such as GPS and TDRSS, even if NASA and military satellite. Spread spectrum communication do not care interference type, has been wildly used in our TT&C as an effective anti- interference technology.

4.4 Adaptive Interference Suppression Technology

(1) Narrow-band interference suppression technology

Formation flying of small satellite adopts spread spectrum system, spreading the power of interference to a wide range of frequency, can achieve the interference suppression in a way. However, when the power of the interference is beyond the anti-jamming ability of DS system, the interference suppression algorithm is requisite to suppress the interference. Narrow-band interference suppression technologies include temporal prediction, transformation domain technology and code aided technology.

Temporal prediction relies on the different predictability of wide-band signal and narrow band interference and the interference is to be deducted from the receiving signal so as to achieve the purpose of inhibiting. Figure 1 shows the structure of linear prediction filter.

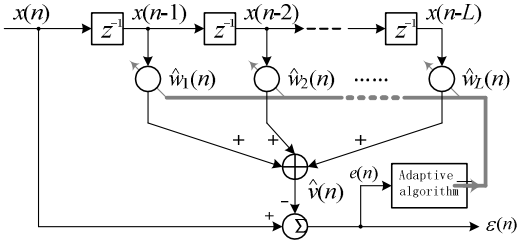


Fig. 1. Structure of linear prediction filter (LPF)

Transformation domain technology maps the receiving signals which include as well as the TT&C signals and the narrow-band interference to different transformation domains. Narrow-band interference can be mapped to function similar to the impulse and TT&C signals have a flat characteristic. So the interference can be recognized in the transformation domain and can be removed by some processing. This technology is suitable for the condition that interference characteristics are unknown and statistical properties are rapidly changing. Figure 2 shows the effect of this method.

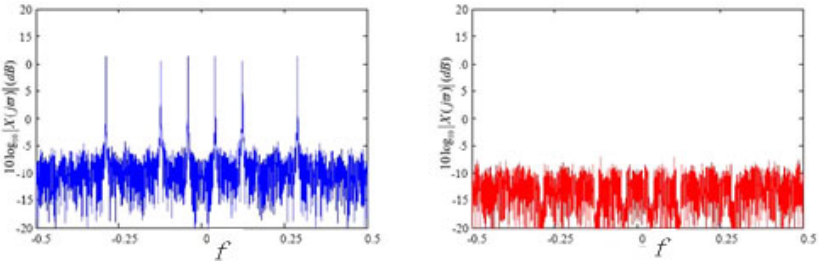


Fig. 2. Spectrum of signals before/after process under the multi-tone interference

Code aided technology, using the method of multi-user detection (MUD), decomposes receiving signal into a set of virtual users that do not overlap each other .The real TT&C signal and virtual users (interference) constitute a new virtual CDMA system then the interference can be handled as the multi-address interference to achieve the narrow-band interference suppression. Figure 3 shows the principle of the code aided technology.

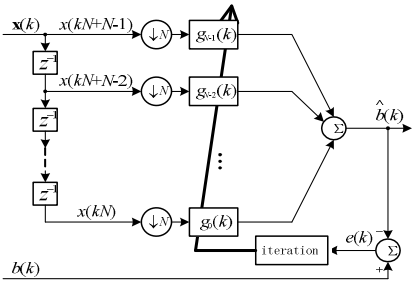


Fig. 3. Principle of the code aided technology

(2) Wide-band interference suppression technology

Wide-band interference includes pulse interference, wide-band barrage interference and linear FM interference. The most common method is to adjust the amplitude of the receiving signal by AGC because the energy of Pulse interference is concentrated in the time domain to adjust the signals whose amplitude changes largely to the constant envelope signals.

Wide-band barrage interference possesses the white noise characteristics and forms, with strong random in time domain and flat spectrum. Against such interference, the most effective method is to increase the power of desired signal and use the gain of the DS system (though dispreading process, the power of interference will be spread and the power of desired signal will be concentrated). With the development of modern signal processing theory, a series of new signal processing method are proposed for it, e.g., Gabor transform, wavelet transform, Radon-Wigner transform, fractional Fourier transform (FRFT), etc. Using these methods, analysis, detection, estimating, synthesis, and suppression become a research hotspot.

4.5 Error Correction Encoding Technology

Error correction encoding technology could effectively improve the reliability of data transmission. The most commonly encoding methods include block code, cyclic code and convolution code. New encoding methods, e.g., Turbo code, LDPC code and concatenated codes (BCH/RS+TC, BCH/RS+CC) are gradually applied to real projects. Digital fountain code (LT code and PAPTOR code) is a rate-independent erasure code for this condition. A serial concatenation encoding scheme which consist of BCH code and digital fountain code can be considered. It could increase the decoded probability of LT code by improving the error correction ability of system though the using of BCH code and it will efficiently improve the symbol recovery ability of LT

code. At the same time, the erasure of the signal processing by using LT code will relatively improve error correction capability of BCH code. The BCH-LT concatenated code can be used, which BCH code is as outer code and LT code as inner code. When BCH code use (127, 64) fountainhead BCH code, LT code input k=2048, output n=3038, PDF use Robust Solition. Emulating result is as figure 4.

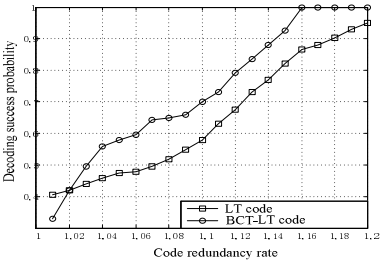


Fig. 4. BCH-LT concatenated code performance

BCH-LT concatenated code can enhance anti-delete ability of LT code, notwithstanding los package. When encode depth is 1.16, BCH-LT serial concatenated code can achieve inerrability communication with LPR=0.05. BCH-LT concatenated code can be studied further, aim at new characteristic.

5 Conclusion

Formation flying of small satellite (FFOSS) is a virtual application and important trend of development for small satellite which can perform many tasks, such as direct counterforce, air fence and information supporting. In this paper, the characteristics of TT&C system for formation flying of small satellite are discussed and the actual interference modes between satellites and satellite- ground are analyzed. The anti-interference technologies are summarized that can provide the reference for the application of TT&C system in the small satellite formation.

References

1. Xu, J.: Key techniques of smal l satellite formation f light and its development. Aerospace Electronic Warfare 23, 24–29 (2007)
2. Jain, P.C.: Onboard Processing in Future Satellite Communication System, Telecommun. Networks Summar, 2–6 (1983)
3. Wang, P.K.C., Hadaegh, F.Y.: Formation flying of multiple spacecraft with automatic rendezvous and docking capability, AIAA Guidance, Navigation, and Control Conference and Exhibit, AIAA (2003)
4. Zhou, D.S., Tian, H.X., Zeng, X.W.: An anti-jamming satellite communication system with high Gp based on OBP de hopping for multi FH as a whole. Journal of Xidian University (1999)

Research on Robust Image Perceptual Hashing Technology Based on Discrete Cosine Transform

Liang Chang[1], Wen-Gang Yan[2], and Wen-De Wang[3]

[1] Information and Electronic Technology Institute, Jiamusi University, Jiamusi, P.R. China, 154007
chang-liang1980@163.com
[2] The International Institute, Jiamusi University, Jiamusi, P.R. China, 154007
[3] Application of Electronic Technology Institute, Jiamusi University, Jiamusi, P.R. China, 154007

Abstract. In this paper, current methods about perceptual hashing were deeply researched and the defects of them were analyzed. Principles and characters of discrete cosine transform, Watson visual model and chaotic model were introduced in details. A new method about perceptual image hashing which combined the characteristics of the three models was designed. This method use discrete cosine transform to feature extraction from image, eigenvector was handled by contrast sensitivity table, Logistic equation was used as chaos sequence generator to encrypt and at last the prediction differential method was used to quantization coding. Experimental results indicated that the method could resist the content-preserving modifications, and possessed strong robustness, safety and retrieval ability, and the collision rate decreases to level 10^{-7}. Therefore, this technique has the applied value in image authentication, copyright protection, image security and content-based image retrieval and so on.

Keywords: Perceptual hashing, discrete cosine transform (DCT), chaotic model, robustness, security.

1 Introduction

Along with the development of information technology and network, a large number of digital images are widely used in daily life and work. But the vast amount of data brought inconvenience to management and maintenance which makes the traditional method powerless. Perception hash (Perceptual Hashing) technology converts image data into hundreds or thousands of binary sequences [1]. To the large number of image database retrieval, it greatly shorten the retrieval time and reduce the cost of storage medium, and brought convenience to management and maintenance.

In addition, in the process of generating hash sequence, secret hash sequence with keys can be applied into the copyright protection and image authentication. In the process of feature extraction with the visual model will reduce the redundant information from data and improve the efficiency of the image compression which greatly cut down the volume of data transmission and time. The feature extraction is

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 799–809.
springerlink.com

the key process of hash sequence generation; there are some existing feature extraction methods: the statistical method based on image, image representation, extraction method based on the relation and roughly said image feature point.

The disadvantages of method based on the statistical is that it can change the contents of image without changing the statistical data [2], and use the wavelet analysis methods to extract sub band[3], but it can't detect malicious attack, too. The method of extracting image hash based on relation characteristics and compress the image by using Discrete Cosine Transform (DCT) has good robustness, but sensitive to geometry Transform [4].

The method based on image rough said with singular matrix decomposition methods get the low rank of the original data[5], this method maintains good robustness under the situation of tiny approach perception modified to the image, but vulnerable to be attacked by the localized tamper, this method using wavelet transform to extract mutating has stronger robustness and high complexity, but uniqueness is uncertain[6].

Above all the feature extraction methods, in order to better satisfy the perception of image hash robustness, safety and uniqueness, using DCT transformation to extract feature is considered as an ideal method.

Currently, the methods used DCT to extract features are mainly as follows:

1) Make use of the statistical methods to generate the hash function to the extraction features which have good retrieval accuracy and computing speed[7], but it can not satisfy the attack requirements of changing the image contents without changing the histogram.

2) The characteristics graph make use of the Hotelling's T-Square(HTS) statistic of block DCT coefficients as the image has good function of tamper with the positioning, but cannot satisfy the irreversibility of the hash perception[8].

3) In the encryption process , generate sequence of encryption y using pseudo random is easy to produce features redundant, and using Huffman coding is easy to produce feature matching dislocation in distance calculations[9].

Based on the above analysis, the author will combine DCT transformation, Watson visual model and chaos theory together to satisfy the perception of the hash on robustness, safety and uniqueness.

Firstly, extracted image features with DCT transformation, and make use of Watson visual model to deal with DCT coefficients; retain one DC coefficient and nine AC coefficients of the partitioned matrix and extract image with perception.

Then according to the iterative no repetitive and initial sensitivity of chaotic area to encrypt feature vector by using chaotic sequence generator based on Logistic equation. Then quantizing and coding feature vector by using quantitative predict difference method.

At last make the matching judgment with the Han marking distance calculating formula.

Experimental results prove that the method can not only satisfy safety and irreversibility but also resist most attacks which have good robustness, compactness and class abstracts.

2 DCT Transformation Principle and Watson Visual Model

2.1 DCT Transformation Principle

DCT transformation is a real number filed transformation; the core of transformation is the cosine functions of real number. Transform the image into even function by using symmetry of Fourier transformation and double edges plait operation and making 2-d discrete Fourier transform to the image. The transformation results will only contain cosine transformation and the important related visual information of image is concentrated in part of the transformation of the DCT coefficients after transforming. Therefore, the DCT is the core of Lossy image compression JPEG image which is also called one of the main "transformation domain (DCT domain)" of "transformation hidden information algorithm" [10].

The two dimension discrete cosine transform formula is as follows:

$$F(u,v)=c(u)c(v)\sum_{x=0}^{M-1}\sum_{y=0}^{N-1}f(x,y)\times\cos\frac{\pi(2x+1)u}{2M}\cos\frac{\pi(2y+1)v}{2N} \tag{1}$$
$$(u=0,1,\cdots M-1;v=0,1,\cdots,N-1)$$

Among them

$$c(u)=\begin{cases}\sqrt{1/M} & u=0\\ \sqrt{2/M} & u=1,2\cdots,M-1\end{cases}$$
$$c(v)=\begin{cases}\sqrt{1/N} & v=0\\ \sqrt{2/N} & v=1,2\cdots,N-1\end{cases} \tag{2}$$

The two dimension inverse discrete cosine transform formula is:

$$f(x,y)=\sum_{x=0}^{M-1}\sum_{y=0}^{N-1}c(u)c(v)F(u,v)\times\cos\frac{\pi(2x+1)u}{2M}\cos\frac{\pi(2y+1)v}{2N} \tag{3}$$
$$(x=0,1,\cdots;M-1;y=0,1,\cdots;N-1)$$

In the formula: x, y --space domain samples values,

u, v --frequency domain samples values.

2.2 Watson Visual Model Based on DCT Coefficient

The visual model based on DCT proposed by Watson is a classic perception combination of comprehensive sensitivity, masking and error model. It replaced multi-channel decomposition with DCT, and can be combined very well with JPEG and some watermarking algorithm [11]. It was composed by four parts: contrast sensitivity table, brightness masking, and contrast masking and perception quality measurement.

$$D(I,I_w)=(\sum_{i,j}|d_{i,j}|^{\beta})^{\frac{1}{\beta}} \tag{4}$$

Only the contrast sensitivity table is used in the paper, therefore the DCT block 8×8 contrast sensitivity is listed, the generating method can be found in reference [13].

Each number in Table 1 represents the minimum DCT coefficients can be perceived under non-exist masking circumstances and considered the coefficient as a noticeable JND (Just Noticeable Difference, criticality perceptible difference). The smaller the number means the more sensitive of people's eye on that frequency.

3 Implement Principles

3.1 Experiment Procedure Process

Experimental procedure process is shown in fig.1.

3.2 Image Normalization

Change the Images into 255 gray image and distinguish ability into $m \times m$ (m usually multiples of 8) by using bi-cubic interpolation with the purpose of unifying the final hash sequence length. Use non-negative matrix to decompose, extract the basis matrix and weight matrix in the image matrix as the later stage dealing data.

Table 1. 8×8 Watson model of contrast sensitivity

1	1	1	1	2.	3.	4.	6.
1	1	1	1	2.	2.	3.	4.
1	1	2	2	2.	3.	4.	5.
1	1	2	3	4.	5.	6.	7.
2	2	2	4	6.	7.	8.	1
3	2	3	5	7.	9.	1	1
4	3	4	6	8.	1	1	1
6	4	5	7	1	1	1	2

3.3 DCT Transformation

Change the $m \times m$ image into small pieces of $n \times n$ and transform each piece with DCT. Each small pieces of image of such image matrix after treatment is decreasing in the direction of main diagonal, one DC coefficient and nine AC coefficients are kept and the rest are placed as zero, dispose the new generating matrix with visual model Watson.

3.4 Encryption

Deal with the matrix with standardization which makes it more robust. According to the iterative and no repetitive of data in chaotic area and initial value sensitivity, encrypt it by using Logistic equation—chaotic sequence generator[11]; a matrix is made by a key, use it to encrypt the DCT coefficient matrix to ensure security of matrix hash function.

Logistic equitation is as follows:

$$x_{n+1} = \mu x_n (1 - x_n) \qquad 3.5699 < \mu \leq 4 \tag{5}$$

3.5 Quantization, Encoding

Inter-block of one-dimensional prognosticate difference quantification was made to DC coefficients in each small pieces, two-dimensional prognosticate difference quantification was made to AC coefficient quantification, and initial value was not kept. Convert the eight DCT coefficients in each piece of quantization matrix to a decimal number, then carries on differential encoding to them. If the matrix data after quantization is still very large, the secondary compression coding can be used on it, the author adopts the second compression to 6 n×n small pieces.

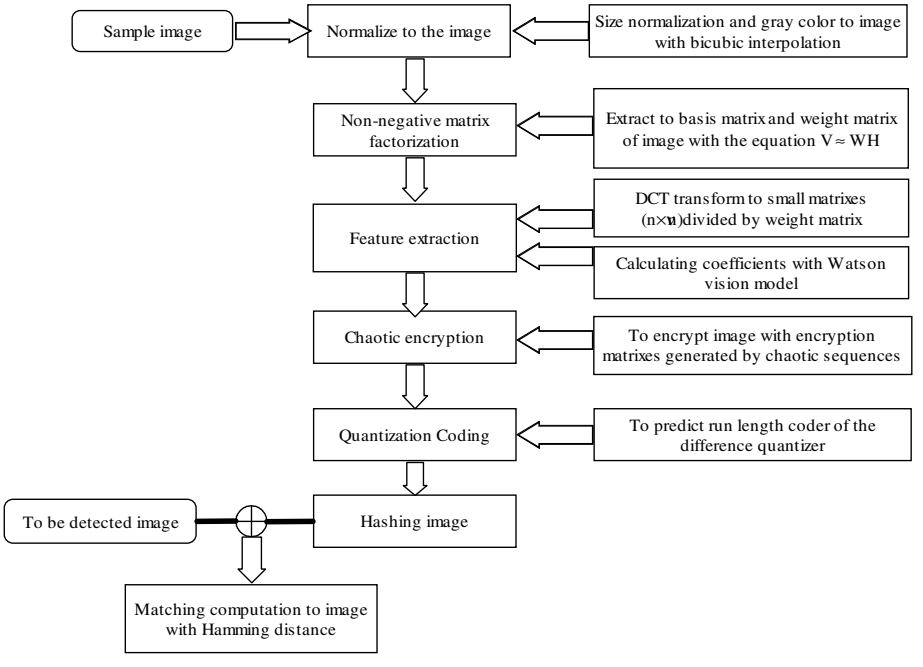


Fig. 1. Flow charts of realization steps

Forecasting difference formula is as follows:

$$\hat{x}_n = \sum_{i=1}^{k} a_i x_i \tag{6}$$

In it, a_i is a fixed constants called prediction coefficient. Pixel per inch k of the predictor is called prediction rank which has a direct impact on the performance of the prediction and the performance of high rank is superior to the low-rank predictor.

3.6 Image Matching

Matching through hamming code distance calculating formula, h_1 and h_2 are two hash sequences.

$$Dis\tan ce = \sum \left| h_1(i) - h_2(i) \right| \tag{7}$$

4 Experiment Results and Analysis

Next, we are going to text this case on crashworthiness, anti-aggressive, safety, Watson visuality and the sequence length.

4.1 Crashworthiness Experiment

The author selected 100 pieces of 128×128 color JPEG images, and gets 4950 group hamming distances in the 100 groups on hash series of alien pairs matching tests. Statistic histogram is shown in fig.2.

In fig.2, we can get that the experimental results is basically obey to normal distribution, mathematical expectation μ=185.78, standard deviation σ = 12.828, the author selects threshold T = 120.

$$P = \int_{-\infty}^{T} \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-(x-\mu)^2}{2\sigma^2}} dx = 1.4654\ e-007 \tag{8}$$

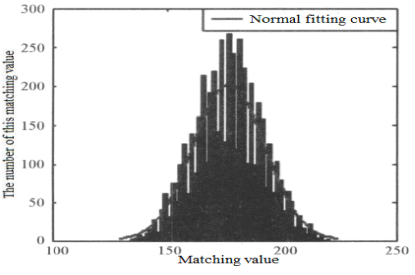


Fig. 2. Collision experiments

From the experiment results we can get that when T = 120, the collision probability is 10^{-7} which satisfy uniqueness.

4.2 Robustness Tests

(1) Robustness tests

With colored Lena.BMP image (512×512) to make the anti-attack test, processing the aggression image with ACDSee tool, the results shows as fig.3.

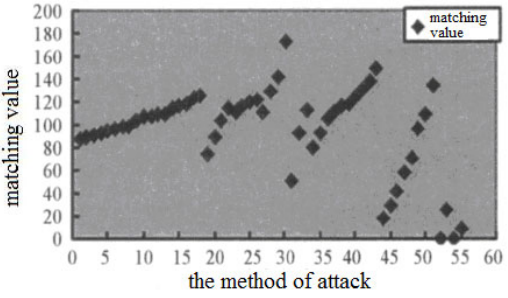


Fig. 3. Robustness tests

1) Select threshold =120, abscissa is 1 ~ 18 for shear degrees2% ~ 30% to make tests;

2) 19-26 is the fuzziness test with Gaussian method, fuzzy degrees is 5 ~ 40;

3) 27-30 is the noise adding test, the color degree of approximation of noise measuring point is about 200, strength is 50-200;

4) 31-33 is for the scale experiment, the scaling is 1/2, 1/4, 1/8;

5) 34-43 for 0.5 ~ 5.0 degrees of image rotation;

6) 44-51 is 10% ~ 80% localized tamper test under the condition of maintain the original image geometric shape;

7) 52-55 is for .png, .ras, .tif, .JPG format conversion testing respectively;

As can be seen from the graph:

a) This case basically meets the shear under 23%;
b) Gaussian fuzziness test below 40 degrees;
c) The noise added below 60 degrees;
d) Satisfy image constrain scaling;
e) Rotation below 3 degrees;
f) Localized tamper below 65%;
g) .png, .ras, .tif, .JPG format conversion.

Since the DCT transform is adopted to extract characteristic, the second method on the resist of geometric attack seems deficient, and we can try to combine the other feature extraction methods and DCT transformation to overcome the sensitivity of geometrical attack.

