Yuhang Yang
Maode Ma
Baoxiang Liu (Eds.)

Communications in Computer and Information Science 392

Information Computing and Applications

4th International Conference, ICICA 2013
Singapore, August 2013
Revised Selected Papers, Part II

Part 2

Springer

Communications in Computer and Information Science 392

Yuhang Yang Maode Ma Baoxiang Liu (Eds.)

Information Computing and Applications

4th International Conference, ICICA 2013
Singapore, August 16-18, 2013
Revised Selected Papers, Part II

Springer

Volume Editors

Yuhang Yang
Shanghai Jiao Tong University
800 Dongchuan Road, Shanghai 200240, China
E-mail: yhyangsjtu@gmail.com

Maode Ma
Nanyang Technological University
School of Electrical and Electronic Engineering
Nanyang Avenue, Singapore 639798
E-mail: emdma@ntu.edu.sg

Baoxiang Liu
Hebei United University, College of Science
46 West Xinhua Road, Tangshan, China
E-mail: liubx5888@126.com

ISSN 1865-0929 e-ISSN 1865-0937
ISBN 978-3-642-53702-8 e-ISBN 978-3-642-53703-5
DOI 10.1007/978-3-642-53703-5
Springer Heidelberg New York Dordrecht London

Library of Congress Control Number: 2013957072

CR Subject Classification (1998): F.1.2, H.3.4, H.3.5, G.3, H.2.7, H.2.8, K.6, C.2.1, C.2.4, J.1, J.3, J.7

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

Printed on acid-free paper

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

Preface

Welcome to the proceedings of the 4th International Conference on Information Computing and Applications (ICICA 2013), which was held during August 16–18, 2013, in Singapore.

As future-generation information technology, information computing and applications become specialized, information computing and applications including hardware, software, communications, and networks are growing with ever increasing scale and heterogeneity, and becoming overly complex. The complexity is becoming more critical along with the growing applications. To cope with the growing and computing complexity, information computing and applications focus on intelligent, self-manageable, scalable computing systems and applications to the maximum extent possible without human intervention or guidance.

With the rapid development of information science and technology, information computing has become the third approach of scientific research. Information computing and applications is the field of study concerned with constructing intelligent computing, mathematical models, numerical solution techniques, and using computers to analyze and solve natural scientific, social scientific, and engineering problems. In practical use, it is typically the application of computer simulation, intelligent computing, Internet computing, pervasive computing, scalable computing, trusted computing, autonomy-oriented computing, evolutionary computing, mobile computing, applications and other forms of computation to problems in various scientific disciplines and engineering. Information computing and applications is an important underpinning for techniques used in information and computational science and there are many unresolved problems worth studying.

The ICICA 2013 conference provided a forum for engineers and scientists in academia, industry, and government to address the most innovative research and development including technical challenges and social, legal, political, and economic issues, and to present and discuss their ideas, results, work in progress and experience on all aspects of information computing and applications.

There was a very large number of paper submissions (665). All submissions were reviewed by at least three Program or Technical Committee members or external reviewers. It was extremely difficult to select the presentations for the conference because there were so many excellent and interesting submissions. In order to allocate as many papers as possible and keep the high quality of the conference, we finally decided to accept 126 papers for presentations, reflecting an 18.9% acceptance rate; 63 papers were included in this volume. We believe that all of these papers and topics not only provided novel ideas, new results, work in progress and state-of-the-art techniques in this field, but also stimulated future research activities in the area of information computing and applications.

The exciting program for this conference was the result of the hard and excellent work of many individuals, such as the Program and Technical Committee members, external reviewers, and publication chairs under a very tight schedule. We are also grateful to the members of the local Organizing Committee for supporting us in handling so many organizational tasks, and to the keynote speakers for accepting to come to the conference with enthusiasm. Last but not least, we hope you enjoy the conference proceedings.

September 2013

Yuhang Yang
Maode Ma
Baoxiang Liu

Organization

ICICA 2013 was organized by Nanyang Technological University, Hebei Scene Statistical Society, IEEE Education Chapter of Singapore, and sponsored by the National Science Foundation of China, Republic Polytechnic of Singapore, Hunan Institute of Engineering, Yanshan University, and Northeastern University at Qinhuangdao. It was held in cooperation with *Communications in Computer and Information Science* (CCIS) of Springer.

Executive Committee

Honorary Chair

Jun Li	Hebei United University, China

General Chairs

Yanchun Zhang	University of Victoria, Australia
Baoxiang Liu	Hebei United University, China
Maode Ma	Nanyang Technological University, Singapore

Program Chairs

Chunfeng Liu	Hebei United University, China
Fengbo Hou	Chengde Petroleum College, China
Leizhen Wang	Northeastern University at Qinhuangdao, China
Jiang Lijun David	IEEE Education Chapter of Singapore, Republic Polytechnic, Singapore

Local Arrangements Chairs

Jincai Chang	Hebei United University, China
Aimin Yang	Hebei United University, China

Steering Committee

Qun Lin	Chinese Academy of Sciences, China
Lorna Uden	Staffordshire University, UK
Nadia Nedjah	State University of Rio de Janeiro, Brazil
Rafal Kapelko	Wroclaw University of Technology, Poland
Wende Zhong	Nanyang Technological University, Singapore
Nadia Nedjah	State University of Rio de Janeiro, Brazil

Gaoxi Xiao	Nanyang Technological University, Singapore
Bingsheng He	Nanyang Technological University, Singapore
Lorna Uden	Staffordshire University, UK
Yiming Chen	Yanshan University, China
Xiangdong Song	Yanshan University, China
Changcun Li	Hebei United University, China
Zhijiang Wang	Hebei United University, China
Rongbo Zhu	South-Central University for Nationalities, China
Guohuan Lou	Hebei United University, China
Xiaoqi Li	Northeastern University at Qinhuangdao, China
Dianchuan Jin	Hebei United University, China
Xiaosheng Wang	Hebei University of Engineering, China
Shenggui Duan	Shijiazhuang University of Economics, China
Xiangdong Song	Yanshan University, China
Fengbo Hou	Chengde Petroleum College, China
Xiaomin Wang	Northeastern University at Qinhuangdao, China

Publicity Chairs

Aimin Yang	Hebei United University, China
Xilong Qu	Hunan Institute of Engineering, China

Publication Chairs

Maode Ma	Nanyang Technological University, Singapore

Financial Chair

Jincai Chang	Hebei United University, China

Local Arrangements Committee

Lihong Li	Hebei United University, China
Shaohong Yan	Hebei United University, China
Yamian Peng	Hebei United University, China
Lichao Feng	Hebei United University, China
Yuhuan Cui	Hebei United University, China

Secretariat

Kaili Wang	Hebei United University, China
Jingguo Qu	Hebei United University, China
Yafeng Yang	Hebei United University, China

Program/Technical Committee

Yuan Lin	Norwegian University of Science and Technology, Norway
Yajun Li	Shanghai Jiao Tong University, China
Yanliang Jin	Shanghai University, China
Mingyi Gao	National Institute of AIST, Japan
Yajun Guo	Huazhong Normal University, China
Haibing Yin	Peking University, China
Jianxin Chen	University of Vigo, Spain
Miche Rossi	University of Padova, Italy
Ven Prasad	Delft University of Technology, The Netherlands
Mina Gui	Texas State University, USA
Nils Asc	University of Bonn, Germany
Ragip Kur	Nokia Research, USA
On Altintas	Toyota InfoTechnology Center, Japan
Suresh Subra	George Washington University, USA
Xiyin Wang	Hebei United University, China
Dianxuan Gong	Hebei United University, China
Chunxiao Yu	Yanshan University, China
Yanbin Sun	Beijing University of Posts and Telecommunications, China
Guofu Gui	CMC Corporation, China
Haiyong Bao	NTT Co., Ltd., Japan
Xiwen Hu	Wuhan University of Technology, China
Mengze Liao	Cisco China R&D Center, China
Yangwen Zou	Apple China Co., Ltd., China
Liang Zhou	ENSTA-ParisTech, France
Zhanguo Wei	Beijing Forestry University, China
Hao Chen	Hu'nan University, China
Lilei Wang	Beijing University of Posts and Telecommunications, China
Xilong Qu	Hunan Institute of Engineering, China
Duolin Liu	ShenYang Ligong University, China
Xiaozhu Liu	Wuhan University, China
Yanbing Sun	Beijing University of Posts and Telecommunications, China
Yiming Chen	Yanshan University, China
Hui Wang	University of Evry, France
Shuang Cong	University of Science and Technology of China, China
Haining Wang	College of William and Marry, USA
Zengqiang Chen	Nankai University, China
Dumisa Wellington Ngwenya	Illinois State University, USA

Hu Changhua	Xi'an Research Institute of Hi-Tech, China
Juntao Fei	Hohai University, China
Zhao-Hui Jiang	Hiroshima Institute of Technology, Japan
Michael Watts	Lincoln University, New Zealand
Tai-hon Kim	Defense Security Command, Korea
Muhammad Khan	Southwest Jiaotong University, China
Seong Kong	The University of Tennessee USA
Worap Kreesuradej	King Mongkuts Institute of Technology Ladkrabang, Thailand
Uwe Kuger	Queen's University of Belfast, UK
Xiao Li	CINVESTAV-IPN, Mexico
Stefa Lindstaedt	Division Manager Knowledge Management, Austria
Paolo Li	Polytechnic of Bari, Italy
Tashi Kuremoto	Yamaguchi University, Japan
Chun Lee	Howon University, Korea
Zheng Liu	Nagasaki Institute of Applied Science, Japan
Michiharu Kurume	National College of Technology, Japan
Sean McLoo	National University of Ireland, Ireland
R. McMenemy	Queen's University Belfast, UK
Xiang Mei	The University of Leeds, UK
Cheol Moon	Gwangju University, Korea
Veli Mumcu	Technical University of Yildiz, Turkey
Nin Pang	Auckland University of Technology, New Zealand
Jian-Xin Peng	Queen's University of Belfast, UK
Lui Piroddi	Technical University of Milan, Italy
Girij Prasad	University of Ulster, UK
Cent Leung	Victoria University of Technology, Australia
Jams Li	University of Birmingham, UK
Liang Li	University of Sheffield, UK
Hai Qi	University of Tennessee, USA
Wi Richert	University of Paderborn, Germany
Meh shafiei	Dalhousie University, Canada
Sa Sharma	University of Plymouth, UK
Dong Yue	Huazhong University of Science and Technology, China
YongSheng Ding	Donghua University, China
Yuezhi Zhou	Tsinghua University, China
Yongning Tang	Illinois State University, USA
Jun Cai	University of Manitoba, Canada
Sunil Maharaj Sentech	University of Pretoria, South Africa
Mei Yu	Simula Research Laboratory, Norway
Gui-Rong Xue	Shanghai Jiao Tong University, China
Zhichun Li	Northwestern University, China

Lisong Xu	University of Nebraska-Lincoln, USA
Wang Bin	Chinese Academy of Sciences, China
Yan Zhang	Simula Research Laboratory and University of Oslo, Norway
Ruichun Tang	Ocean University of China, China
Wenbin Jiang	Huazhong University of Science and Technology, China
Xingang Zhang	Nanyang Normal University, China
Qishi Wu	University of Memphis, USA
Jalel Ben-Othman	University of Versailles, France

Table of Contents – Part II

Internet Computing and Applications

Engineering Management and Applications

Intelligent Computing and Applications

Business Intelligence and Applications

Knowledge Management and Applications

Information Management System

Computational Statistics and Applications

Table of Contents – Part I

Internet Computing and Applications

Engineering Management and Applications

Intelligent Computing and Applications

Control Engineering and Applications

Cloud and Evolutionary Computing

Knowledge Management and Applications

Computational Statistics and Applications

Data Fusion Algorithm Based on Ultrasonic Sensor Network

Huilin Lu

Wuxi Institute of Commerce, Wuxi, Jiangsu Province, 214153 China

Abstract. For observation information from multiple sensors to analyze consolidated. To arrive at decisions and the information needed to estimate task processing, a novel fusion method is proposed based on the approach degree and weights. The method calculate mean and variance based on the measured sensor's data, Using the maximum and minimum approach degree of this fuzzy set, the approach degree of the measured data from various sensors is processed quantitatively, eliminating outlier data by Grubbs method, assigned the weight's of data measured in the fusion process reasonably, so that the final expression of the data fusion is obtained, thus the data fusion of multi-sensor is realized. The results demonstrate that this method can bring higher fusion precision and more suitable for microcontroller and embedded systems applications.

Keywords: Data fusion, Weight, Ultrasonic sensor network.

1 Introduction

With the increasing complexity of the system, relying on a single sensor for monitoring the physical is clearly too restrictive. Therefore, the fault diagnosis system using multi-sensor technology monitor a variety of feature quantity (such as vibration, temperature, pressure, flow, etc.), and these sensor information fusion, in order to improve the accuracy and reliability of the fault location [1-4]. In addition, manual observation is an important source of information troubleshooting. However, this information sources are often inaccurate because inconvenience quantified or being overlooked by people. Information fusion technologies emerge to address these issues provide a powerful tool [5-8]. For the development and application of fault diagnosis has opened up broad prospects. Through information fusion for multiple sensors detect the information and artificial observed facts in a scientific and reasonable comprehensive treatment. That can improve the condition monitoring and fault diagnosis intelligence.

Information fusion using of computer technology will come from multiple sensors or observation information sources to analyze consolidated. To arrive at decisions and the information needed to estimate task processing. Another information fusion is the data integration [9-11]. But its connotation is broader, more precise, more reasonable, but also more general. Includes not only data, but also knowledge of the signal and, as is customary reasons, many literature still using data fusion. Information fusion is the basic principle: take full advantage of the sensor resources. Through a variety of

Y. Yang, M. Ma, and B. Liu (Eds.): ICICA 2013, Part II, CCIS 392, pp. 1–10, 2013.

sensors and artificial observation information, the various sensors in space and time on the complementary and redundant information and according to some optimization criteria or combination of algorithms to produce coherent interpretation of the observed objects and descriptions. The goal is based on the decomposition of the sensor detection information manual observation information [12, 13]. Through the optimization of the combination of the information derived more valid information.

Complex industrial process control is an important area for data fusion shown in figure 1. By the time series analysis, frequency analysis, wavelet analysis, the signal pattern obtained from the sensor to extract the characteristic data, while the extracted feature data is input neural network pattern recognition, neural network pattern recognition feature level data fusion is carried out in order to identify the characteristics of the system data, and input to the fuzzy expert system for decision-making level fusion. Expert system inference when removed from the knowledge base and the database domain rules and parameters to match with the characteristics of the data (fusion) [14-17]. Finally, we make the system operational status, working conditions and equipment failures.

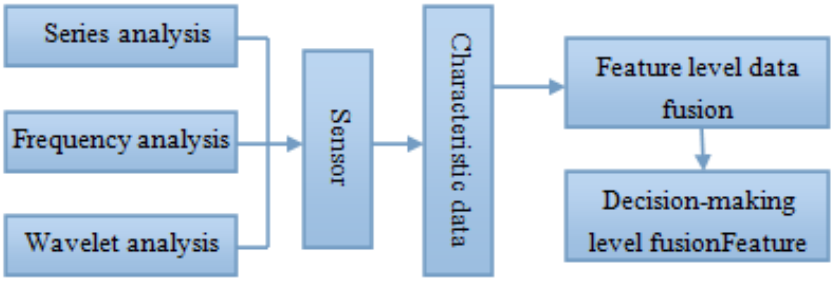


Fig. 1. Data fusion model based on sensor

Therefore, smart devices, sensors; data judged the smart play is crucial. In this paper, the advantages and disadvantages of data integration, improved data fusion algorithm to achieve a high degree of flexibility and high robustness sexual key is in the real environment.

2 Multi-ultrasonic Sensor Data Fusion Principal

Wireless sensor networks typical three types of data fusion algorithms, including those based on the aggregation of distributed database operations, packet data consolidation, and model-driven fusion algorithm in each category, list some typical data fusion method, and these methods were analyzed and compared. energy is important in wireless sensor network resources, and data fusion 's main role is to save energy, therefore, to establish energy model of sensor nodes, data fusion for quantitative analysis of the impact energy of the sensor nodes as well as for wireless life of sensor networks is essential. based on tree network topology, combined with node sleep scheduling mechanism, the establishment of sensor nodes energy model, the simulation

from the energy point of view of a typical data fusion algorithm performance evaluation in the above work based on this monitoring data based on time series models, and time series prediction is proposed for wireless sensor networks based on the predicted time-domain data fusion technology, and thus the wireless sensor network energy optimization paper web temperature data collected as samples by simulation and experimental verification of the method and analysis of simulation results show that the autoregressive prediction algorithm compared with other prediction algorithm has better applicability.

In principal, there are two sensor network traffic [18-20]: a query from the user to the network and from the sensor node to the user perception data. Each sensor node has two effects, might be aware environment, or as a relay node to other nodes to relay the data generated. All data generated by these nodes is likely to block the network and congestion occur, seriously affecting network performance. However, large-scale densely deployed sensor network characteristics result in the majority of these data is invalid, so you can process the data during transmission and reduce invalid data, which appeared within the network data fusion concepts. Wireless sensor network data fusion theory proposed is based on the following principles: Remove redundant data and invalid information, reduce or eliminate poor data reliability, while incorporating process data from different sensor nodes multi-feature information, to reduce network the purpose of the amount of data transmitted. With previous multi-sensor data fusion technology is different, modern fusion technology WSN biggest difference is reflected in the use of technology and the integration of routing protocol layers combine traditional multi- sensor data fusion for the different sources of information as well as various sensor data collected, treated observation target is to achieve better and more intuitive understanding. The wireless sensor network data fusion is not based on this, it is mainly within the entire network by reducing the amount of data transmission to reduce energy consumption, improve data accuracy and reduce the transmission delay to prolong the network lifetime. Therefore, the data fusion in WSN application is necessary, because the value of the data obtained is quite rich, but the useful information is limited, data fusion method application value [21-25]: (1) network energy saving in the wireless sensor network node deployment time, we need to monitor the entire network reliability and accuracy of the information analyzed to ensure that up to a certain precision, which requires redundancy node configuration. Because of this configuration, the high degree of data redundancy, that is monitoring a number of nodes around the area of information collection and reporting of data would be more similar. If we put all the data sent to the sink node, does not make the aggregation node receives more useful information, if the accuracy of the data requirements, or has reached a certain accuracy requirements have been met in the case, it will only consume more network energy. Using the fusion technique, it can send data to the aggregation node before the data processing and analysis, a lot of redundant data will be processed to remove, so that, within the network node energy greatly save resources. (2) improve the accuracy of the information in a wireless network environment, there are a lot of changes, and these changes are often unpredictable and a great impact on the network, the network environment in which it is precisely because of these uncertainties, resulting from sensor nodes the unreliability of the data will be high. At this time, the sensor nodes to the same area of the data collected for analysis, to

improve the accuracy and reliability of the information. (3) improve data collection efficiency 4L Central South University, master's degree thesis first chapter of each node in a network data fusion algorithm step can reduce the data across the network transmission of information, reduce network data transmission delay, thus can reduce the probability and data network paralysis collisions and other phenomena, which greatly improves the quality of data transmission network to solve the congestion and because some of the data generated by different protocol data transfer protocol conflict, to some extent, the user network data collection efficiency has been greatly improved. If you only consider data fusion algorithm pros and cons of various algorithms, regardless of message routing, then it is also not a good overall network performance improvement, we must consider the combination of the two together, on the improvement in network performance to get better the method, or only one is unrealistic to consider, in practice it is relatively poor feasibility. The comprehensive analysis of the current issue of various data fusion technology features, combined with classic LEACH routing protocol and redundancy of information carried out theoretical research, and finally through simulation and existing protocols and algorithms are compared to demonstrate that their good role.

Due to the environment other factors, there are errors between the measured value and the true value of ultrasonic sensors usually documented its basic follow a normal distribution. If ultrasonic probe has n sensors, and the real value of the no i ultrasonic sensor to k time measurement [26-29]. Measured value is $x_{i1,}x_{i2,}...,x_{ik}$,Assumes that the mean of the x_i and the standard deviation is σ_i Let target estimated value x_0 ,The standard deviation for the σ_0 .

Set:

$$x_i = \frac{1}{k}\sum_{j=1}^{k} x_{ij} \quad \sigma_i = \sqrt{\frac{1}{k}\sum_{j=1}^{k}(x_{ij} - x_i)}, i = 1,2,...,n \tag{1}$$

(i represent the number of ultrasonic probe)

$$x_0 = \frac{1}{k}\sum_{j=1}^{k} x_i \quad \sigma_0 = \sqrt{\frac{1}{(n-1)} \bullet \sum_{j=1}^{k}(x_i - x_0)}, i = 1,2,...,n \tag{2}$$

(i represent the number of ultrasonic probe)

Probe transmitting and receiving process is affected by a lot of environmental factors, the need to monitor the invalid data, and to leave the useful data. Here method of Grubbs judgment.

Assume $G_i = \frac{(x_{im} - x_0)}{\sigma_0}$, where x_{im} is max or min data of No i. set x_0 is average. The follow will compare the value G_i of computing with $G_p(n)$ of

Grubbs table, if G_i value more than $G_p(n)$ of Grubbs table threshold, so determine the measured data is an abnormal value, that can be deleted. The threshold $G_p(n)$ have a relationship with two parameters [25]: Check out levels α (associated with confidence probability p) and number of measurements n . normally α is 0.01or 0.05.

Set: A_i , A_0 present the value of normal fuzzy set of measurements and estimated values of the data (NO i probe).that is:

$$\mu_{Ai}(t) = \exp[-((t - x_i)/\sigma_i)^2] \quad \mu_{A0}(t) = \exp[-((t - x_0)/\sigma_0)^2] \tag{3}$$

Its index weighting:

$$\mu_{Ai}(t) = c(k)\mu_{Ai}(t-1) + (1 - c(k))x_{ik} \tag{4}$$

Where $c(k) = (1-b)/(1-b^{k+1})$ and b is forgetting factor, normally set 0.95-0.99, And related to external interference factors.

By the inner product definition $\underline{A_i} \otimes \underline{A_0}$, the inner product of $\underline{A_i}$ and $\underline{A_0}$ is max value of tow fuzzy set $\underline{A_i} \cap \underline{A_0}$, that is :

$$((t - x_i)/\sigma_i)^2 = ((t - x_0)/\sigma_0)^2 \tag{5}$$

There are two solutions to solve for t: $t^{'}$ and $t^{''}$, that is:

$$t^{'} = \left(\frac{\sigma_i x_0 + \sigma_0 x_i}{\sigma_i + \sigma_0}\right) \quad t^{''} = \left(\frac{\sigma_i x_0 - \sigma_0 x_i}{\sigma_i - \sigma_0}\right) \tag{6}$$

Because $t^{''}$ be not in range $x_i \sim x_0$, so delete。

So, inner product:

$$\underline{A_i} \otimes \underline{A_0} = \exp[-((t^{'} - x_i)/\sigma_i)^2] \tag{7}$$

Take the formula(6)the value of $t^{'}$ to formula(7), get:

$$\underline{A_i} \otimes \underline{A_0} = \exp[-((x_i - x_0)/(\sigma_i + \sigma_0))^2] \tag{8}$$

By the definition of outer product, the outer product of fuzzy sets are two fuzzy sets and the minimum product, So: $\underline{A_i} \odot \underline{A_0} = 0$。

Nearness is defined as:

$$S(A_i, A_0) = (A_i \otimes A_0) \wedge (1 - A_i \odot A_0) \tag{9}$$

Take formula(7)to formula(9):

$$S(A_i, A_0) = \exp\left[-\left(\frac{x_i - x_0}{\sigma_i + \sigma_0} \right)^2 \right] \tag{10}$$

The n sensors normalization obtained respective weights:

$$\omega_i = S(A_i, A_0) \Big/ \sum_{i=1}^{n} S(A_i, A_0), (i = 1,2,...,n) \tag{11}$$

3 Algorithm Processes

SCM system to initialize the value of each parameter is as shown in Figure 1. Calculation step is as follows:

1) The order collected six ultrasonic sensor data, each sensor read 4 times
2) Respectively, calculated for each group the mean and standard deviation x_i, σ_i
3) Calculate the mean and standard deviation of the measured data x_0, σ_0
4) The Grubbs abnormal data judgment, excluding abnormal data values in the data.
5) Obtained proximity $S(A_i, A_0)$
6) Calculated for each set of data weights ω_i
7) For $x_i = \sum_{i=1}^{n} \omega_i x_i$, Integration of the system the exact value of the distance

4 Analyses of Results

The experimental with six ultrasonic probes precede four measurements, measured value, mean, standard deviation, the proximity; the relative weights are shown in Figure 2.

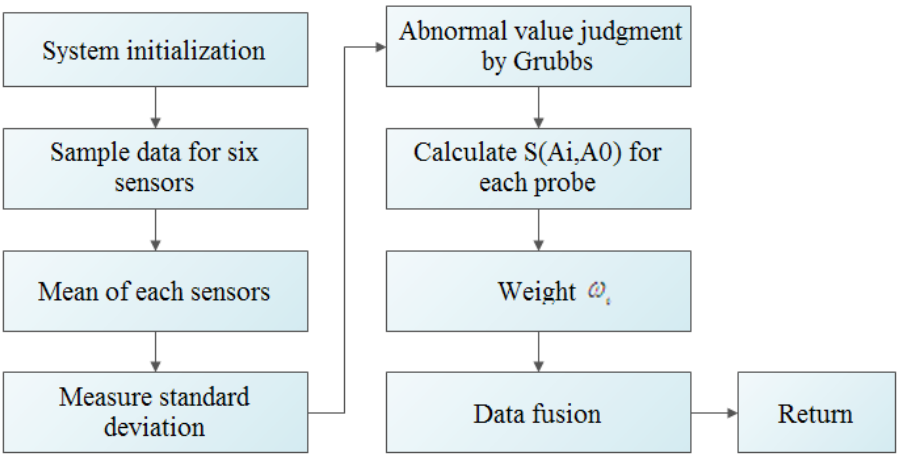


Fig. 2. Data fusion algorithm processes

Table 1. Values, mean, standard deviation, proximity, relative weights of multi-ultrasonic sensor measuring

Number of sensor	1	2	3	4	5	6
No 1(m)	4.15	4.22	4.32	4.16	4.29	4.42
No2(m)	4.18	4.25	4.36	4.18	4.22	4.25
No3(m)	4.16	4.2	4.25	4.16	4.26	4.26
No4(m)	4.12	4.26	4.3	4.14	4.28	4.28
Average(m)	4.1525	4.2325	4.3075	4.16	4.2625	4.3025
Standard deviation	0.025	0.027538	0.045735	0.01633	0.030957	0.07932
Proximity	0.9753099	0.972838	0.955295	0.983803	0.969517	0.923744
Relative weights	6.506E-06	0.47769	0.042968	1.65E-10	0.237101	0.242235

After data fusion, to draw a fusion value of the exact distance between the ultrasonic probe and the obstacle: $x_i = \sum_{i=1}^{n} \omega_i x_i$ =4.260m

Seen from Table 1, the sensor number of 1-6, for the data obtained by each sensor, the stability and reliability of the data of the two sensors, followed by 5,6-sensor data, the sensor number 1,group 4 proximity worst. From which can be concluded that the theoretical calculation of the relative weights consistent with high closeness of all measured values of the mean and fusion. Data from Table 1 illustrate this paper the proximity algorithm relative to the traditional fuzzy reasoning and mathematical statistics, high stability and reliability, and the calculation process is simple, fast and very suitable for embedded microcontroller like system, which suitable for use in portable devices.

5 Conclusion

In principal, there are two sensor network traffic : a query from the user to the network and from the sensor node to the user perception data. Each sensor node has two effects, might be aware environment, or as a relay node to other nodes to relay the data generated. All data generated by these nodes is likely to block the network and congestion occur, seriously affecting network performance. However, large-scale densely deployed sensor network characteristics result in the majority of these data is invalid, so you can process the data during transmission and reduce invalid data, which appeared within the network data fusion concepts. Wireless sensor network data fusion theory proposed is based on the following principles: Remove redundant data and invalid information, reduce or eliminate poor data reliability, while incorporating process data from different sensor nodes multi-feature information, to reduce network the purpose of the amount of data transmitted. With previous multi-sensor data fusion technology is different, modern fusion technology WSN biggest difference is reflected in the use of technology and the integration of routing protocol layers combine traditional multi- sensor data fusion for the different sources of information as well as various sensor data collected, treated observation target is to achieve better and more intuitive understanding. The wireless sensor network data fusion is not based on this, it is mainly within the entire network by reducing the amount of data transmission to reduce energy consumption, improve data accuracy and reduce the transmission delay to prolong the network lifetime. As wireless communication technology and the rapid development of electronic technology, low cost, low power consumption and multifunctional sensor nodes into being, the wireless sensor network is composed of a random distribution of integrated sensors, data processing unit and communication module node through self- organization of a network, and its broad application prospects attracted world attention. In traditional sensor applications, and many times just need to get an accurate description of things, does not need to receive a lot of raw data, data fusion is an important means to achieve this purpose. First, a detailed description of the data fusion concept, role, and in the wireless sensor network applications. By studying the main algorithm of data fusion, classical fusion protocol, the existing wireless sensor network data fusion program by implementing an executable algorithm simulation comparison. According to the monitoring range aims to highlight the size of the advantages and disadvantages of various integration programs. Secondly, based on time series analysis forecasting theory, proposed for wireless sensor networks based on the predicted time-domain data fusion methods. The method will be applied to time series prediction algorithm in wireless sensor networks, users can set their own error threshold to meet different application requirements, in order to ensure the accuracy of the monitoring data. Finally, the data transfer process due to mutual interference of communication, the transmission of data vulnerable to theft, tampering and other problems in the study based on cluster - tree topology data fusion algorithms, they also proposed based on cluster - tree topology. integration of data transmission security algorithms to further enhance the accuracy of the data.

References

1. Hall, D.L.: Mathematical techniques in multisensor data fusion. Artech House, Boston (1992)
2. Liu, C.R.: Discussion on ballistic missile guidance technology. Navy Aeronautical Engineering Institute (2002)
3. Man, S., Yang, H., Peng, Y., Wang, X.: ARM9-based embedded wireless smart home gateway design. Computer Application (2010)
4. Hashemipour, H.R., Roy: Decentralized structures for parallel Kalman filtering. IEEE Transactions on AC (1998)
5. Carlson, N.A.: Federated square root filter for decentralized parallel processes. IEEE T-AES (1990)
6. Niu, Y.-P., Cheng, S.: Based on Ultrasonic Ranging System. Ordnance Industry Automation (April 2005)
7. Su W., Gong B.J., Pan L.: Ultrasonic distance measurement error analysis. Sensor Technology (June 2004)
8. Sun, P.-Y., Li, J.Z., Tang, Y., Liu, Y.: AVR microcontroller based diesel engine oil - water emulsion-line automatic control system. Test and Measurement Technology (February 2005)
9. Lei, J.: LM567 and the ultrasonic detection. Electronic Technology (April 2004)
10. Jian, L.X.: Microcontroller-based gasoline tank of gasoline vapor pressure measurement and control system. Electronic Instrumentation Customer (June 2006)
11. Fang, W.B., Min, L., Haibin, X.: Ultrasonic distance sensor development. Hubei University of Technology (June 2004)
12. Hong X., Wu S.S., TRADITIONAL, Zheng Y.: ultrasonic Driving Circuit. Henan Vocation (March 1999)
13. Wang, L., Lei, S., Yu, J.: Based on PC and AT89C52 single-chip serial communication protocol design. Electrical Engineering Technology (February 2006)
14. Zhang S., Chen J.: Combine harvester automatic alarm system design. Agricultural Mechanization Research (March 2002)
15. Lee M.S.: Ultrasonic distance measurement principle and practice techniques. Practical Testing Techniques (January 1994)
16. Hui, Y.: Based on ultrasonic sensors automatic header height control system of Northwest Agriculture and Forestry University (2007)
17. Chen, J.: Obstacle discrimination based on ultrasonic sensors systematic study. Northwest Agriculture and Forestry University (2005)
18. Chen, Y.: Microcontroller based ultrasonic ranging system. Huazhong University of Science (2004)
19. Wang, Q.: Corn combine harvester yield monitor system intelligence research. China Agricultural University (2005)
20. Jiang, C.: Obstacle discrimination based on ultrasonic sensors systematic study. Northwest Agriculture and Forestry University (2005)
21. Ju, L.: ultrasonic distance measurement principle and its large electrode extrusion process application. Hunan University (2005)
22. Zhou, J.: Dibble fertilizing water planter control system design; Northwest Agriculture and Forestry University (2006)
23. Jiang, X.: Based on the “electronic identity card” model of vehicle identification system research Yanbian University (2007)

24. Han, Y.: Utility vehicle protection device design and implementation. Daqing Petroleum Institute (2006)
25. Qi, H.: Car reversing radar system research. Harbin Institute of Technology (2006)
26. Djenouri, D., Khelladi, L., Badache, N.: A survey of security issues in mobile ad hoc and sensor networks. IEEE Communications Surveys & Tutorials 7(4), 2–28 (2005)
27. Wang, Y., Attebury, G., Ramamurthy, B.: A survey of security issues in wireless sensor networks. IEEE Communications Surveys & Tutorials 8(2), 2–23 (2006)
28. Zhou, Y., Fang, Y., Zhang, Y.: Securing wireless sensor networks: a survey. IEEE Communications Surveys & Tutorials 10(3), 6–28 (2008)
29. Ozdemir, S., Cam, H.: Integration of False Data Detection With Data Aggregation and Confidential Transmission in Wireless Sensor Networks. IEEE/ACM Transactions on Networking 18(3), 736–749 (2010)

Research on OFDM Technology in 4G

Lina Yuan

Department of Mathematics and Computer Science, Tongren University
554300, Tongren, Guizhou, China

Abstract. As 3G (third generation mobile communication systems) is at the stage of increasingly large-scale application in China, multimedia communication services is one of the most prominent features in 3G, the technology research of the next generation mobile communication systems (Beyond 3G-after 3G) or (4G- four-generation mobile communication systems) has long been expanded. This paper is mainly researched by how to use OFDM technology as the core technology to effectively improve the transmission rate, increase the capacity of system, and avoid various interferences caused by high speed in the 4G mobile communication system.

Keywords: 4G, OFDM technology, mobile communication system, interferences.

1 Introduction

Currently the data transfer rates of 3G are up to 2Mbps, the expected data rates of next-generation (4G) mobile communication systems can reach 50Mbps, or even higher [1-5]. In addition to substantially increase in transmission rate, as compared with 3G, 4G has the requirements of performance and functionality [6-11]: It has achieved as 10-20 times as 3G in capacity, the signal coverage of base station is much bigger than third-generation systems, the quality of communication is further to improve, the costs of network overall operating are lower than the third generation, and it supports the next generation internet and all of the information equipments, seamlessly links fixed communication network, and providing personalized service with definition and so on. To achieve this goal, it must realize the technological breakthroughs from exchanging, transmission, networking and access of the communication network and other aspects, particularly in the wireless mobile environment and the conditions of limited radio spectrum resources, how to support high-speed data transmission in the premise of guaranteeing certain quality of communication. Because of its characteristics of good noise immunity, the ability of anti-multipath fading characteristics and high spectrum efficiency, OFDM technology is generally regarded as the core technology of 4G. Many large communications companies and research groups are active in researching the application of OFDM. Currently the developed technologies will be applied for OFDM technology in 4G, that mainly have W-OFDM, V-OFDM and MIMO-OFDM.

Wireless and personal communications in just a few decades has experienced a huge development from analog to digital communications and from the frequency

Y. Yang, M. Ma, and B. Liu (Eds.): ICICA 2013, Part II, CCIS 392, pp. 11–19, 2013.

division multiple access (FDMA) to a code division multiple access (CDMA) [12-14]. At present, there is also the emergence of new technologies, one of which is the orthogonal frequency division multiplexing (OFDM). Compared to Code Division Multiple Access (CDMA) as the core of the third generation mobile communication technology, its applications are more perfect and more advanced technology, which is called the "fourth generation mobile communication technology [15-19]." OFDM technology has great potential in the application of broadband field. Compared with the third generation mobile communication systems, 4G uses a variety of new techniques for OFDM systems, has higher spectral efficiency and good anti-multipath interference and improves the utilization of bandwidth, which can not only increase the capacity of systems, but also more better meets the requirements of multimedia communications that will include voice, data, video and other multimedia services through the wide broadband of information channels to high-quality send them out. In addition, the realization of OFDM technology is simple and its cost is low. With the rapid development of DSP technology, and use fast Fourier transform to bring about a large amount of mutually orthogonal sub-carriers, which provides a great convenience for the realization of high-speed broadband communications system. OFDM will become the mainstream way of realizing future broadband mobile communications, thus is caused to be paid more and more attention and research [20, 21].

2 The Main Contents of Research, Experiment and Trial

Research OFDM technology,and first need to start from studying OFDM, in essence, a form of OFDM follows the process [22-24]: First, fill up zeros behind the symbols that are digitally modulated in N1, constitute the N2 sequence of inputting samples, and then calculate an IFFT. Secondly, the last L1 sequences of outputting samples in IFFT is inserted into the front of the OFDM symbol, and the the most front L1 sequences of outputting samples in IFFT is inserted into the last of the OFDM symbol. Finally, OFDM symbols are multiplied by the promotion and demotion of cosine function in time domain, so the power of outside the broadband in system can be decreased rapidly. Then through the conversion from digital to analog, frequency modulation, and finally send them out. Study details as follows:

2.1 Determine the Parameters

Parameters need to be determined as follows: sub-channel, the number of sub-carriers, the length of FFT, the number of OFDM symbols each use, the level of modulation, symbol rate, bit rate, the length of guard interval, signal to noise ratio, and the number of inserted pilot (basic simulations can be without inserting a pilot and may be become zero).

2.2 Produce Data

Use the generator of a random number to generate an m-sequence.

2.3 Coding and Interleaving

Interleaved codes can be effective against burst interferences. Respectively use encoders and inter-leavers to implement.

2.4 Modulate Sub-carriers

OFDM using four modes of modulation, namely BPSK, QPSK, 16QAM and 64QAM. According to the diagram of constellation, the data on each sub-channel is mapped to be expressed by the complex of the constellation point and are converted to be the in-phase components of Ich and quadrature components of Qch.

2.5 Converse Serial-Parallel and Parallel-Serial

The conversion of serial-parallel converts one high speed data into multiplexed data through the realization of de-multiplexing module. While the conversion of parallel-serial is just the opposite.

2.6 IFFT

In the utilization of reality, the generation and demodulation of signals use the method of digital signal processing to achieve, at this time the signals is to be sampled, and formed discrete-time signals. IDFT can obtain faster processing speeds by adopting fast inverse Fourier transform (IFFT).

2.7 Insert a Cyclic Prefix

Inserting a cyclic prefix means to add guard intervals, the in-phase components and quadrature components of each symbol after the operation of IFFT are separately converted into serial data, and add the length-G of data in the tail of symbols to the head, in order to constitute a cyclic prefix. If an empty interval is added, under the influence of multi-path propagation, it can cause inter-carrier interference(ICI). The length of guarantee intervals(G) should be greater than the maximum length of the expansion in multi-paths. Using the cyclic prefix and adding special OFDM training symbols and other methods can keep better synchronization of time and frequency.

2.8 Windowing

Windowing is to reduce PAPR of the system, and the roll-off factor is 1/32. By this method, it can significantly improve the distribution of high PAPR of the OFDM

communication system, greatly reduce the probability of the peak signals and the requirements of power amplifier, and save costs. The window function often uses the raised cosine window.

2.9 Transmit Channels

Channels can be divided into multi-path experimental channel and Gaussian white noise channel. The direct wave and delayed wave of multipath delay channel decrease progressively about standard time according to a fixed ratio, so the parameters of multi-path delay channel are ratio and the maximum delay time.

2.10 Focus on Synchronization

Synchronization is a very important aspect of determining the performance of OFDM system. The real OFDM system has the process of synchronization. There are three methods of key synchronization, that is, using a pilot, a cyclic prefix and busy algorithms. If the purpose of research is synchronization, it can be achieved this steps in detail. The basic principles can skip this step, assuming the receiver has synchronized with the transmitter.

2.11 Remove the Cyclic Prefix

According to the data obtained by the synchronization, respectively remove the beginning guard intervals with in-phase component and quadrature component of each symbol.

3 FFT is an Inverse Transform of IFFT with Transmitters

3.1 Demodulate the Sub-carriers

The two sets of data of in-phase component and quadrature component after FFT is corresponding to the high points of the constellation. Due to the effects of noise and channel, it is not strictly constellation with transmitters. The obtained points of constellation with the recent principles is judged to the points of original constellation, and according to mapping rules, is restored a set of data . There are many ways of demodulation, for example: differential demodulation, and coherent demodulation.

3.2 De-interleaving and Decoding

According to coding and interleaving, it can obtain the methods of corresponding decoding and de-interleaving. Revert to the original data through the implementation of de-interleaves and decoders. Process error correction, and compare the data generated in step 2 with the received data, and calculate the bit error rate(BER).

4 The Objectives

Firstly, figure 1 shows that the performance of QPSK is in AWGN channel, namely through the relationships of signal to noise ratio (SNR), symbol error rate (Symbol Error Rate, SER) and bit error rate (Bit Error Rate, BER) to determine whether is compliance to the actual situation. In the reality as the increase of SNR, BER and SER of QPSK are all reduced, and BER is less than the corresponding SER.

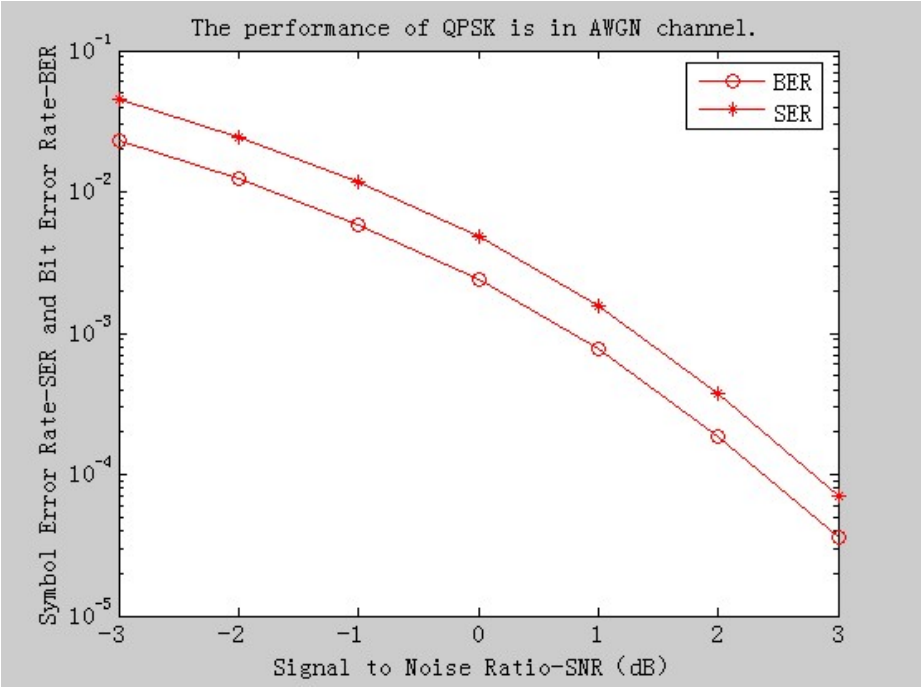


Fig. 1. The performance of QPSK is in AWGN channel

Secondly, study the performance of uncoded and Hamming (7,4) code of QPSK under AWGN channel. Figure 2 shows that when signal noise ratio (SNR) is low($E_b / N_o < 6\,dB$), un-coding bit error rate is better than coding bit error rate, which is because the code may bring the gain of coding , but under the condition of invariable total energy-transmission, due to the reduced energy-transmission of coding bits of each code word and dropped signal to noise ratio(SNR), owing to the signal-to-noise ratio decreased making the bit error rate higher, while the coding gain is little, therefore, the coding results as not coded instead. In high signal noise ratio (SNR), the gain of encoding is greater than dropped signal-to-noise ratio so that the performance losses, therefore, in the case of $E_b / N_o > 6\,dB$, the results of coding are better than un-coding. This paper only adopts the coding method that can correct a bit, if using correct several mistakes bits, the gain of coding is greater.

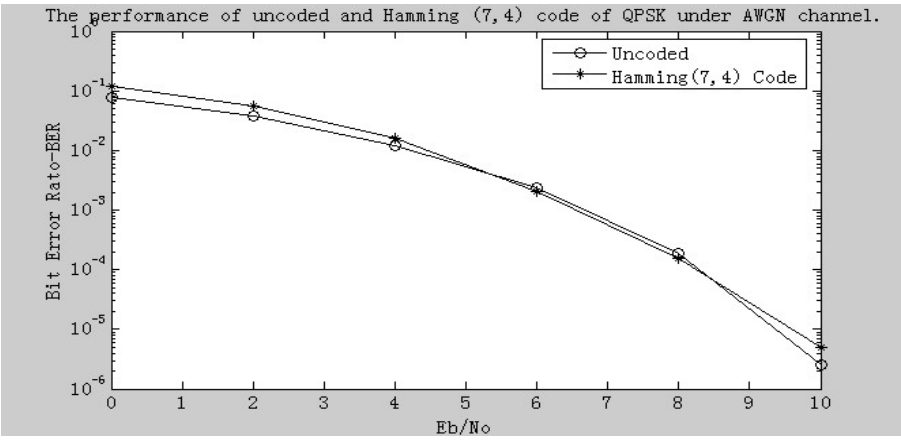


Fig. 2. The performance of uncoded and Hamming (7,4) code of QPSK under AWGN channel

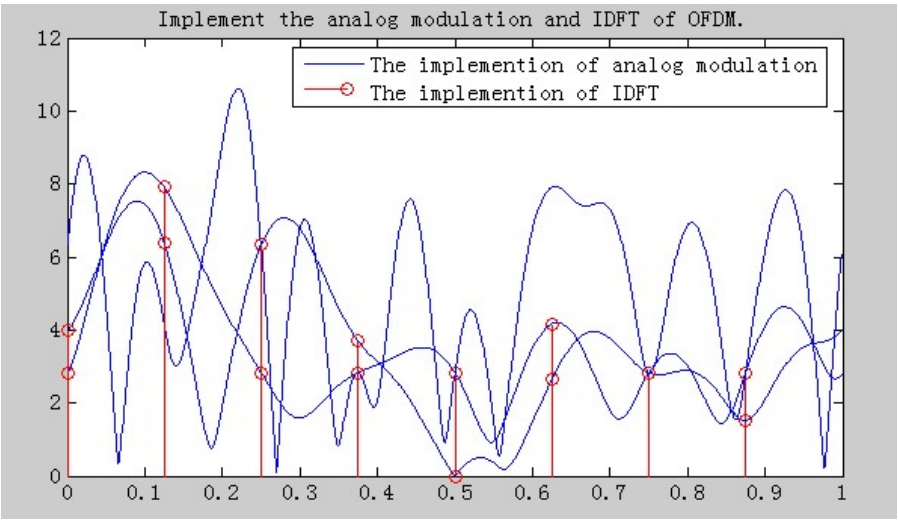


Fig. 3. Implement the analog modulation and IDFT of OFDM

Thirdly, implement the analog modulation and IDFT of OFDM.As seen from the result of the program runs, the methods of implementing IDFT/DFT and analog modulation is completely equivalent. Figure 3 illustrates data symbols that IDFT outputs are obtained by continuous several modulated sub-carriers superposition signals sample.

Fourthly, compare blank prefix with cyclic prefix in OFDM system,their performance is two-path channel and three-path channel. From the view of figure 4 and figure 5, in the case that the time delay of multipath shall not exceed the length of prefix,after equalization, the second symbol added cycle prefix is not affected by the interference of the first symbol and the inter-sub-carriers, while the symbols added blank prefix are affected by the interference of inter-sub-carriers. Therefore, demodulation data is not consistent with the original data. When the maximum time delay of multipath channel is over the length of prefix, the second symbol is affected by the interference of the first symbol.

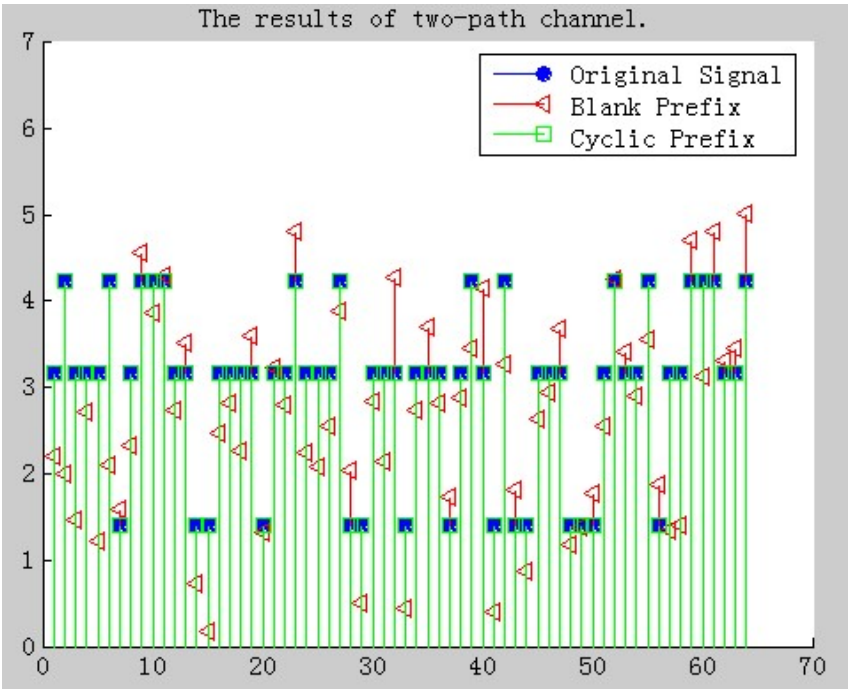


Fig. 4. The blank prefix in OFDM system

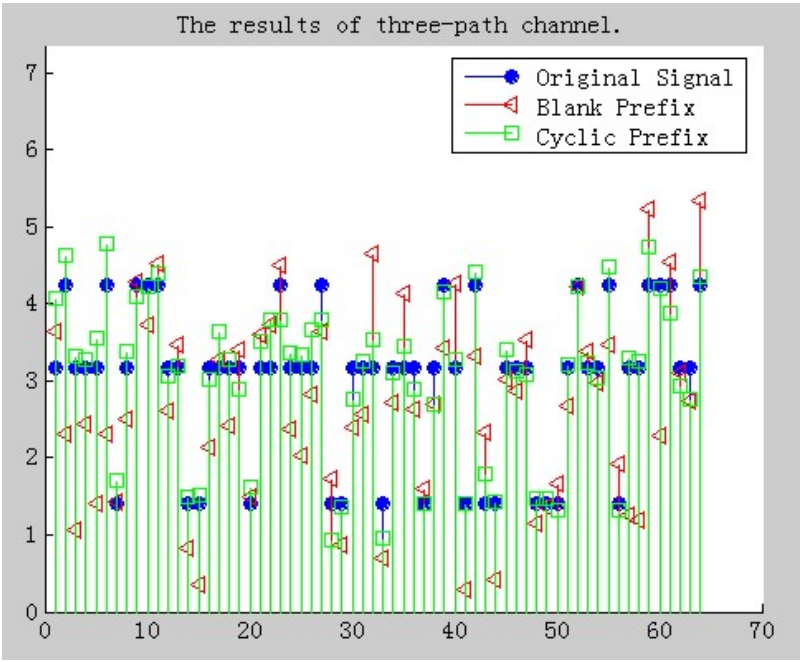


Fig. 5. The with cyclic prefix in OFDM system

5 Conclusions

The key problems of this paper are to overcome frequency selective interference. Because in a broadband wireless communication system, frequency selective interference is one class of the main interferences affecting the high speed of information transmission. It is expressed as the serious attenuation of some frequency components with signals, while other frequency components have higher gains. To overcome this type of fading, a very natural idea is divided into multiple sub-channels in the channel, so that the frequency characteristics of each sub-channel are approximal to be flat, use these separate sub-channels to transmit signals and combine in the receivers, in order to to achieve signal frequency diversity, which is the basic idea of multi-carrier modulation. In wireless communications the most widely applied technique is OFDM multi-carrier modulation, whose each sub-carrier is orthogonal, and promotes the utilization rate of the spectrum. It can be also inserted into a guard interval between OFDM symbols, so that the guard interval is greater than the maximum delay spread of the wireless channel, inter-symbol interferences are eliminated caused by the multi-path as maximum as possible.