(2) Robustness test of comparison algorithm

The safety robust image perception hash technique of Zhang Weike is a new algorithm with good performance[9]; it is used to make comparisons with the algorithm in this paper. His algorithm uses the DCT transformation to extract image characteristic, too. First, the image is normalized; then the image character is extracted by DCT. Keep one DC component and three AC components; normalize the characteristic vector and encrypt it with pseudorandom sequence; and the final hash is generated by using Huffman to compression coding. Select the 0.3 as the threshold and the probability of collision is 0.29902e-8. The robustness test is denoted in fig.4.

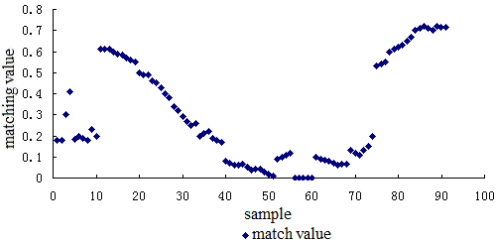


Fig. 4. Offensive test

In fig.4, 1~9 is the affine transformation, 10~39 is the shear of 1%~30%, 40~51 is the JPEG compression. 52~55 is the median filter, 56~60 is the adding noise, 61~71 is the scale variation, 72~91 is the rotation, vertical coordinates is the match value.

(3) Test contrast analysis

From table 2,we can get directly that when the Collision Probability is 0.31e-9, the algorithm in this paper can not only satisfy Gaussian Blur, format conversion, scale variation and select option attack but also has good resistance to localized tamper function. Moreover, it can resist 23% shear attack and adding noise below 60 degree. Therefore, the algorithm in the paper has been greatly improved and enhanced compared with the existent one.

Table 2. Robustness test contrast table

	Algorithm in this paper	Comparison Algorithm
Shear	below 23%	below 10%
Gaussian Blur	satisfy	satisfy
adding noise	below 60 degree	below 50 degree
scale variation	satisfy	satisfy
rotation operation	below 3 degree	below 3 degree
Localized tamper	below 65%	none
format conversion	satisfy	satisfy
Probability of Collision	0.31e-9	0.30e-8

4.3 Safety Test

Making use of the iteration non-repeatability and initial sensitivity of data in chaotic area, we can make encryption by using Logistic equation as a chaotic sequence generator. μ=3.9 take keys with 0.3256 and 0.3255 respectively to encrypt Lena BMP image.

Table 3. Hash sequence segment result from different keys of a picture

…	0	0	1	0	1	0	1	0	0	1	1	0	…
…	0	1	0	0	1	1	0	1	1	0	0	1	…

Through matching calculation, the hamming distance of two hash sequence is 312 which meet the safety requirements of hash sequence. We can't get harsh image without knowing the key; the correct hash sequence can't be obtained even know the

key but not the value of μ and the initial of DPCM coding. We get that using the chaos key method to encrypt is simple and effective [12]. If apply it into copyright protection and certification, the watermarking technique and perceiving hash image can be combined to ensure the safety of image.

4.4 Retrieval Test

In the image retrieval test, all the tested images are downloading from the http://wang.ist.psu.edu/. DD means the matching value between the said image and the original image, two groups of retrieval test results is shown in fig.5 and fig.6.

(a)original image (b)D=140 (c)D=128 (d)D=145 (e)D=129 (f) D=137 (g) D=149

Fig. 5. The first group image retrieval test

(a) original image (b)D=157 (c) D=162 (d) D=160 (e) D=170 (f) D=144 (g) D=147

Fig. 6. The second group image retrieval test

Due to the singleness of retrieval image in the first group, the matching results are ideal. The second group of image retrieval and background is relatively complex, if select threshold=160, we can see the retrieval result can also satisfy the requirements. Basically, the two images that matching value less than 160 can be thought with great perception, and the similarity is decreasing with the increasing matching value. Meanwhile, the selection of image retrieval also decided the retrieval results.

4.5 Watson Visual Test

Use the contrast sensitive chart in Watson visual model to process Lena BMP image. The situation after processing is shown in fig.7.

(a) (b) (c)

(a) Original image (b) Image after compress processing
(c) Image after compress and contrast sensitive chart processing

Fig. 7. Contrast of compress and visual model processing

Only 10 DCT coefficients are retained to make compression processing on the image, after contrasting fig.7 (a) and 7(b) we find that the images compressed didn't change the original image visual content; while contrasting fig.7 (b) and 7 (c), we find the image brightness is low down after the visual model procession. Because the brighter area in the image can be amended greatly without perception, the image after the visual model processing has higher perception.

In the future study, we can try to introduce other three sections in Watson visual model in addition to contrast sensitive chart to improve the perception feature of extraction.

4.6 The Sequence Length

When make the DCT transform to the image, we adopt the 8×8 pieces of image to make optimal region transform. Keep one DC coefficient at the left upper part and nine AC coefficient, then make extraction and the compression ratio is 32:5. and then make DPCM coding and do not retain their initial, after coding the matrix reduce 2 lines and 2 rows, the author regard one piece of eight DCT coefficient characteristics as a decimal number to DPCM coding and compression ratio is 8:1, an image of 512×512 last generation hash sequence length is 4092 bit and compression ratio is 1.56%. If the sequence is over length, we can make the second or third compressed. But compression too much will reduce the uniqueness of hash sequence.

5 Conclusions

Comprehensive of the current research status of the perception hash image, the author combine the DCT transformation, visual model and chaos theory together to generating hash sequence, the test results show that the method can satisfy the safety, robustness and uniqueness. Main jobs and conclusions are as follows:

1) Combining the DCT transformation and Watson visual model to extract image characteristics can effectively remove the redundant data of image and has stronger extraction.

2) Using the theory of chaos encryption which is convenient and effective overcome redundant information of pseudo random sequence encryption.

3) Using the differential forecast coding to generate the hash sequence will not make matching mistake because of tiny attack and abandon of initial value can further enhance the irreversible of hash sequence.

4) This method has a high performance of local anti-attack in the resistance of the content maintenance and high performance of extraction.

Further research work should focus on the following three aspects:

Firstly, study on the feature extraction method. Since the DCT transform extraction characteristic cannot overcome itself inherent defect: sensitive to geometric attacks. Yet there is no effective methods of using DCT transforms and satisfy the geometric attack greatly.

Secondly, study on the characteristics of the research methods. In the processing of lower DCT coefficients extraction, looking for a new method rather than the linear now.

Third, the research on feature matching method; in the feature matching ,use the calculation method reflect the characteristics of image visual rather than current calculation of hamming distance.

Acknowledgments. This work is supported by Science and Technology Research of Heilongjiang Province Educational Committee under Grant No. 12513090.

References

1. Niu, X., Jiao, Y.: An Overview of Perceptual Hashing. Acta Electronica Sinica 7(7), 1406–1410 (2008)
2. Venkatesan, R., Koon, S.M., Jakubow Skim, H.: Robust image hashing. In: Proc. IEEE Conf. on Image Processing, Vancouver, BC, Canada, pp. 664–666 (2000)
3. Mihcak, K., Venkatesan, R.: New Iterative Geometric Methods for Robust Perceptual Image Hashing, pp. 13–21. Springer, Heidelberg (2002)
4. Lin, C.Y., Chang, S.F.: A robust image authentication method distinguishing JPEG compression from maliciousmanipulation. Circuits and Systems for Video Technology 4(2), 153–168 (2001)
5. Kozat, S.S., Mihcak, K., Venkatesan, R.: Robust perceptual image hashing via matrix invariants. In: Proc. IEEE Conf. on Image Processing, USA, pp. 3443–3446 (2004)
6. Sebe, N., Tian, Q., Loupias, E.: Color Indexing Using Wavelet-based Salient Points. In: Proceedings of IEEE Workshop on Content-based Access of Image and Video Libraries, Hilton Head Island, South Carolina, pp. 15–19 (2000)
7. Bian, G., Zhang, X.: Fast retrieval of JPEG images compressed by DCT. Computer Applications 25(7), 1623–1625 (2005)
8. Tang, S.: Research on Multimedia Digital Signature Techniques, Graduate University of the Chinese Academy of Sciences, Doctoral Dissertation, pp. 24-36 (2006)
9. Zhang, W., Kong, X., You, X.: Secure and robust image perceptual hashing. Journal of Southeast University(Natural Science Edition) 37(9), 189–192 (2007)
10. Chen, M.-R., Chen, C.-B., Wei, C.-C.: An approach for solving multiple objective manufacturing problems. Int. J. of Computer Application in Technology 13(3/4/5), 132–137 (2000)
11. Dolen, M., Kaplan, H., Seireg, A.: Discrete parameter-nonlinear constrained optimisation of a gear train using genetic algorithms. Int. J. of Computer Applications in Technology 24(2), 110–121 (2005)
12. Hu, S.: Methodology of Safety Science Research. China Safety Science Journal 13(9), 1–4 (2003)
13. Khorshid, E., Seireg, A.: Discrete nonlinear optimisation by constraint decomposition and designer interaction. Int. J. of Computer Applications in Technology 12(2/3/4/5), 233–244 (1999)

A Message Driven Mechanism Based Distributed IDS Using Developed CUSUM Algorithm

Yingxuan Wang[1], Shuzhen Yao[1], Huobin Tan[2], and He Huang[2]

[1] School of Computer Science and Engineering BUAA Beijing, China
[2] School of Software BUAA Beijing, China
dxtk0812@hotmail.com, {szyao,thbin,huanghe}@buaa.edu.cn

Abstract. In this paper, we proposed a new division method of agent functions and introduced the message driven mechanism into IDS. We developed non-parametric CUSUM algorithm. Unlike previous proposals for usage of CUSUM algorithm in invading detecting, we presented a Deviations Regression Algorithm which could greatly improve performance of CUSUM algorithm. We demonstrated that we can achieve high detection accuracy on different network contexts.

Keywords: Message Driven Mechanism, Non-parametric CUSUM algorithm, Deny of service attacks, Distributed architecture, Deviations Regression Algorithm.

1 Introduction

With the requirements on intrusion detection system (IDS) continuously improving, distributed IDS has become a research hotspot. IDS which implemented based on multi-agent system (MAS) distributed framework can bring many advantages, autonomy and flexible for instance. There are many representative MAS based IDS (MIDS) such as DIDSt, NsTA, IDA, MAIDS, CARDS, NetSTAT, GdDS, AAFID, MADAM, CSM and EMILALD. Each of those IDS has different detection technology or different detecting targets, however, by analyzing the system structure we found that there some common module functions shared by those IDS. The common module functions can be divided into six kinds: Collector, Analyzer, Coordinator, Manager, Responser and User Interface.

Based on the modules above the agents of MIDS can be organized through three Views: Analyzer View (AV), Coordinator View (CV) and Manager View (MV). Each of the three views divided into two methods: centralized method and distributed method. Table 1 is the MIDS system structure classification.

As we can see from table 1, all of current MIDSs have Performance bottleneck. For example, the architecture of AAFID is not a mature structure, and it was implemented without focus on performance and agent security. Moreover, the agent organization mechanism of AAFID needs the data collect agents sending data to up level agents in real-time. This wastes network bandwidth resource, and makes performance of up level agents become the bottleneck of IDS. Moreover, MIDS which based on the system structure above needs the collector agents keep collecting data and the analyzer agents also need to keep analyzing, such workflow will consume system resources continuously.

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 811–819.
springerlink.com

Table 1. Classification of MIDS system structure

MV	DISTRIBUTED(D)			CENTRIALIZED(C)		
AV	D		C	D		C
CV	STATIC	DYNAMIC	/	STATIC	DYNAMIC	/
MIDS	CrIDS;AAFID; MADAM	CSM; EMRALD	/	IDA	MAIDS; NetSTAT CARDS	DIDS NSTAT

2 A Distributed IDS Based on Message Driven Mechanism

In this paper, we proposed a new division method of agent functions—Partition Division Method (PDM). In our method, each agent function is detecting one partition feature of certain intrusion type. Agents are in the same level, each agent's outputs do not represent any intrusion. Instead, the combination output of a group of agents will complete the intrusion detection task. Consequently, agent become independent from each other and this decoupling design makes system more flexible and scalable.

With the PDM based system structure we can apply message driven mechanism to our IDS. We classify agents of IDS by independent detect functions: Trigger agents and Single Feature Detect (SFD) agents. The Trigger agents are the crucial part of our system and used to analysis different statistic data of network traffic. The SFD agents are used to detect independent invasions' features such as network bandwidth, source IP address 6, etc. Fig. 1 illustrates the deployment structure of the IDS. Each agent uses the developed CUSUM algorithm, which will be discussed in next sections, to accomplish their tasks.

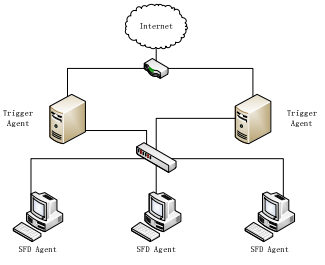


Fig. 1. Deployment structure of the IDS

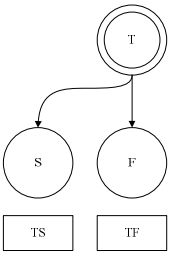


Fig. 2. Logical work mechanism of IDS

Unlike the agent organization mechanism of AAFID, agents in our system work independently, the real-time data transmitting is needed only when the Trigger agents have been triggered by the abnormal of network traffic or single feature network invasion. We define a feature code for each agent. For example, code 'T' represents a Trigger agent used to detect network traffic, code 'S' represents a SFD agent used to detect source IP address and code 'F' represents a SFD agent used to detect the SYN/FIN ratio. Fig. 2 illustrates the logical work mechanism of our IDS. When the Trigger agent 'T' detected network traffic abnormal IDS check the SFD agents according the map in Fig. 2. Then if some of the SFD agents also detected abnormal the IDS generated feature code sequence according the detect results of each agent. Each code sequence represents a kind of network invasion. In Fig. 2 code sequence 'TS' means Distributed Denial of Service (DDos) attack using fake source IP address, and 'TF' means SYNFlooding attack. In this way, each class of agents work independently in our IDS work, and the attacks features library is composed of code sequences, which is a simple character sequence so that we can build efficient index and maintain it easily. With the massage driven mechanism, our IDS do not need every agent keeping work. Agents who are not Trigger agent will not consume system resources until related trigger events come.

3 Nonparametric CUSUM Algorithm

Cumulative Sum 4 (CUSUM algorithm) is a widely used change-point algorithm in the industrial production process. This algorithm accumulates deviations (relative to a specified target) of incoming measurements $\{X_n\}_{n\geq 1}$ and alarms when the cumulative sum gets too large. The algorithm is based on this fact: If a change occurs, the probability distribution of the random sequence will change.

Generally, CUSUM requires a parametric model of the stochastic model being analyzed. However, in data network monitoring is the highly non-stationary characteristic of the data streams. Therefore, CUSUM algorithm needs to be transformed in order to adapt the network monitoring. B.E.Brodsky et.al [4] proposed a kind of non-parametric CUSUM algorithm which could adapt the network monitoring and Haining Wang et.al 5 proposed a mechanism for detecting SYN flooding attacks using non-parametric CUSUM algorithm. However, both of the non-parametric CUSUM mentioned above have a defect that makes deviations regression extremely slow, and this will cause IDS high false negative. We will discuss the defect and propose an effective solution in section 4.

Non-parametric CUSUM algorithm can be easily defined, but before we need some other definitions. We want to analyze a random sequence consisting of the number of new IP addresses in a time interval T. For simplicity, we define the random sequence X_n representing the percentage of new IP address in a time interval T. Usually the pattern presented in Fig. 3 will be observed during a Highly Distributed Denial of Service (HDDoS) attack:

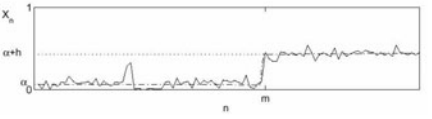


Fig. 3. Behavior of the percentage of new IP addresses during an attack

During the attack there is a step change of the mean value of X_n at m from α to α+h. h will be defined as the minimum increase of the mean value of X_n during an attack (it is not the threshold for the bandwidth attack detection).However, the sequence X_n is related to the network scale and the sampling time. So we need to generalize the sequence. We defined a normalize formula as descript below: Set J is the mean of sequence member in the detecting time window, and this value could change in real-time. The recursive definition of J is:

$$J(n) = \lambda J(n-1) + (1-\lambda) X_n \quad (0<\lambda<1) \tag{1}$$

Then, we set $\Delta n = X_n / J$. New sequence Δn is no longer connected with network scale and the sampling time. One of the assumptions for the non-parametric CUSUM algorithm is that mean value of the random sequence is negative during normal conditions, and become positive when a change occurs. Consequently, it is necessary a transformation of Δn into a new sequence Z_n. The transformation is given by $Z_n = \Delta n - \beta$, where β is a constant such that $a = \alpha - \beta < 0$. In Fig. 4, it is possible to observe the comon pattern of a Z_n sequence during an attack, the major part of the values of the sequence must be negative during normal conditions. The parameter β must be set according to the network normal conditions.

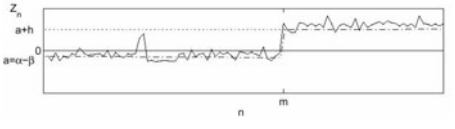


Fig. 4. Behavior of sequence Z_n

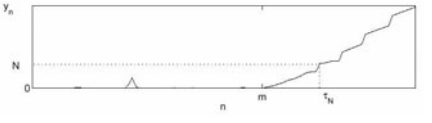


Fig. 5. recursive version of the CUSUM

The problem of the online detction of attacks is solued using a recursive version of the non-parametric CUSUM. It can be defined using another new sequence y_n:

$$\begin{cases} y_n = (y_{n-1} + Z_n)^+ \\ y_0 = 0 \end{cases} \tag{2}$$

Where $(y_{n-1}+Z_n)^+$ means that if $y_{n-1}+Z_n>0$ the $(y_{n-1}+Z_n)^+ = y_{n-1}+Z_n$, otherwise $(y_{n-1}+Z_n)+ = 0$. Fig. 5 illustrated the application of the recursive version of the non-parametric CUSUM in the previous analyzed data. A large value y_n is a strong indication of an attack, and a constant N must be adjusted according to the network condition s as a threshold for attacks detection.

4 Deviations Regression Algorithm

One characteristic of the algorithm mentioned above is that it shows the increase of the increase of the number of new IP addresses but hte scheme can not detect accurately the end of the attack. This behaviour would not be very important unless the fact that the Internet traffic has a lot of unexpected events. For example, suppose that we have only for an instant an impulsive increase of the number of the number of new IP addresses. So, the values of the sequence y_n will increase. Now, let's suppose the abnormal event occurs again. The CUSUM algorithm decreases very slowly and because of the second event it will possibly detext a false alarm of intrusion. This can be illustrated in the Fig. 6 and Fig. 7, where it is possible to see that because of the first event the second event seems to have a greater number of new IP address than actually it has.

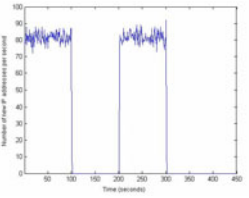


Fig. 6. Number of new IP addresses

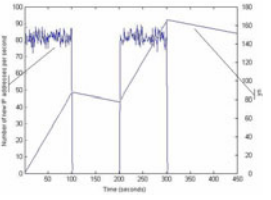
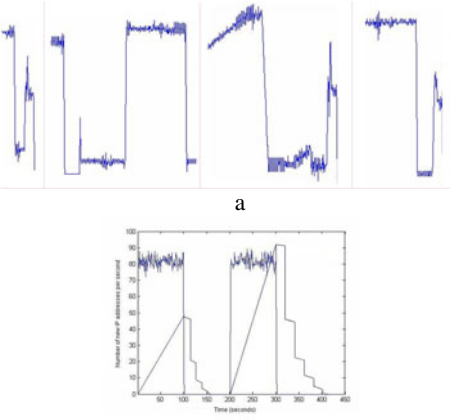


Fig. 7. Slow decrease of the algorithm

We researched four kinds of typical DDoS attacks' traffic trace in order to find a method to accelerate the decreasing of CUSUM algorithm with out any side effects. The Fig. 8 illustrates traffic traces with constant rate attack, pulsing attack, increasing rate attack and gradual pulse attack during the end of these attacks. As we see in the picture, the traffic traces will experience a dramatic decline after the DDoS attacks stroped. The traffic trace decline slope during this period is very close to tan(-π/2) as we researched. Therefore, we definet a slope threshold γ (γ<0), and the formula of γ is:

$$\left|\gamma-(-\frac{\pi}{2})\right|\le\varepsilon\left(\varepsilon>0\right) \tag{3}$$

We calculate the decline slope of the traffic trace in a time window, and if the value of decline slope is less than the slope threshold continuously we could guess that the attacks may be stop. Then we half the y_n tentatively, if the attacks really have stopped we already accelerate the regression of y_n. Otherwise, the value of y_n will become greater if the attacks is not finished, and the action of regression accelerating would not bring false positives to IDS.



a

b

Fig. 8. a. Traffic traces of four kinds of typical attacks; b. Accelerated decrease of the algorithm

In conclusion, we propose a Deviations Regression Algorithm (DRA) as an add-on algorithm of CUSUM algorithm. The idea of the DRA is similar to the congestion control method used in TCP protocol.We set a time window for y_n (YnWindow), and calculate the slope change rate of the y_n function in this window. If y_n function-change-rate slope appears negative in Q interval ΔT_k and the negative slope value is less than slope threshold γ (γ<0), we half the y_n value, the formulas are as below:

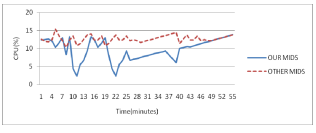
$$\Omega_k=(y_k-y_{k-1})/\Delta T_k(k=0,1,2...)\qquad Count_k\ge Q\Rightarrow y_n=y_{n-1}/2 \tag{4}$$

$$Count_k=\begin{cases}Count_{k-1}+1(\Omega_k\le 0 and\,|\,\Omega_k-\Omega_{k-1}\,|\ge\gamma)\\ 0(k=0 or\,|\,\Omega_k-\Omega_{k-1}\,|<\gamma or\Omega_k>0)\end{cases}(\gamma\ge 0,k=0,1,2...) \tag{5}$$

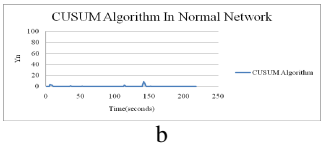
DRA could accelerate the deviations regression efficiency, and it could improve the CUSUM algorithm obviously. Fig. 8.b shows the performance of CUSUM algorithm with Deviations Regression Algorithm. DRA makes CUSUM algorithm detect finish-point of attacks more precisely, especially in distinguishing the pulsing attack. The performance of CUSUM algorithm with DRA is the practical application will be discussed in next section.