In the future broadband wireless communication system, there exists two most serious challenges: multipath fading channel and bandwidth efficiency. The ability of anti-multipath interference and frequency selective fading in OFDM is strong, and the frequency selective multipath fading channel in the frequency domain is converted to a flat channel, reducing the effects of multipath fading. Because OFDM system spreads data on many sub-carriers, which greatly reduces the symbol rate of each sub-carrier, it weakens the effect of multipath propagation. If using the method of adding cyclic prefix as a guard interval again, or even it completely eliminates inter-symbol interference. At the same time OFDM adopts the allocation technology of dynamic sub-carriers to enable the system to achieve the maximum bit rate. Selecting each sub-channel, the number of bits per symbol, and the allocation of power to each sub-channel makes the total bit rate achieve maximum. That requires the distribution of each sub-channel information should follow the information theory of "pouring water theorem", which means that high-quality multi-channels transmit more, poorer channels transmits little, and inferior channels do not transmit. It can promote the development of the communications industry, increase the number of users, and produce good social and economic benefits.

Acknowledgment. Serial Number: A Scientific Research of China's Tongren City Marked by NO. (2013)9-13.

Sustentation Fund: The project of Science and Technology Bureau in China's Tongren City.

References

1. Sodnik, J., Stular, M., Milutimovic, V., Tomazic, S.: Mobile Communications: 4G. Medio-Mobile Comm. Wireless and Mobile Communications, pp. 634–642 (2008)
2. OFDM Forum, `http://www.ofdm-forum.com` (accessed July 12, 2007).

3. Liu, X., et al.: Explain the Modeling and Simulation of Communication Systems with MATLAB/Simulink in detail (July 2012)
4. The Book of Visions 2000 Visions of a Wireless World, vol. 1.0 (2000) Available: IST Wireless Strategic Initiative
5. Bi, Q., Zysman, G.I., Menkes, H.: Wireless Mobile Communications at the Start of the 21st Century. IEEE Commun Mag., 110–116 (2001)
6. Lu, W.W.: Compact Multidimensional Broadband Wireless: The Convergence of Wireless Mobile and Access. IEEE Commun. Mag., 119–123 (2000)
7. Zysman, G.I.: Bell Labs Tech. J. 5(1), 107–129 (2000), 3G Wireless 01 Conf. (2001), http://www.delson.org/3gwireless01
8. Haardt, M., Mohr, W.: The Complete Solution for Third Generation Wireless Communications: Two Modes on Air One Winning Strategy. IEEE Pers. Commun. 7(6), 18–24 (2001)
9. Molisch, A.F.: Wideband Wireless Digital Communications. Prentice Hall PTR (2001)
10. Verdu, S.: Multiuser Detection. Cambridge Univ. Press (1998)
11. Mitola, J.: The Software Radio Architecture. IEEE Commun. Mag., 26–38 (1995)
12. IEEE Communications Mag. Special Issue on Software Radio 37(2), 82 –112 (1999)
13. IEEE Commun. Mag. Feature Topic on Design Methodologies for Adaptive and Multimedia Networks 39(11), 106–148 (2001)
14. Andrews, J., Choi, W., Heath Jr., R.: Overcoming Interference in Spatial Multiplexing MIMO Cellular Networks. IEEE Wireless Commun. 14(6), 95–104 (2007)
15. Castellanos, C.: Performance of Uplink Fractional Power Control in UTRAN LTE. In: VTC Spring, pp. 2517 -2521 (2008)
16. Pischella, P., Belfiore, J.C.: Power Control in Distributed Cooperative OFDMA Cellular Networks. IEEE Trans. Wireless Commun. 7(5), 1900–1905 (2008)
17. Uplink Power Control with Fractional Frequency Reuse for EUTRA (2007)
18. Radio Performance Evaluation Phase 2 Report (2008)
19. Elayoubi, S.-E., Haddada, O.B., Fourestie, B.: Performance Evaluation of Frequency Planning Schemes in OFDMA-based Networks. IEEE Trans. Wireless Commun. 7(5), 1623–1633 (2008)
20. Simonsson, A.: Frequency Reuse and Intercell Interference Co-ordination in E-UTRA. In: VTC Spring, pp. 3091–3095 (2007)
21. Svedman, P.: Opportunistic Beamforming and Scheduling for OFDMA Systems. IEEE Trans. Commun. 55(5), 941–952 (2007)
22. Hu, H.: Radio Resource Management for Cooperative Wireless Communication Systems with Organized Beam-Hopping Techniques. IEEE Wireless Commun. 15(2), 100–109 (2008)
23. Barbero, L.G., Thompson, J.S.: Performance of the Complex Sphere Decoder in Spatially Correlated MIMO Channels. IET Commun., 112 -130 (2007)
24. Choi, W., Andrews, J.: The Capacity Gain from Base Station Cooperative Scheduling in a MIMO DPC Cellular System. In: ISIT 2006, pp. 1224–1228 (2006)

Network Resource Search System Based on Cluster Architecture

Chen Yandong

Network Information Center of Zhangjiakou Education College Zhangjiakou Hebei 075000, China

Abstract. The paper proposes network resource search system applying cluster architecture. Resource search system uses incremental update to control the cluster, which can realize collection and updating maintenance of network resource. The paper focuses on describing operating principle of the system, based on which the paper proposes the design and implementation of the system. Based on hash method of vector space, the system realizes load balance of cluster assignment, dynamically achieves running state of computing nodes by balanced CPU, MEMEORY, I/O and NET, and realizes dynamic monitor and change of task assignment scheduling policy. And the paper improves the efficiency of network resource search system by combing angle cosine vector method and incremental update.

Keywords: Web data mining, parallel gripper, incremental updating strategy.

1 Introduction

Network resource search system is the front end of Web search engine and collects the information resources needed for Web. The characteristics that cluster models and computing nodes are combined to improve processing capacity assigns update tasks to appropriate computing nodes for operation, which can process the update of network resources [1-5]. And adding and reducing computing nodes can automatically make task allocation adjustment, which can improve extensibility, based on which the paper introduces incremental updating technology to make the system estimate the updating rate of web page, which not only reduces work load of network resource search system updating web pages, but also improves the freshness of pages [6-10].

From the implementation of the system and the final operation results, we can see that the network resource search system using cluster structure is effective and can improve the performance of network crawlers [11-16].

With increasing website traffic, most network services are based load balancing cluster approach to provide services, along with the expansion of the cluster, the original stand-alone server -based management model has not been able to meet our needs, new needs must be centralized, grouping, volume, automated management of the server can batch execution plan tasks.

In distributed server management system software, there are some relatively good software, which is an ideal Cfengine [17-21]. It can group servers; different groups can

Y. Yang, M. Ma, and B. Liu (Eds.): ICICA 2013, Part II, CCIS 392, pp. 20–29, 2013.

customize the system configuration files, respectively, scheduled tasks and other configurations. It is based on C / S structure, all server configuration and management scripts are stored in Cfengine Server, while the managed server running Cfengine Client program, Cfengine Client to connect via SSL encryption periodically sends a request to the server to Get the latest configuration files and management commands, scripts, patches, installation and other tasks [22-26].

With Cfengine This centralized server management tool, we can achieve efficient management of large-scale server clusters, managed servers, and Cfengine Server can be distributed in any location, as long as the network can be connected will be able to achieve rapid automated management.

2 Multithreading Searcher

Collection efficiency is a very important performance index of crawlers. There are at least thousands of crawlers processing pages, so how to improve the performance and efficiency of program is very important. It is very common to take multithreading technology to improve collection efficiency in system implementation.

Thread synchronization must be considered under multithreading condition. In the system, URL queue to be accessed is multithreaded operation. While the thread accesses the queue, the other thresholds can't insert, extract and modify the queue. Figure 1 shows concrete steps of thread synchronization of URL queue to crawl.

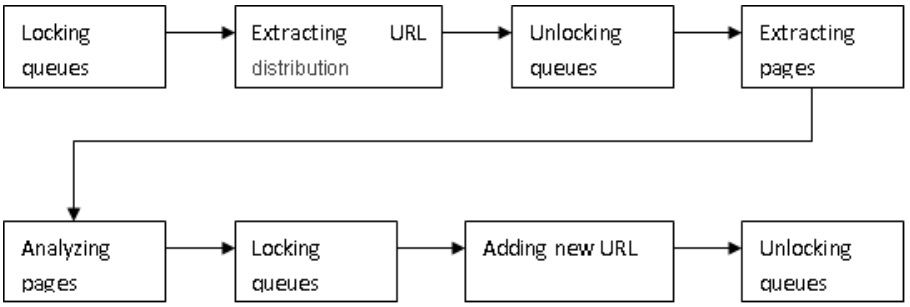


Fig. 1. Thread synchronization

Crawlers are multithreaded, so one method is needed to assign the tasks in different threads. The job is completed by Spider Worker object. The object maintains thread pool, which can undertake the task of Spider creating and destroying thread objects.

Defined function of Spider Worker object:

```
Public class Spider Worker extends Thread
{
    ……
}
```

It indicates that Spider Worker object subclasses Thread object. Each execution calls start () function of Spider Worker, and the function calls run ().Spider Worker object in Spider, as follows.

Public Spider (I Spider Reportable manager, String URL, HTTP http, int pool Size, I Work load Storable w)

```
{
    _manager = manager;
    _world Spider = false;
    _pool = new Spider Worker [pool Size];
    For (int i = 0; i < _pool.length; i++)
        HTTP hc = http.copy();
    _pool[i] = new Spider Worker (this, hc);
}
```

The codes on the last line generates Spider Worker object, which means that Spider Worker object is generated to manage the thread and thread pool is generated while constructing Spider objects.

Spider Worker object is executed in run () method of Spider object at the beginning.

```
Public void run()
{
    if(_halted)
        return;
    for(int i = 0; i < _pool. length; i++)
        _poll[i].start ();
}
```

After start () method of Spider Worker object is called, run () method of Spider object begins to be executed, which prepares for processing the pages. And the codes are as follows.

```
Public void run()
{
    for(;;)
    {
        _target = owner. get Work load ();
        if (_target == null)
            return;
        _owner. get Spider Done().worker Begin ();
        process Workload();
        owner. Get Spider Done ().worked End ();
    }
}
```

synchronized public void workerBegin()

Based on initializing thread and finding URL queues to be processed, Spider Worker calls worker Begin () method of Spider Done to notify spider thread which has been initialized. Worker Begin() method of Spider Done is shown in the following codes.

```
{
    _ active Threads ++;
    _started = true;
    notify();
}
```

Calling notify () method of threads.

Process Workload () method is called to enter actual executing codes of processing pages and transfer into class Get Site. Process Workload() method is over, which means that the thread fulfilled its historical task and executes the process of processing pages, so it should give away the resources for the other threads. Therefore, the thread should be closed. Worker End () method of executing Spider Done class is as shown in the following codes.

```
synchronized public void workerEnd()
{
    _active Threads--;
    notify();
}
```

3 Incremental Updating Strategies of Network Resources

3.1 Basic Model

In the current study, the change of web pages is seen as Poisson process [37, 38], which means that starting from the time 0, X (t) is used to record the times of the page changing at time t. The change of pages is independent, and the changing frequency is λ. According to the definition of Poisson process,

$$\Pr\{X(s+t)-X(s)=k\}=\frac{(\lambda t)^k}{k!}e^{-\lambda t},k=0,1,\cdots \qquad \text{(Formula 1)}$$

If the time of the page changing next time is T, probability density function of T is

$$f_r(t)=\begin{cases}\lambda e^{-\lambda t}, & t>0\\ 0, & t\le 0\end{cases} \qquad \text{(Formula 2)}$$

When k=1, we can know from formula (1) that the interval of the page changing of two times obeys index distribution which is formula 2. If the collection system has the same content of the local page with that of Web, the page is up to date. The timeliness of the page ei maintained by the collection system is defined as follows.

$$F(e_i;t)=\begin{cases}1,\ if\ \ e_i\ is\ up-to-date\ at\ time\ t\\ 0,\ 0\ otherwise\end{cases} \quad \text{(Formula 3)}$$

The timeliness of page ei also can be defined in a different way, that is, the age of the web page. And the definition is as follows.

$$A(e_i;t)=\begin{cases}0,\ if\ \ e_i\ is\ up-to-date\ at\ time\ t\\ t-LMT(e_i),\ otherwise\end{cases} \quad \text{(Formula 4)}$$

In the formula, LMT (ei) means the last updating time of page ei before t. According to the definition of formula 3 and formula 4, the average timeliness and age of set S composed of N pages is as follows.

$$F(S,t)=\frac{1}{N}\sum_{i=1}^{N}F(e_i,t),\ A(S,t)=\frac{1}{N}\sum_{i=1}^{N}A(e_i,t) \quad \text{(Formula 5)}$$

Incremental collection system maintains a page set S which concerns the average timeliness and age of S at a certain time. And the average value can be taken in formula 5 for measurement.

$$\overline{F(S)}=\lim_{t\to\infty}\frac{1}{t}\int_0^t F(S,t)dt,\ \ \overline{A(S)}=\lim_{t\to\infty}\frac{1}{t}\int_0^t A(S,t)dt \quad \text{(Formula 6)}$$

The average timeliness and age of single page can be achieved by combing formula 1 and formula 4 to calculate the corresponding expected value.

If the average change frequency of the page is λi, the probability of it changing in interval I=(0,t) is

$$\Pr(T\le t)=\int_0^t f_T(t)dt=\int_0^t \lambda_i e^{-\lambda_i t}dt=1-e^{-\lambda_i t} \quad \text{(Formula 7)}$$

Therefore, the average timeliness and age of page in interval I is

$$\left.\begin{aligned}E[F(e_i;t)]&=0*\Pr\{T\le t\}+1*(1-Pr\{T\le t\})=e^{-\lambda_i t}\\ E[A(e_i;t)]&=t\left(1-\frac{1-e^{-\lambda_i t}}{\lambda_i t}\right)\end{aligned}\right\} \quad \text{(Formula 8)}$$

The above definition not only establishes basic mathematical model for the change process of page, but also defines the timeliness and age of evaluating the quality of incremental collection system page. The average timeliness and age of the page set S maintained by incremental collection system in interval I can be achieved according to formula 5, formula 6 and formula 7.

4 Implementation of System

4.1 Structure of Web Crawlers

There are two ways of constructing crawlers. One way is that it is designed to be recursive program, and the other way is to write a non-recursive program which

maintains a web list to be visited. The premise of choosing the way is that the constructed program can visit large Web sites.

(1)Recursive program

Recursion is the design technology of calling its program. When the same basic task needs to be done repeatedly, recursion is very useful such as the following codes.

```
void Recursive Spider (String URL)
{
    ……download URL……
    ……process URL……
For each URL, Recursive Spider method begins to be called
circularly.
End loop.
    ……process the downloaded page data……
}
```

In the program codes, a single task of Web page is put in Recursive Spider method. Although recursion is a reasonable construction scheme, it is not suitable for the design of the system, the reason for which is that each recursion needs to be pressed in the stack when the recursive program operates, and the design of the system uses multithread which allows multiple tasks to be operated in one time. But multithread is not compatible with recursion, the reason for which is that each thread has its own stack in the process, but when one method calls itself, they need to use the same stack, which means that recursive crawler system can't use multithread.

(2) Non-recursive construction

The second method of constructing the system uses non-recursive way. When the crawler finds new pages, it uses the method that doesn't call itself. The method uses a queue. Non-recursive program uses the queues. In order to be processed by the system, the newly discovered pages must wait in the queue.

When non-cursive method is used, a page to be visited is given to the crawler, and it can add the page to its queue. When new pages are discovered, it adds them to the queue. After processing the current page, the crawler finds the next page to be processed in the queue.

In the design of the crawlers of the system, the program uses four queues, waiting queue, processing queue, error queue and completion queue.

4.2 Robot Protocol

Obeying Robot protocol is a special necessary task of search engine.

Search engine uses crawlers to visit the pages of internet automatically and achieves web information. While visiting a website, the crawlers firstly check if there is a plain text file which is called robots.txt under the root domain of the website. A plain text file which is called robots.txt can be created in the website. And the document should declare the part that can't be visited by robot and the specific part that only can be recorded by search engine.

Robots.txt document is under the web root. For example, while visiting a website such as http://www.abc.com, the crawlers firstly check if there is http://www.abc.com/robots.txt in the website. If the crawler finds out the document, based on which it determines the scope of access right. Table 1 shows the locations of several websites.

Table 1. Examples about the location of robots.txt in web site

Website URL	Corresponding URL of robots.txt
http://www.w3.org/	http://www.w3.org/robots.txt
http://www.w3.org:80/	http://www.w3.org:80/robots.txt
http://www.w3.org:1234/	http://www.w3.org:1234/robots.txt
http://w3.org/	http://w3.org/robots.txt

Only when the website includes the content which doesn't want to be recorded by search engine, robots.txt document can be used. If the search engine is wanted to record all content of the website, robots.txt document shouldn't be established and an empty robots.txt document should be created.

robots.txt document includes a or more records which are separated by null string (CR, CR/NL and NL are used as end mark, and the format of each record is as follows.

<field>:<optional space><value><optional space>。

can be used to note in the document, and the concrete method is the same to the practice of UNIX. The record in the document often begins with one line or more lines of User-agent, and some Disallow and Allow lines are added.

5 Experimental Conclusions and Analysis

5.1 Comparison of the Number of Pages between Parallel Crawler and None-parallel Crawler

Table 2 is the comparison of the number of pages between parallel crawler and none-parallel crawler using cosine vector method in unit time. Two crawling nodes, three crawling nodes and four crawling nodes are used in the experiment. We can see that the proportion of the number of pages of single node crawler and parallel crawler using Cosine vector method is 0.707, 0.704 and 0.702.

There is difference between crawling system based on none-parallel crawler and parallel crawler for the average efficiency of crawling pages, the reason for which is that the system has overhead to detect and collect real-time information of crawlers, and there is overhead of the communication between crawlers and controllers. W e c an see that the increase of the number of crawlers from 2 to 4 has no great influence on average operation efficiency of single crawler, which indicates that the system has good expansion capability and the performance can't be reduced greatly.

Table 2. The comparison between parallel crawler and none-parallel crawler

Parallel crawler						
	1h	2h	3h	4h	5h	6h
Parallel crawler	341	351	339	329	312	331
None-parallel crawler						
	1h	2h	3h	4h	5h	6h
Crawling nodes of 2 machines	483	495	483	461	438	476
Crawling nodes of 3 machines	722	739	718	689	652	712
Crawling nodes of 4 machines	959	982	953	915	868	947

5.2 Status of System Updating Pages Continuously within Continuous Time

(1) From the condition of the system monitoring page update within continuous time, we can know that the frequency of pages changing after 18:00 increases, which accords with the habit of students surfing on the internet.

(2) The peak value of cluster system updating pages falls back after midnight. We can know from the data that the reason is that the number of people visiting websites after 18:00 increases and network congestion degree of campus network increases, which limits the speed of crawlers updating. So the updating time interval is delayed.

6 Conclusion

The paper discusses the network resource search system using cluster structure in detail. And the paper starts from crawling task assignment task of cosine vector matching algorithm and combines incremental updating thought, which not only improves the efficiency of the system, but also improves the freshness of page database.

The paper firstly introduces search engine, based on which the paper introduces the concept of network crawler. Then, the paper describes technique relating to network crawler, system design and implementation. And the paper expounds the construction, the load balancing policy and the incremental updating strategy of the system, and describes the relevant algorithms, which is the core of the cluster system. Lastly, the experimental results indicate:

(1) The characteristics of using parallel computation model to calculate nodes and assigning update tasks to suitable computation nodes for operation can process the updating problem of network resources. And adding and reducing computing nodes can automatically make task allocation adjustment, which can improve extensibility,

(2)The paper introduces incremental updating technology to make the system estimate the updating rate of web page, which not only reduces work load of network resource search system updating web pages, but also improves the freshness of pages.

References

1. Jin, H.: Networked storage technology and curriculum thinking. Chongqing Institute of Technology (Social Sciences), 15 (2011)
2. Ma, T.: $ 20,050,000 educating people - “EMC- Fudan Network Storage Training Center” was established. eWeek, 29 (2003)
3. Cattle farming. Networked storage security technology analysis, Silicon Valley (February 2009)
4. Stars. Green Storage. Encyclopedic knowledge (2010)
5. Chen, J., Li: Network storage technology and library selection. Digital and Microfilm Images (January 2004)
6. Hongbo, C., Jie, L., Gui, H.: Networked storage technologies in GIS Application. Surveying & Spatial Information Technology (June 2006)
7. Video surveillance storage technology development and research. Chinese Security (May 2009)
8. He system Chau, Hu, Y.: ISCSI-based network storage technology. Yunyang Teachers College (March 2005)
9. Sun, D., Min, L.: Public security information network-based storage system data integration. Computer Technology and Development (August 2006)
10. Zhang, Wangfa, J., Sun, K., Jia, H.-P., Liang, L., Xu, X.: Networked storage technologies in the medical network information system application. Information of Medical Equipment (March 2007)
11. Li, J.-W., Meng, Z., Zheng, Z., Li, C.: Mass satellite data network storage. In: China Meteorological Meteorological Society 2007 Annual Meeting Proceedings Venue Comprehensive Detection Technology (2007)
12. Cao, J., Gulian, X., Yang, L., Zeng, E.: Simulation-Based Network Storage System Performance Research method. In: 2008 System Simulation Technology and Application Conference Proceedings (2008)
13. When the Tao, Broadcast system two cache memory architecture discussion. In: Digital TV Industry and Triple Play Conference Proceedings (2009)
14. When the Tao, broadcast system two cache memory architecture discussion. independent innovation and sustainable growth of the Eleventh Annual Meeting of China Association for Science and Engineering (4) (2009)
15. Yong, W., ZHU fresh: iSCSI-based IP SAN network storage technology and safety study. In: 15th National Information Storage Technology Conference Proceedings (2008)
16. Jin, Liu, X., Tao, Zhao, K., Jiao, Y.: IP network storage technology analysis. In: China Academic Forum Information Technology and Applications Proceedings (two) (2008)
17. Jiang, B., Zhang, Bing, H.: Television program production and broadcasting high-definition network architecture news production and broadcast business system storage architecture design and implementation of programs. China News Technology Federation Five Years of Secondary and Academic Council Proceedings (Part I) (2010)
18. Ping, Z., Tao, Z., Gao, J., Hu, J.: HPC storage key technologies. In: 2010 16th National Information Storage Technology Conference (IST2010) Proceedings (2010)
19. Linmao, Feng, Z., Xu, T., Tayier, Liu, L.: Build demand for seismic exploration and development of petroleum PC cluster storage system. In: Tenth National IT Industry Application Annual Conference Proceedings (2005)

20. Yang, L.: Massive geographic information data storage backup technology development and prospect. Surveying and Mapping. Mapping scientific frontiers Technology Forum Abstracts (2008)
21. Guo, J.-F.: Digital library information storage strategy study. China Petroleum University (2007)
22. Chen, D.: IP SAN network storage technology in the rail transportation video surveillance system application research. Anhui University (2011)
23. Shen, C.: P2P-based network storage system research. Wuhan University (2006)
24. Wang, J.: WAN copy storage service data management strategy study. Huazhong University of Science (2007)
25. Qi, L.: Small network storage system design and implementation of Beijing University of Posts and Telecommunications (2009)
26. Zhang, J.: Nonlinear editing network research and implementation of East China Jiaotong University (2009)

Network Sharing Mode in the University Libraries

Du Yu

Shenyang Aerospace University Library, Shenyang, Liaoning, 110136

Abstract. In the era of knowledge economy, the construction of the Information Commons is an inevitable trend of the further development of university libraries in the new era. The construction of the Information Commons of university libraries has important significance, which not only meets the needs of information on students and teachers, but also improves the educational function of the libraries. Therefore, this paper analyzes the practice of the network sharing mode of university libraries based on a large number of literature documents, and probes into some problems related to the design and implementation, so as to improve the application level and the degree of information sharing in university libraries.

Keywords: Network sharing mode, university libraries, design, realization, practice, content.

1 Introduction

With the expansion of all kinds of information resources, and the contradiction between the diversity of user requirements and the limited service the libraries can provide, the traditional service mode of university libraries is facing severe challenges. Information Commons (IC) further integrates resources and services now available of the libraries, creates a free and easy sharing environment for the users, provides one-stop information service, promotes the information literacy education, and is conducive to ease the contradiction of university libraries to a certain extent [1-8].

Information Commons is the library service mode developed in the university libraries in the United States in the early nineteen nineties accompanied by the development of computer technology, network technology, modern communication technology and the change of the information retrieval mode. The research of the IC started late at home, and the study was also from the recent years, and is often an overview of the IC or the analysis of the foreign successful cases [9-14]. But the new library service model from the beginning of the introduction has been attached great importance to by the domestic libraries circles and the education circles from all walks of life. There are many domestic university libraries which have or began to build the information commons, and received many praises. At the same time, the traditional library services also face new challenges [15-18]. For example, the librarians for the libraries, and the entry rate of the libraries decreased year by year; the electronic resource demands increased sharply; the library service requirements of readers rise. The traditional library services need to inject new vigor to achieve better development prospects.