5 Performance Evaluation

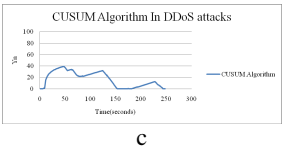
To evaluate the efficacy of our detection scheme, we conducted the following experiments in the network topology described in Fig. 1 with a 100Mbps internet access link. Agents are the vector of our developed CUSUM algorithm. Firstly, we compared the CPU occupation rate of our MIDS with other MIDS. As Fig. 9.a illustrates, the solid line represents our MIDS' CPU occupation rate and the dotted line represents other MIDS' data. It is obviously that our MIDS can save more system resources.



a



b



c

Fig. 9. a.CPU occupation rate ; b. Behavior of algorithm in normal; c. during a pulsing attack

As we discussed in section 3, the value of y_n is the kernel indicator of CUSUM algorithm for attack detection. Consequently, we could evaluate the algorithm's efficacy by tracing the change of y_n during the execution of the IDS with CUSUM algorithm. Secondly, we use the algorithm in normal network traffic, and the value of y_n changes around 0 (as the Fig. 9.b shows) which means the network traffic is in a safe-state. Thirdly, we use the algorithm in the pulsing attack context, and as the Fig. 9.c illustrates that the trace of y_n increasing continuously when the attack is processing.

Finally, in order to evaluate the Deviations Regression Algorithm conperhansively we run CUSUM algorithm with DRA (CUSUMD) and CUSUM algorithm without DRA (CUSUMN) in two agents respectively. Both of the agents are running in the same service and detect same source of the network traffic simultaneously. Moreover, we test two algorithms in four kinds of network context, including constant rate attack, pulsing attack, increasing rate attack and gradual pulse attack. Fig. 10.a, Fig. 10.b, Fig. 10.c and Fig. 10.d presented the trace of y_n in four kinds of network context mentioned above. The solid line represents the CUSUMD, and the dotted line represents the CUSUMN.

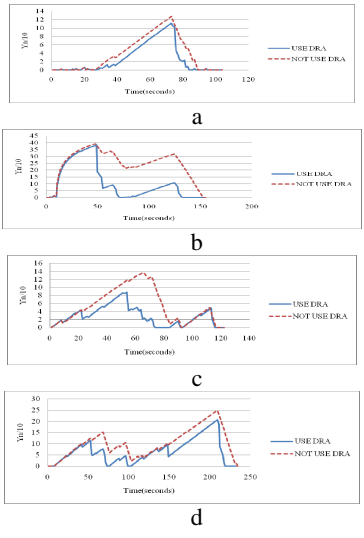


Fig. 10. Behavior of algorithm during a. constant rate attack; b. pulse attack; c. increasing rate attack; d. gradual pulse attack

In Fig. 10.a we can see that attack stopped at second 78, and the CUSUMD cancel the attack alarm at second 82, which is 50% earlier than the CUSUMN. In Fig. 10.b, the second attack takes place at second 80, 30 seconds after the first one. This picture clearly shows the advantage of CUSUMD which could detect attacks precisely, but the CUSUMN takes two attacks as a single one. Fig. 10.c illustrates that CUSUMD detected 3 attacks in increasing rate attack correctly, the CUSUMN on the other hand only detect 2 attacks. Fig. 10.d shows there is traffic disturbance at second 75. The CUSUMD could distinguish it from the attack. The CUSUMN, on the contrary, detect it as an attack.

6 Conclusion

In this paper we proposed a scheme of distributed IDS using non-parametric CUSUM algorithm. Moreover, we presented a Deviations Regression Algorithm which could greatly improve the performance of non-parametric CUSUM algorithm. The experimental results show that using DRA as an add-on algorithm of CUSUM algorithm could exactly detect the end of the attack, and the improvement is greater especially in the pulsing attack, increasing rate attack and gradual pulse attack context.

References

1. Jiang, J., et al.: A Survey of Network Security Intrusion Detection. Journal of Software 11(11), 1455–1460 (2000)
2. Ballasubramaniyan, J.S.: e.tal. Autonomous agents for intrusion detection. CT Report (1998)

3. Bellifemine, F.: Developing Multi-Agent Systems with JADE. John Wiley & Sons, Ltd. (2007)
4. Brodsky, B.E., et al.: Nonparametric Methods in Change-point Problems. Kluwer Academic Publishers (1993)
5. Wang, H., et al.: Detecting SYN flooding attacks. In: IEEE Infocom 2002 (June 2002)
6. Tao, P., et al.: Detecting distributed denial of service attacks using source IP address monitoring (November 2002)
7. FIPA, Foundation for Intelligent Physical Agents, `http://www.fipa.org`

High-Quality Resolution Reduction Method for Halftone Image

Kaoru Ishito[1], Kazuki Katagishi[1], and Kazuo Toraichi[2]

[1] Graduate School of Systems and Information Engineering, University of Tsukuba, 1-1-1 Tennodai, Tsukuba-shi, Ibaraki, 305-8573, Japan
[2] Fluency R&D Labs., 3-1-1-1214, Azuma, Tsukuba-shi, Ibaraki, 305-0031, Japan
ishito@wslab.risk.tsukuba.ac.jp

Abstract. This paper proposes a new resolution reduction method for halftone images. Recently, smartphone or tablet PC has become widespread. When viewing comics on mobile devices as e-books, image resolution conversion is required, because there are various resolution devices. If comic images are reduced with conventional resolution conversion methods using interpolating functions such as bicubic interpolation, moiré which leads to deterioration in image quality is generated, because comic images include halftones with periodic structures. In this paper, we propose a resolution reduction method for halftone images which preserve periodic structures of original images. Experiments show that a moiré is not generated with our proposed method, and the method is superior in subjective image quality and computational time to the conventional bicubic interpolation.

Keywords: Resolution conversion, halftone, moiré, interpolating function, bicubic, fluency DA functions, bifluency interpolation.

1 Introduction

Recently, as the popularity of smartphone or tablet PC, the opportunity to view comics as e-books on portable devices has increased. When viewing comics on mobile devices, image resolution conversion is required because of distribution of various resolution devices in the market.

The methods using interpolating functions such as nearest-neighbor, bilinear, bicubic and Lanczos interpolation [1][2] have been conventionally used to convert image resolution. However, if comic images are reduced with the conventional resolution conversion methods, a moiré that leads to deterioration in image quality is generated after resolution reduction, because comic images include halftones, which simulate continuous tone imagery. Therefore our goal is to propose a high-quality resolution reduction method for halftone images that does not generate any moirés.

If conventional resolution conversion methods using interpolating functions are used, one cycle pattern of an original halftone image might not correspond to that of a low-resolution image. In this case, the periodic structure of the original image collapses, and a different pattern, which means moiré, is generated in the low-resolution image.

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 821–828.
springerlink.com

To propose a high-quality resolution reduction method for halftone images, it is necessary to preserve a periodic structure of an original image. Thus, we focus that halftone images are composed of periodic array of one cycle pattern. In this paper, we propose a method generating one cycle pattern of a low-resolution image from that of an original image, and composing a low-resolution image by arranging it periodically. When generating one cycle pattern of a low-resolution image from that of an original image, we use bifluency interpolation [3] using Fluency DA Function of Degree 2 [4] based on "Fluency Information Theory" [5]. Bifluency interpolation provides less staircase noise than bicubic interpolation. Through experiments, it is demonstrated that with our proposed method, a moiré is not generated for an arbitrary magnification. And it is shown that our proposed method is superior in subjective image quality and computational time to conventional bicubic interpolation.

2 Mechanism of Generation of Moiré with Conventional Method

2.1 Basic Matters of Halftone

The halftone [6] is a pattern of dots to simulate continuous tone imagery by binary in the offset printing. The area ratio of a binary pattern is pseudo-perceived as continuous tone because of a human visual performance.

The halftone is divided into two classes, AM (amplitude modulation) screen and FM (frequency modulation) screen. Former represents tone by changing area of dots, latter by changing the number of dots per unit area. The AM screen is commonly used to create comics. The AM screen as a digital image can be considered to be composed of a periodic array of one size ($P \times Q$ pixel) images. An example of a halftone image and its one cycle pattern are shown in Fig. 1.

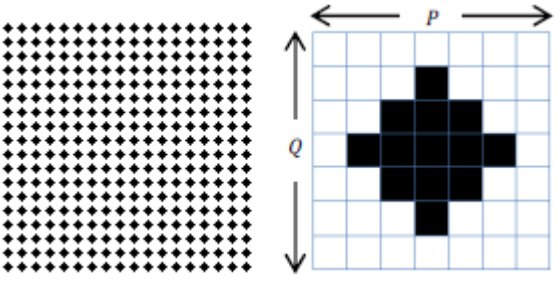


Fig. 1. Halftone image at left, its one cycle pattern at right

2.2 Mechanism of Generation of Moiré

When a halftone image having periodicity is reduced with conventional methods using interpolating functions, a moiré which is a pattern different from the original image is generated. In this section, we describe that mechanism.

The conventional resolution conversion methods using interpolating functions convert the pixel coordinates of the output image to these of the original image and

decide the pixel values of the output image by interpolation using the interpolating functions. There are some interpolating functions, nearest-neighbor, bilinear, bicubic or Lanczos. Bicubic interpolation is most commonly used.

The halftone is composed of the periodic array of one cycle pattern. We assume the size of one cycle pattern of a halftone image to be $P \times Q$ pixel, and the reduction magnification to be $\alpha = b/a$. If both P and Q are integral multiples of a, since both αP and αQ can be integers, the $\alpha P \times \alpha Q$ pixel-pattern is generated after resolution reduction. Therefore it corresponds to the original $P \times Q$ pixel-pattern and a moiré cannot be generated. In contrast, if either P or Q is not an integral multiple of a, since either αP or αQ cannot be an integer, $\alpha P \times \alpha Q$ pixel-pattern cannot be generated, because coordinates of a pixel must be integers. Therefore, the pattern different from the original image, which is moiré, is generated after resolution reduction.

An example is shown in Fig. 2. If a halftone image (a) composed of a pattern (c) (7×7 pixel) is reduced to half along each dimension with methods using interpolating functions, since $\alpha P = \alpha Q = 7/2$ (not integers), the 7/2×7/2 pixel-pattern cannot be generated. Instead of that size pattern, the 7×7 pixel-pattern (d) which is twice along each dimension bigger than it and includes four patterns of original image (c) is generated. Thus, the low-resolution image (b), which has a moiré, is generated.

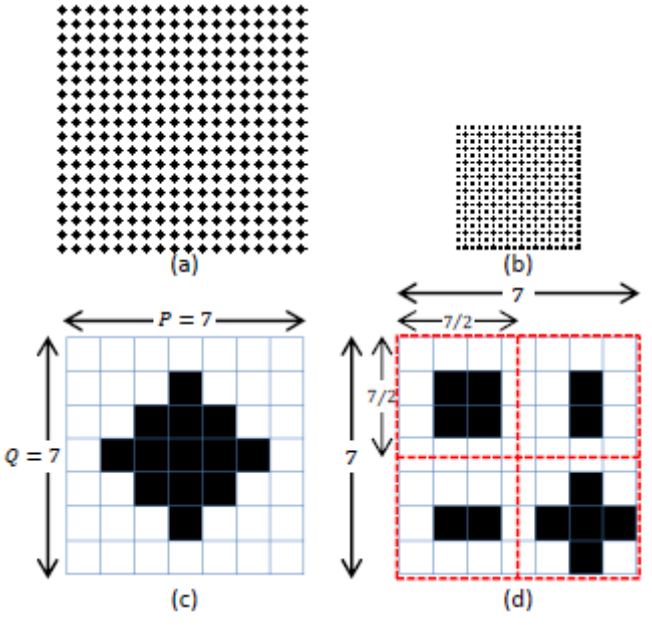


Fig. 2. (a) Original halftone image, (b) Reduced image with bicubic interpolation (Magnification is 1/2 along each dimension.), (c) One cycle pattern of original image, (d) One cycle pattern of reduced image (Dashed lines show the areas corresponding to one cycle pattern of original image.)

3 Proposed Method

In this chapter, we describe our proposed method according to the list of symbols Table 1. Where $\lfloor a \rfloor$ is the maximum integer which is not greater than a.

When using conventional method, if either αP or αQ is not an integer, a $\alpha P \times \alpha Q$ pixel-pattern cannot be generated and a moiré is generated. Thus, we focus that a halftone image is composed of a periodic array of one cycle pattern. In our proposed method, we generate one cycle pattern of a low-resolution image from that of an original halftone image, and generate a low-resolution halftone image by arranging it periodically. When generating one cycle pattern of a low-resolution image from that of an original image, we use bifluency interpolation that our research group has proposed. It is one of the resolution conversion methods using interpolating functions.

Bifluency interpolation uses Fluency DA Function of Degree 2,

$$ {}^{3}_{[c]}\psi(t) = \begin{cases} -\frac{t^2}{4\tau^2} - \frac{t}{\tau} - 1, & -2\tau \le t < -\frac{3}{2}\tau \\ \frac{3t^2}{4\tau^2} + \frac{2t}{\tau} + \frac{5}{4}, & -\frac{3}{2}\tau \le t < -\tau \\ \frac{5t^2}{4\tau^2} + \frac{3t}{\tau} + \frac{7}{4}, & -\tau \le t < -\frac{\tau}{2} \\ -\frac{7t^2}{4\tau^2} + 1, & -\frac{\tau}{2} \le t < \frac{\tau}{2} \\ \frac{5t^2}{4\tau^2} - \frac{3t}{\tau} + \frac{7}{4}, & \frac{\tau}{2} \le t < \tau \\ \frac{3t^2}{4\tau^2} - \frac{2t}{\tau} + \frac{5}{4}, & \tau \le t < \frac{3}{2}\tau \\ -\frac{t^2}{4\tau^2} + \frac{t}{\tau} - 1, & \frac{3}{2}\tau \le t \le 2\tau \\ 0, & otherwise\ , \end{cases} \tag{1} $$

Table 1. List of Symbols

Fluency DA Function of Degree 2	${}^{3}_{[c]}\psi(t)$
Two dimensional fluency DA function	$\psi(x,y) = {}^{3}_{[c]}\psi(x) \cdot {}^{3}_{[c]}\psi(y)$
The size of the input halftone image	$I \times J$ pixel
Pixel values of the input image	$L(i,j)$, i=0,1,…,I-1, j=0,1,…,J-1
The size of one cycle pattern of the input image	$P \times Q$ pixel
Pixel values of the one cycle pattern of the input image	$H(p,q)$, p=0,1,…,P-1, q=0,1,…,Q-1
Reduction magnification	$\alpha = b/a$
The size of output (low-resolution) image	$\lfloor \alpha I \rfloor \times \lfloor \alpha J \rfloor$pixel
Pixel values of the output image	$\tilde{L}(\tilde{i},\tilde{j})$, $\tilde{i} = 0,1,\cdots,\lfloor \alpha I \rfloor - 1$, $\tilde{j} = 0,1,\cdots,\lfloor \alpha J \rfloor - 1$
The size of one cycle pattern of the output image	$\lfloor \alpha P \rfloor \times \lfloor \alpha Q \rfloor$pixel
Pixel values of one cycle pattern of the output image	$\tilde{H}(\tilde{p},\tilde{q})$, $\tilde{p} = 0,1,\cdots,\lfloor \alpha P \rfloor - 1$, $\tilde{q} = 0,1,\cdots,\lfloor \alpha Q \rfloor - 1$

based on "Fluency Information Theory" as an interpolating function. Where τ means a sampling interval. For the remainder of this paper, τ is set to 1. Fluency DA Function of Degree 2 is a sort of sampling function which consists of piecewise quadratic polynomial. Since this function converges to 0 at 2 point after and before the sampling point, no truncation error is provided in applications. Actually, this function is applied to digital-analog converter of CD player and DVD-Audio player known as "Fluency Audio". Due to this function, Fluency Audio can generate supersonic component, which conventional sinc function cannot generate. Therefore Fluency Audio has won many awards.

The steps of our proposed method are as follows.

Step 1: Generation of One Cycle Pattern of a Low-Resolution Image

First of all, cut one cycle pattern $H(p,q)$ (P×Q pixel) out of original image. Where the values of P and Q are given manually. Then, generate a $\lfloor \alpha P \rfloor \times \lfloor \alpha Q \rfloor$ pixel-image $\tilde{H}(\tilde{p},\tilde{q})$ from it. That is, reduce the P×Q pixel-image $H(p,q)$ to $\lfloor \alpha P \rfloor / P$ and $\lfloor \alpha Q \rfloor / Q$ times along each dimension with bifluency interpolation as follows.

$$\tilde{H}(\tilde{p},\tilde{q}) = \sum_{p=0}^{P-1}\sum_{q=0}^{Q-1} H(p,q)\psi\left(\frac{\tilde{p}}{\frac{\lfloor \alpha P \rfloor}{P}} - p, \frac{\tilde{q}}{\frac{\lfloor \alpha Q \rfloor}{Q}} - q\right), \quad \tilde{p} = 0,1,\cdots,\lfloor \alpha P \rfloor - 1, \tilde{q} = 0,1,\cdots,\lfloor \alpha Q \rfloor - 1 . \tag{2}$$

Step 2: Construction of Low-Resolution Image

Regard the $\lfloor \alpha P \rfloor \times \lfloor \alpha Q \rfloor$ pixel-image $\tilde{H}(\tilde{p},\tilde{q})$ generated in step 1 as one cycle pattern of a low-resolution image. Then compose a low-resolution image by arranging it periodically. That is, the output pixel values $\tilde{L}(\tilde{i},\tilde{j})$ are decided as follows.

$$\tilde{L}(\tilde{i},\tilde{j}) = \tilde{H}(\tilde{i} \bmod \lfloor \alpha P \rfloor, \tilde{j} \bmod \lfloor \alpha Q \rfloor), \quad \tilde{i} = 0,1,\cdots,\lfloor \alpha I \rfloor - 1, \tilde{j} = 0,1,\cdots,\lfloor \alpha J \rfloor - 1 . \tag{3}$$

Where a mod b means the remainder on division of a by b.

4 Experiments and Results

4.1 Experimental Condition

We reduced two images shown in Fig. 3 (200×200 pixel, binary) with our proposed method and bicubic interpolation. The size of one cycle pattern of the image-1 is 7×7 pixel and that of the image-2 is 12×12 pixel. Reduction magnifications are 80%, 60% and 40% along each dimension. We implement these algorithms with Microsoft Visual C++ 2008 and execute on the environment with Intel Core2 Duo E8400 CPU, 998MB RAM.

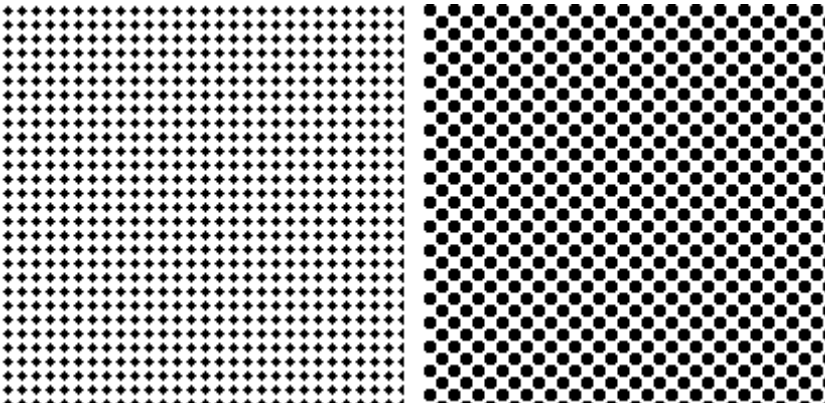


Fig. 3. Original image-1 at left, original image-2 at right

4.2 Experimental Results

Figures 4 and 6 show result images with our proposed method. Figures 5 and 7 show result images with bicubic interpolation. A moiré is not generated in any reduction magnification for each image with our proposed method. In contrast, a moiré is generated in all the reduction magnification for both images with bicubic interpolation. Especially a moiré stands out in the reduction magnification 40%. Our proposed method is superior in subjective image quality to bicubic interpolation.

Table 2 shows the comparison of computational time between our proposed method and bicubic interpolation. The computational time of our proposed method is about 1.7%-3.2% compared to that of bicubic interpolation. The reason why our proposed method is faster than bicubic interpolation is that the former reduces only one cycle pattern of a halftone image, in contrast, the latter reduces the entire halftone image.