Y. Yang, M. Ma, and B. Liu (Eds.): ICICA 2013, Part II, CCIS 392, pp. 30–38, 2013.

Currently, the library community is in service mode from the traditional to the modern service mode transition is a transition period; the new service model is still in the start-up phase, needs to be further improved and perfected. As we all know, due to historical reasons and aspects of objective conditions, the current university library service model transformation and innovation, there are many areas need to be strengthened [19-23]. Combined with the current actual work, I believe that in the digital and networked environment, good service pattern of university libraries transformation and innovation, but also continue to strengthen the work of the following aspects:

Digital and networked environment, the formation of Academic Libraries raised new and higher requirements, university libraries to boldly break past that the " possession -based" non- human management model to construct a "people -oriented" management philosophy books, that user as a leader, service-oriented, the implementation of personalized service to win user acceptance and trust. And the implementation of high-tech personalized service, both need to take a scientific approach, but also need to create a good cultural environment [24-27]. Thus, in the 21st century university library management model should be completed as soon as possible from the traditional to the modern scientific management of human management changes. Scientific management based on traditional target management and total quality management (TQM) -based management approach, emphasizing the effectiveness of services - products; technology-intensive human management is shaping the cultural background of university libraries and the human spirit departure to the staff for the management, the main emphasis is service - people. Library staff and students to be oriented to serve them wholeheartedly, and this are our next major university library management development ideas, without this development ideas, then everything is emptiness. Library management as soon as possible to achieve scientific and humane combination is the current and future for a long period of University Library service model transformation, innovation and a very important guiding ideology [28-30]. In the digital and network environment, the user demand for diversified multi-level literature and knowledge needs of individual, educational and cultural development of industrial and intellectual capital of the competition, the traditional university library management presents a challenge. Traditional management model of rapid development in the information age, due to its own reasons, not long-term development, has been difficult to adapt to the new era of network development needs. There is no market there will be no competition, no market will lose survival value. Since the special functions assumed by reason of University Library Management System mode by public institutions for management, through administrative means to maintain long-term non-market and non-competitive. To break these shackles, we must bold reform and innovation. Enhance quality and improve service efficiency is good university library primary task. To achieve this, we are required to change the traditional concept of other universities to learn from the advanced management experience, bold reform the traditional management system model, ongoing system innovation, the work of the library service has a higher quality of service and social benefits.

Campus network connected to the information superhighway brought us a lot of valuable information. Through the network not only to understand the latest scientific and technological achievements and scientific management methods, and can get a

variety of advanced production technology and product supply and demand information. How to search on the Internet, dump data, and to provide proactive internal and external users of information services, university library under the network environment is one of the functions. This effective approach is: According to the school and social development needs, especially in the information needs of university teaching and research, etc., to determine the scope and focus of information search, and the development and utilization of network information means and methods to develop practical information development and utilization planning and implementation. Not only should pay attention to the information search. Greater emphasis should be development and utilization of information, should be based on the school, for the community to find information users in order to maximize the development and utilization of information resources.

At present, China Academic Library System (CALIS) the establishment of the University Library Information Resources Sharing a reality. CALIS aim is: to build information resources of knowledge sharing, in order to maximize the social and economic benefits for China's higher education services. Its more than 100 members of the House have begun to introduce foreign joint electronic information resource database, union catalog, the joint construction of libraries. According to statistics, the library in terms of higher education, will own LAN is connected to the campus network, which in turn is connected to the Internet has more than 300 libraries. Digital and networked environment to build, break the inter-library mutually independent, closed situation, making the exchange of information, data acquisition unimpeded, tangible, independent library concept will inevitably be invisible, generalized books Pavilion Concept replaced. In this case, if the lack of a shared library concept, the development of other libraries to develop over something that will is a huge waste of resources. On the other hand, most sites do not provide scientific and technical aspects of information resources. University libraries can make use of the advantages of education and research network, combined with their own actual situation, the establishment has a museum featuring special literature database, Title of Thesis text database, research and information on the database and multimedia database, you can also collect domestic online database form has its own characteristics information products for teaching and research and a particular user community to provide professional services.

Digital and networked environment university libraries, librarians have higher requirements of its analysis, management and development information. Therefore. Librarians not only has a wealth of academic knowledge and new literature information management expertise, but also have a certain level of foreign language and a certain ability to operate a computer and Internet skills, should be able to skillfully use a variety of modern information technologies and means for information gathering, organization, processing, publishing, and information, education, has a network of knowledge and the ability to use network software. This requires that librarians can consciously accept various forms of continuing education, active learning, and constantly update their knowledge, improve the knowledge structure, to understand the development trend of information resources, discover and collect valuable information resources, and better for the majority of the school division health services.

2 The Practices of the IC in Domestic Universities

In recent two years, some university libraries in Taiwan and Hongkong have been planning to build the IC. In October 2004, Lingnan University constructed the IC, and in June 2007 it changed its name to "Jiang Zhen Information Commons". There are 4 discussion rooms, equipped with 12 multimedia workstations, desks, reference libraries and coffee shops. The butterfly created by the students is one of its characteristics. This is one of the successful cases to create China's first construction of the IC.

In 2005, in the "Forum of the innovation and development of the library services of the two sides of the Taiwan Straits", the curator of University of Hong Kong Anthony W Fergusn addressed the report entitled "10 years expectation of the libraries of University of Hong Kong". He described the process of establishing service model of the libraries of University of Hong Kong, IC and planning for specific facilities, layout, and personnel of the IC.

In recent years, the mainland university libraries also attempt at the beginning. For example, the Tsinghua University libraries have changed the library facilities of individual reading rooms, and gradually transit to IC. Jinan University libraries are planning a personal research area, with wireless internet access, and the main exhibition in the museum building, in order to prepare for the construction of the IC.

In the mid September 2006, College of Visual Arts of Shanghai Fudan University constructed IC Libraries on the second floor of the libraries, which is divided into multiple areas of the information desk, group learning (discussion) area, multi-functional demonstration rooms, library information area, multimedia audio-visual area, print and scan the area, leisure area etc. In addition, in the next door of the IC there is a small lounge. IC is in its embryonic form, but on a smaller scale, it has not fully played all the functions of the IC, but this is the libraries that the Chinese mainland will be the first to put the concept of IC into practice. In October 2007, Peking University Libraries Information Commons was officially opened, and the IC is called the "multimedia service sharing space", consisting of the multimedia learning center, audio-visual appreciation area, multimedia seminar rooms and audio-visual classrooms together.

In March 2007, the information sharing space of Shanghai Normal University was built, and the "IC" is in cooperation with the Institute of Finance and the Institute of education, taking the discipline resources of the two schools as the key discipline of the construction, and the readers of the discipline as the center, and divided it into 8 areas. It is reported that this space is currently the largest IC which is already in substantial operation.

In 2008, the IC of the libraries on Jiangwan campus of Fudan University was opened. The IC is provided by the general reference desk area, electronic classrooms, flexible partitioning of open access resources areas (including user individual learning workstations and collaborative learning workstations), body workstation, notebook computer wireless Internet zone covering the entire IC. In the same year the new book information center built in the Minhang campus of Shanghai Jiao Tong University has reference to the concept of IC in the design space, with quiet reading rooms, equipped with multimedia facilities, and the service functions of E-reading, interactive learning space, meeting rooms and the lecture hall, comfortable leisure space for the enlightening meaning, cultural exhibition etc.

3 The Design and Implementation of the Mode of Network Sharing of University Libraries

University libraries adhere to the "people-oriented" service concept. The construction of IC can not only provide collection technology and good environment of services in one to attract users to study and research in the libraries, but also can provide personalized information service for the various needs of different users, to use the network to serve for each user.

3.1 Analysis of the Feasibility of Building IC

First of all, the collection conditions improve, to protect the substantive space construction. With the expansion and consolidation of many colleges in recent years, all the universities should expand or build the library buildings. The increased construction area of the libraries provides good conditions of the physical space for the university libraries to have a good planning of the entity collections and the information communication areas of the libraries, especially the establishment of new service projects.

Secondly, good information resources and traditional service accumulation provide a realistic basis for the construction of IC. At present, the university libraries have become the focus of domestic research of information resources organizations, and have full strength to meet the information needs of research users. With the development of information technology, the quality of librarians also gets a boost, and provides the service infrastructure and the personnel support for the construction of IC. The improvement of software and hardware facilities is conducive to the university libraries to construct the IC according to the actual needs of the museum, the functional extension and integration of the libraries, and improving the environment, to realize the innovation of information service as each user's "learning center".

3.2 The Contents of Constructing the IC

IC develops based on the traditional library reference service, and makes the service more open and modernized through the application of modern information means. IC is the floorboard of the information service environment and the user information space which is composed of library infrastructure, information resources, the library's culture, and the users. It includes not only the information resource space (real and virtual spaces), but also the cultural spirit of libraries and user information environment. In addition to functional integration to strengthen and support the information technology resources, the IC construction also should take full account of all environmental factors which may face the university libraries.

3.2.1 Physical Space Construction

The IC entity space is the physical space ensuring the exchanges and cooperation in the learning. According to user needs and goals, the entity space should follow the "change according to the needs" principle, with specific focus on the rational layout and landscaping, and enhance the IC information exchange, cooperation and "one-stop"

services. According to the actual needs of the university libraries, the IC entity space generally includes open access areas, leisure areas, the exchange of learning areas, discussion rooms promoting the team research, the information desk improving the user information technology and the multimedia studio etc. The IC physical layer can be divided into information service platforms, the public areas of learning, individual learning rooms, group collaborative discussion rooms, multimedia rooms and leisure areas, aided by various computer configuration, network equipment and other hardware devices, and the information consultancy librarians, information professionals, scholars and student assistants and other components of the human resources, to meet the information demand of various users.

3.2.2 Virtual Space Construction

In addition to the entity space support, the IC also needs the support of the building of a virtual space. Virtual space is the learning portal providing the user exchange and share of knowledge based on modern network tools. It provides online consultation, online collaborative learning community, classroom and digital library resources. The users not only can search the information through the virtual space, but also can do real time communication or share knowledge with other users through it, and the IC information consulting librarians will use a variety of effective information resource integration, to provide help to the user retrieval. With the development of network Web 2.0 technologies and the mobile Internet, the expansion of the native network generation and the emergence of instant communication tools, blog, knowledge base and other software tools as the times require, IC provides new technological medium to provide users with the information exchange and the sharing of knowledge. IC is connected to the social network information by constructing this platform of the virtual space, and can strengthen the information communication and cooperation between librarians and others and can close contact between museums and galleries, users, and finally embodies the ubiquitous learning model between ICs.

3.2.3 The Construction of Human Resources

The Information Commons is a new service model formed by the integration of a variety of services. Therefore, it needs many experts and staff support, which puts forward higher requirements on the quality of librarians. In the design of the staff, we should integrate information retrieval technology of the consulting table, computer technology, and the professional knowledge of library management. The library staff should be very familiar with various business processes, and can solve all problems by service users. Specifically, the librarians can be subdivided into resource librarians, subject librarians, technique librarians and reference librarians, and we must specify the main businesses and responsibilities of different stuff, which are targeted to improve their professional levels and comprehensive qualities. In addition, because the IC cannot do without the wide application of computers and the network, therefore, the libraries must possess a batch of computer technicians, mainly to provide technical support, and ensure the normal operation of the equipment such as the computer, and the multimedia equipment. While the users encounter computer technical problems, they can help them to correctly use the various hardware and software equipment and also they shall be responsible for the maintenance of the copier, the printer, the scanner, etc.

3.2.4 Establish the Mechanism of Service Evaluation and Information Feedback of IC

In the establishment of the system of service evaluation and information feedback of IC, it is necessary to clarify the purpose, the principle of evaluation, and the way of the evolution of constructing the explicit service evaluation mechanism. Because the construction goal of the IC of libraries is to continue to meet the needs of users' learning and research, and to play the educational function of the libraries, only by understanding the user's further demand can we know the service deficiencies, thus making improvements based on the feedback, and improving the user satisfaction. The university libraries construct the evaluation system of the IC in accordance with the requirements, establish feedback channels of the IC, reflect the users' evaluation of the quality of service, and improve information sharing service according to the feedback information of the users, thereby forming the service model of high quality of library information commons. The most popular service evaluation tool is currently the Lib-QUAL+TM. University libraries can use this index system to form electronic questionnaires or paper-based questionnaires, and transplant the electronic questionnaires in the terminal interface of the IC, and the users can fill in them while entering the terminals. The paper questionnaires can be put on the desk areas, convenient for the users to take. According to the actual situation of the pavilion, the university libraries can design operation evaluation index of strong operation, and let users mark them. The libraries regularly recycle these questionnaires, and do statistical analysis, to gradually improve the quality of the services.

3.2.5 Optimize the User Information Environment

The user information environment is a new collaborative, user oriented service mode and service mechanism established for the scientific research and teaching with the users as the core, mainly by relying on the subject librarians and the university libraries and public information resources, for the specific organizations and users, to provide strong guarantee and support for the teaching and research. According to the use frequency and time for users, the IC should expand the services popular with the users, to adjust or abandon the services of low frequency in the usage, and actively optimize the user information environment, and ensure the highest utilization rate of the IC.

The service locations where the IC optimizes the user information environment are not limited to the libraries, but more are the user's environment, especially the network platform. The service location should be from the libraries to the users, and the service depth has been beyond the document services and the information services, and should stress the knowledge services. The service contents should no longer be the consulting and training, but should optimize the user information environment, and develop an incentive the user needs, and organize the user information activities, to enhance the ability of the users to use the information. IC comprehensive utilization, coordination, organization and innovation of resources and services breached the concept and paradigm of "museum", to promote the transformation and change of the libraries.

4 Conclusion

The present era is the information age. As the center of information collection, storage, processing and transmission, the libraries should adapt to the needs of the new situation, and we must adopt a new management system to realize the automation of management and scientific decision. The management information system should have the most advanced management ideas and design systems, and the introduction of it into the management of university libraries will inject new vitality into the development of libraries.

Digital and networked environment university libraries, librarians have higher requirements of its analysis, management and development information. Therefore. Librarians not only has a wealth of academic knowledge and new literature information management expertise, but also have a certain level of foreign language and a certain ability to operate a computer and Internet skills, should be able to skillfully use a variety of modern information technologies and means for information gathering, organization, processing, publishing, and information, education, has a network of knowledge and the ability to use network software. This requires that librarians can consciously accept various forms of continuing education, active learning, and constantly update their knowledge, improve the knowledge structure, to understand the development trend of information resources, discover and collect valuable information resources, and better for the majority of the school division health services.

In summary, the digital and networked environment, China University Libraries industry should take effective measures to meet the challenges and seize the opportunities, and actively making library service model transformation and innovation, the better for teaching, research work services. In other words, the university library industry must establish a "people -oriented" guiding ideology, bold reform the traditional operation system, and continuously enhance the overall quality of librarians, increase digitization, networking, investment in information technology, and its systems, in-depth research. Only in this way, our industry to adapt to the university library digitization and networking of the international environment in order to gain more space for development, in order to create a more brilliant tomorrow.

References

1. Ping, Z., Zhu, P.: Teacher Community of Practice: A New Perspective of Teacher Professional Development. Teacher Education Research (February 2009)
2. Zhang, K.Y., Zheng, T., Li, K.: Ontology-based network questionnaire generation system OntoSurvey. Computer Engineering and Applications (September 2006)
3. Xu, J.: On Construction provincial public libraries under the network environment a new relationship with the University Library. New Century Library (May 2003)
4. Wang, D.-B.: Digital Library Federation and China Features Digital Library Federation building strategy. New Century Library (January 2004)
5. Luo, X.: Domestic Studies on teachers' professional development stages. Educational Research (July 2006)
6. Li Zhi pure: Library Technology Collaborative Study. Library (March 2005)
7. Ci, W., Li, F.: Regional information resources sharing Protecting System. Library (June 2005)

8. bility,Tao, Z.: JALIS of Universities of Jiangsu Construction of Digital Library. Library and Information (February 2006)
9. Tao, Z.: Technology Alliance: collaborative digital library construction new development. Library and Information (September 2008)
10. Chen, Xi: Strengthen the strategic management of technology university libraries Thinking. Library Science (September 2005)
11. Zhong, Q.: Digital libraries and building understanding and thinking. Library (March 2000)
12. Chengdu Municipal Party Task Force: Meet the challenges of digital libraries. CPC Chengdu Municipal Party College (June 2001)
13. Zhou Y.: Analysis of Agricultural Colleges Digital Library Construction Problems and Countermeasures. Journal of Library and Information Sciences (May 2002)
14. Mysterious Dongmei: About China Digital Library Reflections. Jixi University (2002)
15. Marina Red Eagle ; Library under the New Situation Information Service. Library and Information Forum (April 2002)
16. Fang, F., Qing, L.: Digital library service environment Reengineering. Library Science (June 2005)
17. Li, Y.: Higher Vocational Colleges Digital Library Construction and teaching at different levels. Science and Technology Information (Academic Edition) (March 2006)
18. CRITICAL, Xiaomin: On the present vocational colleges digital library construction. Computer and Telecommunications (July 2007)
19. Du, D.-H.; Tan Digital Library and Copyright Protection. Era Literature (bimonthly Edition) (January 2007)
20. Ling, Y.: Digital Library Copyright Fair Use System Challenges and Countermeasures. Library Science (September 2007)
21. Xi, Y.M.: Construction of Digital Library. In: Fujian Provincial Library Association 2009 Annual Conference Proceedings (2009)
22. Wang, Y., Bing, Ping: Scientific development concept to guide the construction of digital library. In: National University Social Science Information Research the Sixth Congress of the 13th Symposium Proceedings (2010)
23. MATERNAL ; I Museum in Digital Library Construction Challenges and Countermeasures. In: 2011 National Hospital Information Library Management Conference Proceedings (2011)
24. Pu, D.: Fujian College of Digital Library Construction Survey. In: Fujian Provincial Library Association 2009 Annual Conference Proceedings(2009)
25. Zhang, A.-M.: Digital library construction problems and countermeasures. Library Theory and Practice (2002)
26. Hong, D.: Tan digital library and my school digital library construction. In: Fujian Provincial Library Association 2002 Annual Conference Proceedings (2002)
27. Cheng, H., Li, H., Baiyu, S., Xiao, Z.: CNKI based digital library construction in university teaching and research in applied research. North China Association of Universities Figure Twenty-fourth Annual Conference Papers (articles) of the 2010 (2010)
28. Fu, G.: Higher Vocational Colleges Digital Library Construction. Fujian Provincial Library Association 2008 Annual Conference Proceedings (2008)
29. Rhodes one; Polytechnic current situation and prospects of digital library construction analysis. Library Theory and Practice (2002)
30. Pan, X., Liu Kui Wave: Digital library service system construction and implementation of information integration. In: Strait Information Service Conference and Education Symposium New Directions Exchange of Papers (2002)

Interaction Multi-station Network Based on Android Platform

Jialin Gang and Shi Liu

Dalian Neusoft University of Information, Dalian, Liaoning 116023 China

Abstract. This paper put forward to design mobile site applications of multi-station managed client for mobile phones, the other one is that Social + Local + Mobile (SLM) mode is introduced to android mobile management platform to design, and provide mobile users a new experience of social networking sites integration interface.

Keywords: Android platform, Mobile network, Multiple station management, Cross design.

1 Introduction

Development in the mobile social networking site, users of mobile social networking sites also exist multiplicity, the average of user to use the number of mobile social networking sites between 3-5 [1-6]. For many mobile social networking sites, there is a corresponding the APP (Application, third-party smart phone applications) and widget. For mobile phone users, social networking sites are each separate community, click into the program is a separate application, social networking sites in the switch when it is inconvenient community view [7-11]. Based on the user's point of view, this article will be in the Android mobile platform environment for the management of social networking sites to design, based on the phone itself developed a social networking site integration of multiple clients within this interface can toggle view different social networking sites and other operations [12-16].

Mobile site management and design meet the needs of users of mobile phone users to bring service experience, speaking from the commercial value of the mobile phone manufacturers also give a profit. Mobile site management mainly involves services as: Android platform for mobile users and mobile phone with social networking sites and mobile phones provide the link between the three management clients. Phone as the fifth media has become always accompany people's life partner, handset manufacturers teamed up with social networking sites will increase the phone's applications [17-20]. At present, the number of mobile phone social networking site account management needs to be improved, and social networking sites and mobile phone module on the degree of integration is not mature enough, do not let the phone be truly social mobile [21-25].

The rise of Android phones to promote the development of mobile social networking sites, social networking sites also provides a new selling point for Android phones [26-28]. Because social networking sites being used by the PC side to the mobile

Y. Yang, M. Ma, and B. Liu (Eds.): ICICA 2013, Part II, CCIS 392, pp. 39–47, 2013.

terminal development, mobile phones and mobile phone operators to combine social networking sites to launch social networking phone, in order to improve the user on the phone's favorite, the number of mobile social networking sites integrated design industry has become increasingly worthy of study and exploration.

2 Analysis of the Android Platform Interactive Behavior

Android is an open source operating system Linux-based handheld mobile devices, Android platform today is the most popular mobile platform development objects. Android platform with five Advantages Features is mainly reflected in: openness, break the shackles of operators, rich hardware selection, without any restriction of the developers to seamlessly integrate Google Apps. Google and Open handset alliance, through mutual cooperation to create a standardization open mobile platform for mobile devices, that designed mobile Internet industry to establish an open ecosystem and the environment.

As people increasingly high expectations of mobile terminals can be completed only voice communications, messaging and management simple mobile terminals began to lose market. Including multimedia, colorful animation and mobile commerce applications, including new wireless mobile terminals will be put forward higher requirements for the mobile terminal from the provision of voice -based communications services to the progressive development of data-centric communications services and people work, live and communicate significant impact. Interactive to the user through the exchange of information with the mobile terminal. Therefore interactive experience is directly related to the performance of the system can give full play to be able to make users accurate, efficient, easy and enjoyable to operate the system is essential. At present, software developers in the software design process very focused software development technology and its business function has the while ignoring the needs of users of interactive software, thus affecting the ease of use, friendly interface design studies for also focused on interface design technology, design the means. Software developers with experience in the design of the lack of reference for understanding the actual user display requirements. To know the software -friendliness, ease of use with the user's "features" closely related to the same software interface different users may have diametrically opposed assessment.

Sharing is a major feature of Android, just two client support, users can at any two local client for the sharing of information, such as the micro -Bo saved to the local Evernote application, selection in the gallery By WhatsApp to send a picture to a friend.

Android's sharing feature, can be arbitrarily two local clients, he can be more efficient to meet different user task flow. Such as sharing pictures, Android can be used to set various filters after taking Camera360 directly from any local microblogging client sent to microblogging. Take pictures, and then share, also in line with the user's operating habits. By contrast iOS is more trouble, trouble or developer to put the filter function to do microblogging client, either the user the trouble of the first camera settings saved to the gallery filters, and then send microblogging interface to pick and choose from the gallery.

Android's sharing mechanism, but can effectively avoid the client function expansion, helps developers save effort, but also to meet a variety of usage scenarios and user. The new phones are now joined the three-dimensional rules accelerators accelerometer so gestures, movement detection is becoming part of the development of mobile applications. Users are subject to rotating his Android phone browsing the

content they wish to make appropriate adjustments along. Of course, if your application does not support the direction of the change in the majority only in specific interfaces such as full-screen video playback and display only specific direction device does not respond to changes in the direction it is reasonable. Note change in direction is responding to the user end and the sound, feel the same depending on the user terminal settings. Can also be detected by shaking the extent to let the phone do something such as Shake mobile phone to switch programs. Or you want to erase a good picture can paint a slight shake your phone can erase the image of the painting Bristol also more fun application is completed the project schedule notes gently shook automatically deleted. By moving direction and intensity of the sensing applications can be flexibly implement various functions. Interactive behavior refers primarily action between the user and the operating range of the object of physical or psychological. This process includes three levels of information: the user layer, presentation layer and system layer. User layer is the core of user requirements and user behavior; presentation system is to connect the bridge layer and user layer, the layer functions of the system for presenting information to the user a way; system layer, the main function is the product and the internal structure can be realized works. Android platform, the biggest feature is its interactive gestures, Tap, Pinch and Spread, Drag, Flick, Press and Double Tap and other operations, including long press is the most significant feature of gesture interaction Andrews, for example, long press brings up the standby interface widget, long press the screen option brings up a single entry in its context menus.

3 Analysis of Web Trend for the Android Platform Mobile Phone

2011 SLM (Social + Local + Mobile) concept proposed illustrates the development trend of mobile social networking site, that is So namely Social, social; Lo namely Local, local location; Mo namely Mobile, the mobile network, SLM mode now is Exploration of the object in the major technology companies.

Social have become ubiquitous, from Facebook to Twitter, and then to the recent Google+, social networking sites have proved the prospect trend.

Local emphasized that local location to location-based services represent(Location Based Service, LBS) based on various positioning and sign, which is Foursquare or roadside, including Face book Places and all report. Internet turned out a number of areas when it broke through the territorial restrictions (such as print media, shop, etc.). But now the Internet and the concept of localized around the back. Everyone's life there is a certain service circle regional, localized variety of services.

Mobile is a smart-phone brings a variety of mobile Internet applications, the trend with 3G applications are increasingly popular. We are transferred from the PC to Mobile; Mobile will be the future of the inevitable trend of the Internet, and now under the influence of SLM mode mobile social networking sites are growing rapidly.

The current domestic markets are more representative of mobile social networking sites are: everyone, happy network and Sina Weibo. By analyzing these three basic information about social networking sites, mobile social networking website summed modules trends. Now more and more social networking sites, each period has their favorite social networking sites, multiple social networking sites simultaneously will cause some trouble, but each social networking sites will exist a point of interest or people. I also registered several social networking sites, so he met the management

issues, in particular, how to synchronize between them questions.So this paper carried out on the management of exploration and design.

The management of social networking site designed to allow users to understand "their social circle at the latest developments", you can easily view all the social networking sites updated and personal information released simultaneously to different social networking site platforms. By monitoring multiple social networking sites page information in a unified platform to manage. Users can share the same platform. For mobile phones mobile social networking sites, social networking sites on the social networking site is not just content, but also that social networking sites and cell phone contact. The first is the phone's global positioning system (GPS) location-based social networking sites and services, integrated design, followed by the mobile phone address book contacts and friends social networking sites integrated design, function modules for mobile phones, such as camera, music, pictures and social networking sites are also inter-linked design a bright spot.

4 Interaction Design for Website Management

4.1 The Service Module Analysis

The design consists of two main parts, the first binding process is a registered design and interface features, followed by combination with phone functions widget, contacts, instant photo upload design, interface design features which combine SLM mode mobile social website design, divided into four functional modules: Home (Friends News), information (and I related information), chat and locations shown in Figure 1.

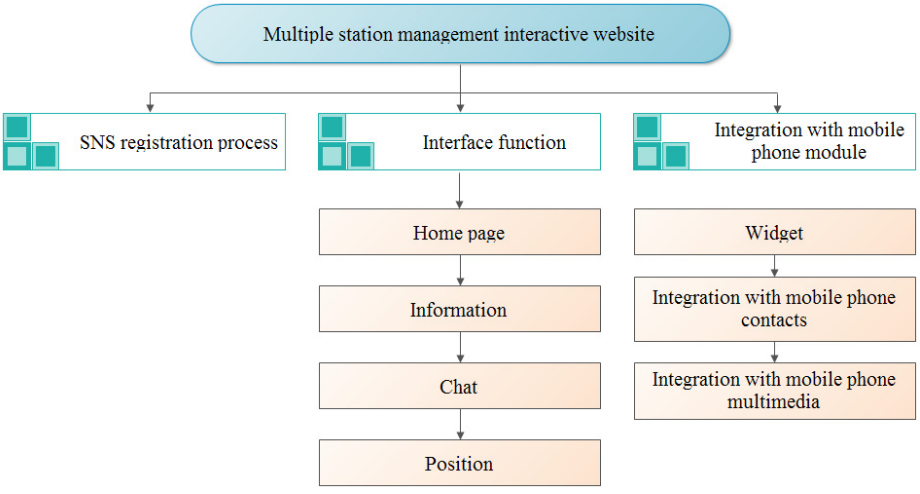


Fig. 1. Analysis of the management service module

(A) The process of registering binding social networking site.In the beginning of using the client, you must first register the user's account, and then bind your mobile phone to log in social networking sites. Phone built-in current popular mobile social networking sites, you simply login to your client held username and password, you can synchronize binding.

(2) Interface functions.Home module presents information primarily dynamically update Friends, the different social networking sites integrate vary information, including text, images and video sharing interaction. Social networking sites managed services integrate multiple different social networking site information through the client, which can enhance multiple social networking sites and user interaction, but also enhances the user's contacts with various social networking sites relationships. Home module is the user Share Friends Friends dynamically important channel of information.

(3) Platform integration with the phone function modules

A) Platform integration with the phone idle screen, Widget presented mainly includes two parts: the release of the new state; friends dynamic real-time push.

B) Platform integration by social networking sites and favorite contacts, including the following: adding contacts in the contact list screen icons held social networking sites; interface in the group put into different social networking sites social networking sites according to friends grouping; interface shows that the active interface import contacts friend's activities.

C) Platform and photos, audio, video integration, timely synchronization of the multimedia content to social networking sites.

4.2 Information Architecture Diagram

According to the logical relationships between them and establishment of inter-module usage information architecture diagram, that is shown in Figure 2.

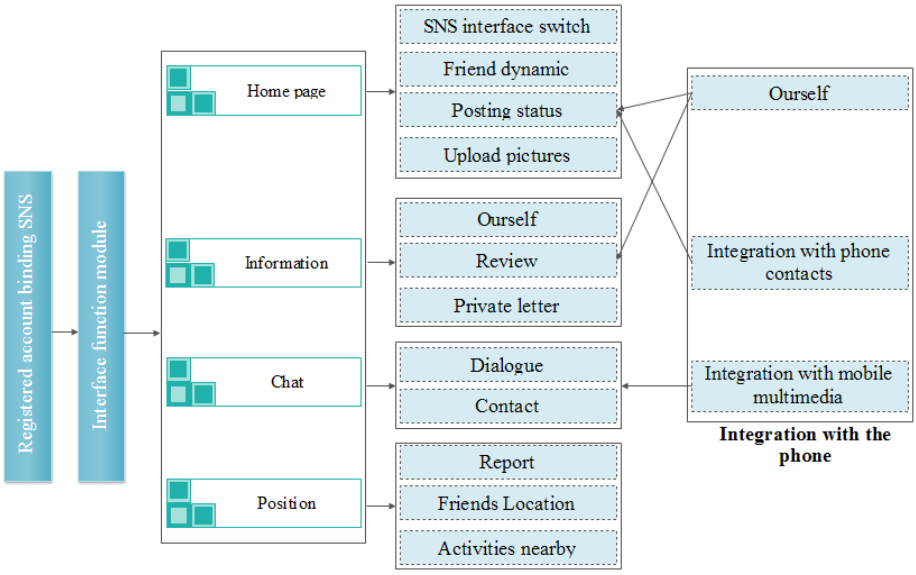


Fig. 2. Information architecture diagram of platform management

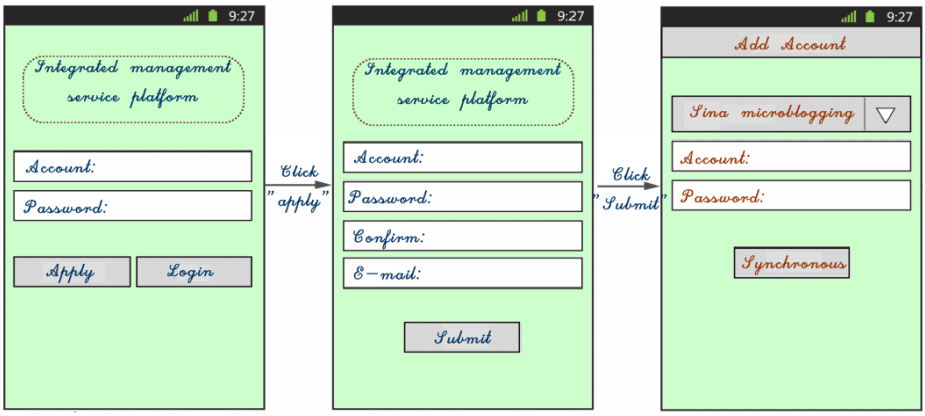


Fig. 3. Interactive interface of register and binding site

4.3 Design Sketches

According to information architecture diagram, combined with the basic operation of Android applications, Figure 3 shows the Android platform mobile social networking site management sketch design.

Figure 4 is a social networking site client registration binding interaction processes.

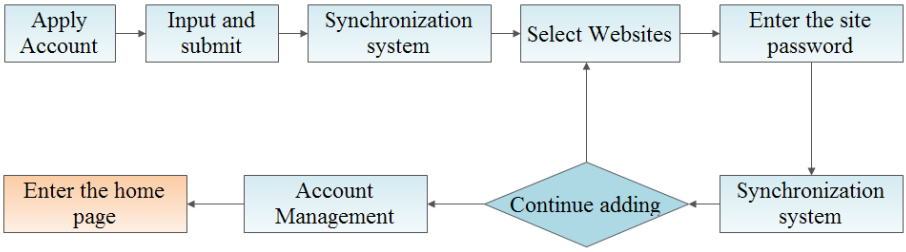


Fig. 4. Interactive process of client registration and binding social networking site

5 User Experience Evaluations

Innovation of management software has three main aspects: First, product innovation: less is more to follow the design principles and emphasizes differences and novelty of the product; followed humane interaction design; Once again the unity of the visual effects: Propagation product concept (including mobile product concept) and stressed the interface clean and standardize and unify. As shown in Table 1, the user experience score sheet.

Table 1. User Experience score sheet

	Useful	Satisfaction	Reliable	Obtain
User 1	8	9	7	9
User 2	9	9	8	9
User 3	8	8	6	8
User 4	8	7	6	8
User 5	9	8	8	8
User 6	9	8	8	8
User 7	9	8	9	9
User 8	8	9	8	9
Average	8.5	8.25	7.75	8.5
	Available	Convenient	Value	
User 1	8	8	9	
User 2	7	8	8	
User 3	8	6	8	
User 4	7	8	7	
User 5	7	9	8	
User 6	8	8	7	
User 7	9	7	9	
User 8	8	7	8	
Average	8.25	7.625	8	

6 Conclusion

Android's sharing feature, can be arbitrarily two local clients, he can be more efficient to meet different user task flow. Such as sharing pictures, Android can be used to set various filters after taking Camera360 directly from any local microblogging client sent to microblogging. Take pictures, and then share, also in line with the user's operating habits. By contrast iOS is more trouble, trouble or developer to put the filter function to do microblogging client, either the user the trouble of the first camera settings saved to the gallery filters, and then send microblogging interface to pick and choose from the gallery.Android's sharing mechanism, but can effectively avoid the client function expansion, helps developers save effort, but also to meet a variety of usage scenarios and user. The new phones are now joined the three-dimensional rules accelerators accelerometer so gestures, movement detection is becoming part of the development of mobile applications. Users are subject to rotating his Android phone browsing the content they wish to make appropriate adjustments along. Of course, if your application does not support the direction of the change in the majority only in specific interfaces such as full-screen video playback and display only specific direction device does not respond to changes in the direction it is reasonable. Note change in direction is responding to the user end and the sound, feel the same depending on the user terminal settings. Or you want to erase a good picture can paint a slight shake your phone can erase the image of the painting Bristol also more fun application is completed the

project schedule notes gently shook automatically deleted. By moving direction and intensity of the sensing applications can be flexibly implement various functions.

Mobile social networking site management evaluation is very high on Android platform. Starting from the actual experience of the user, through its quantitative assessment of the management to make an objective and honest feedback. Practice has proved that demonstrate management platform meets the needs of users, is worth promoting.

References

1. Jin, J., Zhang: Android-based campus streetscape Research and Implementation. In: 2010 Annual Communications Theory and Signal Processing Annual Conference Proceedings (2010)
2. Mobile webserver to the Android platform. In: Proceedings of 2010 International Colloquium on Computing, Communication, Control, and Management (CCCM 2010) vol. 2 (2010)
3. Meng, X., Chen, C.: Android platform for mobile calls and SMS firewall key technologies. In: Proceedings of 2010 the 3rd International Conference on Power Electronics and Intelligent Transportation System, vol. 6 (2010)
4. Bo, D., Liu, X.: Android phone help system design and implementation. In: Proceedings of International Conference of China Communication and Technology, ICCCT 2010 (2010)
5. Meng, X., district cheung: One kind of J2ME applications to the Android platform transplant program. In: Proceedings of 2010 International Conference on Future Information Technology and Management Engineering (FITME 2010), vol. 3 (2010)
6. Liu, Summer Ice, Zhang Y., Xu, Y.: Android city guide system design and implementation. In: Proceedings of 2010 Asia-Pacific Conference on Information Theory, APCIT 2010 (2010)
7. Zhao, F., Sun, M.: Android-based mobile phone system QR Code Recognition Technology. Chinese Society of Agricultural Engineering Electrical Information and Automation Committee, Chinese Society of Electrical Engineering Branch of Technology and Education for Rural Electrification Special Committee in 2010 Academic Year Meeting Paper Abstract (2010)
8. Android Based Wireless Location and Surrounding Search System Design. In: Proceedings of the Ninth International Symposium on Distributed Computing and Applications to Business, Engineering and Science (2010)
9. Yan, T., meter HUIWEN: Natural products based TUI interface design method. In: 18th National Conference on Multimedia (NCMT2009), 5th National Conference on Human-Computer Interaction (CHCI 2009), 5th National Pervasive Computing Conference (PCC2009) Proceedings (2009)
10. Fan J.: SUNNY. Digital home systems IGRS and UPnP dual stack Research and Implementation. In: Third Joint Conference on Harmonious Human Machine Environment (HHME 2007) Proceedings (2007)
11. Ye, P.-F.: Android operating system migration and key technology research. Jinan University (2010)
12. Cheng, P.L.: Google Android-based smartphone platform research and application of Anhui University (2010)
13. Yue fax; Android system migration and application development. Fudan University (2010)

14. Wang, L.: Harbin CMB Android phone bank exchange rates module design and implementation of Jilin University (2011)
15. Hu, S.: Android-based touch-screen system design and implementation. Nankai University (2011)
16. Liu, M.: Mobile terminal Android porting and application design. Xi'an University of Electronic Science and Technology University (2011)
17. Lan, S.: Based on the Android platform, network sharing research and implementation of East China Normal University (2011)
18. Lee, T.: Mobile devices HCI design study. East China Normal University (2010)
19. Li, Z.: Android browser application development platform for research. Beijing University of Posts and Telecommunications (2011)
20. Zhu, G.: Android-based Camera module design and implementation of Xi'an University of Electronic Science and Technology University (2011)
21. Xia, X., Lei, F., Tu speech, Sung, L.: Based on Android and Google Maps life support system design and implementation Chongqing University of Posts and Telecommunications (Natural Science) (February 2012)
22. Huang, Y., Van Morrison, Yue, L.S.: Based on Google Earth geological seismic information display. Geomatics (February 2011)
23. Lin literature. Hotel information on Google Map search service system and its implementation. Computer Development & Applications (May 2012)
24. Wan, T., Xiong, W.: Dalian City Tourism Information System Design and Development. Geospatial Information (June 2011)
25. Hong, D., Glow , Feng, D.: World map in the seismic industry Preliminary application. Seismological Research (April 2011)
26. Zhou, L., Ji, T.: Virtual campus based on Web GIS Integrated System. Computer Development & Applications (September 2012)
27. Xu, G.Y., Songting new , Zhu, Q.: Google Maps-based logistics and real-time monitoring in transit Display Technology. Software Tribune (January 2012)
28. Hu, W., Mahui, J., Gao, Y.-B., Mahan, C.: 51ditu diffusion -based dynamic display network Mashup study. Computer Technology and Development (July 2012)

Driving Program Design of Wireless Sensor Node

Ju Jinwu[1] and Luo Renjia[2]

[1] School of Computer Science of Sichuan University of Science and Engineering, Zigong, 643000
[2] Chongqing Nanfang Translators College of Sichuan International Studies University, Chongqing, 401120

Abstract. The key component of wireless sensor network (WSN) is wireless sensor node. This paper introduces the basic structure of the WSN node, which is composed by CC1100, explains the hardware interface circuit and the characteristics of CC1100, and describes in detail in the process of implementing the CC1100 driver within the embedded operating system of Windows CE, which is based on S3C6410, including CC1100 initialization, receiving and sending interruption, and the interface with the kernel of the operating system kernel etc.

Keywords: wireless sensor network, Windows CE, embedded system, Device driver, CC1100, S3C6410.

1 Introduction

With the rapid development of network technology and sensor technology, the traditional wired sensor networks have been replaced by wireless sensor network [1] (Wireless Sensor Network, WSN) [1-5]. WSN is a wireless network composed of a large number of wireless sensor network nodes through self-organizing modes, it is a highly dynamic network, in which some nodes may go out, and some may join at any time, the network topology can be dynamically changed [6-11]. The WSN node s can be divided into several modules such as sensor, wireless communication, processor, and wireless communicative module works for wireless communication with other nodes, thus the wireless sensor network is formed [12-16]. The Design of wireless communication module determines distance among nodes and the transmission reliability in the system.

The wine cellar control system is responsible for the collection of parameters in the microbial fermentation environment of a plurality of pits, and at the appropriate time, controls the external environment of the pits, making the fermentation process in the best condition, improving the efficiency of production in the wine cellar pools. In this system, there is a decentralized state between two pits, each pit realizes its data acquisition and controlling through a WSN node, and a wireless sensor network comes into being, the main engine of the embedded operating system [17-22] based on Windows CE now works as the control center of the network. Both the control center and nodes realize wireless communications through CC1100 RF transceivers and

Y. Yang, M. Ma, and B. Liu (Eds.): ICICA 2013, Part II, CCIS 392, pp. 48–56, 2013.

construct WSN nodes, which run the protocol for wireless sensor networks, and all the nodes make up of a wireless sensor network, then the host will realize both data acquisition and control of the whole system by WSN [23-29]. In the Windows CE embedded system, to achieve the driving procedure of CC1100 chip is used to support the operation of the protocol in wireless sensor networks.

2 Hardware Structure of the System

The host of Control center is based on ARM11 S3C6410 processor of Samsung Corp, which is a kind of embedded RISC processor with many advantages such as high performance, low cost, low power consumption, small volume etc. The S3C6410 is a 32 bit processor whose working frequency may go up to 800MHZ, with a large number of hardware components built in it, and it adopts the 64/32 bus structure including the 16KB instruction / data cache, MMU storage manager, digital camera interface, video codec, AC-97 audio interface, IIC and SPI serial bus, 4 channel UART, a real-time clock RTC, ADC, interfaces of USB host and devices, LCD controller, general I/O [30, 31], so it is very convenient to use and greatly simplifies the overall structure of pit control system.

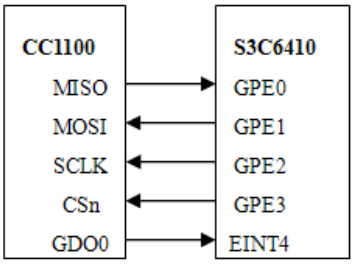


Fig. 1. The connection between CC1100 and S3c6410

2.1 CC1100 Hardware Interface

The CC1100 is a low power RF transceiver [4], which works during the 315/433/868/915MHZ UHF frequency bands, and the transfer rate of the chip data goes up to 500Kbps, the sensitivity -110dBm, and RF output power +10dBm, what's more, the built-in 64 bytes of RX and TX data FIFO, with signal strength indicator (RSSI) output, can support the function of electromagnetic wave (WOR) and the checksum function of built-in hardware CRC, it also supports the transmission mode of a variety of data packets and the modulation of 2-FSK/GFSK/MSK. And the whitening treatment of optional data can significantly reduce the interference of background noise on the data. The working power of CC1100 is 3.3V, and the receiving operating current mode is only 15.6mA, the operating current transmission mode is 28.8mA (+10dBm), the chip is closed through QLP4*4mm, with 20 pins, only a few external components connected, can form the RF transceiver of good performance.

CC1100 uses a 4 line SPI interface to connect with the processor, including SI, SO, SCLK, CSn. The chip is selected when CSn is low, and chip configuration and data transmission are both carried out through the SI, SO, SCLK. The connection of CC1100 and S3C6410 can be shown as Figure 1, because the SPI bus interface of S3C6410 has been used to connect SD card, S3C6410 is connected to the 4 wire SPI interface of the CC1100 through GPIO simulation, and output GDO0 of CC1100, as an interrupt request signal, is connected to the external interrupt EINT4 of S3C6410.

2.2 The Structure of System Hardware

The overall structure of pit control system is shown as figure 2. The control system, as the whole wireless sensor network center, realizes the connection of each WSN node and the wireless communication channel of 433MHZ. In the control system, the LCD display module is used for the real-time monitoring of working state of each node, and the touch screen module realizes its operation and control by receiving the screen hits, NAND Flash module is used to store Windows CE kernel and the file system, Nor Flash module is used for storing BootLoader module, RAM module is the main memory, Ethernet module system is used for the network expansion, the USB module is connected with the external U disk, which is used to save log data of the system monitor, CC1100 wireless module is used to construct the wireless sensor network with lower levels of each WSN node.

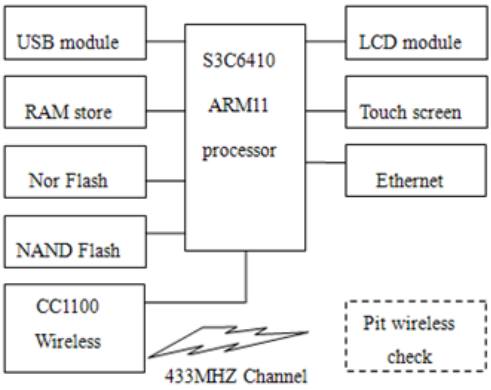


Fig. 2. Overall structure of system

3 Design and Implementation of the Driver

In the operating system of Windows CE, the device driver is the interface between the kernel and hardware circuit, it is an extension of the operating system kernel, which completes device initialization and unloading, transfers the data between the kernel and hardware equipment, between kernel and upper level program, and checks and deals with the errors. The driver shields the operation details of the hardware, while providing the

interface functions to the kernel. The upper application system, through interface functions provided by the file system, indirectly operates the hardware circuit. The CC1100, which is a serial data stream mode based on the chip, should write Windows CE for its stream interface driver. And the driver program of stream interface abstracts the hardware as a device file, whose applications are operated like ordinary device file, and operation functions are standard functions, such as open, close, read, write, control equipment file, and these functions are provided by the operating system. The design of Linux operating system under the character device driver is similar to the design of stream interface driver program under Windows CE[5].

The operating system of the liquor pits controlling system is Windows CE, which is ported to run on S3C6410 and its version is 6 RC3. To design CC1100 driver in Windows CE operating system can make it visit the CC1100, and realize the configuration of CC1100, and complete wireless communication. The CC1100 driver, which is designed into a stream interface driver, can derive some standard primitives from the operating system, and the primitives correspond to the API function, including: CC1100_Init initialization function, CC1100_Open equipment function, CC1100_Read equipment function, CC1100_Write function functions, CC1100_Close equipment function, CC1100_IOCTL controlling command function and so on. The driver must implement initialization functions such as S3C6410 SPI interface, CC1100, CC1100 interrupt service routine.

3.1 Initialization of SPI Interface

In the pit controlling system, the SPI interface connecting CC1100 is simulated by GPIO of S3C6410 through IO interface, before the driver is initialized, priority should be given to the initialization of the IO interface, the main task is to configure the GPIO of S3C6410. The initialization function of the SPI interface is shown as follows:

```
Void SPI_init () {
GPECON = 0xAAAAAA55; / / GPECON[7:0]=010101b
/ / set the GPE0, GPE1, GPE2, GPE3 into GPIO output mode
GPEUP = 0x0000FFFF; / / close up
GPEDAT = GPEDAT | 0x0000000B; / / set the MISO, MOSI, CSn
high
GPEDAT = GPEDAT & (~ (1<<2)); / / SCLK is set to low
GPFCON= (GPFCON & 0Xffffffcff) | (0x02<<8); / / set the GPF4
pin for external interrupt EINT4
EXTINT0= (0x2<<16); / / falling edge trigger
PRIORITY=0X00000000; / / default fixed priority
INTMOD=0X00000000; / / all interrupts are IRQ interrupts
INTMSK= ~ 0X00000010; / / enable EINT4 interruption
EINTMASK= ~ 0X00000010;
```

3.2 CC1100 Initialization

The initialization of CC1100 is mainly the configuration of the CC1100 internal registers, which has 14 command filter registers and 47 configuration registers, many of which are used for chip testing, so 34 registers need to be configured when initializing. The configuration operation runs through the SPI interface, first writes the internal address of the register in CC1100, and then writes the configuration register value. Because the number of registers to configure is so big and complicated, the register configuration values can be obtained through the TI provided by the SmartRF Studio software, users may choose the working mode, frequency, transmission mode, transmission rate, bandwidth, checking methods of CC1100, and the configuration register value can be automatically generated. In order to simplify the code, the configuration data can be stored in the array RF_SETTINGS, the register addresses are stored in the array RF_REGISTER_ADD, which corresponds to each other. The following code is the CC1100 initialization function:

```
void CC1100_init () {
int i;
for (i=0; i<37; i++) {
halSpiWriteReg (RF_REGISTER_ADD[i], RF_SETTINGS[i]);
/ / halSpiWriteReg function is used to write the
configuration data into the specified CC1100 register
```

3.3 The Service Routine of CC1100 Interruption

When CC1100 succeeds in receiving or sending data, or RX FIFO and TX FIFO overflow occurs, an interruption request signal will come out at the GDO0 foot. The CC1100 driver has corresponding interrupt service program, with which it can handle the interruption request, and the program reads the CC1100 status register to obtain relevant information, and deals with it. In initialization code of the driver, the interrupt service routine must be registered in the system. The interrupt service program works as the following code.

```
Static void CC1100_interrupt (int IRQ, void *dev_id, struct
pt_regs *regs)
{char status;
status=halSpiReadStatus (CCxxx0_PKTSTATUS); / / read the
CC1100 status register to understand the reason for
interruption
if (receive data triggered) {
Read CC1100, and receive data.
}else if (data sent) {
The transmit interrupt event
}else if (buffer overflow) {
Error processing
}
```

3.4 The Program Interface of Upper Protocol

Like other hardware devices such as I2C driver, Serial program, CC1100 driver provides interface function of [6] to the upper protocol program, mainly the reading and writing interfaces, which are used to read CC1100 data received. These interface functions can be called by wireless sensor protocol layer.

3.4.1 Data Sending

The process of sending data to CC1100 is shown in figure 3. The function takes 2 parameters, one is the length of data to be transmitted, the other is a pointer to the data to be transmitted. The program firstly calls the function of halSpiWriteReg, writes the length of data to be sent to CC1100; then calls the function of halSpiWriteBurstReg, sends the data to CC1100; and then calls the function of halSpiStrobe, set the internal state of CC1100 machine into sending state, CC1100 is automatically packaging the data, and then sends out. When the transmission is completed, CC1100 will issue an interrupt request, notifying the driver. If it has not yet issued when time is out, the timeout sign appears.

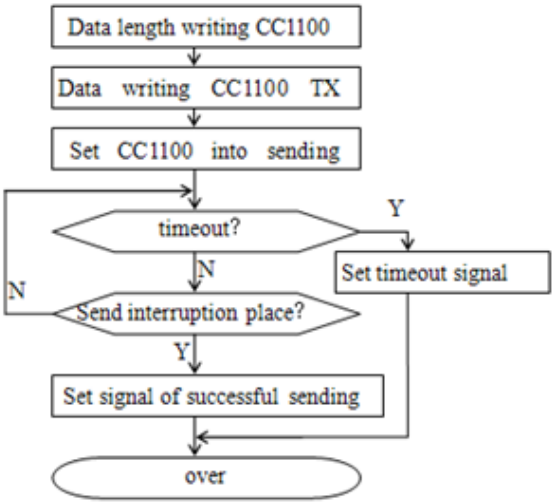


Fig. 3. The process of data sending

3.4.2 Data Receiving

Data-Receiving process is shown in figure 4. The program first calls the function of halSpiStrobe, and sets the internal state of CC1100 into receiving state; and then calls the function of halSpiReadStatus to read CC1100 data length received, and if it is consistent with the predefined length, then calls the function of halSpiReadBurstReg to read data from the RX FIFO of CC1100 to the buffer area; if the length is not consistent, data error, the data is discarded. Finally, the program calls the halSpiStrobe to send a RX FIFO clearing command to the CC1100, clearing RX FIFO and preparing for the next reception.