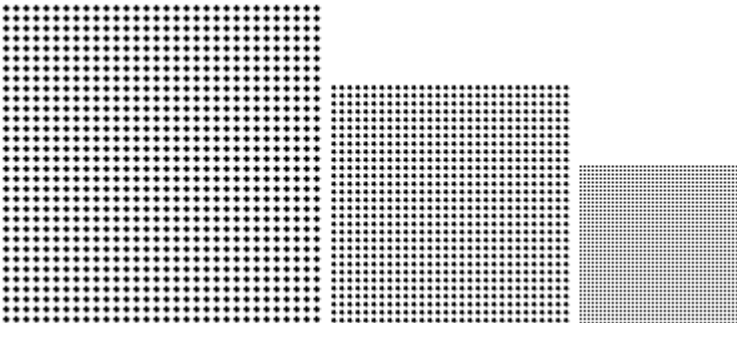


Fig. 4. Reduced images of image-1 with our proposed method, 80% at left, 60% at center, 40% at right

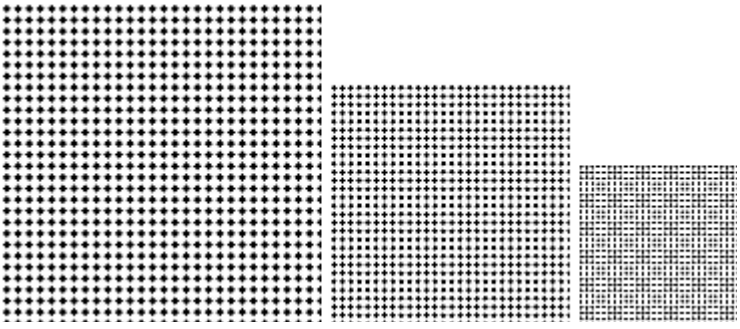


Fig. 5. Reduced images of image-1 with bicubic interpolation, 80% at left, 60% at center, 40% at right

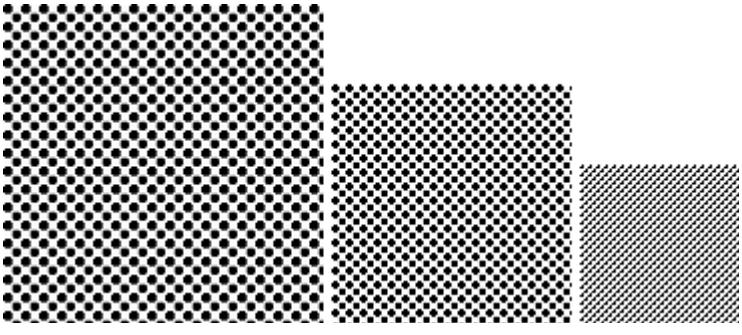


Fig. 6. Reduced images of image-2 with our proposed method, 80% at left, 60% at center, 40% at right

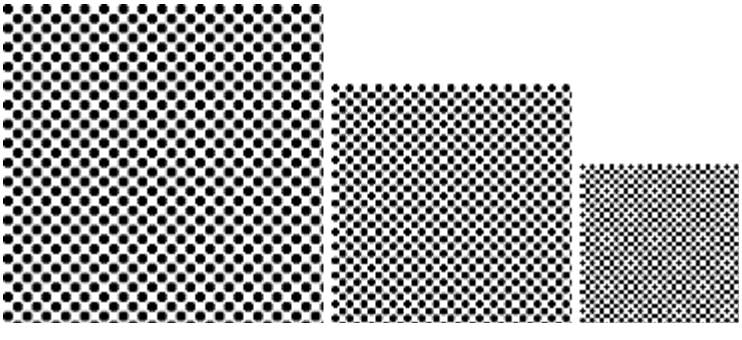


Fig. 7. Reduced images of image-2 with bicubic interpolation, 80% at left, 60% at center, 40% at right

Table 2. Comparison of computational time (msec)

	Image-1			Image-2		
Reduction Magnification	80%	60%	40%	80%	60%	40%
Proposed	1.7	1.0	0.6	1.9	1.1	0.3
Bicubic	59.7	61.8	18.1	107.3	37.8	18.1

5 Conclusion

In this paper, we proposed a high-quality resolution reduction method for halftone image which does not cause generation of moiré. Conventional resolution reduction methods using interpolating functions cause generation of moiré that leads to deterioration in quality. The reason of generation of moiré is that one cycle pattern of an original halftone image does not correspond to that of a low-resolution image. In our proposed method, we generate one cycle pattern of a low-resolution image from that of an original halftone image and compose a low-resolution image by arranging it periodically. Through experiments, it is shown that our proposed method is superior to bicubic interpolation because our proposed method does not generate a moiré and computational time of our proposed method is faster than that of bicubic interpolation.

In our proposed method, the size of one cycle pattern of an original halftone image $P \times Q$ pixel is given manually. It is necessary to recognize them automatically to apply this method to an actual application. Furthermore, our proposed method targets only a halftone image whose density is uniform. We should consider how to apply this method to a halftone image whose density is not uniform.

References

1. Park, S.K.: Image Reconstruction by Parametric Cubic Convolution. Computer Vision, Graphics, and Image Processing 23, 258–272 (1983)
2. Turkowski, K.: Filters for Common Resampling Tasks. Graphics Gems I, 147–165 (1990)
3. Fukui, D., Katagishi, K., Ohmiya, Y., Toraichi, K.: Image Interpolation Technique Using The Space-variant Compactly Supported Fluency DA Function of Degree 2. In: Proceedings of The Eighth IASTED International Conference on Signal Processing, Pattern Recognition, and Applications, pp. 173–179 (2011)
4. Toraichi, K., Nakamura, K.: Sampling Function of Degree 2 for DVD-Audio. The Transactions of the Institute of Electrical Engineers of Japan 123-C(5), 928–937 (2003)
5. Nakamura, K.: Fluency AD/DA Functions-based Signal Processing and Their Applications (in Japanese). Doctoral Dissertation of University of Tsukuba (2007)
6. Wesner, J.W.: Screen Patterns Used in Reproduction of Continuous-Tone Graphics. Applied Optics 13(7), 1703–1710 (1974)

Resource Pool-Oriented Resource Management for Cloud Computing

Zhang Hengxi[1,2], Li Chunlin[1], Shi Zhengjun[2], and Zhang Xiaoqing[1]

[1] Department of Computer Science, Wuhan University of Technology, Wuhan, P.R China
[2] Department of Fundamental Courses, Xuzhou Air Force Academy, Xuzhou Jiangsu, China
{Zhanghengxi,szjkjhx}@126.com, chunlin74@yahoo.com.cn, zxqtzy@whut.edu.cn

Abstract. The capacity setting of the resource pool is the key issue in the resource cost-oriented computing resource management. The problem of cloud computing resource management exploring resource pool is discussed based on the queue theory and the global optimization theory, and a computing method of the optimal capacity of the resource pool in cloud computing is presented.

Keywords: cloud computing, resouce management, resource pool, Poisson process.

1 Introduction

Cloud computing users choose and gain the resource in largely depending on the resource availability and the price. This has certain things in common with the maximum- benefit principle which is mainly concerned in the market economy model for both sides of users and resource suppliers, a resource is chosen according to the resource cost [1]. Based on this, this paper discusses cloud computing resource management for the resource pool from a theoretical point of view.

Cloud computing resources are heterogeneous and geographically distributed, and can be provided by multiple organizations, which increases the difficulty of resource management [2]. To bonding together the resource providers scattered around, the concept of resource pools is introduced in this paper, which is allocated and scheduled in a unified way by a specific management module of cloud computing system [3]. The mode of scheduling and allocation of available resources in the resource pool is not fixed. It can be allocated and scheduled according to the specific system requirements, or following certain guidelines to establish priorities for certain resources, and the allocation and management of resources can be further optimized according to some guidelines.

The resource pool is a centralized strategy for resource allocation and management. With the concept of the resource pool, the whole system can be taken into account further and can be comprehensive measured under the current state of the resource utilization, quality of service and load balance, etc.

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 829–832.
springerlink.com

Resource users want to minimize costs, and resource providers want to maximize the return on investment in the cloud computing environment [4]. In theory, the behavior of the resource pool in cloud computing can be described as follows. Suppose that there are S units of resources (such as S virtual machines) in the resource pool of the cloud computing. When the system is running, the resources are constantly consumed, and the users can add orders from other suppliers to meet their needs. Due to the demand of the resource quantity and the arrival time of the order, the number of the resources in resource pool change randomly. How to set the capacity of the resource pool? If configured a too large resource pool capacity, the resources are bound to be wasted, and the construction cost is high, simultaneously, the corresponding storage cost becomes higher with more storage resources. On the contrary, if configured a too small resource pool capacity, resource utilization becomes low with too little storage capacity, which led to system losses. Therefore, the setting of the resource pool capacity becomes a key issue.

2 Construction of Resource Pool

Internet resources have the natural characteristics of dynamics, growth and autonomy. If only the minimum amount of necessary resources is used to provide services for the application, it is hard to guarantee the quality of service for the application submitted by users because of the resource autonomy and network instability. And if all resources are used to provide services for an application, which is contrary to the original intention that cloud computing is used to share the resource and improve its utilization. And secondly, it may affect the efficiency of whole system because of the surge in traffic between resources. Therefore, it is need to select an appropriate scale set of resources, that is, resources pool, to provide services for users. This not only will cause too much communication overhead in the system, but also have some fault-tolerance and deformation-tolerance, which will support to provide different services for different users.

It is allowed to establish resource pools for multiple applications accordingly in cloud computing environment. In order to facilitate the management of the resource pool, the description of the resource pool should include the following content: resource pool identifier (unique), scale of resource pool, information table of resources and statement of resource pool, etc. The scale of resource pool is decided by both the service type and the request-level of users. The information describing each resource pool is saved in the special unit of cloud computing centers, and to be maintained and updated periodically.

3 Optimal Capability of Resource Pool

Cloud resource requirement process is a Poisson process [5]. Assuming the Poisson intensity is λ, the quantity demanded is 1, with its probability q_1, the quantity is 2, with its probability q_2,…, and the quantity is N, with the probability q_N (N is the maximum quantity allowed to apply). Assuming the maximum storage amount of resource is S. At the initial time t=0, the resource pool is filled with S units of resource.

And then, whenever a user request occurs, as long as the resource pool have a reserve meeting the needs of the request, that is to be met, otherwise waits in the queue. Meanwhile, whatever there are some resources or not in the resource pool, as long as the user request occurs, assuming that the quantity requested is K units, the K-unit order should be issued immediately to supplement the reserves of resource pool or wait for request. The ordering service is achieved by a specialized service system. The storage mode of resource pool is shown in figure 1.

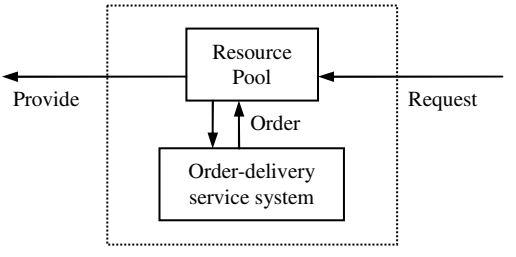


Fig. 1. Storage mode of resource pool

For simplification, only four kinds of resource costs of cloud computing are considered. The storage cost per resource unit in resource pool is e_1, the maintenance cost per resource unit in resource pool is e_2, the shortage cost of per resource unit within shortage time unit is e_3, the shortage cost of per resource unit is e_4. Let $f(S)$ be the total expectation cost in per unit time, p_n is the probability of the number of resources ordered but not delivered n, then [6][7]

$$f(S) = e_1 \sum_{n=0}^{S} (S-n)p_n + e_2 S$$
$$+ e_3 \sum_{n=S}^{\infty} (S-n)p_n$$
$$+ \lambda e_4 \sum_{k=1}^{N} kq_k \sum_{n=S}^{\infty} p_n$$

To find the optimal resource pool capacity means to find S* which makes a minimum worth $f(S)$.

When S* exists, an algorithm for the calculation of S* can be as follows: with the initial S=0, calculate the value of $\Delta f(S)$ one by one until $\Delta f(S)$ change into a non-negative for the first time, then the S which makes $\Delta f(S)$ non-negative is the global optimal storage capability S* of resource pool.

4 Conclusion

With the above model, further researches on the global optimization and quantitative analysis of the resource pool capacity can be expanded based on queuing theory and global optimization to get different algorithm for resource pool capacity according to different application in cloud computing. The binding of the scattered cloud computing resources through the resource pool is valuable for the improvement of the utilization efficiency, load balance and robustness of cloud resources.

Acknowledgments. This work was supported by National Key Basic Research Program of China (973 Program) under Grant No.2011CB302601, National Natural Science Foundation of China No.60970064, New Century Educational Talents Plan of Chinese Education Ministry No.NCET-08-0806, the Fundamental Research Funds for the Central Universities No.2010-YB-19.

References

1. Buyya, R., Yeo, C.S., Venugopal, S.: Cloud computing and emerging IT platforms: Vision, hype, and reality for delivering computing as the 5th utility. Future Generation Computer Systems 25(6), 599–616 (2009)
2. Liu, Y.-B., Jian, Y.: A novel abstract grid owner resource management model. Journal of Chongqing University of Posts and Telecommunications(Natural Science Edition) (6) (2008)
3. Xicheng, L., Huaimin, W., Ji, W.: Internet-based virtual computing environment (iVCE): Concepts and architecture. Science in China (Series F: Information Sciences) 49(06), 681–701 (2006)
4. Maurer, M., Emeakaroha, V.C., Brandic, I.: Cost-benefit analysis of an SLA mapping approach for defining standardized Cloud computing goods. Future Generation Computer Systems (in press, corrected proof)
5. Roy, S., Mukherjee, N.: Efficient resource management for running multiple concurrent jobs in a computational grid environment. Future Generation Computer Systems 27(8), 1070–1082 (2011)
6. Zhou, F.: New Characterizations of the Poisson Process. Journal of Tsinghua University (Science and Technology) (03) (1987)
7. Sun, R., Li, F.: The Parameters Algorigthm of the Superpositionof Interrupted Poission Processes. Acta Electronica Sinica (04) (1998)

The Application of Visual Speech Synthesis in Computer-Assisted Pronunciation Training

Yanyan Chen[1] and Hao Li[2]

[1] College of Foreign Languages, South China University of Technology, 510600 Guangzhou, China
[2] College of Electronic Science and Engineering, National University of Defense Technology, 410073 Changsha, China
{cyyscut,lihao7183200}@163.com

Abstract. To promote the efficiency of English pronunciation training, the Computer-Assisted Pronunciation Training (CAPT) system is drawing more and more attention of language teaching, but the current CAPT systems for Chinese second language learners are still quite dissatisfactory. The paper briefly introduces the CAPT technology and its realizing process, and discusses the application of visual speech synthesis technology in CAPT to enhance its friendliness in detail. Comparable teaching experiment demonstrates that the CAPT system improved by visual speech synthesis can significantly improve the efficiency of English pronunciation training.

Keywords: visual speech synthesis, computer-assisted training, English pronunciation, language learning.

1 Introduction

Speech is an important part of languages and the three elements of languages----sound, form and meaning are all indispensable. With the increasingly frequent exchanges in the international society, the importance of oral communication skills has been given more attention, which is also reflected in the latest reform in China's college English teaching [1]. Currently, the teaching of English pronunciation in China is mainly carried out in the initial stage of English learning. At university level, in addition to English majors, basically English speech training is no longer open for learners. However, the actual speech level of English language learners is not optimistic. The limitations of the quantity and quality of teachers make speech teaching in small classes cannot be achieved, while large classes with 50 to 60 people attending is bound to greatly reduce the effect of speech teaching. In college classroom, English learners in general still think that the time of speech teaching is not enough and lack of effective feedback and guidance [2]. Thus, using technology to improve the quality and efficiency of speech teaching is now imminent. With the constant development of Computer-Assisted Language Learning (CALL) technology, the technical means of Computer-Assisted Pronunciation Training (CAPT) gradually generate and develop. Applying CAPT system to build self-learning environment for

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 833–838.
springerlink.com

English learners and carry out English speech training, evaluation and correction learners is an important means to improve the speech level of the learners and spread the speech test.

2 The Development of CALL and CAPT Technology

CALL technology is to use computer-assisted methods in language teaching, and improve the efficiency and quality of language learning through the application of computer technology. The emergence of CALL brings innovation to foreign language education methodology, which is beneficial to promoting the transition from the traditional teacher-centered teaching model to the student-centered learning model. In doing so, learners can reduce the dependence on teachers, making them acquire languages more freely in an environment close to nature. This meets the objective requirement to promote the formation of unique learning method of students and the development of self-learning ability of them, and therefore the application of computer technology has become a priority of English teaching reform [3]. After being through the behavioral phase of alternating repetitive work experience and the communication phase focusing on language use, the current CALL technology lies in the phase of integrating the computer technology into the actual needs of language teaching [4]. CALL technology gives increasing attention to the actual needs of language teaching and linguistic research support.

CAPT technology belongs to a branch of CALL technology, which is a targeted practical application of the speech teaching in linguistics. Ideal CAPT system provides a relaxed and stress-free speech learning environment for language learners [5], characterized by few speech input restrictions, freedom to control the learning rhythm, and real-time feedback of the training effect. Currently, quite a few influential CAPT systems have emerged in the international stage. All these systems have supplemented the consideration of language teaching effect in addition to the simple speech test at early stage, achieving the two functions of evaluation and feedback.

CAPT systems target for Chinese English learners have achieved a certain degree of development and application, but overall, more researches are on the Chinese CAPT technology and less attention given to the second language teaching, without specific English speech teaching for Chinese learners. At present, the e-learning systems in some of the undergraduate English textbooks have added the support of speech-learning modules, such as the "Real communication, listening and speaking, new standard" from the Foreign Language Teaching and Research Press allows the learner to repeat pronunciation and record it, and then generate Contrast speech waveform. "Recite the word in New Oriental School ", "learn English from the scratch" and "100e English listening and speaking training platform," etc. can also provide the waveform and spectrum of speech for learners. PowerWord from Kingsoft provides small-scale video library of the vocabulary for repeat and imitation.

3 The Visual Development of CAPT

The CAPT system is implemented generally as shown in Figure 1. If the evaluation law has been pre-set to determine the evaluation parameters, the standard pronunciation can be trained to build the model library (hidden Maldives model is usually applied, that is, the HMM). After learners inputting the test speech, the computer can complete the quality score and error detection by making comparison between the result and standard pronunciation, and get the feedback score and give some suggestions on how to correct it combined with the pre-set knowledge of speech. To test the performance of CAPT system, it is necessary to establish a non-standard testing speech library; computer and language specialists respectively utilize CAPT software to score the test speech and carry out the error detection manually, finally calculating the correlation between the two results. The correlation between the scores of some outstanding CAPT software and experts can reach more than 90% and the accuracy rate of error detection is above 90% [6].

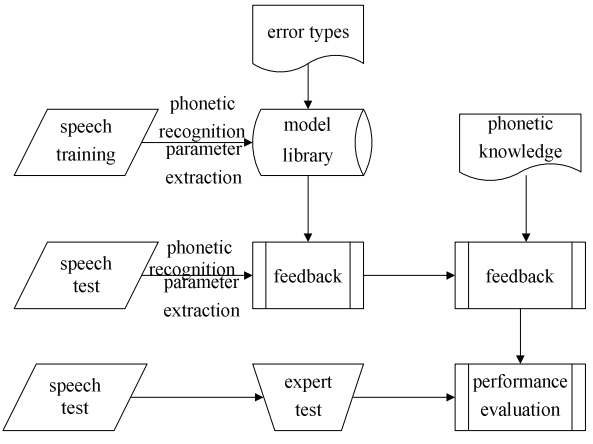
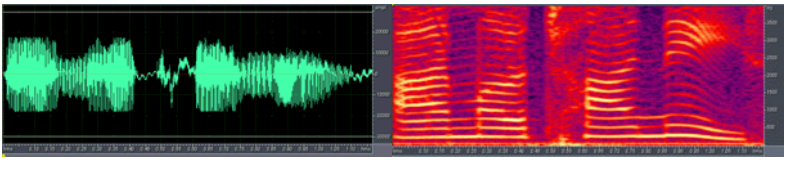


Fig. 1. The flow chart of CAPT system

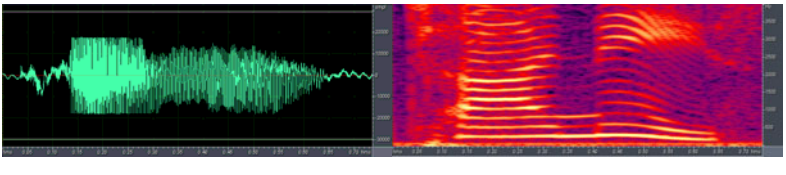
To achieve a better teaching effect, CAPT system should take full account of the characteristics of speech teaching, especially the dual-mode feature of speech [7], which is shown as sound in the auditory field with the basic unit of phoneme (phoneme); in the visual field, it is shown as the mouth shape and the continuous variation of facial expression with the unit of basic mouth shape, that is, the viseme. The speech teaching is obviously target-orientated and skillful. For the grammar training and vocabulary training, the corresponding text questions can be designed, and comments on the answers and error correction can also be preset. While the pronunciation training aims at the pronunciation imitation of learners although there still exists fundamental teaching and theories, including listening, imitation, the minimal pair training, vowel displacement training, even imitation as well as repetition of sentences. Psychological research shows that 3 main modes that human beings communicate with each other directly in the physical world are text

transmission, speech transmission (hearing) and facial expression transmission (visual), which is further summarized as that emotional expression = 7% text+54 +39% speech+54 % facial action. [8] The dual-mode characteristics of speech information determine that the interaction of CAPT system should be at the dual-mode state, that is, audio and video.