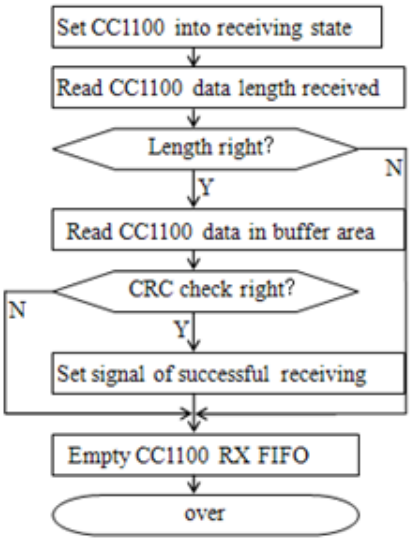


Fig. 4. Process of data receiving

4 Conclusion

This article gives a detailed analysis on the process of implementing the CC1100 RF transceiver driver in the Windows CE embedded operating system, analyzes the application of the hardware interface circuit, software structure and the corresponding code. The Methods to achieve are universal and practical. And the driver has been applied in the project of "white wine cellar controlling system", and achieved good results. The methods introduced in this paper enjoy practical reference value in the design, development of Windows CE driver.

On Windows CE supports multi- threaded, multi- tasking, preemptive embedded operating system. With the release of Windows CE 6.0, a significant upgrade of its core performance and the source code is open, will lead in the consumer electronics, industrial control, mobile communications and other fields has been widely used. Usually given in Platform Builder supports multiple CPU common device drivers, such as LCD drivers, mouse drivers , USB drive , serial port drivers, etc. ; But sometimes the platform adopted other specific hardware device , its driver in Platform Builder does not is given , then the user will need to develop their own against the actual hardware to meet individual needs. This involved the SPI interface driver belongs to this category.

A driver is a physical device or virtual device abstraction function software, drivers, management and operation of these devices feature to export the device to the operating system and applications. According to the different interfaces exported by the driver, windows CE to drive can be divided into native device drivers (native device driver) and streaming driver (streams device driver). Native device driver for Windows CE-based platform integrated into the device. These device drivers are required for some of the hardware is created by the original equipment manufacturer , and used to drive such as a keyboard , LED, touch screen and so on. Native driver is statically

linked at compile time to GWES, system run by GWES loaded. Streaming driver, also known as "Installable Driver" is determined by the device manager (device.exe) dynamically loaded user-mode DLL. Convection driver, whether it needs to control what type of equipment , all streaming drivers are using the same interface and export the same set of functions - stream interface functions. Streaming driver applies any logically be considered as a data source or data storage of the I / O device.

In Windows CE, user mode, each process with other processes memory space occupied by the virtual memory mechanism isolation, a process can not be achieved directly between visits. Drivers and applications in the communication process, the driver is located device.exe process space, due to the above reasons, the driver initiates the application -way communication difficult. The usual solution is in kernel space shared synchronization objects, or through the establishment of message queue pointer mapping to complete. These methods can not avoid the process of the data replication process, and therefore applies only to a small number of data transmission. For large amounts of data or time-critical situations, in the virtual address is Ox4200 0000 ~ 0x7fff ffff space create a named memory mapping to implement shared memory, which can avoid data replication in the process space . Related API is CreateFileMapping and MapView () fFile.

After the driver interface functions written in the form of its interface to export the dll , and then write a simple registry file , specify the path to the driver installation , the prefix and the index . Thus driver design work is complete, the drive and the registry to the current platform can be. This paper describes the SPI Interface Platform Builder driver design and driver interacts with the application implementation. Driver routines have been successfully applied to embedded systems based on nRF2401 wireless communication among the high reference value.

Acknowledgments. Sichuan University of Science and Engineering Incubation Project (2012PY12).

References

1. Qing, Y.: ECal Window CE Driver Development; Science and Technology Information (October 2011)
2. Liu, Y., Li, Z., Ma, Z.: Based on Windows CE USB Key Driver Design; Microcomputer Information, 23 (2008)
3. Wang, J., Wang, J.-G.: Windows CE driver under the Bluetooth protocol stack implementation; Electronic Measurement Technology (May 2006)
4. Chang, Si, W.: WinCE-based stream interface driver development. Industrial Control Computer (June 2008)
5. Liu, W., Zhang, Z.B., Yuan, L., Yun, H., Liu, B.: Windows CE.Net environment Stream Interface Driver Development; Science Technology and Engineering, 20 (2006)
6. Mu, X.: Windows CE USB device under flow-driven development and design; Hebei Academy of Sciences (April 2010)
7. Yao, Z.: Liddy, Lai B cases; embedded technology with Windows CE portable railway ticketing machines; Computer Applications (2003)
8. Kai, P., Zhu name day, Zhang L.: ARM9-based image acquisition system design sucrose crystals. Computer Systems & Applications (July 2010)

9. Tao, Sun F., Sun A.: Windows CE stream interface driver design method. Ship Electronic Engineering (June 2006)
10. Bin, Hong Y.: Automotive navigation SD card Driver Design. electromechanical product development and innovation (January 2008)
11. Zhang, Y.C., Liu, J., Qiang, Ling, B.: Windows CE-based research network image transmission. In: 2010 National Communications Security Conference Proceedings (2010)
12. Fu, Y., Zhang, Q.: S3C2440A-Windows CE 5.0 -based embedded system platform design. In: Second National Joint Conference Proceedings Embedded Technology (2007)
13. Xu, T., room ZONG, Cao, J.: Based on Windows CE Embedded γ spectrum analysis software design . In: 13th National Computer Network in the Field of Application of Modern Science and Technology Conference Proceedings (2007)
14. Jiao, H.: Windows CE under interrupt-based matrix keyboard driver design. In: 2010 China Instrument Academic, Industry Conference (Proceedings 1) (2010)
15. Xing, C., Fu Jie shadow : Based on Windows CE serial communication field meter. In: National New Automation Technology Symposium Conference Proceedings(a) (2005)
16. Lee decided, Li, S.: Car Terminal Based on Windows CE.NET touchscreen shared with the ADC Research. In: 2008 Western China Youth Communication Conference Proceedings (2008)
17. Du, D., He, Q., Tang, B., Chen, F.: SPI bus speed A / D converter Window CE driver implementation. In: 2009 Chinese Control and Decision Conference Proceedings (2) (2009)
18. Jing, J., Yang, W., Gu, S.: Pocket PC voice command system to achieve. In: 2000 Chinese Control and Decision Conference Proceedings (2000)
19. Han C.; Wang K.: S3C24A0 and Windows CE- based Camera Driver Development. In: 2006 Beijing region of Graduate Symposium - Communications and Information Technology Conference Proceedings (under) (2006)
20. Wang, J., Liang, L.: Fanzhe meaning, Zhou, Z., Liu, Z.: LabVIEW Embedded Windows CE-based platform for application development. In: Second National Joint Conference Proceedings Embedded Technology (2007)
21. Bin, based on Windows CE memory card driver design and implementation of Xiamen University (2008)
22. Teng snow, wireless sensor network routing algorithm design and implementation. Jilin University (2008)
23. Gong, M.: Windows CE-based platform for embedded navigation systems research. Wuhan University of Technology (2010)
24. Cheng, L.: Windows CE device drivers and thin client USB-Expander study design. East China Normal University (2005)
25. Xiaoyun: Healthcare Research and Design of Instrument; Nanchang University (2007)
26. Jing, K.: WinCE-based ultrasonic testing automatic interpretation algorithm. Huazhong University of Science (2007)
27. Zheng, L.: Soccer robot based on embedded technology design and implementation. Chongqing University (2007)
28. Liang, Y.: WinCE-based VPN client and Implementation. Huazhong University of Science (2008)
29. Zhang is, ARM9-based gas station level meter console design. East China Normal University (2009)
30. Wang: Based on CAN bus intelligent heating control study. Hebei University of Technology (2005)

Using Behavior of Social Network Sites Based on Acceptance Model

Ya-Kang Chiu and Chien-Wei Huang

Department of Communications Management, Shih Hsin University

Abstract. The population of social network sites (SNSs) users in the world is growing rapidly, and the growth rate is beyond imagination. In light of this, this study conducting in-depth interviews to discover the relevant clues and factors that may affect users' adoption of social network website. The results indicated that attitude has a significant influence on behavioral intention, and intrinsic motivation such as perceived playfulness has a positive influence on the creation of behavior intention. However, perceived usefulness shows no significant influence on behavioral motivation. Moreover, perceived playfulness, perceived ease of use, perceived usefulness, and relationship development all have significant influence on behavior and attitude.

Keywords: Social Network Sites, Technology Acceptance Model, Theory of Motivation, Perceived Playfulness, Perceived Convenience, Relationship Development.

1 Introduction

In recent years, the rapid increase of internet users and the flourishing development of Web2.0 resulted in many prosperous Internet–related activities and markets, and triggered the speedy development of network technology applications, which in turn changed people's style of life [1-6]. In other words, people's information acceptance model has changed from the traditional passive acceptance to active acceptance, and is further marching towards an interactional model through indirectly contributing to the emergence of social network sites (SNSs) with personal traits, engaging in social activities with other users online [7-11]. In other words, in a virtual world created by the Internet, individuals can interact and communicate with others, or engage in information exchange and transmission, which is the same as what they do in the real life. Specifically, under the condition of a high Internet prevalence rate, users' behavior of using the Internet at home is increasing, where the increase rate of social network site usage is most rapid [12-17]. Current popular SNSs are Facebook, Plurk, Myspace, and Twitter. Despite the relatively short development history, SNSs have accumulated an enormous number of users, revealing their market foresight and market development potential [18-20]. In a nutshell, SNSs, propelled by the Internet, has become a global technology commodity that can bypass language barrier, reduce regional difference, and thread users all around the world.

Y. Yang, M. Ma, and B. Liu (Eds.): ICICA 2013, Part II, CCIS 392, pp. 57–66, 2013.

Davis (1986) proposed the Technology Acceptance Model (TAM) to study users' acceptance of new technology, and most studies pointed out that perceived usefulness and perceived ease of use have significant influence on acceptance and adoption of technology systems. However, previous studies either mainly focused on the quality of systems or were from an organizational perspective, yet a perspective from a user experience or motivation was hardly seen [21-23]. In light of this, this study attempts to find out why users actively participate in SNSs through adopting the TAM framework as the research scheme, discounting organizational influence, and trying to include other factors into this model. Besides, the Theory of Motivation comments that TAM mainly focused on the extrinsic motivations such as usefulness and ease of use, whereas little was said about users' intrinsic motivation in the technology using behavior. Therefore, this study, with TAM and Theory of Motivation as the theoretical framework, and structural equation modeling (SEM) as the analysis instrument, aims to investigate whether individuals are also subject to personal intrinsic motivation, in addition to extrinsic motivation, when coming to and adopting a new technology product [24, 25].

Moreover, observation of the practical world and references to theories has discovered that there is not much research and investigation into recently emergent SNSs. Therefore, to ensure relevance and reliability, this study conducted a qualitative survey on users of SNSs to understand features and differences of the relevant using behaviors prior to running a quantitative survey to verify and analyze the conceptual framework.

2 Literature Review

(1) Technology Acceptance Model

TAM was propounded by Davis in 1986. As an extension based on the Theory of Reasoned Action (TRA) by Fishbein & Ajzen (1975), its purpose was to forecast and interpret human's user behavior and the deciding factors in adopting technology systems. It offered a model to explore the influence of external factors on the inner belief, attitude, and behavior motivation of individuals, and the resulting situations of use.

The TAM framework used perceived usefulness and perceived ease of use as the two main constructs as influence on people's adoption of technology systems. Perceived usefulness was defined by Davis (1989) as "the degree to which a person believes that using a particular system would enhance his or her job performance", while perceived ease of use was defined as "the degree to which a person believes that using a particular system would be free from effort." Davis (1989) suggested that users' attitude in technology use is subject to perceived usefulness and perceived ease of use, while attitude would affect behavior motivation, which would subsequently determine happening of the actual behavior of use. According to Barrett, Thornton & Cabe (1968), perceived usefulness and attitude indeed have positive correlation, and perceived ease of use has significant influence on attitude (Carroll & Thomas 1988; Davis 1989).

According to Bagozzi(1981)'s research, both TAM and TRA frames focused on the interrelation of attitude and behavior intention, i.e. the attitude of the individual has a positive influence on his behavior intention (Davis, Bagozzi & Warshaw, 1989). However, the difference between the concepts of TAM and TRA lies in the fact that TAM assumed that not only attitude affects behavioral intention, but perceived usefulness also has direct influence on behavior intention. Davis et al. (1989) assumed that attitude should be seen as a partial mediation that affects perceived usefulness and behavioral intention. In addition, perceived usefulness has positively relevance to both attitude and behavioral intention. This is to say, behavioral intention is not solely determined by attitude, and perceived usefulness can also lead to formation of behavioral intention. Other associated studies also point out correlation does exist between belief and behavioral intention. In other words, TAM empathizes that the construction of human's behavioral intention to adopt technology systems is a cognitive evaluation. For instance, employees choose to use technology systems due to a self-improvement aspiration or a conception that it can enhance their work performance. Specifically, in addition to attitude, TAM puts more emphasis on the fact that an individual's perceived usefulness in technology system would affect their behavioral intention. In addition, Davis et al. (1989) suggested that perceived ease of use has direct positive influence on perceived usefulness, which means when individuals assume a technology system is easier to use, and more performance can be achieved to enhance personal efficiency with the same effort, individuals would lean toward a perception that this technology is useful, and will have a positive attitude toward the technology system, which would further affect individual's behavioral intention and usage behavior. Although Davis et al. (1989) did not mentioned clearly what the external variables that affect perceived usefulness and perceived ease of use are, a number of research (Adams, Nelson & Todd, 1992; Chin & Todd, 1995; Morris & Dillon, 1997) have used TAM as the research basis to investigate factors that affect the adoption of new technology system. These factors and variables, however, would vary depending on the technology feature, users, and content. Because of this, many other scholars have expanded on the basic framework of TAM to investigate the relevant factors that affect users' adoption of new technology systems.

(2) Theory of Motivation and Perceived Playfulness

Underneath an individual's behavior lies the influence of personal motivation, and therefore the occurrence of actual behavior. Deci (1975) pointed out that motivation is an intrinsic factor that propels people to engage in all sorts of activities. In general, the Theory of Motivation categorizes factors that affect people's behaviors into extrinsic motivation and intrinsic motivation (Calder & Staw, 1975; Deci, 1975; Mathieson, 1991). Deci (1975)'s definition of the two types of motivation is as follows: extrinsic motivation refers to an individual's perception that a task performance can help the individual to achieve a valuable result, such as enhanced work performance or raised salary; intrinsic motivation means that an individual cares about the process of engaging in a task. Intrinsic motivation arises from an individual's internal need, and is not driven by any external reasons.

Perceived playfulness is seen as an important concept of intrinsic motivation. According to Davis, Bagozzi, & Warashaw (1992), perceived playfulness means in addition to the resulting benefits of using information technology, an individual also feels a pleasant feeling from using information technology. Perceived playfulness is therefore categorized as intrinsic motivation, and perceived usefulness is categorized as extrinsic motivation. Ridings & Gefen (2004) pointed out that the entertainment offered by the Internet virtual communities is one reason that leads to people's participation. Chen, Wigand & Nilan (1999) used TAM as the fundamental theory to investigate influence of intrinsic motivation such as perceived playfulness on users when they use the Internet. Shang, Chen & Shen (2005) studied the intrinsic motivation of consumers, and found that perceived playfulness has influence on the users' on-line shopping behavior. To summarize the above, an individuals' perceived playfulness is an important factor that attracts them to use computer technology systems.

(3) Perceived Convenience and Relationship Development

It can be found from the previous research that quite few research have discussed about perceived convenience and a clear definition was unseen (Yoon & Kim, 2007). However, some scholars still postulate that convenience is a multi-dimensional concept (Brown, 1990). Berry, Seiders, & Grewal (2002) suggested that any convenience that can reduce consumers' time and energy can be called service convenience. In conclusion, the main concept of convenience lies in saving users' time or energy whereas the purpose of the convenience offered by technology is to improve the efficiency of users' information acquisition, reduce the spent time and energy of users, and help users acquire relevant information instantly and rapidly.

According to Teri & Connie (2002), self-disclosure is a feature of individuals revealing information relating to oneself to others through verbal communication, and the information disclosed by individuals includes personal thoughts, feeling, and experience (Derlega & Grzelak, 1979; Derlega, Metts, Petronio & Margulis, 1993). Jourard (1959) stated that individuals, through self-disclosure, can take the initiative to share personal information with others, letting others know the storyteller. Derlega & Grzelak (1979) mentioned that self-disclosure is a voluntary behavior, and the disclosed information is hard to obtain elsewhere. Berg & Archer (1982) disassociated personal genuine information and evaluative information, pointing out that disclosing personal true and everyday information is not only the foundation of relationship development at the initial stage, but also the key factor in maintaining a relationship in the future. In conclusion, self-disclosure refers to the process of an individual making others understand oneself by expressing and narrating personal thought, emotion, or experience. Henderson & Gilding (2004) pointed out that moderate self-disclosure is conductive to the development of personal relationships in the Internet. In sum, in the behavior of internet use, correlation does exist between self-disclosure and relationship development, which is particularity obvious in peer interaction behavior. The development of Internet applied technology not only provides people a channel to gather information, but also becomes a shortcut for people to distribute information.

3 Pretest and Pilot Study

Before the questionnaires were distributed, this study conducted an initial survey on college students' use of SNSs to understand this main user group's situation and habit in using SNSs. A total of 74 valid questionnaires were collected after deleting those seriously incomplete ones. Analysis found that it is true that Facebook is very popular among college students. Facebook far surpasses the counterparts Plurk and Twitter in the degree and frequency of users' interaction with friends and relatives, the degree of friends and relatives' communication and interaction with others, the number of friends, the average duration of using time, and how many times users log in. It can be observed that college students' acceptance to Facebook is higher than the acceptance to other SNSs, which is reflected on the actual behavior of use.In light of this, this study chose Facebook as the topic for an in-depth investigation to further find out the key factors behind users' acceptance of SNSs.

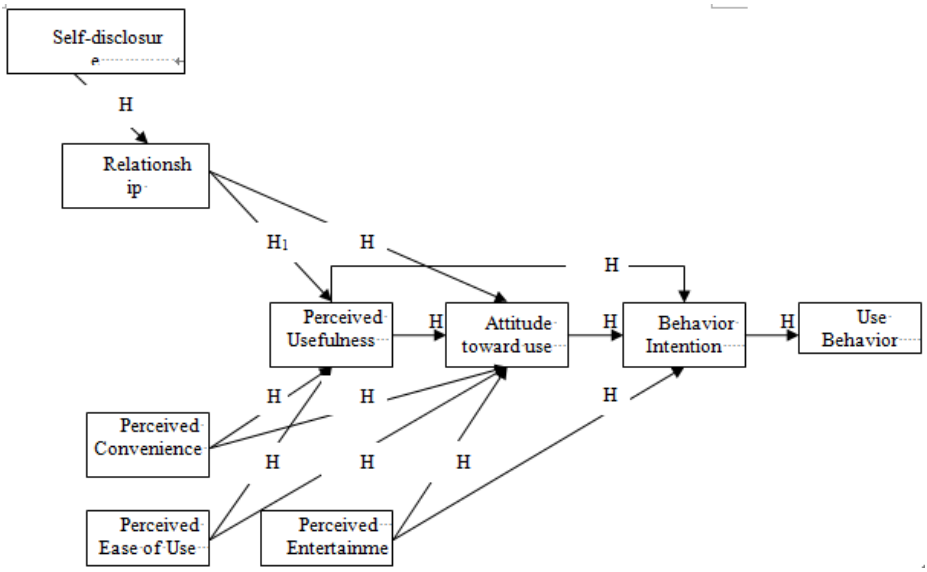


Fig. 1. The conceptual framework of this study

After confirming the research subjects, the qualitative research instruments were used to examine the features of the relevant usage behaviors. All users were categorized into two groups: light users and heavy users, to differentiate the different behaviors of social network site users. A total of seven college students with an average age of 21 years old were interviewed. The average interview time was one hour. This interview attempted to understand the relevant clues to college students' use of SNSs, and extract the relevant factors that affect individual's use of SNSs.

4 Research Framework

In the study, technology acceptance model is used as the main conceptual architecture, and five concepts are proposed for deeply exploring the behaviors of users to use social network sites in combination with the pilot study, and then the conceptual framework of this study is shaped. The overall study framework is as shown in figure 1.

5 Hypotheses Deployment

According to the TAM proposed by Davis (1989), perceived usefulness and perceived ease of use are the main factors affecting the behavioral attitude of individuals toward science and technology system. Therefore, it is reasoned in this study that individuals will hold a positive attitude toward social network sites if they think the sites will provide relevant information beneficial to them. Therefore, the following hypothesis is proposed in this study.

H_1:「perceived usefulness」 will positively influence 「behavioral attitude」

H_2:「perceived ease of use」 will positively influence 「behavioral attitude」

According to Davis et al. (1989), perceived usefulness and perceived ease of use will not only influence behavioral attitude in common, and also have a correlation between each other. That is, perceived usefulness will be positively influenced by perceived ease of use. Therefore, the following hypothesis is proposed in this study.

H3:「perceived ease of use」 will positively influence 「perceived usefulness」

Moon & Kim (2001) study shows that TAM can effectively explain the behaviors of individuals for technology system indeed, and user's behavioral intention will be influenced by behavioral attitude. Therefore, the following hypothesis is proposed in this paper.

H4:「behavioral attitude」 will positively influence 「behavioral intention」

H5:「behavioral intention」 will positively influence 「use behavior」

In addition, Davis et al. (1989) argues that behavior intention is not entirely decided by attitude, and perceived usefulness may also influence the existence of behavioral intention. Accordingly, a conclusion is drawn up in this study that the behavioral intention of individuals to use social network sites will be positively influenced if the perceived usefulness for the sites is higher and higher. Therefore, the following hypothesis is proposed:

H_6:「Perceived usefulness」 will positively influence 「behavioral intention」

The following hypothesis is proposed:

H_7:「perceived convenience」 will positively influence 「perceived usefulness」

H_8:「perceived convenience」 will positively influence 「behavioral attitude」

According to the survey of Collins & Miller (1994), self-disclosure can promote information receiver to like the disclosure person, while the receiver will produce a sense of responsibility for providing a kind response. Individuals may think that the friendships with other people can be effectively developed using social network sites, and subsequently think that it is useful for them to use social network sites, and then will trust the advantages of using a social network site, so they will hold a good attitude toward social network sites. Therefore, the following hypothesis is proposed in this study.

H_9:「self-disclosure」 will positively influence 「relationship development」

H_{10}:「relationship development」 will positively influence 「perceived usefulness」

H_{11}:「relationship development」 will positively influence 「behavioral attitude」

In this study, the author thinks that the purpose of individuals to use social network sites may source from inner psychological motivation. For example, perceived entertainment aims at enjoying the pleasure of using a social network site, so individuals hold an initiative attitude toward social network sites. Based on the above mentioned, the following hypothesis is proposed.

H_{12}:「perceived entertainment」 will positively influence 「behavioral attitude」

H_{13}:「perceived entertainment」 will positively influence 「behavioral intention」

There are 9 constructs totally in this study, and they are measured using Likert five-point scale and are scored from "agree very much" (5) to "disagree very much" (1). In the aspect of use behavior, open question items are used for asking about the objects' actual circumstances about the use of social network sites.

In self-disclosure, Derlega & Gerzelak (1979) are used as reference, including a total of 5 items for measuring the self-disclosure degree of users. In relationship development, Derlega & Margulis (1983) are used as reference, including a total of 9 items for measuring the relationship development degree thought by users for using social network sites. In perceived entertainment, Van der Heijden (2004) and Davis et al. (1992) are used as reference, including a total of 5 items for measuring the perceived entertainment degree of users to use social network sites. In perceived convenience, Yoon & Kim (2007) are used as reference, including a total of 5 items for measuring the perceived convenience degree of users to use social network sites. In perceived usefulness and perceived ease of use, Davis (1989) is used as reference, including a total of 5 items for measuring the perceived usefulness and perceived ease of use to use social network sites. In behavioral attitude and behavioral intention, Moon & Kim (2001) are used as reference, including a total of 5 items for measuring the behavioral attitude and behavioral intention of users to use social network sites. In behavioral intention, a total of 8 items are adopted for measuring the intention of users to use social network sites. Finally, use behavior is developed by the author, including a total of 9 items for measuring the actual use behavior of users to use social network sites.

Study results

6 Test on Model Fit

Study data was collected using a network questionnaire survey, and there were 432 valid questionnaires after the questionnaires filled by the same persons were deleted.

Table 1. The test on the overall fit before and after the items in questionnaire are deleted

	Indicators of model fit						
	χ^2/df	P-value	GFI	AGFI	CFI	NFI	NNFI
Ideal value	<3	0.000	>0.8	>0.8	>0.95	>0.9	>0.9
Before questionnaire are deleted	4.28	0.000	0.68	0.65	0.96	0.95	0.96
After questionnaire are deleted	2.89	0.000	0.85	0.82	0.98	0.97	0.98

Note: The above data is processed by the author in this paper.

In this study, the model fit was tested using LISREL software first. Confirmatory factor analysis (CFA) was made and then question items were deleted based on modification indices, so as to improve the fit of the model. The model fit before the

deletion of question items did not achieve the ideal value standard. Therefore, the questionnaire before the deletion of question items was corrected, and the question items were deleted according to the modification indices, so as to improve the model's explanation ability. Later, the model fit was test again using CFA according to the questions after the deletion of question items, and the chi square/df (χ^2/df) after the deletion of question items was 2.89 (reduced to less than the ideal numerical 3), and also GFI, AGFI, CFI, NFI and NNFI were all more than 0.8. The numerical values of all model fit pointers passed the test, suggesting the fit of the model was good.

7 Hypothesis Test Results

The model after deletion of question items was tested using SEM's path analysis for testing the causal relationship between variables. According to the test results, ten hypotheses (H_1, H_2, H_4, H_5, H_7 and H_9, H_{10}, H_{11}, H_{12}, and H_{13}) were supported and significantly correlated, except that H_3, H_6 and H_8 did not reach the significant level. In addition, an overall path analysis was made to the non-established hypotheses, even though H_3, H_6 and H_8 hypotheses were not significant in the direct path testing. Namely, perceived ease of use, perceived usefulness, and perceived convenience did not play a significant effect on perceived usefulness, behavioral intention, and use attitude, respectively, but perceived convenience, perceived ease of use, and perceived usefulness showed a significant effect in the overall path test, as shown in table 2.

Table 2. The overall path test results of the study's hypotheses

Overall path	Internally contained paths	Estimated coefficient	t-value
Perceived convenience →Use behavior	Perceived convenience →Perceived usefulness →Use behavior →Behavioral intention →Use behavior		
	Perceived convenience →Perceived usefulness →Behavioral intention →Use behavior	0.04	3.41***
	Perceived convenience →Use attitude →Behavioral intention →Use behavior		
Perceived ease of use →Use behavior	Perceived ease of use → Perceived usefulness →Use attitude →Behavioral intention →Use behavior		
	Perceived ease of use → Perceived usefulness →Behavioral intention →Use behavior	0.04	4.13***
	Perceived ease of use → Use attitude→ Behavioral intention →Use behavior		
Perceived usefulness →Use behavior	Perceived usefulness →Use attitude→ Behavioral intention →Use behavior	0.07	2.95**
	Perceived usefulness→ Behavioral intention →Use behavior		

Note: *=P<0.05; **=P<0.01; ***=P<0.001

8 Conclusion and Suggestions

Based on the above results, it is concluded that the architecture advocated by TAM is very applicable to discussing the behaviors of using social network websites. In this study, it was found that perceived entertainment and relationship development will influence the generation of behavioral attitude and then behavioral intention besides perceived usefulness and perceived ease of use will positively influence behavioral attitude. The significant effect of perceived convenience on behavioral attitude was not tested in this study, but it was found that perceived convenience and relationship development can positively influence behavioral attitude via perceived usefulness. In addition, the significant effect of external motivation's perceived usefulness on behavioral intention was not tested in this study, but it was found that the significant effect of perceived entertainment on the behavioral intention of users to use social network sites exists. Finally, from the study results, a positive significant correlation can be tested between behavioral intention and use behavior.

References

1. Barrett, G.V., Thornton, C.L., Cabe, P.A.: Human Factors Evaluation of a Computer Based Storage and Retrieval System. Human Factors 10, 431–436 (1968)
2. Berry, L.L., Seiders, K., Grewal, D.: Understanding Service Convenience. Journal of Marketing 66(3), 1–17 (2002)
3. Brown, L.G.: Convenience in Services Marketing. Journal of Services Marketing 4(1), 53–59 (1990)
4. Calder, B.J., Staw, B.M.: Self-Perception of Intrinsic and Extrinsic Motivation. Journal of Personality and Social Psychology 31(4), 599–605 (1975)
5. Carroll, J.M., Thomas, J.C.: Fun. ACM SIGCHI Bulletin 19(3), 21–24 (1988)
6. Collins, N.L., Miller, L.C.: Self-Disclosure and Liking: A Meta-Analytic Review. Psychological Bulletin 116(3), 457–475 (1994)
7. Davis, F.D.: A Technology Acceptance Model for Empirically Testing New End-User Information Systems: Theory and Results. Doctoral Dissertation, Sloan School of Management. Massachusetts Institute of Technology (1986)
8. Davis, F.D.: Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. MIS Quarterly 13(3), 319–340 (1989)
9. Davis, F.D., Bagozzi Richard, P., Warshaw, P.R.: User Acceptance of Computer Technology: A Comparison of Two Theoretical Models. Management Science 35(8), 982–1003 (1989)
10. Davis, F.D., Bagozzi, R.P., Warashaw, P.R.: Extrinsic and Intrinsic Motivation to User Computers in the Workplace. Journal of Applied Social Psychology 22(14), 1111–1132 (1992)
11. Deci, E.L.: Intrinsic Motivation. Plenum Press, New York (1975)
12. Derlega, V.J., Grzelak, J.: Appropriateness of Self-Disclosure. In: Chelune, G.J. (ed.) Self Disclosure: Origins, Patterns and Implications of Openness in Interpersonal Relationships, pp. 151–176. Jossey-Bass, San Francisco (1979)
13. Derlega, V.J., Margulis, S.T.: Loneliness and Intimate Communication. In: Perlman, D., Cozby, P.C. (eds.) Social Psychology, pp. 207–226. Holt, Rinehart and Winston, New York (1983)

14. Derlega, V.J., Metts, S., Petronio, S., Margulis, S.T.: Self-Disclosure. Sage, Newbury Park (1993)
15. Fishbein, M., Ajzen, I.: Belief, Attitude, Intention and Behavior: An Introduction to Theory and Research. Addison Wesley, Boston (1975)
16. Henderson, S., Gilding, M.: I've Never Clicked This Much with Anyone in My Life: Trust and Hyper Personal Communication in Online Friendships. New Media & Society 6(4), 487–508 (2004)
17. Liao, Z., Cheung, M.T.: Internet-Base E-Banking and Consumer Attitudes: An Empirical Study. Information & Management 39(4), 283–295 (2002)
18. Mathieson, K.: Predicting User Intentions: Comparing the Technology Acceptance Model with the Theory of Planned Behavior. Information Systems Research 2(3), 173–191 (1991)
19. Moon, J.-W., Kim, Y.-G.: Extending the TAM for a World-Wide-Web Context. Information and Management 38(4), 217–230 (2001)
20. Morris, M.G., Dillon, A.: How User Perceptions Influence Software Use. IEEE Software 14(4), 58–65 (1997)
21. Ridings, C.M., Gefen, D.: Virtual Community Attraction: Why People Hang Out Online. Journal of Computer-Mediated Communication, 10 (1), Article 4 (2004)
22. Shang, R.-A., Chen, Y.-C., Shen, L.: Extrinsic versus Intrinsic Motivations for Consumers to Shop On-Line. Information and Management 42, 401–413 (2005)
23. Teri, Q., Connie, S.: Academic Self-Disclosure in Adolescence. Genetic, Social, and General Psychology Monographs 128(1), 47–75 (2002)
24. Van der Heijden, H.: User Acceptance of Hedonic Information Systems. MIS Quarterly 28(4), 695–703 (2004)
25. Yoon, C., Kim, S.H.: Convenience and TAM in a Ubiquitous Computing Environment: The Case of Wireless LAN. Electronic Commerce Research and Applications 6(1), 102–112 (2007)

Detection and Service Security Mechanism of XML Injection Attacks

Zhao Tao

Yunyang Teachers' College, Shiyan Hubei, China, 442000

Abstract. The paper studies two classical XML injection attacks, structure-altering XML injection attack and parameter-error XML injection attack. For XML injection attacks, the paper proposes a detection method based on SOAP message verification tree. For the characteristics of parameter-error XML injection attack, the paper proposes a method that detects if there are nodes with the same name under a node of SOAP messages. Simulation results indicate that two detection schemes proposed in the paper is effective.

Keywords: net platform, XML injection attack, SOAP message, attack detection.

1 Introduction

With wide application of Web service, Web service plays an important role in software system including critical business of enterprises. If the security of Web service is weak, it may cause serious consequences [1-3]. Therefore, the security of Web service becomes an important problem restricting the application and development of Web service, and attracts people's attention [4]. WS-Security provides quality assurance for the transmission of SOAP message by message integrity, message confidentiality and individual message authentication, which protects the security of Web service at SOAP message level [5-7]. On net platform, the paper studies and detects structure-altering XML injection attack and parameter-error XML injection attack.

2 XML Injection Attack Scenario

2.1 Attack Scenario Altering XML Structure

Structure-altering XML injection attack means that the attacker writes the unauthenticated data into SOAP request message. After receiving SOAP message, Web service couldn't detect the abnormity of SOAP message and resolve the incorrect data. The XML injection attack modifies nest relation of SOAP message, which means that the other element is embedded into an element to realize the attack. For SOAP message in Fig. 1, 3 is regarded as parameters to be embedded into the element a. WSDL document of the Web service provides that the element a is an int data. It is obvious that the SOAP message doesn't meet the requirements of Webs service on SOAP message sent by the requester side, and it should be abandoned. In fact, on net

Y. Yang, M. Ma, and B. Liu (Eds.): ICICA 2013, Part II, CCIS 392, pp. 67–75, 2013.

platform, the SOAP message is not abandoned, and the resolved result of the SOAP request message is a=3, b=0. Therefore, the attack updates XML nested structure to make XML injection attack, which causes unpredictable results such as access restricted data.

```
<Envelope>
    <Body>
        <sum>
            <a><b>3</b></a>
            <b>2</b>
        </sum>
    </Body>
</Envelope>
```

Fig. 1. Strucure-altering XML injection attack SOAP message

2.2 Parameter-Error Attack Scenario

The writer gives a XML document fragment as shown in Fig. 2, and points out that in the process of using SAX parser to parse the XML document fragment, when id label appears for the first time, the value of the label will be assigned to id attribute of variable user, and the id attribute value of variable user is 0728224. But XML attacker transmits the incorrect parameters including Shanghai</addr> <id>0</id> <addr>China to addr, which makes SAX parser parse id label after parsing the first addr tag. And SAX parser assigns the value of attribute id which is 0 to user, the value of attribute id of user is rewritten, and the attribute id value of variable user is 0. We can see that in the process of using SAX parser to parse XML document, the values of the posterior nodes are allowed to cover the values of the preceding nodes. If the values of key nodes are covered, it may make illegal users steal advanced permissions, and the attacker pretends to be legal user to steal confidential data, which makes data information leak.

```
<user>
    <id>0728224</id>
    <name>Tracy</name>
    <email>tracy@hotmail.com</email>
    <addr>Shanghai</addr><id>0</id><addr>China</addr>
    <phonenumber>02188888888</phonenumber>
    <zip>200093</zip>
</user>
```

Fig. 2. XML document fragment

Alex Stamos and Scott Stender of ISEC Partners Company quotes the example in Table 4-1 in the report of 《Attacking Web Services: The Next Generation of Vulnerable Enterprise Apps》.

In the experiment, we find that using SAX parser to parse XML document fragment in Fig. 2 has the phenomenon that the node data is covered, and the value of the next id node, 0728224, covers the value of the preceding id node. But we find from the experiment results under net platform and Axis2 platform that SOAP request message like Fig. 2 is delivered to Web service, and there is no the phenomenon that the node data is covered in the process of Web service parsing. In the process of parsing SOAP message on net platform and Axis2 platform, the value of the first id node, 0728224, is only taken. When SOAP request message shown in Fig. 3 is delivered under net platform and Axis2 platform, in the process of parsing the SOAP message, the value of the first id node, 0, is only taken, and the value of effective node id, 0728224 is ignored, which causes XML injection attack.

```
<Envelop>
    <Body>
        <getUserInfo>
            <user>
                <addr>Shanghai</addr><id>0</id><addr>China</addr>
                <id>0728224</id>
                <name>Tracy</name>
                <email>tracy@hotmail.com</email>
                <phonenumber>02188888888</phonenumber>
                <zip>200093</zip>
            </user>
        </getUserInfo>
    </Body>
</Envelop>
```

Fig. 3. SOAP request message

3 Attack Simulation Experiment Based on .Net

3.1 Attack Simulation Experiment of Altering XML Structure

(1) Creating Web service

The experiment needs to create a Web service which provides service for service requester. Firstly, a class Service is established, and Service inherits System.Web.Services.Web Service. A Web method sum() is established in Service class. The return value of sum method is int type. Calling the method needs two int-type parameters. The function of the method is to add two parameters and return the acquired result to service request side, as shown in Fig. 4.

```
[WebService(Namespace = "http://tempuri.org/")]
[WebServiceBinding(ConformsTo = WsiProfiles.BasicProfile1_1)]
public class Service : System.Web.Services.WebService {
    [WebMethod]
    public int sum(int a, int b){
        return a+b;
}
```

Fig. 4. Web service providing addition method

(2) Creating service request side

A Web service request side is created for the experiment. A Web service proxy class is created in the solution, and it is used to request addition service. The codes of service request side are shown in Fig. 5.

```
sumService.ServiceWse ss = new sumService.ServiceWse();
int result = ss.sum(3,2);
```

Fig. 5. Service request side

In order to simulate XML injection attack, custom policy assertions need to be deployed at service request side. Firstly, a custom assertion class ConsoleAssertion is established. Its main function is to modify SOAP message, which means to embed the element b into the element a, which can generate a SOAP request message causing XML injection attack. Then, the custom policy assertions are deployed to the service request side. A policy assertion class object policy is generated in the codes of service request side, and an example ca of custom assertion Console Assertion generates. Ca is added to the policy, and the policy is deployed at service request side ss. And the service request side after deploying custom policy assertions is as shown in Fig. 6.

```
sumService.ServiceWse ss = new sumService.ServiceWse();
Policy policy = new Policy();
ConsoleAssertion ca = new ConsoleAssertion();
policy.Assertions.Add(ca);
ss.SetPolicy(policy);
int result = ss.sum(3,2);
```

Fig. 6. Service request side after deploying custom policy assertions

After deploying custom policy assertions, service request side requests Web service. And SOAP message which is sent at this time is processed by policy assertion Console Assertion, and the element b is embedded into the element a. SOAP request information after modification of custom policy assertions is shown in Fig. 1.

3.2 Attack Simulation Experiment of Parameter Error

(1) Creating Web service

In the experiment, a Web service is created to serve service request side. Firstly, a service class Service is established, and it inherits System.Web.Services.WebService. A Web method getUserInfo () is established in Service class. Calling getUserInfo method needs a User-type parameter. GetUserInfo method extracts the information of User-type parameter to compose a character string and return to user. Web service codes shown in Fig. 7.

```
[WebService(Namespace = "http://tempuri.org/")]
[WebServiceBinding(ConformsTo = WsiProfiles.BasicProfile1_1)]
public class Service : System.Web.Services.WebService
{
    [WebMethod]
     public string getUserInfo(User user) {
         return user.id + " " + user.name + " " + user.email + " " + user.addr + " " +
                  user.zip + " " + user.phonenumber;
     }
}
```

Fig. 7. Web services code

User class used in the experiment is a Bean class passing data, and it includes six data fields such as id, name, email, addr, zip and phonenumber. The data fields exist in SOAP message with the form of tag.

(2) Creating service request side

In the experiment, in order to inject incorrect parameters for addr node in SOAP message, custom policy assertion is deployed at service request side. Firstly, a custom assertion class Console Assertion is established. And its main function is to inject incorrect function Shanghai</addr><id>0</id><addr>China into addr tag. Then, custom policy assertion is deployed at service request side. A policy class objective is generated at service request side. And a case ca of custom assertion Console Assertion is generated. Ca is added to the policy, and the policy is deployed at service request side ss. Service request side after deploying custom policy assertion is shown in Fig. 8.

```
sumService.ServiceWse ss = new sumService.ServiceWse();
Policy policy = new Policy();
ConsoleAssertion ca = new ConsoleAssertion();
policy.Assertions.Add(ca);
ss.SetPolicy(policy);
User user = new User("0728224","Tracy","tracy@hotmail.com ","Shanghai",
"200093","02188888888");
string result= ss. getUserInfo(user);
```

Fig. 8. Service request side after deploying custom policy assertion

Service request side uses getUserInfo method of Web service side, and SOAP message to be sent is processed by policy assertion Console Assertion. Incorrect parameter information Shanghai</addr><id>0</id><addr>China has been injected into addr tag. SOAP message which is sent to service request side is shown in Fig. 3.

4 Attack Detection Schemes

4.1 XML Injection Attack Detection Algorithm

Deterministic process of Validate () method is as follows.

The first step: Judging if the node a is the same to the node v. If they are not the same, it returns to false. If they are the same, it returns to the second step.

The second step: If the number of children of the node a and the node b is not 0, it enters the second step. If the number of children of the node a is 0, and the number of children of the node v is not 0, it enters the fourth step. If the number of the children of the node a and the node v is 0, it returns to true. And if there are other situations, it returns to false.

The third step: the child nodes with the same name of the node a and the node v are used as the parameters to call validate () method. When the number of children of the node a and the node v is the same, it returns to true, or it returns to false.

The fourth step: Judging if the data type of the value of the node a is legal. If it is legal, it returns to true, or it returns to false.

According to SOAP message verification tree, SOAP request message which is sent by request side needs to be verified, which means to verify nested formats of elements in SOAP message. If SOAP request message which is sent by the request side is validated, it is allowed to access the Web service. If the verification fails, Email address of SOAP message header is extracted and recorded in journal file, which forms a blacklist. When there is SOAP request message being sent to the server side, it is necessary to check if Email address of SOAP message header is in blacklist.

4.2 Attack Detection Schemes of Parameter Error

Test method of judging if there are nodes with the same name under the node a is as follows.

The first step: Judging if the node a has child node. If it has no child node, it returns to true, or it enters the second step.

The second step: Judging if the children under the node a have the nodes with the same name. The child node of the node a is used as parameter call test method.

The third step: Comparing if all child nodes of the node a have the same name. If they have the node with the same name, it returns to false, or it returns to true.

The algorithm uses recursive way to judge if there is the node with the same name under the node a.

On .net platform, test Net method of judging if there is the node with the same name under the node a is shown in Fig. 9.

```
bool testNet(XMLNode a){
    if(a.child.count==0)
        return true;
    int count = a.child.count;
    bool flag[] = new bool[count];
    for(int i=0;i<count;i++){
        flag[i]=false;
        flag[i]= testNet (a.child[i]);
     }
    bool result = true;
    for(int j=0;j<count;j++)
        result&=flag[j];
    if(result == false)
        return false;
    bool f = true;
    for(int k=0;k<count-1;k++){
        for(int m=k+1;m<count;m++){
            if(a.child[k].name==a.child[m].name){
                f=false;
                break;
               }
          }
        if(!f)
            break;
    }
    return f;
}
```

Fig. 9. Detection method test Net under .net platform

5 Simulation Experiment Analysis

The detection methods proposed in the paper receive simulation experiment, and the experiment results are as follows.

5.1 Experiment Results and Analysis of Attack Detection of Altering XML Structure

Custom policy assertions are deployed at Web service side and detect SOAP request message. If the service request side sends SOAP request information in Fig. 1, the workflow of SOAP message verification algorithm in custom policy assertions is as follows.

According to SOAP message verification tress, SOAP detects SOAP request message, and the detection process is shown in Fig. 10.

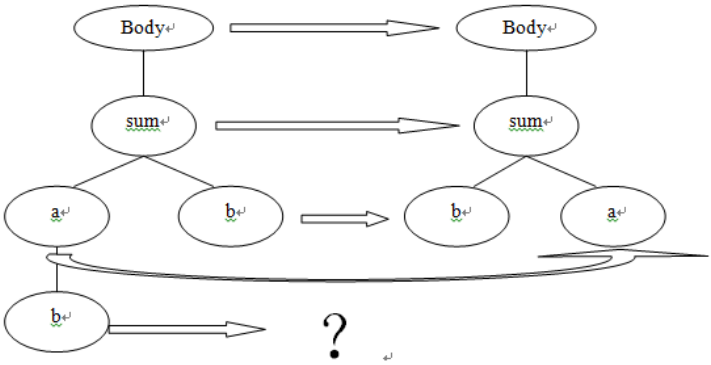


Fig. 10. Process of detecting SOAP message

In Fig. 10, the nodes are compared from Body node one by one. From the arrow indication, we can seethe nodes of SOAP message such as Body, sum and a. And the corresponding elements are found out in SOAP message verification tree, which indicates that these elements are legal and appear in the appropriate hierarchical structure. In SOAP message verification tree, the element corresponding to element b under SOAP message element a couldn't be found, which indicates that it is illegal for the element b to appear here or the element b can't appear under the element b. In Fig. 10, the location with red question mark indicates that the element b is illegal and can't pass the detection of SOAP message detection algorithm, and the detection result is false. Web service rejects to provide service for service requester sending SOAP request message, and sends a false prompt message for service request side.

5.2 Experiment Results and Analysis of Attack Detection of Parameter Error

In the simulation experiment, SOAP message received by Web service side is shown in Table 2. We can see from Fig. 11 that there are two node id and two node addr under user node.

When SOAP request message is detected at Web service side, the recursive way is used to check if there is the node with the same node under the same node of SOAP message tree. In the detection process, there is the node id with the same node under user node, which means that there is the node with the same node under a node of SOAP message. And we can determine that the SOAP request message may cause XML injection attack. Web service determines that SOAP message is illegal according to detection results, and refuses to provide service for service request side sending SOAP request information.

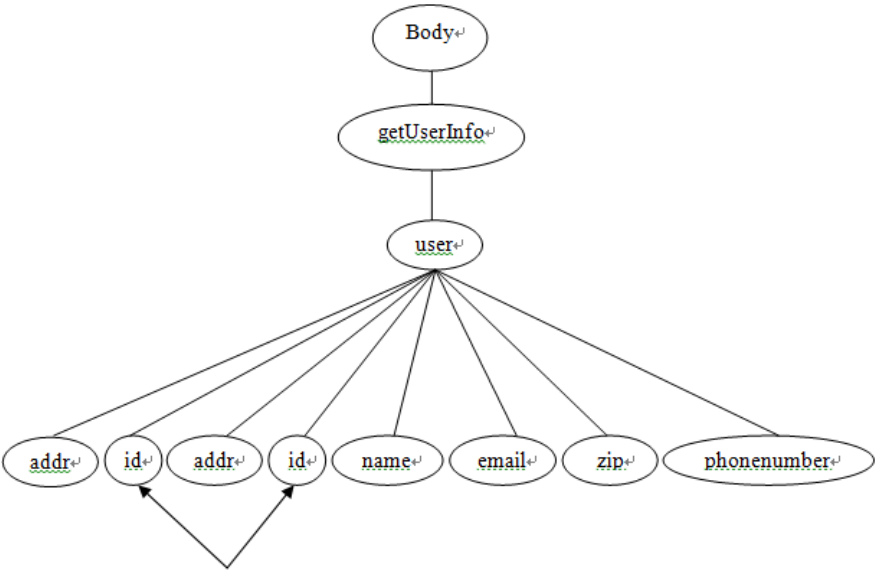


Fig. 11. Tampered SOAP message structure

6 Summary

On .net platform, the paper studies and detects structure-altering XML injection attack and parameter-error XML injection attack. The simulation experiment results indicate that the detection schemes proposed in the paper are effective. Web service security technology has wide application prospect. With the development of Web service, new specifications and relevant technologies emerge in endlessly. Applying new specifications and technologies to solve security problems and protect the security of Web service is a continuous research process.

References

1. Jin, J., Zhang, H., Liang, J.: Analysis on safety of Web service. Microellectronics and Computer 21(3) (2004)
2. Yang T., Liu J., Tan H.: Research on Web service security infrastructure. Computer Applications 26(6) (June 2006)
3. Zhang G., Song B., Wang P.: Security policies of network information based on SOAP. Journal of Nanjing University of Science and Technology 31(1) (February 2007)
4. Yu, X.: Study on Web Services security model based on SOAP. Microcomputer Applications 25(3) (2009)
5. Zhang, Y., Zhou, M., Yu, K.: Study on XML engine security model based on Web services. Computer Application Research 25(7), 2161–2163 (2008)
6. Vorobiev, A., Han, J.: Security Attack Ontology for Web Services. IEEE (2006)
7. Jensen, M., Gruschka, N., Herkenhoer, R., et al.: SOA and Web Services: New Technologies, New Standards-New Attacks. IEEE (2007)

Web Database Based on Data Mining

Wu Yang-bo

Xinyu University, Jiangxi 338000, China
wuyagbo3423@163.com

Abstract. In the vast data ocean, discovering and using the valuable information has become the key technology. The data mining is the powerful tool to solve this problem. In this paper, the commonly used data mining technology is introduced, and the current popular four Web database technologies are analyzed, and the data mining model that is suitable for comprehensive Web database is put forward finally. To sum up the above, it has certain theoretical research and practical application value.

Keywords: Mass Data, Data Mining, Web Database, Knowledge Discovery.

1 Introduction

In recent years, with the continuous development of computer network technology, and the rapid growth of the informationization level, the data and information grows with the explosively rapid speed. In the vast data ocean, people urgently need to improve the ability of data collection and information exploration. Finding the useful knowledge timely and improving the utilization ratio of information has become the key to related fields. Data mining technology can extract the potentially useful information and knowledge from the mass, incomplete, noisy, and fuzzy, random data. Data mining is a kind of process that data offers possible knowledge, but the database query is only the way to inquires the interesting questions or information from the user's point. In contrast, the data mining is more able to provide the general law of knowledge that the user didn't notice and easy to be ignored. In the first, this article tells the commonly used data mining technology, and then introduced the current popular four Web database technologies, after that, a kind of data mining model that is suitable for Web database is put forward by integration.

2 Data Mining Technology

Data mining is also known as knowledge discovery in database, which is a process that generates the effective, novel, useful, understand model from mass data [1].

2.1 Classification of Data Mining

According to the difference of analysis methods, the data mining can be divided into six aspects respectively [2]: classification, valuation, prophecy, relevance group or

Y. Yang, M. Ma, and B. Liu (Eds.): ICICA 2013, Part II, CCIS 392, pp. 76–84, 2013.

association rules, aggregation, description, and visualization. According to the processing mode, the six analysis method and can be divided into two categories: direct and indirect data mining data mining. The former builds a model with the available data, which describes the rest of the data with a specific variable. The typical methods includes classification, valuation, prophesy. The latter has not adopted a specific variable to describe the model, but set up a relationship in all of the variables, which includes correlation group or association rules, aggregation, description, and visualization.

2.2 Data Mining Methods

2.2.1 Neural Network Method

This method has good robustness, self organization adaptive, parallel processing, distributed storage and high fault tolerant characteristics, and is very suitable for solving the problem of data mining. So it has more and more attention in recent years. This method is mainly divided into three categories [3]. The first is the feed-forward neural network model which takes perceive machine, BP back propagation model and function type network as a representative and is used for classification, prediction and pattern recognition. The second is the feedback neural network model which takes the discrete model and continuous model of Hopfield as a representative and is used for associative memory and optimization calculation. The third is the self-organizing mapping method which takes ART model and Koholon model as a representative and is used for clustering. The defect of this method is the "black box", which is difficult to understand the network learning and decision making process for people.