The key point of dual-mode interaction of CAPT system is to read the video information from the input speech and correct error from the output results, respectively involving the recognition and synthesis of visual speech. However, no matter for recognition and synthesis, it is necessary to study the mouth shape of English pronunciation and establish the library of mapping law between phonemes and the mouth shape, as well as the mapping library of common pronunciation errors of Chinese learners and mouth shape, which can reflect the mapping law of phoneme and viseme. Especially for the error feedback, most of the current teaching systems select the intuitive acoustic method to demonstrate the difference between the learners' speech and the standard speech, mainly including acoustic oscillogram and spectrogram, but ignore the pronunciation problems caused by mouth shape. Speech oscillogram shows the variation of acoustic amplitude with the time, in addition to the magnitude of speech in the oscillogram, the information about error correction which can be provided for learners is extremely limited. Spectrogram is the two-dimensional diagram that spectrum energy varies with the time, which demonstrates the process that the fundamental frequency resonance varies with time. It can provide the information to distinguish the subtle elements in acoustic research, but have a limited feedback and too professional for the common learners. Figure 2 shows the oscillogram and spectrum when learners respectively input the speech "I'm a Chinese" and "Chinese". In order to achieve a better teaching effect, an accurate evaluation on learners should be provided and three-dimensional mouth shape should be utilized to intuitively point out the problem and how to correct it. Nowadays, researches on such kind of mapping law in CAPT system are mainly conducted by engineers and technicians utilizing very limited sample data, so it is difficult to intuitively and comprehensively demonstrate the mapping relationship between English pronunciation and mouth shape of English learners in China, which has fundamentally restricted the dual-mode interaction of CAPT system.



a. the oscillogram of "I'm a Chinese." b. spectrogram of "I'm a Chinese."



c. the oscillogram of "Chinese" d. spectrogram of "Chinese"

Fig. 2. The demonstration of oscillogram and spectrogram

4 Analysis on the Teaching Effect

This study compares the speech training effect of visual CAPT system and non-visual CAPT system. The non-visual system is the speech training software paired with "Real communication, listening and speaking, new standard" from the Foreign Language Teaching and Research Press. Based on it, the visualization system adds conjecture face and correct errors using three-dimensional animation. The experimental subjects are 100 students enrolled-in 2009 whose major is Electronics and Information System in National Defense University. In the experiment, first the speech test (pre-test) was conducted to divide students into 2 groups: experimental group and control group with 50 in each group according to the test results. The teaching hours of two groups are exactly the same, with the total of 8 hours for 4 weeks. After the experimental teaching, speech test was carried out again (posttest). The 2 tests adopted 10-point scale, the data were analyzed using t test, and statistical analysis was completed using SPSS10.0.

Table 1. Comparison of the test result between experimental group and control group

Test	number of people	Mean	Std	t	P
pretest	Experimental group	4.9500	1.0288	0.900	0.929
	Control group	4.9335	1.7564		
posttest	Experimental group	6.1335	0.9255	5.998	0.004
	Control group	5.1335	1.6742		

T-test of the groups in table 1 shows that there is no significantly difference between experimental group and control group in the enrollment speech test. However, in the second speech test, the difference is obvious ($P < 0.05$). This indicates that visual CAPT system can better satisfy the requirements of speech training, for it can conduct speech training, evaluation and correction for English learners more effectively.

5 Conclusions

Global society has put forward higher requirements of English teaching quality and efficiency, especially the demand for application capacity of listening and speaking is increasing. In response to this demand, computer-based English language teaching and testing has become the focus of English teaching reform. This paper probes into the application of visual speech technology in CAPT system and the application in English speech training. Comparative teaching experiments show that the visualized CAPT system can carry out evaluation and feedback on the input speech information of the learner through the virtual human face. The friendly man-machine interaction enhances the communication comfort, and the realistic three-dimensional presentation also greatly improves the acceptability of learners on the speech correction. In language teaching, visual speech teaching software learning from the basic rules of applied linguistics will be the future trend of development. The combination of CAPT system with translation software can better achieve the conversion between multiple

languages, thus overcoming the shortcomings of giving too much attention to the text messages in traditional English teaching and effectively improving learners' listening and speaking ability, which is more in line with requirements of language learning objective.

References

1. Cai, J.: Research on the teaching reform and countermeasures of college English. MIFLI (133) (2010)
2. Zhang, J.: Research on the pronunciation teaching in university. Shandong Foreign Language Teaching (3) (2010)
3. Department of Higher Education. College English Curriculum Requirements. Tsinghua University Press, Beijing (2007)
4. Barson, J., Debski, R.: Calling back CALL: technology in the service of foreign language learning based on creativity, contingency, and goal-oriented activity. Second Language Teaching and Curriculum Center, Honolulu (1996)
5. Yukari, H.: Computer Assisted Pronunciation Training for Native English Speakers Learning Japanese Pitch and Durational Contrasts. Computer Assisted Language Learning (3) (2004)
6. Wan, J., Xiao, Y., Ye, W.: Review on the application of error detection in speech teaching. Journal of Chinese Information (4) (2009)
7. Waters, K., Rebg, J., Loughlin, M., Kang, S.B., Terzopoulos, D.: Visual Sensing of Humans for Active Public Interfaces. Digital Cambridge Research Lab TR 95/6 (1995)
8. Mehrabian, A.: Communication without words. Psychology Today 2(4), 53–56 (1968)

Microstrip Antenna Design Based on Left-Handed Materials

Song Chengtian[1], Song Zhiliang[2], and Pan Xi[1]

[1] State Key Labortory of Mechatronical Engineering and Control, Beijing Institute of Technology, China
[2] Dalian Airforce Communication, NCO Academy Dalian, China
{songct,panxi}@bit.edu.cn, zhiliang404@gmail.com

Abstract. An review on the development and extrodinary properties of left-handed materials is first coverd.Based on the current research, left-handed materials is applied to the microstrip antenna design.Two miniaturized antenna are designed by using the left-handed materials, which makes futhur miniaturization of microstrip antenna possible.

Keywords: Left-handed materials, Microstrip antenna, Antenna design.

1 Introduction

With the rapid development of wireless communication technology, the antenna has been widely used in various kinds of areas. Among the antennas, microstrip antenna is the one that has been used most frequently. It has many advantages, such as small size, light weight, stable structure, low profile, flexible feed mode, low cost and the ability to generate linearly polarized wave and the circularly polarized easily. At present, the microstrip antenna has been widely used in mobile communication, satellite communication and other related areas. With the development of the communication technology, the requirement of the microstrip antenna also becomes stricter. Small antenna and high directivity is the objective of the development of antenna in the future. The proposal of the left-handed materials provides a new possibility to propel the development of antenna. By using its extraordinary physical properties, the performance of the antenna will be greatly improved.

2 The Proposal of Left-Handed Materials

Left-handed material is a media. Its dielectric constant ε and the magnetic permeability μ are negative. Therefore, it is also called the double negative media. Its obvious feature is that the electric field, the magnetic field and the wave vector form the left-hand relationship. The left-handed materials have many extraordinary properties, such as backward wave effect, negative refraction effect, inverse Doppler Effect and the inverse Cerenkov radiation.

The concept of left-handed material was firstly proposed by V. G. Veselago in 1968, who was a physicist from Soviet Union. V. G. Veselago predicts that in left-handed

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 839–846.
springerlink.com

materials, the direction of the phase transmission and direction of the energy transmission are opposite. Besides, when refraction rate is negative, the direction of the wave vector is opposite to the direction of energy transmission. k、E、H rom the left-hand law. And this is the property of the left-handed materials predicted by V. G. Veselago. In recent years, the left-handed materials have attracted more and more attention in scientific area. In 2003, the left-handed material had been labeled as one of the ten greatest breakthroughs in technology by magazine Nature. In 2006, the invisible electromagnetic wave fulfilled by gradient media which is based on left-handed materials was seen as one of the ten greatest breakthroughs in technology by magazine Nature.

3 Extraordinary Electromagnetic Properties of Left-Handed Materials

3.1 Backward Wave Property

According to the Maxwell equations:

$$\nabla \cdot B = 0 \qquad \nabla \cdot D = \rho$$
$$\nabla \times E = -\frac{\partial}{\partial t} B \qquad \nabla \times H = J + \frac{\partial}{\partial t} D \tag{1}$$

The LHM constitutive relation is:

$$B = \mu H \qquad D = \varepsilon E \tag{2}$$

The Helmholtz wave equation is:

$$\nabla^2 E + k^2 E = 0 \tag{3}$$

Then the dispersion relation is:

$$k^2 = \omega^2 \mu\varepsilon = \omega^2 n^2 \tag{4}$$

In the equations, n is the refraction rate. If $\mu\varepsilon$ >0, the dispersion relation is applicable. Thus the wave equation is solvable. According to this, both the value of the dielectric constant and the value of the magnetic permeability of the left-handed materials are less than zero. This situation does not violate the Maxwell principle. When electromagnetic wave transmits in the left-handed materials, the solution of the wave equation is quite different.

When the plane monochromatic wave is transmitted in the medium, the Maxwell equation can be simplified as follows:

$$\nabla \times H = \frac{\partial}{\partial t} D \qquad \nabla \times E = -\frac{\partial}{\partial t} B \tag{5}$$

$$E = E_0 \exp(-jk \cdot r + j\omega t) \tag{6}$$

$$H = H_0 \exp(-jk \cdot r + j\omega t) \tag{7}$$

By combing the equation 2 and equation 5, the wave quation is:

$$k \times E_0 = \omega_0 \mu H_0 \quad k \times H_0 = -\omega \varepsilon E_0 \tag{8}$$

Obviously, when ε >0, μ >0, the direction of the vector k 、 E 、 H complies with the right-hand rule. Poynting's vector $S = E \times H$ has the same direction as that of the vector k . It is the forward wave. When ε <0, μ <0, the direction of the vector k 、 E 、 H complies with the left-hand rule. The direction of vector S is opposite to that of k . It is the backward wave. When ε <0, μ >0 or μ <0, ε >0, vector k is imaginary number. There is no transmitted wave, only evanescent wave exists.

3.2 Negative Refraction Property

Negative reflection is the most evident property of left-handed materials. It is also a key research area. In nature, when incident light goes through the interface of two different kinds of media, the phenomenon of reflection and the refraction will be generated. This is called the 'positive refraction'.

3.3 Reversed Doppler Effect

In left-handed materials, the direction of the phase velocity and the direction of the group velocity are opposite. It means the transmitted direction of the energy and that of the phase are opposite. The frequency shift shows inverse Doppler Effect. In the traditional media, when the detector is approaching to the light source, the detected frequency is higher than ω_0 . Otherwise, the detected frequency is lower than ω_0 . But in the left-handed materials, the result is opposite.

3.4 Reversed Cerenkov Radiation

The charging particles moving in a constant speed in the vacuum does not radiate electromagnetic wave. But when charged particles moving in a constant speed in media, it will generate the induced current and therefore form a series of second wave source to generate the second wave. When the velocity of the particles exceeds the light speed in the media, the second waves interfere with each other and then radiate the electromagnetic waves. This is called the Cerenkov radiation. In traditional media, the wave front generated after the interference is a cone. The energy of the electromagnetic wave is transmitted along the norm of the cone. The direction of the energy radiation and the moving direction of the particles form an angle θ . θ is defined by the following equation:

$$\cos\theta = \frac{c}{nv}$$

v epresents the velocity of the moving particle. But in the left-handed materials, the direction of the energy transmission is opposite to that of the initial velocity. Therefore, the direction of the energy radiation is opposite to the direction of the moving particle. The two directions form a forward cone angle.

4 Applying the Left-Handed Materials in the Antenna Design

The property of backward wave of left-handed materials enables it to be widely used in the micro wave antenna. By combining the backward wave effect of the left-handed materials and the forward wave effect of the conventional media, the resonance cavity which is less than half wave length can be designed. Moreover, the size of the resonance of the cavity is no longer limited by the resonant frequency. Under this circumstance, the left-handed materials function as a phase compensator. The phase difference which is generated when electromagnetic waves are transmitted in traditional media can be compensated by the backward wave effect of the left-handed materials. According to this property, two kinds of small antenna in left-handed materials are designed in this paper.

4.1 Microstrip Antenna Using Complementary Circular Left-Handed Materials

It uses the FR4 as its base and its size is 18mm*14mm*10mm. Fig 4.1 and Fig 4.2 show its three dimensional structure and geometric structure respectively. The complementary ring resonator unit which is a kind of left-handed materials is placed between the ground and the radiation SMD. Simulation of the loss of the backward wave is shown in the Fig 4.3. The antenna pattern is presented in Fig4.4 and Fig 4.5. According to the simulation, the antenna gain meets the practical requirements. Compared with the antennas which have the same gain, the size of this antenna is about two third of their size.

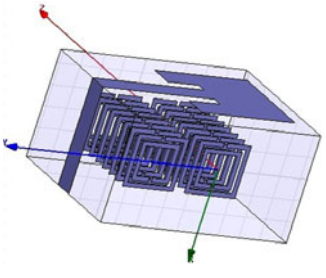


Fig. 4.1. The three dimensional structure of the microstrip antenna based on complementary ring left-handed materials

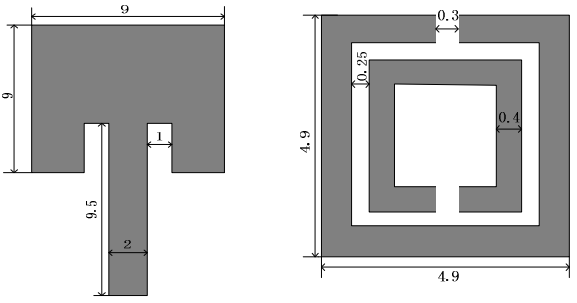


Fig. 4.2. The geometric structure of the microstrip antenna based on complementary ring left-handed materials

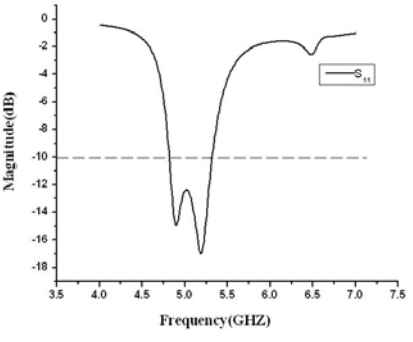


Fig. 4.3. The simulation of the echo loss of the microstrip antenna based on complementary ring left-handed materials

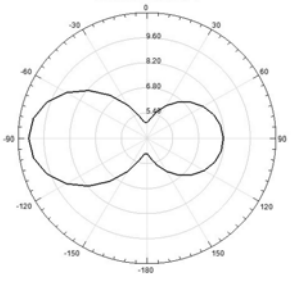


Fig. 4.4. The antenna pattern when f=5.1 GHz, Phi=0°

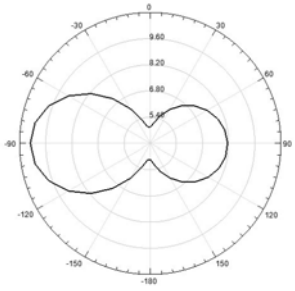


Fig. 4.5. The antenna pattern when f=5.1 GHz, Phi=90°

4.2 Microstrip Antenna Using Spiral Square Ring Resonator Arrays as the Radiation Uint

It uses the Rogers RT/duroid 5880 as its base and its size is 50mm*50mm*175mm. This antenna adopts the double resonant left-handed materials to establish the transmission line. The two arms of the microstrip transmission line and the corrosion consist of five spiral square ring resonator, which is shown in the Fig 4.7. This represents the small and compact structure of the microstrip antenna. Fig.4.6 shows the three-dimensional structure of the microstrip antenna. The Fig 4.8 reflects its geometric structure. Fig 4.9 is the echo loss chart. Fig 4.10 and Fig 4.11 describe the antenna pattern when f=8.5GHz.

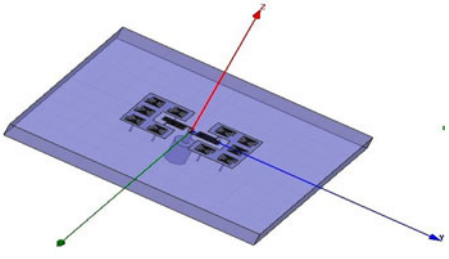


Fig. 4.6. Three dimensional structure of the microstrip antenna

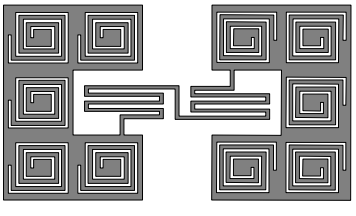


Fig. 4.7. Radiation SMD structure of the microstrip antenna

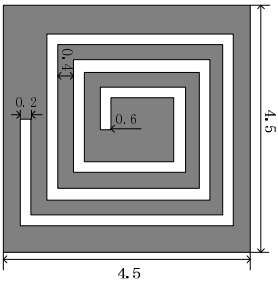


Fig. 4.8. The unit size structure of the microstrip antenna

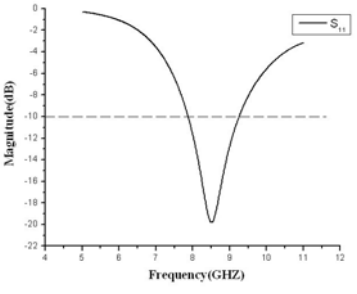


Fig. 4.9. Simulation result of the echo loss of the microstrip antenna

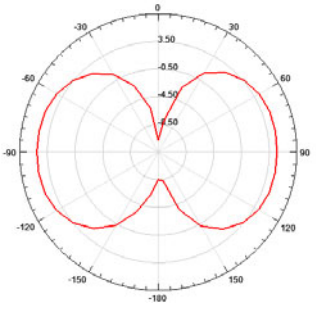


Fig. 4.10. The antenna pattern when f=8.5 GHz, Phi=0°

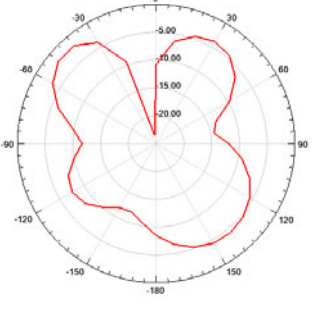


Fig. 4.11. The antenna pattern when f=8.5 GHz, Phi=90°

The two types of antenna adopt split ring resonator structure and the spiral ring resonator structure respectively. Both of them fulfill the objective that designing the small and ultra-wideband microstrip antenna. Compared to the other kinds of antenna, the two antenna are more compact and have better properties.

5 Conclusion

Left-handed materials have special electromagnetic properties and potential prospect. It is still a new research area that using left-handed materials in the microstrip antenna design. With the development of the left-handed materials research, many existing technologies will get new development even breakthroughs. Nowadays, left-handed materials have already become a hot research area which attracts scientists and researchers from all over the world. This kind of material has vital influence on the wireless communication technology in the near future and needs to be further developed.

References

1. Smith, D.R., Padilla, W.J., Vier, D.C.: Composite medium with simultaneously negative permeability and permittivity. Physical Review Letters 84(18), 4184–4187 (2000)
2. Veselago, V.G.: The Electrodynamics of Substances with Simultaneously Negative Values of Permittivity and Permeability. Soviet Physics Uspekhi. 10, 509–514 (1968)
3. Valentine, J., Zhang, S., Zentgraf, T.: Three-dimensional optical metamaterial with a negative refractive index. Nature 84, 377–379 (2008)
4. Yen, T.J., Padilla, W.J., Fang, N.: Terahertz magnetic response from artificial Materials. Science 303(5663), 1494–1496 (2004)
5. Caloz, C., Sanada, A., Liu, L.: A Broadband Left-Handed coupled-line backward coupler with arbitrary coupling levels. In: IEEE-MTT Int'l Symp., vol. 1, pp. 317–320 (2003)
6. Pendry, J.B.: Negative material. Contemporary Physics 2, 1–13 (2004)
7. Saha, C., Siddiqui, J.Y., Guha, D.: Square Split Ring Resonators: Modelling of resonant frequency and polarizability. In: IEEE Applied Electromagnetics Conference AEMC, December 1, pp. 19–20
8. Chair, R., Lee, K.M., Lee, K.F.: Miniature Multilayer Shorted Patch Antenna. Electronics Letters 36, 3–4 (2000)
9. Sharawi, M.S.: Use of Low-Cost Patch Antenna in Modern Wireless Technology. IEEE Potential 25, 35–47 (2006)
10. Row, J.S.: A simple impedance-matching technique for patch antennas fed by coplanar microstrip line. IEEE Transactions Antennas and Propagation 53(10), 3389–3391 (2005)

Measurement and Control System Development of Performance Research Test Bench for Diesel Engine Injector Based on Labview

Suyun Luo and Yuanming Gong

Automotive Engineering College,
Shanghai University of Engineering Science, 201620 Shanghai, China
lsyluo@163.com, GongYuanMing@tsinghua.org.cn

Abstract. A measurement and control system based on Labview was developed for the performance research of high-pressure common rail diesel engine injector of marine. The developed system provides smooth high-pressure oil source for injector with just $\pm 5MPa$ deviation to aimed pressure. It also function with data collection and calibration, fuel and hydraulic oil temperature controlling, pressure curve and injection rate displaying, data limits surpassing setting and alarming. The paper detail introduced the electronic components of the control system, software design and the PID closed-loop control of injection pressure.

Keywords: Electronic Injector, Measurement and control system, High-pressure common-rail pressure control, PID.

1 Introduction

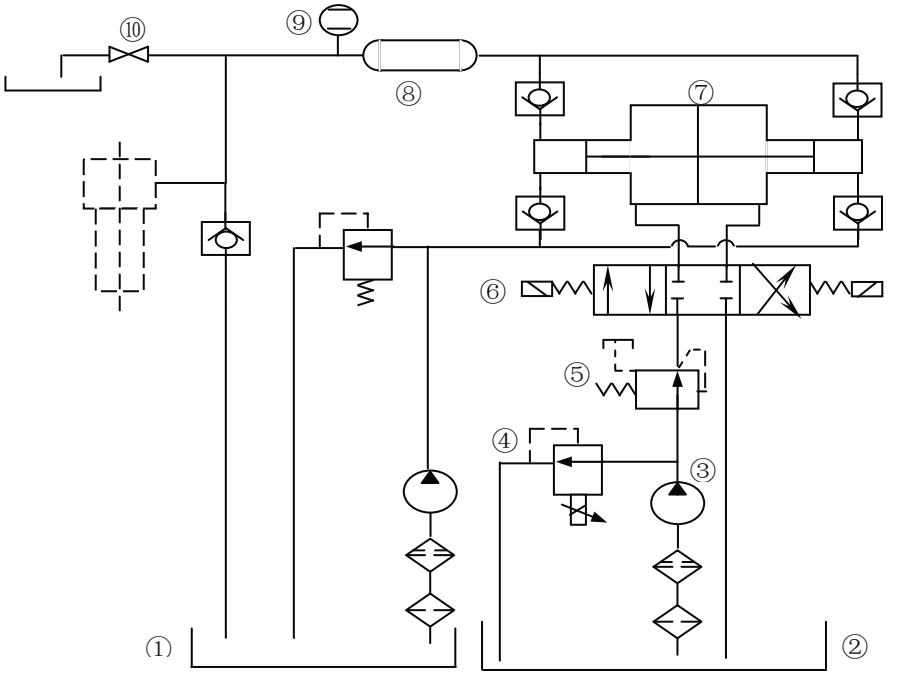
Electronic injector performance research Test-bed simulates the operating environment of electronic injector, provides smooth high-pressure common-rail (HPCR) fuel pressure, is an important platform for performance research, design and manufacturing, quality measurement of electronic injector. The measurement and control system introduced in this paper is based on marine diesel engine electronic injector performance research test-bed. The test-bed is mainly used for the performance research, analysis and measurement of new type injector or modified injector, or used for spot-check of products. In order to assure the system a maximum $200MPa$ fuel pressure and enough fuel flow, hydraulic booster is used in the test-bed. Continuously adjustable fuel pressure between $[25MPa, 200MPa]$ is provided by the system to meet the needs of different type of injector. The measurement system is developed based on visual graphical programming software Labview8.5, make up of electronic hardware and software.

2 High-Pressure Oil Circuits of Test-Bed

The main function of this measurement and control system is to start test-bed, assure a reliable, smooth running and measure, record, display system data etc. The operation

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 847–854.
springerlink.com

of the whole test-bed includes the start and regulate running of three types of oil circle: (1) Hydraulic circuit, the main function of this circuit is build itself an enough pressure, then affect fuel circuit and prompt it build high enough pressure for injector, an aimed fuel pressure value between $[25MPa, 200MPa]$ could be set. Hydraulic circuit is started by hydraulic pump and the motor, hydraulic pressure and fuel pressure is controlled through proportional regulator and proportional relief valve. (2)Fuel circuit, it is affected by hydraulic circuit and build itself a suitable pressure for the injection of electronic injector, this circuit is started and regulated by fuel pump and motor, through circulate hydraulic and fuel separately into pressurized big cylinder and small cylinder of turbocharger and control the pressure of big cylinder, the fuel pressure of small cylinder is adjusted to aimed value. (3) Cooling circuit, this circuit is responsible for the cooling of hydraulic circuit and fuel circuit and assure the system a regulate operation. Fuel warming and cooling solenoid valves and hydro cooling solenoid valve are needed to make sure the temperatures of circuits are in the range. Figure 1 shows the theory of turbocharger oil circuit.



①Fuel Tank ②Hydraulic Tank ③Hydraulic Pump ④Proportional Relief Valve
⑤Proportional Regulator ⑥Directional Control Valve ⑦Turbocharger
⑧Accumulator ⑨Pressure Sensor ⑩Safety Valve

Fig. 1. Turbocharger oil circuit theory of high-pressure common-rail electronic injector performance research test-bed

3 Theory of Measurement and Control System

Measurement and control system controls the works of all components comprising of test-bed, is make up of three function parts. They are signal acquisition, industrial PC and actuators, as showed in figure 2.

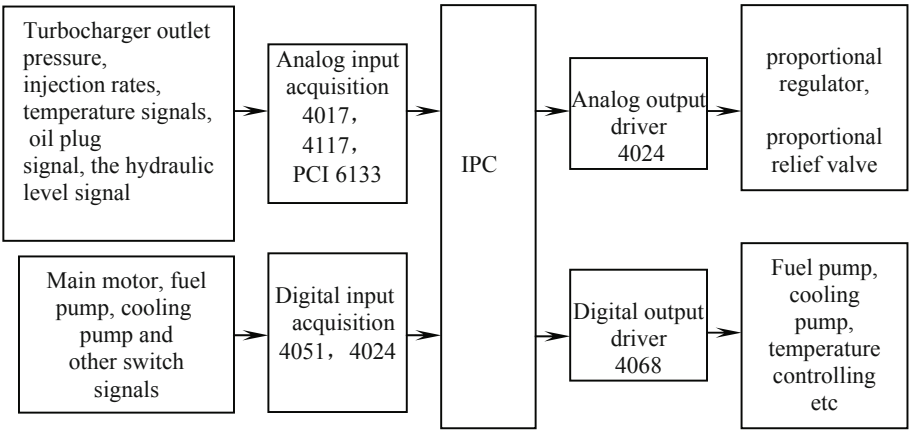


Fig. 2. The theory of measurement and control system based on electronic injection performance test-bed

3.1 Sensors Signals Acquisition

Signals should be collected are: Injector needle lift, injector solenoid coil current, pressures in the circuits, temperatures of hydraulic, fuel and cooling oil, levels of fuel tank, hydraulic tank and cooling oil tank etc. The details are lists in table 1.

Table 1. Parts of sensor signals measured by measurement and control system

Signal	Signal type	Signal output	Range
Common-rail pressure	Analog, voltage	[0, 10] V	[0, 200] MPa
Fuel pressure	Analog, voltage	[0, 10] V	[0, 3.5] MPa
Fuel tank temperature	Analog, voltage	[0, 10] V	[-20, 200] ^{o}C
Hydraulic Tank Temperature	Analog, Current	[4, 20] mA	[-200, 450] ^{o}C
Fuel level	Analog, voltage	[0, 10] V	[0, 10] KPa
Fuel coarse filter blockage detection	Analog, voltage	[0, 10] V	[0, 200] KPa
Coolant coarse filter blockage detection	Analog, voltage	[0, 10] V	[0, 200] KPa
Protective door	Switching signal		
Motor Start signal	Switching signal		
Fuel pump start signal	Switching signal		

Sensor signals are gathered through Adam4117, Adam4017, Adam4051, Adam4024 modules and PCI 6133 board.

3.2 Industrial PC

Industrial PC(IPC) calculates and analyzes all data acquired and output control signals according to specific control strategies.

3.3 Actuators

The measurement and control system includes actuators such as proportional regulator, proportional relief valve to control fuel injection pressure, fuel pump motor, coolant pump motor and relays controlling cooling valve and heater etc. Adam4068 module is used to drive actuators. SCB-68 and Adam4520 module are used to be isolated converter arranged between all the sensors, actuators and the IPC.

4 Software Development of Measurement and Control System

Visual graphical programming software Labview8.5 is used to develop the measurement and control system. State machine, message queue, user interface events and master/slave structure modes are combined during the development. The software's function of the whole system is divided into 3 parts: data collection, data processing (including data A/D transform, data calculation, data display, alarming and data saving), actuators controlling (including common-rail pressure PID control, temperature PID control, and motors control etc).

Table 2. Some commands used for signals collection

Command	Function	Syntax	Description	Module
#AA	Analog In	Data#AA(cr)	The command will return the input value from a specified(AA) module in the currently configured data format.	Adam4017
#AAN	Read Analog Input Channel N	#AAN(cr from)	return the input value from one of the eight channels of a specified (AA) module in the currently configured data format.	Adam4017
$AA6	Digital In	Data$AA6(cr)	requests the specified (AA) module to return the status of its digital input channels and feedback value from its digital output channels.	Adam4051
#AABB	Digital Out	Data#AABB(data)(cr)	sets a single digital output channel or sets all digital output channels simultaneously	Adam4068

4.1 Data Collection

Advantech company modules of Adam4117, Adam4017, Adam4051, Adam4024 and NI company board PCI 6133 were applied to collect signals of the system. All of the modules and board have 8 or 16 channels. Each channel could be connected to a sensor. Some channels are used for analog signal, while others for pulse signal or switching signal. Advantech company provides commands for user to communicate

with the module, so programmer could access signals and data the modules get from sensors or configure the module. Following commands shown in table 2 are parts of commands be used in measurement and control system developing.

AA (range 00-FF) represents the 2-character hexadecimal address of the output value. BB is used to indicate whether all channels will be set or a single channel will be set. In the last case, BB will also indicate which channel it is.

A data queue is used to hold all the data collected from sensors, each one data is labeled with an address to distinguish it. So it could be processed correctly in data processing phase.

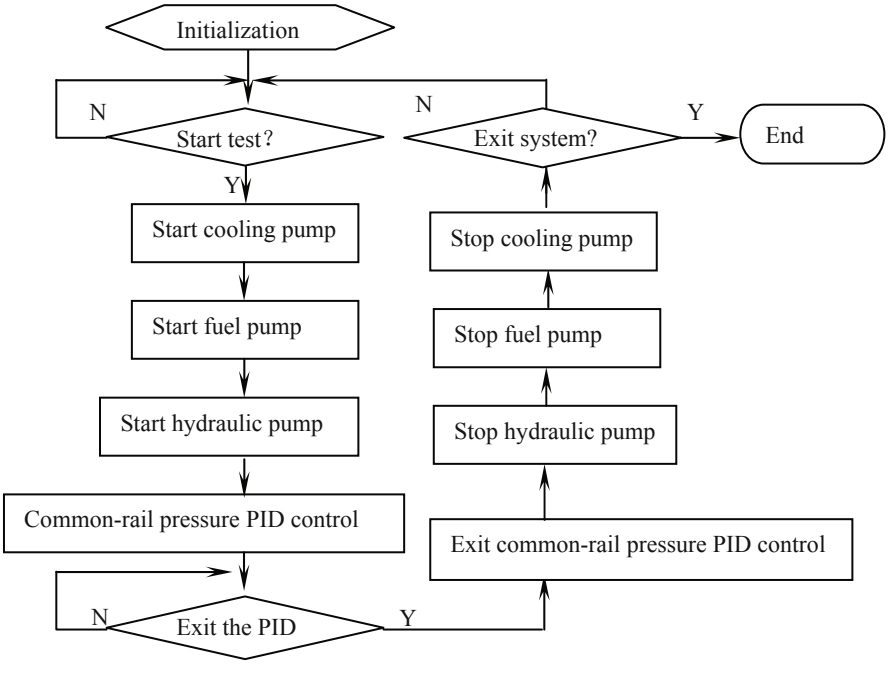


Fig. 3. Main flow chart of measurement and control system

4.2 Data Processing

Data A/D transform, data calculation, data display, data overtake alarming and data saving are included in data processing. All data processed here come from the data queue built in data collection phase. Data A/D transform should be processed according to signal's output and range. Based on upper and lower limits user set via system interface data overtake judgment is done.

4.3 Actuators Controlling

High-pressure common-rail pressure PID control and temperature PID control are the main jobs in this phase, what's more motors controlling for fuel pump, hydraulic pump and coolant pump is also included in this phase.

The overall control process is demonstrated in figure 3.

Common-rail pressure is controlled through proportional regulator and proportional relief valve. IPC output voltage signals between 0-10V by two channels of Adam4024 to proportional regulator and proportional relief valve. Proportional regulator and proportional relief valve are initialized to be zero and opened, 2-4 seconds after the starting of hydraulic pump motor, a voltage signal is send to proportional relief valve. The value of the voltage signal is between 0-10V proportional to $0\text{-}200MPa$ and the aimed pressure. Set the signal to be stable and then adjust the value of voltage signal to proportional regulator in closed-loop, thus the real common-rail pressure is adjusted to aimed pressure value.

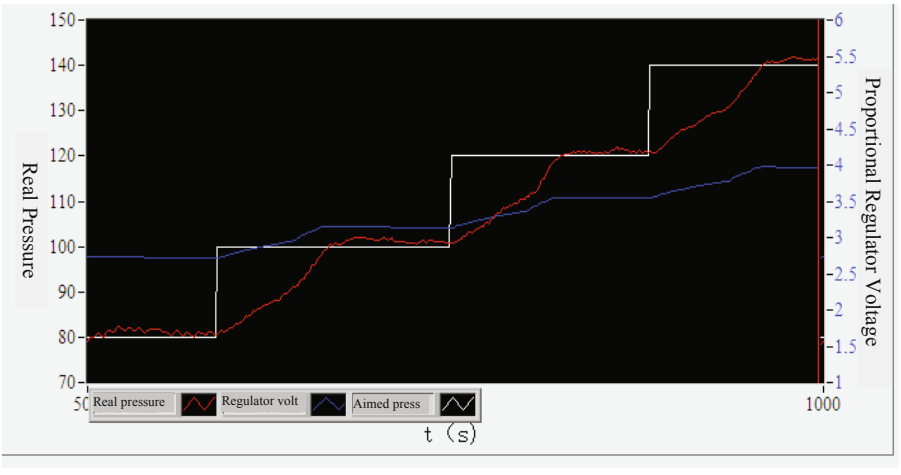


Fig. 4. Controlled common-rail pressure curve at aimed value of $80MPa$, $100MPa$, $120MPa$, $140MPa$

Experiments show that proportional regulator and proportional relief valve response with lag to voltage signal sent to them and if voltage is lower than about 1 volt, there are no or very slow response coming from the two actuators, what's more, common-rail pressure fluctuates with pulse, and rising slowly decline rapidly when it fluctuate up and down along aimed pressure. The above factors make common-rail pressure difficult to control. Phases PID control is adopted base on experiments and analysis and it is divided into ten phases according to the change and speed of change of e and Δe. Each phase has it's own PID parameters based on experience and experiments. Figure 4 and figure 5 show the system control results: developed system has $\pm 5MPa$ precision.

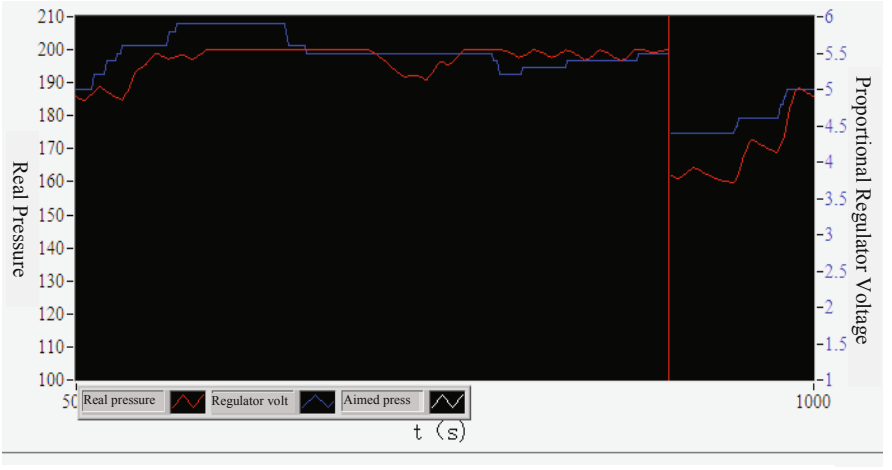


Fig. 5. Controlled common-rail pressure curve at aimed value of $200MPa$

5 Conclusion

Developed measurement and control system provides a stable, reliable platform to electronic injector performance test-bed and assure it be operated smoothly. The common-rail pressure PID control provides stable and continues adjustable fuel pressure between $25MPa$ - $200MPa$ that tested injector needs. PID control on temperatures is realized either.

Many papers and experts have given efficient methods on how to decide PID parameters on temperature control, but not so many for common-rail pressure control on test-bed. More accurate PID parameters should be found or methods be studied on, such as cooperate PID with fuzzy control or PID with neural network, so we could have $\pm 3MPa$ or $\pm 1MPa$ or better precision. The measurement and control system interface could be designed to be more human-oriented. More functions such as data analysis on injector performance parameters could be merged into the system. Further research and works need to be done in the future.

References

1. Nie, J.-J.: Trends and present status of the high-pressure common rail fuel injection system for diesel engine. Internal Combustion Engine 8, 6–9 (2009)
2. Liu, Y., Zhang, Y.-T.: Optimization Research for a High Pressure Common Rail Diesel Engine Based on Simulation. International Journal of Automotive Technology (11), 625–636 (2010)
3. Lee, J., Min, K.: Effects of Needle Response on Spray Characteristics in High Pressure Injector Driven by Piezo Actuator for Common-Rail Injection System. Journal of Mechanical Science and Technology (19), 1194–1205 (2005)

4. Doudou, A., Maslouhi, A.: A Macro-Microscopic Investigation of High-Pressure Sprays Injected by a Common Rail System. Journal of Mechanical Science and Technology (21), 1284–1292 (2007)
5. Xiao, Q., Yan, F.: Exploitation of test dais for flow characteristics and test analysis of electronic-controlled injectors. China Mechanical Engineering (16), 1419–1424 (2005)
6. You, L., An, W.: Rail pressure control method in experiment system of common-rail high pressure pumps. Journal of Jiangsu University (Natural science edition) 4(13), 432–438 (2010)
7. Guo, L.-Y.: Automatic hydraulic switching pressure gaining return and application. Coal Mining Machinery 3, 90–91 (2005)

Modeling the IEEE 802.11 DCF with Hidden Stations

Jielin Fu[1,2], Zheng Zhou[1], Lin Zheng[2], and Zhenghong Liu[2]

[1] Wireless Network Lab,
Beijing University of Posts and Telecommunications,
Beijing 100876, China
[2] Key Lab. of Cognitive Radio & Information Processing, The Ministry of Education,
Guilin University of Electronic Technology, Guilin 541004, P.R. China
fujielin@gmail.com

Abstract. Many works have been put on analyzing the performance of the IEEE 802.11, but hidden station problem was seldom involved. In practice, the existing of hidden stations will lead heavy drop of system throughput. A novel model is proposed in this paper, which can be used to analyze the throughput of 802.11 DCF in presence of hidden stations under saturation condition. Different scenarios are used in simulation to validate it, and the results show that the model is accurate.

Keywords: IEEE 802.11, DCF, Saturation, Hidden stations.

1 Introduction

During the last few years, the IEEE 802.11 [1] has become the standard protocol for Wireless Local Area Network (WLAN) and has been widely deployed. The Medium Access Control (MAC) technique of 802.11 is called Distributed Coordination Function (DCF). Many models have been established to analyze the throughput of 802.11 DCF. Most research of DCF modeling focused on the maximum throughput or the saturated throughput.

Bianchi [2] presented a Markov chain model to compute the throughput under saturation condition, i.e., each station always has a frame available for transmission and the transmission queue of each station is assumed always nonempty. One of the main assumptions is that all stations are in the same radio proximity. This model then has refined in different ways, but the problem of hidden station was seldom dealt with. However, stations do not typically operate under the same radio proximity, and maybe some stations cannot hear the transmissions of other ones. RTS/CTS mechanism can reduce the effect of hidden stations [3], but it cannot eliminate it. So, recently, some research efforts have been devoted to the problem. In [4], an analytical model was derived from Bianchi's Markov model and Tobagi's hearing graph framework [5]. However, its results show that it is a little sketchy. Based on [2], Vassis [6] presented a model introducing the transmitting time, within which the transmitting station may experience collision because of hidden stations. Nevertheless, its theoretical results under usual number of stations do not match up to the simulation. Paper [7] proposed a model focusing on ad hoc networks with hidden

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 855–860.
springerlink.com

terminals. The results showed in the paper are good, but the author gave some equations without clear proof. In this paper, we extend Bianchi's model and present a model applicable to the hidden station problem.

2 Saturated Model with Hidden Stations

In this section, we present a model suitable for analyzing the IEEE 802.11 DCF with hidden stations under saturated condition. In our scenario, total n stations around the AP are divided into J groups. Group j, $j \in [1, J]$, has n_j stations. Stations in one group are in the same radio domain, and they can hear each other. However, they cannot hear stations in other groups, and which indicates the hidden station problem. All stations can be heard by the AP, and they always have frames to send to the AP.

The main assumptions follow those in [2]. We assume that the channel is ideal (i.e. no link errors), and stations in same group are equally likely to access the channel. We also make the most of the same notations used in [2]. Equations (7) and (9) in [2] represent the saturated model without hidden stations, and they are rewritten here as equations (1) and (2) for purposes of completeness.

$$\tau = \frac{2(1-2p)}{(1-2p)(W+1)+pW(1-(2p)^m)} \tag{1}$$

$$p = 1-(1-\tau)^{n-1} \tag{2}$$

In equation (1), transmission probability τ is expressed as functions of the conditional collision probability p. This equation is suitable for the scenario with hidden stations, and the only difference is using τ_j and p_j for group j. But equation (2) is not tenable when hidden stations are present. To deal with the hidden station problem, we now extend Bianchi's model and introduce more variables.

For group j, let Es_j be the expected time spent per state.

$$\begin{aligned}
Es_j &= Pi_j\sigma + Pa_jTa + Ps_jTs + Pc_jTc + Pb_jTb \\
Pi_j &= (1-\tau_j)^{n_j}\left(1-\sum_{k\neq j} Ps_k \frac{Es_j}{Es_k}\right) \\
Pa_j &= (1-\tau_j)^{n_j}\sum_{k\neq j} Ps_k \frac{Es_j}{Es_k} \\
Ps_j &= (1-(1-\tau_j)^{n_j})(1-p_j) \\
Pc_j &= 1-(1-\tau_j)^{n_j} - n_j\tau_j(1-\tau_j)^{n_j-1} \\
Pb_j &= (1-(1-\tau_j)^{n_j})p_j - Pc_j
\end{aligned} \tag{3}$$

In equation (3), σ is the time of an idle slot, and Pi_j is the probability that station in group j experiences idle, that is, none of the stations in the group transmits and

AP does not send anything either. Pa_j is the probability that all stations in group j do not transmit but they can hear an ACK (for basic access mechanism) or a CTS (for RTS/CTS mechanism) from AP caused by successful transmission in other groups, and Ta is the time taken for an ACK or a CTS. Ps_j is the probability that at least one station in group j transmits and succeeds, and Ts is the time taken for a successful transmission. Pc_j is the probability that more than one stations in group j are transmitting simultaneously, and Tc is the time taken for a collision between stations in same group. Pb_j is the probability that at least one station transmits and collides with transmission from other groups, and Tb is the time taken for a collision between stations in different groups.

We now consider the expected time Et_j between two successive transmission of DATA (for basic access mechanism) or RTS (for RTS/CTS mechanism) frames for a station in group j.

$$
\begin{aligned}
Et_j &= Qi_j\sigma + Qa_jTa + Qs_jTs + Qc_jTc + Qb_jTb \\
Qi_j &= \sum_{i=0}^{m} \frac{W_i}{2} {p_j}^i \Big/ \sum_{i=0}^{m} {p_j}^i \\
Qa_j &= \sum_{k\neq j} Qs_k \frac{Et_j}{Et_k} \\
Qs_j &= n_j(1-p_j) \\
Qc_j &= n_j(1-(1-\tau_j)^{n_j-1})/2 \\
Qb_j &= n_j(p_j-(1-(1-\tau_j)^{n_j-1}))
\end{aligned}
\tag{4}
$$

Qi_j is the expected number of idle slots between two successive transmission. According to the different back-off stage it may stay, the expected number of idle slots is $W_i/2$ with probability ${p_j}^i$, where $i=0,\ldots,m$. Qi_j is then calculated by the weighted average of them. Qa_j is the expected number of ACKs or CTSs caused by successful transmissions in other groups. Qs_j is the expected number of successful transmissions in the group. In Et_j, all stations in the group may transmit once on average. So totally n_j attempts of transmitting may happen. But some of them may collide with probability p_j; thus Qs_j equals to $n_j(1-p_j)$. Qc_j is the expected number of collisions with stations in same group. In n_j attempts of transmitting, some of them may fail because other stations in the group begin transmitting at the same time. When two stations in the group collide, they share the time of collision, so a coefficient 1/2 occurs in the expression of Qc_j. We ignore the probability of more than two stations transmitting at the same time because it is relatively smaller. Qb_j is

then the expected number of collisions with stations in other groups. Finally, Et_j is calculated by summing up them.

For transmission probability τ_j is scaled in Es_j and Et_j is the time between two successive transmitting, this leads to

$$\tau_j = Es_j / Et_j \tag{5}$$

Equations (1), (3), (4) and (5), which describe the model, represent a nonlinear system with p_j, τ_j, Es_j and Et_j. It can be solved by numerical computation. In our study, we solve the model using the method *fsolve* implemented in the MATLAB optimization toolbox.

The model is suitable for both basic access and RTS/CTS mechanism. It is necessary only to specify corresponding values of Ta, Ts, Tc and Tb when different payloads and mechanisms are employed, and they are calculated in the same way as in [2].

3 Model Validation

To validate the model, we have compared its results with that obtained with NS-2. The results are based on basic access mechanism with 11 Mbps data rate and 500 bytes packet payload. Two scenarios, including different kinds of grouping schemes, are used for validation.

The first scenario includes three grouping schemes, i.e., n stations are equally divided into 2, 3 and 4 groups. The results are shown in Figs. 1 and 2. Fig. 1 shows how Et increases with n. Fig. 2 shows how normalized total throughput drops with n. From the figures, we see that, though stations transmit more quickly when they are divided into more groups, total throughput drops anyway. Smaller group lets stations get more chances to transmit, whereas experience more collisions because of more hidden stations.

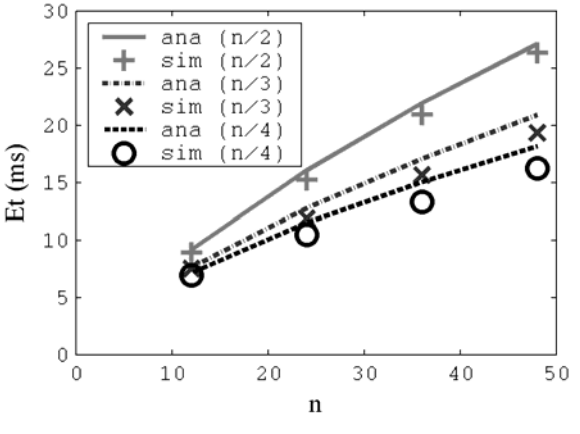


Fig. 1. Et vs. n

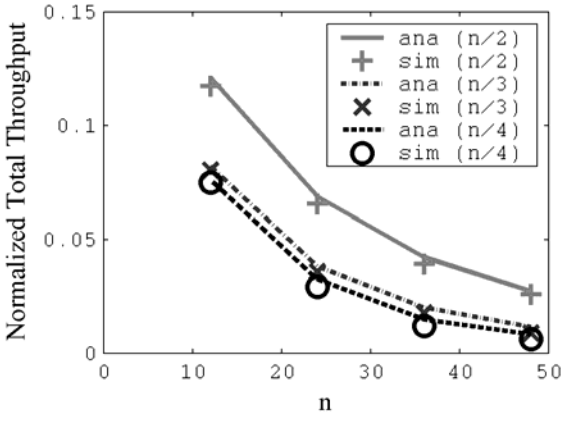


Fig. 2. Throughput vs. n

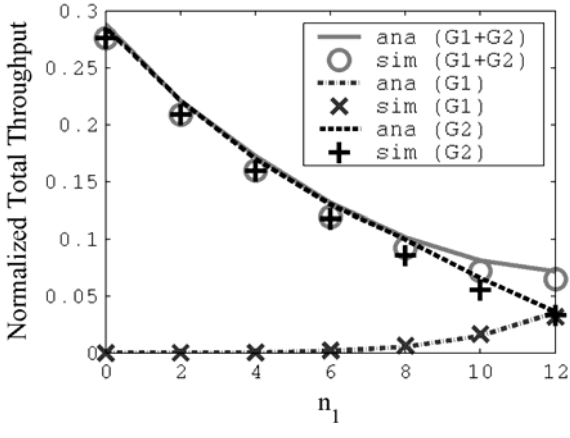


Fig. 3. Throughput vs. n_1

In the second scenario, 24 stations are divided into 2 groups, G1 and G2, which has n_1 and $24-n_1$ stations, respectively. Throughputs of G1 and G2, as well as their sum, are plotted in Fig. 3 versus the value n_1. It shows that group's throughput is not in proportion to its number and smaller group's throughput is extremely low when there is a big difference between the numbers of two groups. This is to some extent unfair to the minority. Moreover, the degradation of total throughput from no hidden stations case (n_1 = 0) is obvious, and it descends further as the proportion of hidden stations goes high.

4 Conclusions

In this paper, we have presented a novel model to analyze 802.11 DCF in presence of hidden stations under saturation condition. The model's analytical results are validated against simulations, and it shows that the model is extremely accurate. Future works will address RTS/CTS mechanism and real machine evaluation.

Acknowledgments. This work is supported by The National Natural Science Foundations of China(No.60962001,No.61071088) and Guangxi Natural Science Foundations (0991018Z).

References

1. IEEE standard for Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications, ISO/IEC 8802–11: 1999(E) (1999)
2. Bianchi, G.: Performance Analysis of the IEEE 802.11 Distributed Coordination Function. IEEE J. Sel. Areas Communications. 18(3), 535–547 (2000)
3. Khurana, S., Kahol, A., Gupta, S., Srimani, P.: Performance Evaluation of Distributed Coordination Function for IEEE 802.11 Wireless LAN Protocol in Presence of Mobile and Hidden Terminals. In: Proc. MASCOTS 1999, pp. 40–47 (1999)
4. Rahman, S.: Throughput Analysis of IEEE 802.11 Distributed Coordination Function in Presence of Hidden Stations. Tech. Rep., Stanford Univ. (2003)
5. Tobagi, F., Kleinrock, L.: Packet Switching in Radio Channels, Part II - The Hidden Terminal Problem in Carrier Sense Multiple Access and the Busy Tone Solution. IEEE Trans. on Communications 23(12), 1417–1433 (1975)
6. Vassis, D., Kormentzas, G.: Throughput Analysis of IEEE 802.11 Ad Hoc Networks Under the Hidden Terminal Problem. In: Proc. CCNC 2006, pp. 1273–1276 (2006)
7. Hou, T.C., Tsao, L.F., Liu, H.C.: Analyzing the Throughput of IEEE 802.11 DCF Scheme with Hidden Nodes. In: Proc. VTC 2003, pp. 2870–2874 (2003)

Study of EMC Testing Tools

Bo Yuan and Zhiwei Wang

Department of Computer
Tangshan Vocational College of Science and Technology
Tangshan Hebei China
{skyinwater,tskjwzw}@yahoo.cn

Abstract. EMC is always top topic in the power electronic products. Although EMC has the theory of Maxwell to support it, however, the EMC design engineer feels hard to design and solve problems by guide of theory.They just can improve the ability of solving EMC problems by designing much more products. So EMC make engineer to feel vague,mysterious and hard. We need do some basic studies about EMC to improve this situation.

Keywords: EMC, EMI, testing, interference sources, shaking frequency.

Foreword

EMC defines the ability of the equipment can work normal in the electromagnetic environment and has no any effect on the other equipments in the system. EMC includes two areas:one means the equipment can not exceed certain limits, as we can say it EMI. The other means the equipment can work normal condition in the electromagnetic interference environment, as we can say it EMS. From the up two explanations, We can induce EMC has three factors: the source of interference, the mode of interference, the tools used to test the interference. In this paper, firstly we study the test tool, secondly we study the jittering frequency by simulation.

Now we do some explanations about the parameters of the EMI receiver.

1 Structure of Quasi-Peak Detector

The quasi-peak receiver is mainly made up of three parts which are Mixer、Filter (Resolution Bandwidth)、Detector. So We can understood the principle of the receiver by studying the three parts.

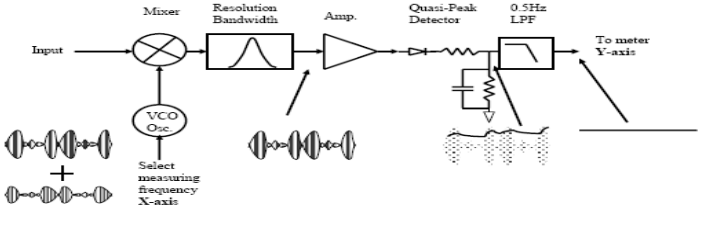


Fig. 1. Structure Of Quasi-Peak Receiver

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 861–867.
springerlink.com

1.1 Mixer

Selecting signal frequency is the basic function of the mixer. The function can be realized by scanning the local oscillator signal. The detailed realization is that the testing signal frequency is firstly transformed to intermediated frequency signal ,and then,we can obtain the signal magnitude by dealing with the intermediated frequency signal using the mixer.

1.2 Filter

We can using the figure 2 to demonstrate the system frequency response characteristic of the intermediated frequency filter.

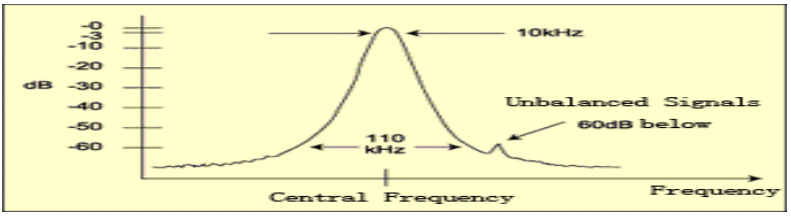


Fig. 2. Intermediated Frequency Filter

1.3 Detector

The detector is the equipment which can detect some useful information about the signal. The detector can identify if the wave ,oscillation is present. It is usually used to extract the information which the signal carries. Usually the frequency spectrum analyzer performance such as: frequency resolution, dynamic range and scan time, is related to the detector.

The frequency spectrum analyzer's characteristics such as frequency resolution, dynamic range and scan time, characteristics are interrelated. Frequency resolution can be regarded as an unknown signal band "Scan Window" waveform, The waveform of scan window is similar to the figure 2. The frequency spectrum analyzer can provide the optional Frequency resolution which can be called the resolution bandwidth RBW (ResolutionBandwidth). RBW can be regard as the-3dB band-pass bandwidth of the intermediated frequency amplifier channel.The frequency resolution is crucial to measure the signals, because the frequency spectrum analyzer can only distinguish those signals which the frequency is closely by frequency resolution.

In some interference applications, the signals have a completely different magnitudes. In this case, "optional" will be an important pointer to the weak signal in two signals . The reason is the weak signal will likely be annihilated in the filter edge of the strong signal. We can explain the reason as following description:

We know the waveform factor of spectrum analyzer defines the IF amplifier ratio which is the-60dB bandwidth divide with -3dB bandwidth. We can know the conventional waveform factor of 10kHz RBW filter of the figure 2 is 11:1. So the signal magnitude of 110KHz is -60dB ,but the magnitude of the under 60KHz signal will be decreased to half the value. When the the interval frequency between the two signals is

60KHz, but the magnitude of the weak signal is 60dB lower than the strong signal. This condition will result that the weak signal annihilates in the edge of the main signals.

The measurement precision of any frequency spectrum analyzer is depend on the precision of the various components made up for frequency spectrum analyzethe . We should think about the measured precision of the frequency spectrum analyzer when we compare the magnitude of the testing signal with the limits of the standard specification. During testing a typical interference. tester often need to determine the number of ratios, such as the margin of the desired signal compared with the interference signal in the same working frequency. So the relative precision is more important than the absolute precision.

We can know the waveform factor is: 11:1, the frequency resolution is 10kHz (decay-3dB) by reading the setup value of the frequency spectrum analyzer (in fact we can use the resolution and waveform factor to describe the specific performance of a filter).

EMI receiver has different frequency resolution under the condition of different product testing standards and frequency bands. As shown by figure3.

QP-detector	*10 - 150kHz*	*0.15 – 30MHz*	*30 – 300MHz*	*0.3 – 1GHz*
6dB bandwidth	*0.2kHz*	*9kHz*	*120kHz*	*120kHz*
Charge time	*45mS*	*1mS*	*1mS*	*1mS*
Dizcharge time	*500mS*	*160mS*	*550mS*	*550mS*

Fig. 3. Test

It can be seen from the figure, the frequency resolution is 9KHz(attenuation is -6dB) and the waveform factor is 10:1 of the EMI receiver under the condition of CE test. In the other words, the frequency is 90KHz when the magnitude is attenuated with 60dB.

The reason why we set the frequency resolution using 9KHz not 10KHz or 20KHz is we can better distinguish more specific frequency signal using the lowest value. As the black line shown of the figure 4 We can't obtain the 60KHz or 70KHz signal if the bandwidth is 10KHz but using the 3KHz bandwidth. We can also use the 1KHz bandwidth to distinguish.interference signals.

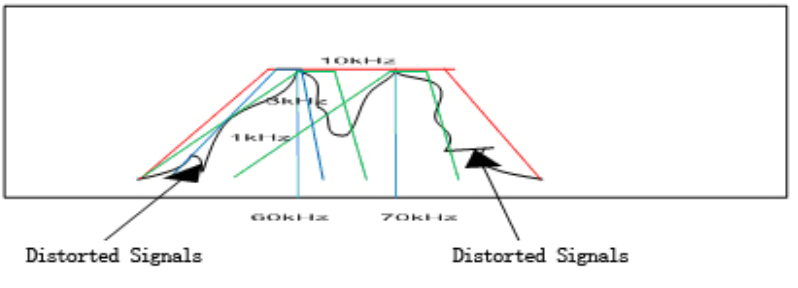


Fig. 4. Different Resolutions Signals

We can use the HP3577 which usually is used to test loop to explain the figure4 (In fact, the frequency spectrum analyzer has some differences with the EMI receiver, however, the explanation is still useful).

For example, we can input one sinusoidal signal to the input channel of the HP3577 (frequency is 10KHz, magnitude is 10mV) by using the Agilent signal generator. See the following figures（scan time is 10 seconds, start frequency is 5Hz, stop frequency is 30MHz）.

enter an Agilent signal generator 10mv of 10kHz sinusoidal signal, this signal is direct access to the general frequency analyzer to analyze the receiving end, the results are as follows (scan time of 10s, from 5Hz-30Mhz):

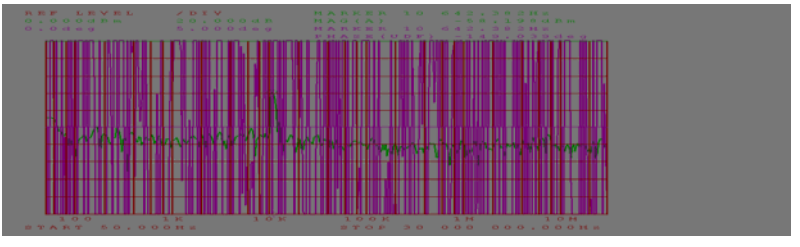


Fig. 5. Range 10mvpp, Frequency 10khz, 10Hz Resolution Bandwidth

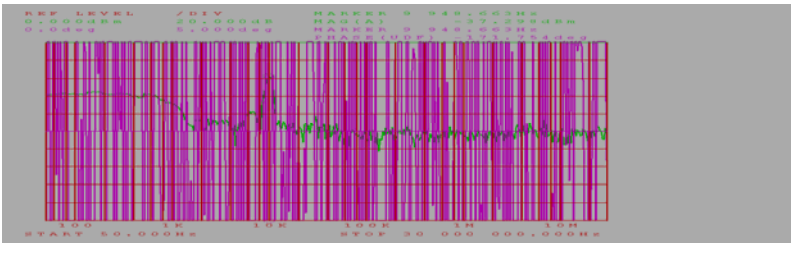


Fig. 6. Range 10mvpp, Frequency 10khz, 1khz Resolution Bandwidth

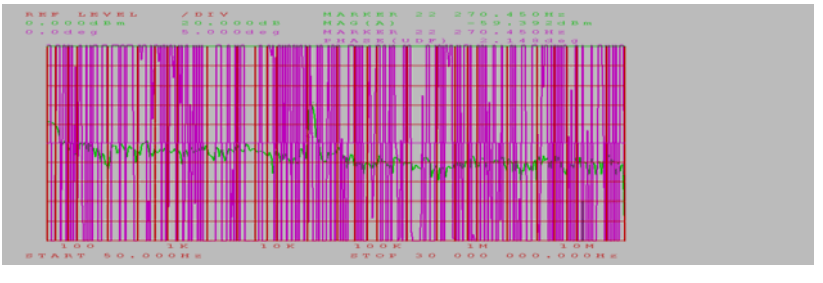


Fig. 7. Range 10mvpp, Frequency 20khz, 10Hz Resolution Bandwidth

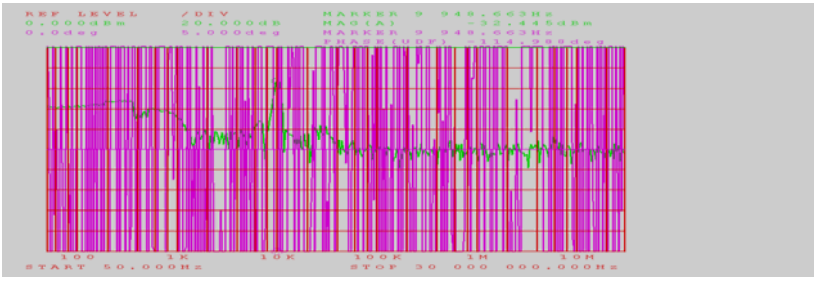


Fig. 8. Range 10mvpp, Frequency 10khz, 1khz Resolution Bandwidth

We can higher the frequency resolution by reducing the bandwidth, but this will result in longer the scan time. So we can set up the scan time by handing ,and then this mode will speed up the scan process to improve the measuring speed. However, changing of scan time will cause error of the measurement.

Reduce the resolution bandwidth can improve the resolution, but the measurement is a lower resolution will increase the scan time. Then we can set the scan time for people to speed up its scanning speed and improve the measurement speed. However, due to changes in scan time can cause measurement errors on the specific frequency is increased, while the rate decreased.

In fact, we can use a formula (the same to EMI receiver)to demonstrate the relationship of the resolution bandwidth, scan time, start frequency ,stop frequency.

$$T_{charge} = \frac{BW}{f_{stop} - f_{initial}} T_{sweep}$$

We can obtain the conclusion that the frequency analyzer not only used to test loop, but also used to test the EMI.

2 The Shaking Frequency Simulation

After explaining the some terms about the EMI receiver ,now ,we can explain the jittering which is used to "cheat" EMI receiver to make the product meeting the standard by simulation.

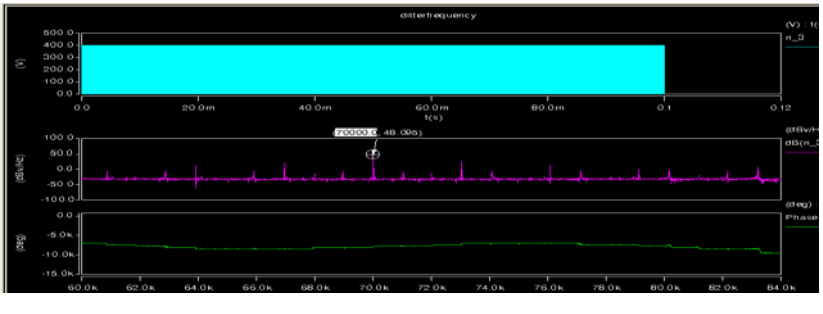


Fig. 9. Switching Frequency 70khz, 400V, Duty Cycle 0.5 Fourier Analysis

We can obtain the fundament value by Fourier transform, the value is 251 volt(48dB volt). From the figure 9, the simulation result is the same to the value by Fourier transform. So we can analyze the result by simulation instead of mathematical transformation.

So, we can also analyze the jittering frequency by simulation.

The key of the jittering algorithm is jittering step and jittering amplitude. We can simulate the condition of the center frequency is 70kHz and the jittering step is 250Hz.following is the simulation result.

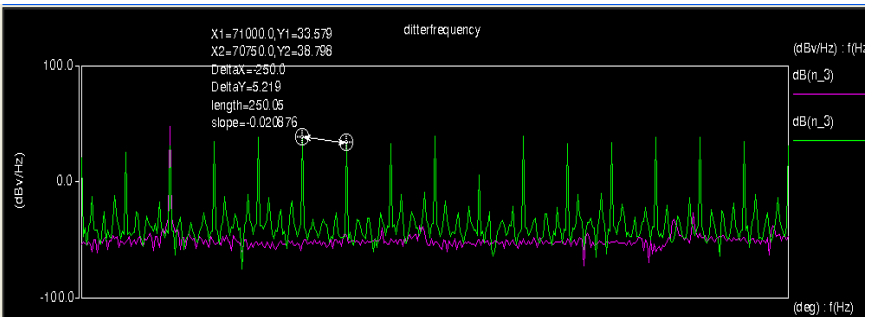


Fig. 10. The Red Line Is Not Shaking Frequency, Shaking Frequency For The Green Line

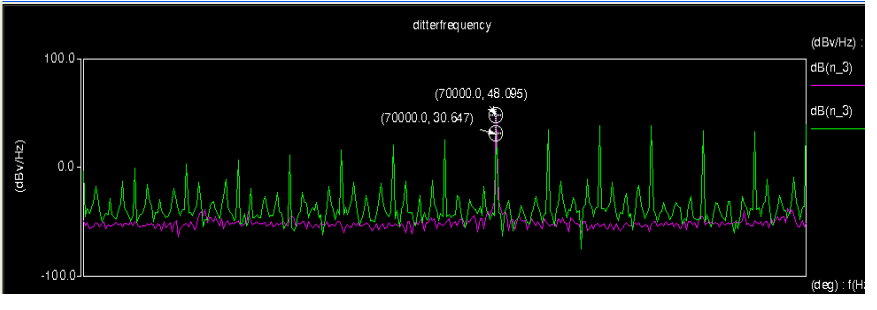


Fig. 11. The Red Line Is Not Shaking Frequency, Shaking Frequency For The Green Line

As shown the simulation result,We can reduce the amplitude of 70kHz(about 18dB) by jittering frequency, however, this will higher the amplitude of the frequency 70KHz+-n*250Hz.

In conclusion, the jittering can “cheat” EMI receiver to make the product meet the limits of the standard if the reasonable jittering step is designed.

References

1. Qu, C.-Y., Jiang, Q., Lv, R. (eds.): Electromagnetic emission and susceptibility measurements. Southeast University Press, Nanjing (1988)
2. Zhang, J., Sun, Z.: Hyundai Automotive EMC theory and design basis. Beijing Jiaotong University Press, Beijing (2009)

3. Ma, W., Zhang, L.: Meng Jin of the independent power systems and power electronics EMC, pp. 19–24, 176–209. Science Press, Beijing (2007)
4. JW Molyneux-Child, EMC shielding materials. Oxford; Boston Newnes (1997c)
5. Williams, T.: EMC for product designers, Witerwoof Inc. (2007)
6. Tesche, F.M., Ianoz, M.V., Jason, T., Karlsson: EMC Analysis Methods and Computational Models

A Hybrid Routing Tree to Avoid the Energy Hole Problem in Wireless Sensor Network

Yuhua Liu[1], Wenwen Dai[1], Kaihua Xu[2], and Meirong Zheng[1]

[1] Dept of Computer Science,
Huazhong Normal University, Wuhan, China
[2] College of Physical Science and Technology
of Huazhong Normal University, Wuhan, China
yhliu@mail.ccnu.edu.cn,
{wwdaiwh,tixadt}@163.com, Xukh@phy.ccnu.edu.cn

Abstract. In order to solve the problem of "energy hole" in wireless sensor networks, this paper proposed a Hybrid Routing Tree Based on Ring Network to avoid this problem. The routing tree base on ring network mode uses the threshold E_s and the middle ring to control the communication style among nodes. The simulation results show that the routing algorithm is effective to avoid the "energy hole" problem and maximum extension the network lifetime.

Keywords: WSN, Communication styles, Energy hole, Hybrid Rout Tree algorithm.

Introductions

The wireless sensor network is a self-organizing network system, which is formatted by multiple hops radio communications among the tiny sensor nodes .Those sensor nodes are deployed in the monitoring area to perceive, collect and dispose the information from the monitoring objects in the network, and send the information to observer. [1-2] However, the energy, the ability of calculation and communication of those sensor nodes are limited. Those sensor nodes are random displayed in the so bad environment that we cannot help its energizing. So, design a rote algorithm with high-efficiency and low-energy consumption to prolong the lifetime of networks is very important.

At present, there have been many energy routing algorithms are designed. Most of them adopt single-hop or multi-hop as its communication style. But those algorithms cause the “energy hole” [3-4] easily. When use multi-hop communication, the nodes near the sink must consume excessive energy to forward the data from other nodes, and thus the formation of energy hole. Meantime, when uses single hop communications, the nodes far away from the sink node will consume excessive energy for the long-distance communication. Those two ways can lead to the premature death of network. In this paper, we propose a hybrid routing tree algorithm to balance the two means of communication; the algorithm can balance the remaining energy and prolong the network lifetime.

M. Zhu (Ed.): Business, Economics, and Financial Sci., Manag., AISC 143, pp. 869–876.
springerlink.com

1 Related Work

Related works show that the sensor network has the phenomenon of uneven energy consumption. A study on this problem developed by literature [5] suggests that there is remaining more than 90% of the initial energy when the network died. Many research studies have been carry out on this topic. Literature [6]'s model focuses on the deployment. In this model, near the sink with higher node density to bear a large number of energy consumption. However, this approach causes the deployment of nodes is so hard that it difficult to achieve in the real environment. Literature [7]'s model draws attention on mobile sensor nodes, those mobile sensor nodes will avoid the enormous cost of artificial deployment and solve the "energy hole" problem. The approach taken by literature [3] is based on the ring network model, which avoids the "energy hole" problem by analyzing the sequence of network traffic.

This article bases on the literature [3]'s network model, building the minimum cost hybrid routing tree by Kruskal, and making the communication among nodes switch between single hop and multi-hop under the effect of energy threshold and distance threshold. This algorithm avoids the "energy hole" phenomenon by sharing the energy consumption to each layer from outside to inside gradually.

2 Wireless and Network Model

2.1 Network Model

Shown in Figure 1, this paper uses the minimum cost routing tree which is based on ring network model and uses hybrid routing idea to solve the "energy hole" problem.

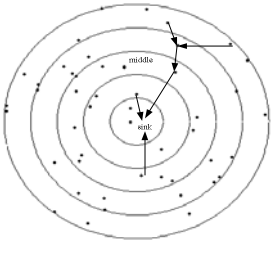


Fig. 1. Ring network model

This article assumes that the sensor network has the following properties: (1) the sensors will work independently and stationary once deployed; (2) the energy of sensor nodes can not be supplied except the sink nodes; (3) the width of each ring is equal to; (4) each node has a unique identification ID; (5) transmission strategy is so perfect that there is no conflict and retransmission; (6) all of the nodes have simple data fusion capabilities and the transmission power is adjustable, this paper assumes that there always have data to be sent.

Definition (Network Lifetime). Li and Mohapatre in literature [6] give the definition of network lifetime, which is the time from the beginning until part of the sensor

nodes die. A ring dies when the ring can not forward any data or sends their own data in sensor network. In this paper, the definition of network lifetime is the time from beginning to a ring in the network has been dead.

2.2 Energy Consumption Model

A typical sensor node includes data collection, data processing and data sending or receiving. We only consider the energy consumption of receiving, sending data and data processing, ignoring the energy consumption of data collection because of this part consumes less energy in the energy model. This article uses the formula of energy consumption in literature [6] in our analysis and simulation:

$$E_{tr} = (\beta_1 + \beta_2 d^{\alpha})L \tag{1}$$

$$E_{rec} = \beta_3 L \tag{2}$$

E_{tr} is the energy consumption of transmitting data, E_{rec} is the energy consumption of accepting data, L represents the frequency of generation and transmission data, d^{α} is caused by the consumption path, when the distance is greater than d_0 ,it uses multi-path model ($\alpha = 4$), when the distance is less than d_0 , the free space model is used ($\alpha = 2$). According to the literature [6], this paper uses the following parameters:

$$\beta_1 = 45 \times 10^{-9} J/bit,$$
$$\beta_2 = 10 \times 10^{-12} J/bit/m^2 (\alpha = 2),$$
$$\text{or } \beta_2 = 0.001 \times 10^{-12} J/bit/m^4 (\alpha = 4),$$
$$\beta_3 = 135 \times 10^{-9} J/bit$$

The energy of data fusion $E_{DA} = 5nJ(bit.message^{-1})$.

At the phase of data transmission, the energy threshold is used to control the communication styles. And the energy threshold can limits the remaining energy of forwarding sensor nodes to protect the forwarding node, it allows the sensor node to forward data only when $E_{ever} \geq \partial E_s$, which E_{ever} is the average residual energy of all nodes in forwarding ring, E_s is the initial average energy of all nodes in forwarding ring, ∂ is weighting factor $(0 < \partial < 1)$.

Literature [8] gets the relations of ∂ and the simulation time by a large number of simulations, the network performance is better when $0.7 < \partial < 0.8$. To ensure the forwarding ring has enough energy to forward the data from other nodes and will not run out of energy, this paper uses the average value: $\partial = 0.75$.

3 Protocol Descriptions

3.1 Initialization of the Minimum Cost Routing Tree

Sensor nodes send a short message to sink node after the deployment of network. The message includes the residual energy of node and the node's ID. Sink node generates

the topology distribution of the whole network and produces a minimum cost spanning tree with sink node as the root of the tree. Finally, sink node sends "notification route" message to all the nodes, and requires each node to confirm. In this process, we need to focus on the problem as follows:

① Determine the number of middle ring. As shown in figure (2), the network performance well when the middle ring number is 5. So the middle ring number is 5.

② The intermediate forwarding node must record the signal strength RSSI when forwarding data.

③ In the process of establishing the routing tree, it must ensure that the parents of all nodes in the ring C_i are in the ring C_{i-1}. That is all the nodes in ring C_i are in the same layer of routing tree to ensure that energy consumption is carried out by ring.

④ There are at least three options in the parent nodes of each node's routing table and the probability to be parent node is reducing in turn, that is it considers the use of the second node in the routing table as parent node only when the first parent node can not be used, and so on. This ensures that the link can't be broken by the energy consumption or natural causes.

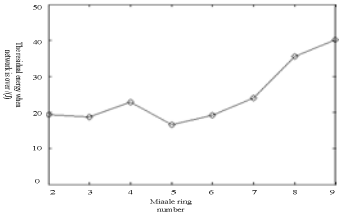


Fig. 2. Number of middle ring and the remaining energy comparison

3.2 Data Transfer Phase

After initialization, it considers the nodes in the ring C_i and the ring C_{mid}. It uses multi-hop manner to send data to sink node when $i > mid$ while the nodes send data to sink node directly when $i < mid$.

All nodes in the middle ring must carry the residual energy E_{rec} of the node when forwarding data. The sink node calculates the average residual energy E_{ever} of all the nodes in the middle ring according to the formula (3) and compares it with the energy threshold E_s after receiving and processing the data. If $E_{ever} < E_s$, the sink node sends a message of replacing the middle ring to the ring whose ring number is C_{mid-1}, the original middle ring will uses the established routing tree to forward data rather than directly sending data to the sink node. Sink node does not send any messages and continue to use the original way of communication between the nodes when $E_{ever} >= E_s$.

$$E_{ever} = \frac{E_1 + E_2 + ... + E_n}{1 + 2 + ... + n}, n << N \tag{3}$$

Which, $E_1, E_2, ...E_n$ stand for the residual energy of nodes in the middle ring, N is the total number of sensor nodes in the network.

3.3 The Maintenance of Routing Tree

Minimum cost routing tree must consider the situation of nodes' death caused by energy consumption and natural damage in practical applications. Hence, the new agreement needs to add the necessary route maintenance strategy. The method is that each node sends an inquiry message to its parent nodes periodically at a certain time. If the node does not receive the response from its first parent node in time, it indicates that the link to the parent node has been destroyed. At this point, the node will find its routing table in sub-optimal node as its parent node.

3.4 Algorithm Description

Based on the above ideas, the algorithm will work as follows:

Step1: Initialize the network. 200 nodes are deployed in a ring area randomly, the width of each ring is same to each other and the ring which is close to sink has high-density nodes.

Step2: Sink node obtains the distribution of the whole network and generates the minimum cost routing tree.

Step3: Sink node sends route messages to all sensor nodes according to the minimum cost routing tree.

Step4: Sink node calculates the number of middle ring and sends message to all nodes in middle ring and tells those nodes to forward the data from other ring.

Step5: Sensor nodes need determine its ring number before sending data to sink:

If the ring number is lager than middle ring: all nodes in the ring communicate by multi-hop. Otherwise, nodes send the monitored data and its own residual energy information to sink node directly.

Step6: Sink nodes deal with the collected information simply and calculate the average residual energy value E_{ever}

Step7: Compare E_{ever} and E_s :

If $E_{ever} < E_s$, the sink node sends a message of replacing the middle ring to the outer ring;

Otherwise: sink node does not send any message.

Step8: Nodes determine whether its link is connected. If it is not connected, it will change its parent node, otherwise, turn into step 2.

4 Simulation Analyses

In this experiment, 200 nodes are randomly deployed in the region of $200m \times 200m$, each node has $0.5J$ of energy at beginning and sends out $2 \times 10^3 bit$ data every time. Other parameters are shown in table 1. This paper compares the single-hop protocol Direct algorithm and multi-hop protocol Multi algorithm with our energy hole avoiding algorithm HRTBR.

Table 1. The parameters in simulation

parameter	value	parameters	value
Monitoring area	(-100, -100)-(100, 100)	The number of middle ring	5
Base station	(0,0)	∂	4
N	200	θ	0.75
The initial energy of node	0.5 J	β_1	$45 \times 10^{-9} J / bit$
The package' number	2000 bit	β_2	$10^{-15} J / bit / m^4$
The ring number	10环	β_3	$135 \times 10^{-9} J / bit$

4.1 Network Lifetime

This article defines when a certain ring of death that all the nodes that the network killed. A round is named after one communication process among all nodes in the network. As can be seen from Figure 3, the network died at 110 rounds with Multi algorithm and there are remaining nearly 150 nodes. And, the network died at 100 rounds with Direct algorithm and there are remaining nearly 190 nodes. However, the proposed algorithm can extend the network to nearly 200 rounds while remaining the only 100 nodes. So, the algorithm is effective to extend the network lifetime.

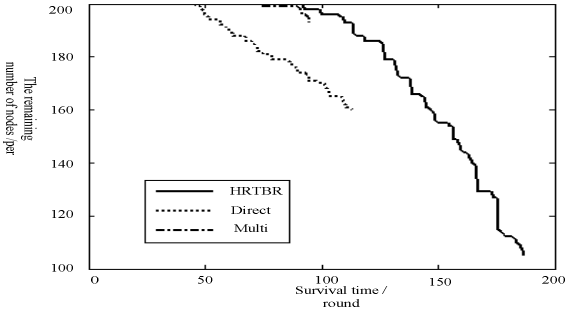


Fig. 3. Network Lifetime

4.2 Energy Efficiency

As shown in Figure 4, the energy efficiency of Direct algorithm is 0.3463 and Multi algorithm is 0.5442, while the HRTBR algorithm is only 0.8370. The HRTBR algorithm can solve the shortage of the two algorithms. It makes each node in the network uses energy equally and extends the network lifetime and improves the utilization of energy in the network.

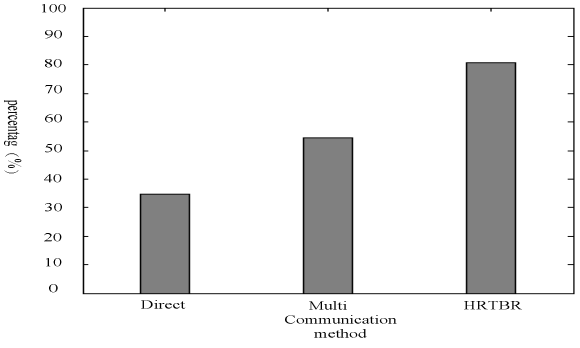


Fig. 4. Energy efficiency

It can be seen from experimental results that this routing algorithm not only effectively improve the energy efficiency but also prolong the network lifetime extensively. The reason is that the nodes far from the base station must transmit data over long distances and consume a large amount of energy when using Direct mode routing policy, which leads to premature death of the network. And Multi mode routing policy is used, the nodes near the base station will overload for forward data from other nodes and consume their own energy largely, which is leading to the end of the network. Contrast to Multi policy and Direct policy, the HRTBR algorithm proposed in this paper combines the advantages of these two means of communication. So, the nodes far from the base station use multi-hop forwarding strategy to avoid long-distance transmission and nodes near of base station send data directly to the base station in order to avoid node load is too large to premature death. This paper uses energy threshold and middle ring to control the balance of these two means of communication, not only take into account the residual energy of nodes, but also consider the long-distance communication and different energy consumption. Therefore, the proposed algorithm, comparing with the above routing algorithm, can improve energy efficiency and prolong network lifetime at a great advantage.

5 Summaries and Outlook

The phenomenon of “energy hole” in wireless sensor networks always causes the network to die early. This paper finds a proper distance between single hop and multi-hop base on ring model and proposes the idea of hybrid routing tree to solve the "energy hole" problem. The simulation analyses tell us this algorithm can effectively

avoid this problem and prolong the network lifetime. The algorithm has good scalability and adapt to the physical environment. Therefore, we should consider using this method in the clustering sensor network in our next study. We hope it can be applied on a larger scale network.

References

1. Song, C., Liu, M., Gong, H.-G., et al.: ACO-based algorithm for solving energy hole problem in Wireless Sensor Networks. Journal of Software 20(10), 2729–2743 (2009)
2. Gavrilovska, L., Krco, S., Milutinovic, V., Stojmenovic, I., Trobec, R.: Application and Multidisciplinary Aspects of Wireless Sensor Networks: Concepts, Integration, and Case Studies, p. 282. Springer-Verlag New York,Inc. (2010)
3. Olariu, S., Stojmenovic, I.: Design. guidelines for maximizing lifotime and avoiding energy holes in sansor networks with uniform distribution and uniform reporting. In: Domingo-Pascual, J. (ed.) Proc. of the IEEE INFOCOM, pp. 1–12. IEEE Communications Society, New York (2006)
4. Li, C.-F., Chen, G.-H., Ye, M., Wu, J.: An Uneven Cluster-Based Routing Protocol for Wireless Sensor Network. Journal of Computers 30(1), 27–36 (2007)
5. Lian, J., Naik, K., Agnew, G.: Data capacity improvement of wireless sensor networks using non-uniform sensor distribution. Int'I Journal of Distilbuted, Sensor Networks 2(2), 121–145 (2006)
6. Li, J., Mohapatra, P.: Analytical modeling and mitigation techniques for the energy hole problems in sensor networks. Pervasive and Mobile Computing 3(3), 233–254 (2007)
7. Shine, H.Y., Yu, G.J., Sheu, J.P.: Energy hole healing protocol for surveillance sensor networks. In: Proc. of the Workshop OIl Wireless, Ad Hoc, and Sensor Networks, WASN (2005)
8. He, G.-Y., Chen, D.: High Energy-Efficient Routing Protocol Based on Optimal Cluster-Heads for Wireless Sensor Network. Journal of Sensor and Actuators 21(10), 1739–1743 (2008)

Author Index