2.2.2 Genetic Algorithm

This method is a random search algorithm that is based on the biological natural selection and genetic mechanism [4], which carries out global optimization based on bionics. It has implicit parallelism and easy to combine with other model, so it has a certain application in data mining.

2.2.3 Decision Tree

This method is used to forecast, which can find some valuable and potential information with the objective classification for mass data. Its main advantage is the simple description, classification speed, especially suitable for large-scale data processing. The most influence and the decision tree method is put forward by the famous Quinlan based on information entropy ID3 algorithm [5]. But, this method also defects, mainly displays in the following. Firstly, ID3 is incremental learning algorithm. Secondly, ID3 decision is a university decision tree and is difficult to express the complicated concepts. Thirdly, the relationship between the emphases is insufficient. Finally, the anti-noise property is very bad. Therefore, some researchers have improved the ID3 algorithm in a certain extent. Schlimmer and Fisher designed ID4 incremental learning algorithm [6]. Ming Zhong and Wenwei Chen have put forward the IBLE algorithm [7], etc.

2.2.4 Rough Set

This method is a mathematical tools to study the imprecise and uncertain knowledge [8]. It has the following advantages: no need to give additional information, simplify the input information expression space, simple, easy to operate. The processing object of this method is the information sheet that is similar to the two-dimensional relation table. Information table data mainly comes from mature relational database management system and data warehouse management system, form attribute is continuity. But the mathematical basis of rough set is set theory, which is difficult to deal directly with continuous attributes. So, the discrimination of continuous attributes is the difficulty that has restricted the application of rough set.

2.2.5 Including Positive Example and Rejecting Counter One

This method finds the rules by including all positive examples and rejecting all negative cases. Firstly, it selects a seed randomly in the positive cases, and compares it with all the elements in the negative set. Secondly, the field value selector compatibility is eliminated, on the contrary is retained. Finally, According to this thought cycle all positive cases of seeds, will get positive cases of rules (select the son conjunction). The typical algorithms include the AQ11 method of Michalski [9], the improved AQ15 method [10] and AE5 method of Jiarong Hone.

2.2.6 Statistic Analysis

There are two kinds of relationships among the fields of database: function and correlativity. The former can be expressed by function formula, and the latter couldn't be, but the correlative relation is existing in nature. The statistical methods can be used to analyze the two relations [11]. That is to say, the information in the database can be analyzed with the principles of statistics. The commonly used method includes: common statistics (calculating the maximum, minimum, sum, average of a large number of data, etc.), regression analysis (representing the quantitative relation with regression equation between variables), correlation analysis (measuring the degree of correlation between variables with a correlation coefficient), variance analysis (determining whether there is difference between the overall parameters from the sample statistics of the difference).

2.2.7 Fuzzy Sets

This method use fuzzy set theory for fuzzy evaluation, decision, pattern recognition and clustering analysis in practical problems [12]. The system of the higher the complexity, fuzziness is stronger, the general fuzzy set theory is used to depict the membership degree of fuzzy things under uncertainty. Deyi Li etc have proposed the qualitative and quantitative uncertainty conversion model - cloud model based on traditional fuzzy theory and probability statistics. The famous cloud theory is formed in this view.

3 Web Database Technologies

Web database is integrated with database technology and Web technology, which makes database become an important part of Web. Web database set the Web technology and

database technology advantage, make both qualitative changes happened. On the one hand, Web pages from static Web development become by the database driver of the dynamic Web pages. On the other hand, database to realize the development environment and application prospect of bad separation, the client can use unified browser realize cross-platform and multimedia services. The combination of them not only takes advantages of them, but also makes full use of a large number of database information resources in Web browser, which make users search and browse the contents of database easily. At present, Web database technology has become a hot spot in the study of database technology.

3.1 Three Development Stages of Web Database Technology

The first generation of Web database provides static access and static content application. The Web database manage and access static document, the programmer use HTML Web page to write according to the database content, users accessing the database is visiting the static HTML document in fact. The weakness of mode is insufficient real-time [13-15] When the inside database information is updating, the corresponding static file must be changed to keep the equality of the client information and the database information, which is lead to the mass workload in database maintenance.

The second generation of Web database provides static access and dynamic content application, which has realized the dynamic document management and access of database. The technology using CGI programming will database and Web server connected directly, which will be reflected in real time dynamic database information on the page. Among them, the CGI technology according to the contents of database and automatic update of the relevant departments of the static page, provide the end user. The user access is static HTML document, but the document content with database change and dynamic update. However, this technology will be washed out gradually because it can't keep database connection state, performance bottleneck, the lack of scalability and security condition, etc.

Based on the second generation of Web database, the third one also provides online transaction processing (OLTP) ability, which realizes the dynamic and personalized communication and interaction between client and server. In this mode, the simple CGI program evolved into a strong function of the database application server, it has both facing Web server interface, and the interface to the database server. When the Web server receives the Web client visit dynamic data content of the request, if need be and database connection, through the application server set up database server and the connection between the Web server, so that customers can access both Web database form dynamic pages, and can complete OLTP function (insert, update, and delete database, etc). This method carries out the all operation of database(add, delete, modify) through standard Internet browser interface, which is especially suitable for the end user and more adapt to the requirement of Internet technology development and network interconnection.

3.2 Four Key Web Database Technology

Currently, the commonly used Web database technology mainly includes four kinds: Perl, ASP, PHP, and JSP.

3.2.1 Perl

This is an early scripting language, which is used for the original Web application. Perl is similar to C, which are both flexible and convenient in file operation. However, the flexibility and redundant grammar of Perl results in the difficulty in reading and maintenance of Perl code, which seriously limits its application. In addition, Perl costs lots of CPU resources, and has low efficiency and maintenance.

3.2.2 ASP

This is the script language of Microsoft Windows IIS system itself, can carry out the dynamic Web service applications, and its grammar is very similar with Visual BASIC. It is the most simple in these four scripting language learn the development of language, its code easy, combined with the HTML code, can be quick to complete the application of the website. The upgrade version of ASP is ASPX, which provides the stronger function. The simplicity of the ASP has limited its function. The COM/DCOM technology provided by Microsoft can greatly expand the application range of the ASP, which makes ASP have the unlimited scalability. However, ASP couldn't support cross-platform because it is based on Windows operating system, and doesn't do well in the development and maintenance of large-scale project.

3.2.3 JSP

This is a dynamic web technology designed by Sun Corporation, which takes Jave as a script language. In the implementation methods, PHP and JSP, ASP has essential difference. PHP and ASP program execution is interpreted by a language engine code, but the JSP code is compiled into Servlet and executed by the Java virtual machine. Because of this operation is only on the JSP page requesting occurs for the first time, so it is generally agreed that the implementation of the JSP efficiency is higher than the PHP and ASP. JSP is a server-side script language, and the biggest advantage is to develop high efficiency. JSP Web based on MVC framework for development, the application is divided into three parts model, view and controller. Among them, the model is to point to the application's data, and the operation of the data; View refers to the user interface; Controller is responsible for the user interface and data synchronization between the programs. This development framework can weaken the coupling between each part, and can separate business logic, data processing and Web page. When the code of one module has changed, it doesn't affect the normal operation of the other modules. Therefore, many foreign large-scale enterprise systems and business systems all use MVC framework to support the highly complex large-scale Web application.

3.2.4 PHP

This is a scripting languages embedded in the HTML page. Its grammar is similar to the one of C and Perl, which enables developers to write fast dynamic pages by combined with the characteristics of PHP itself. This is completely free, so as to Apache and MySql. Thus, the dynamic website system can be quickly built by combining them. Generally speaking, the performance of APACHE+PHP+MYSQL is better than the one of IIS+ASP+ACCESS, and the latter also need to pay to Microsoft. PHP is very similar with Perl in the grammatical structure. But the number of functions of PHP is far more than the one of Perl. In addition, we must try to avoid the name conflict of modules, because PHP has no namespace. This innate defects makes PHP only is used for the small and medium-sized amateur web site.

4 Data Mining Technology of Web Database

The main processes of Web data mining include three stages: data pretreatment, pattern recognition, mode analysis [13]. As shown in figure 1 and figure 2.

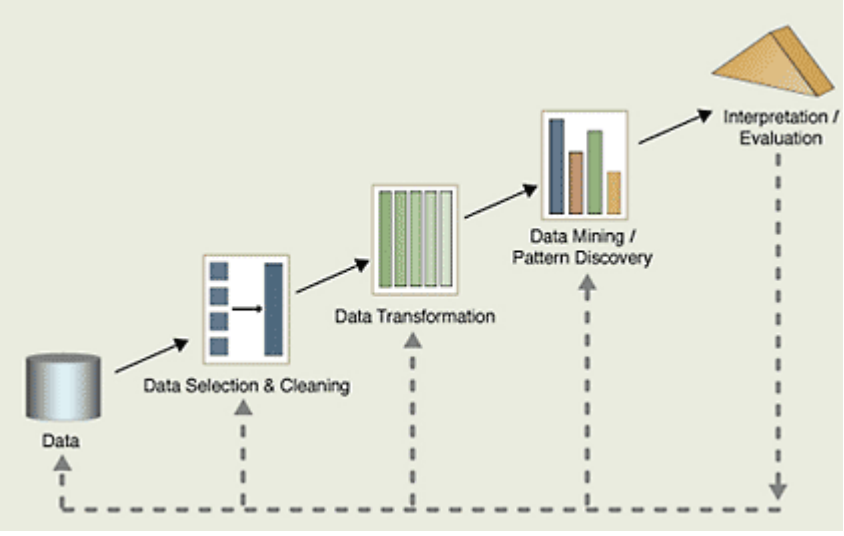


Fig. 1. Data mining structure

4.1 Data Preprocessing

The acquisitive data often has all kinds of problems because the real data is "dirty", which include incomplete (some interesting attributes are lack of value), noise (some data contains errors or abnormal), inconsistent (there are some differences in the code and name). Therefore, it is necessary to carry out data pretreatment, which can improve the knowledge quality of the data mining mode. The processes in this phase include data cleaning, user identification, session, and transaction recognition [9].

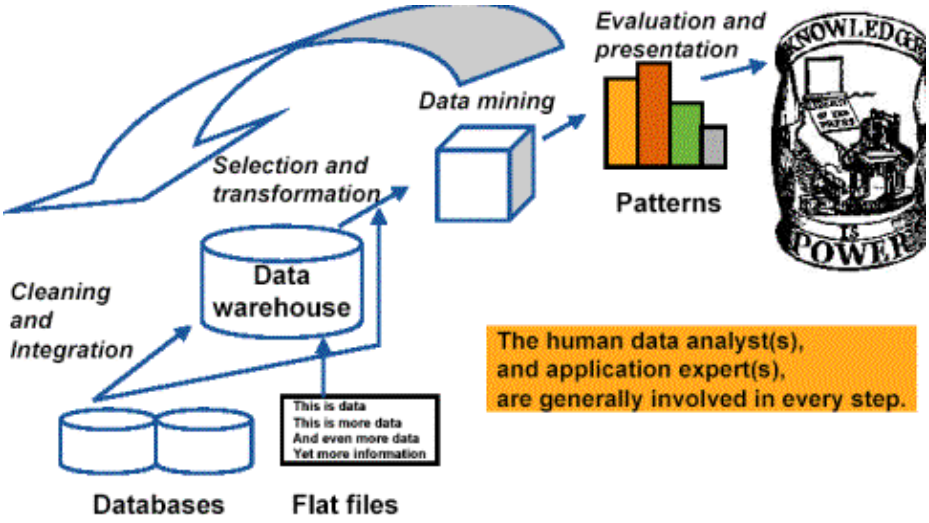


Fig. 2. Data mining with web datebases

4.1.1 Data Cleaning

The main tasks in this stage include: filling in the vacancy, smoothing noise data, identification, deleting the outlier, solving the inconsistency, etc.

4.1.2 User Identification

The main task of this stage is distinguishing each visit website users[15]. This work would become complicated because of some disguise facilities for user, such as local cache, firewall, and proxy server. In order to solve this problem, the user behavior may be tracked. Generally speaking, the most commonly used method is based on the log/site, also including some heuristic rules to identify a user. That is to say, it is considered to be the same user, if there is no evidence to express the difference between these users.

4.1.3 Session Identification

User session is a sequence of Web pages that are accessed by user, which express the sequent Web page request to server. User session recognition is divided the multiple access logs into some conversation. Currently, the commonly used user session method includes two kinds. One is expressing the conversation as the set of accessed Web pages, which is simple and intuitive. The other is expressing user conversation as the page sequence with the addition accessing time information, which is based on the former.

4.1.4 Transaction Identification

The user session is only object that has the natural affairs characteristics in Web data mining. But it is too thick for the task of association rules mining. So, it is necessary to divide the user session into the smaller one with the specific algorithm. The specific meaning of break up affairs is following. User clicks on the page sequence to obtain

meaningful information, namely the user session every time forward to the first page back to the previous page of path. Affairs recognition hypothesis is commonly used in business is the last page of the content page (also known as the biggest forward quote), before the page is auxiliary, the biggest to the reference path (MFP) defined affairs before. For each user session, taking the start page as a starting point, each maximum prior reference path is an issue.

4.2 Pattern Recognition

This stage uses the methods of statistics, machine learning to find the results, rules and modes. The algorithm can be a general statistical method, such as the visiting number of each page, the most frequent accessing page and the average browsing time of each page. It can also include other mining results, such as path analysis, statistical analysis, association rules and sequence pattern recognition.

4.3 Pattern Analysis

In this stage, the appropriate techniques and tools are used to carry out analysis, interpretation, and visualization for the user behavior model. Its target is filtering the useful model to improve the understanding level. The common used mode analysis methods include the query mechanism based on SQL(such as WebMiner system), on-line analytical processing (OLAP), and the visualization technology(such as Piktow people WebViz system, it can carry out visualization for mining model). In addition, the content and structure information also can used to filter out the specific patterns. Now, there are more than 30 kinds of commercial Web log analysis tools. The following information can be gotten through these tools: the user's lifestyle, hobbies, purchase frequency, user groups and their common features, visiting information of Web page, and advertisement clicks. But the analysis results mostly are about clicks, transmit data, access frequency, error rate, etc, which is very simple. So, there are many shortcomings in the comprehensive and depth of analysis, reasonableness and legality of the results, speed, flexibility, maintainability and inclusiveness.

5 Conclusion

With the rapid development of network information and the mass data, analyzing data and discovering information and knowledge has become the key point and the difficulty in the application of Web technology. Data mining is an important tool in analyzing the massive data and finding the efficient information. In this paper, the commonly used data mining technology and Web database technology are introduced, and a kind of data mining model that is suitable for Web database is proposed. This model is easy to operation and implementation, which has a certain theoretical research and practical application value. In the subsequent research, this data mining model must be applied and modified according to the actual case, which is useful to improve its scope and efficiency.

References

1. Han, J., Kamber, M.: Concepts and Technologies of Data Mining, vol. 13(4), pp. 75–79. Mechanical Industry Press, Beijing (2001)
2. Kantardzic, M.: Data mining-concept, model, method and algorithm, vol. 13(5), pp. 244–249. Tsinghua University Press, Beijing (2004)
3. Steinberg, D., Cart, C.P.L.: Tree-structured nonparametric data analysis. Salford Systems 6(7), 46–52 (1995)
4. Lv, A., Lin, Z., Li, C.: Technique Methods of Data Mining and KDD. Science of Surveying and Mapping 25(4), 36–39 (2000)
5. Kantarcioglu, M., Clifton, C.: Privacy-Preserving distributed mining of association rules on horizontally partitioned data. IEEE Trans. on Knowledge and Data Engineering 16(9), 1026–1037 (2004)
6. Warmer, S.L.: Randomized response: A survey technique for eliminating evasive answer bias. Journal of the American Statistical Association 60(309), 63–69 (1965)
7. Evfimievski, A.: Randomization in privacy preserving data mining. ACM SIGKDD Explorations Newsletter 4(2), 43–48 (2002)
8. Vaidya, J., Clifton, C.: Privacy preserving association rule mining in vertically partitioned data. In: Proceedings of the Eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, vol. 45(4), pp. 57–62. ACM Press, New York (2002)
9. Yen, S.J., Chen, A.L.P.: A graph-based approach for discovering various types of association rules. IEEE Transactions on Knowledge and Data Engineering 13(5), 839–845 (2001)
10. Breslow, L.A., Aha, D.W.: Simplifying decision trees: A survey. The Knowledge Engineering Review 12(1), 31–40 (1997)
11. Shang, W.: Statistical Analysis Frame of Complex Data Mining and MART & MCMC Applications in CRM. Sichuan University, Chongqing, Sichuan 34(5), 56–62 (2004)
12. Zhang, H.: Data Mining Study Based on Rough Sets and Fuzzy Neural Network. Tianjin University, Tianjin 8(2), 75–83 (2006)
13. Wang, W.: Design and Realization of Long-Distance Education Based on Web. China University of Geosciences, Beijing 63(4), 766–771 (2008)
14. Zhang, S.: The system design and data mining of education website based on Web technology. Shandong University, Jinan, Shandong 54(8), 65–69 (2005)
15. Stephens, R., Plew, R.R., He, Y., et al.: Design of database, vol. 55(6) pp. 3–14. Mechanical Industry Press, Beijing (2001)

Training Data of City Tunnel Traffic Situation Awareness

Lin Li[1], Weiguo Wu[2], and Luo Zhong[1]

[1] School of Computer Science and Technology, Wuhan University of Technology
{cathylilin,zhongluo}@whut.edu.cn
[2] School of of Transportation Wuhan University of Technology
mailjt@163.com

Abstract. The network of city Tunnel monitoring systems has accumulated a great deal of multi-source heterogeneous monitoring data, which mainly consists of video data, traffic data, environmental data, sensor data and so on. This paper discusses how to use the monitoring data to realize city tunnel traffic situation awareness. First, we need to mine a set of strong associated characteristic groups according to the association rules, and then determine the corresponding traffic awareness for each characteristic group to build our training data set for machine learning methods. Finally we can aware the situation of city tunnel traffic by using machine learning methods. We solve the problem of how to build the training data in the background of complex multi-source data, which is a prerequisite for many machine learning methods.

Keywords: SVM, HMM, Training data, City tunnel, Traffic situation awareness.

1 Introduction

The network of city tunnel monitoring systems has accumulated a large amount of monitoring data, such as environmental data, traffic data, video data and so on. In such a complex environment of multi-source data, how to realize traffic situation awareness is an important research problem. The technology of situation awareness is that taking the system security state as continuous variation, we can do their state assessment and trend analysis, so as to eliminate security risks and protect system security. Traffic situation awareness is derived from the studying on situation awareness, Endsley[1] defined the situation awareness as following: under the certain conditions of time and space, it does the job of getting and understanding the environmental factors, then makes the forecast of the future states, as shown in Fig 1.

Situation awareness (SA) is increasingly becoming a popular research topic. In a dynamic and complex environment, decision makers need the aid of situation awareness tools to display the current continuous changes of environment to accurately make decisions. Hui Wang [2] and Stanton [3] discussed the relationship between situation awareness and security. Endsley et al. [4] analyzed the measurement method of situation awareness capability in air traffic control. Endsley [5] studied the influence factors of operating errors in the traffic control process of pipeline air. Fang Lanet al.[6] analyzed the problems in the network security situation awareness.

Y. Yang, M. Ma, and B. Liu (Eds.): ICICA 2013, Part II, CCIS 392, pp. 85–94, 2013.

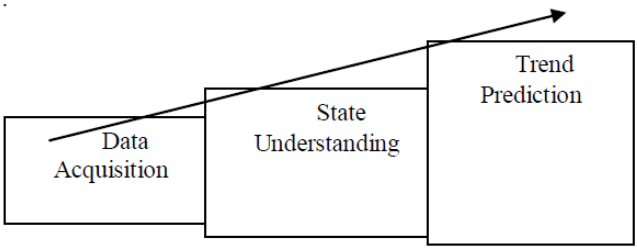


Fig. 1. Three levels of situation awareness

We can see that scholars in the world have made a deep research for the theory of situation awareness, and focus its study on situation awareness in the practical application of aviation, military, network security. However, there are few researches about situation awareness in the field of city tunnel traffic. In addition, machine learning is one of the effective solutions. However, a key issue of machine learning is that we need training samples, especially in the large amount of data that from complex cases. Therefore, completely relying on human judgment to build training samples is more difficult. Our approach is effective to solve this problem.

2 The Brief Introduction of Our Traffic Monitor System

Our traffic monitor system consists of several sub-systems including video, broadcast, traffic, environment, ventilation, lighting, draining and so forth, as shown in Figure 2. The sub-systems can collect various kinds of data to monitor the traffic situation of a city tunnel and while they send control orders to equipments for reacting to the current situation. The types of data are a lot and their magnitudes are totally different. To put them into a same scale, in the next section we will introduce a method to transfer different data to discrete magnitudes (areas).

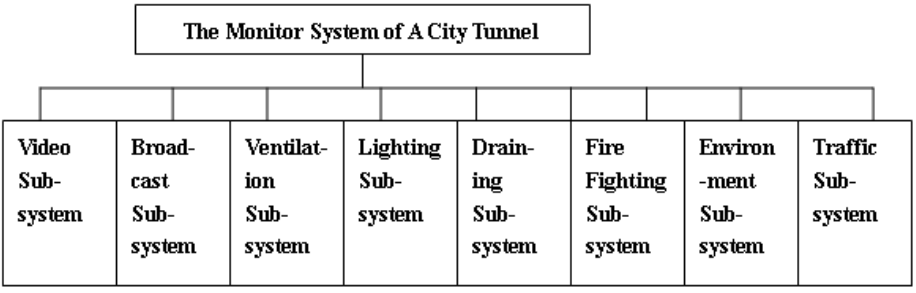


Fig. 2. The eight sub-systems of our traffic monitor system

3 Building Training Data Sets

The building of training data sets id divided into three steps: data preprocessing, strongly correlated combination mining and situation judgment. The specific process is shown in the following. In the process, we need to note that the selection criteria of data discretization and the setting of critical threshold in Apriori algorithm are based on experience and historical data to determine the threshold value [7].

3.1 Data Preprocessing

Based on the predetermined thresholds, the monitoring data is converted from numeric to its belonging area. Different thresholds and the number of area are selected according to different types of data. Then continuous values are converted to discrete magnitudes (areas), so that the next step for association analysis of discrete magnitude is to identify strongly correlated.

3.2 Strongly Correlated Combination Mining

Apriori algorithm for mining strongly correlated combinations mainly consists of two stages. First stage is that identifying all large itemsets from the data set. Large frequency means that a frequency relative to all of the records in a itemset must reach a certain level. The frequency of an itemset is called Support. If the support value is greater than or equal to the threshold value of Minimum Support, then we can call {A,B} as a large itemset. A k-itemset that meets minimal support can be called Frequent k-itemset, generally expressed as Large k or Frequent k.

Apriori algorithm thinks that the large k+1 is produced by large k in a itemset, until you can no longer find more large itemsets' group. In the second stage, a association rules is produced by these large itemsets. These rules are generated by making use of the previous step rules that large k-itemset, which is under the minimum confidence.

Our experimental data come from one urban tunnel of Wuhan and its collected time is 2012-01. The total records are 88388, and the min sup is set to be 2500. Some data are listed as follow.

```
CLL_H;   26538
CLL_L;   36300
CLL_M;   25550
……
CLL_H;CS_H;   3954
CLL_H;CS_L;   13897
CLL_H;CS_M;   8687
CLL_L;CS_H;   17244
CLL_L;CS_L;   7134
……
CO_H;CLL_H;CS_L;   3013
CO_H;FS_H;CLL_H;   3602
CO_H;FS_H;GQ_H;   2520
```

CO_H;GQ_H;CLL_H; 2803
CO_H;VI_M;CLL_H; 4809
CO_H;VI_M;CS_L; 3453
CO_H;VI_M;FS_H; 4161
……
CO_H;VI_M;CLL_H;CS_L; 2778
CO_H;VI_M;FS_H;CLL_H; 3327
CO_H;VI_M;GQ_H;CLL_H; 2752
CO_L;FS_H;GQ_M;CLL_M; 2832
CO_L;FS_L;CLL_L;CS_H; 13041
……
CO_L;FS_L;GQ_L;CLL_L;CS_H; 2979
CO_L;FS_L;GQ_M;CLL_L;CS_H; 9835
CO_L;FS_L;GQ_M;CLL_L;CS_L; 3783
CO_L;FS_L;GQ_M;CLL_L;CS_M; 6539
CO_L;VI_M;FS_H;GQ_M;CLL_M; 2592
CO_L;VI_M;FS_L;CLL_L;CS_H; 12658
……
CO_L;VI_M;FS_L;GQ_L;CLL_L;CS_H; 2971
CO_L;VI_M;FS_L;GQ_M;CLL_L;CS_H; 9461
CO_L;VI_M;FS_L;GQ_M;CLL_L;CS_L; 3533
CO_L;VI_M;FS_L;GQ_M;CLL_L;CS_M; 6196
CO_M;VI_M;FS_H;GQ_H;CLL_H;CS_L; 2773
CO_M;VI_M;FS_H;GQ_M;CLL_H;CS_L; 2566

From frequent item set, we can get the association rules of different data items. For example, if we want to know the association rules between CO and other data, Equation 1 and Equation 2 are used to compute the values of Support and Confidence.

$$\text{Support}(X \Rightarrow Y) = \frac{\text{Sup}(X \cup Y)}{|D|} \qquad \text{Equation (1)}$$

$$\text{Confidence}(X \Rightarrow Y) = \frac{\text{Sup}(X \cup Y)}{\text{Sup}(X)} \qquad \text{Equation (2)}$$

The computation steps are as follows:
Sup (CO_L^VI_M^FS_L^GQ_M^CLL_L^CS_H) = 9461;
Sup (VI_M^FS_L^GQ_M^CLL_L^CS_H) = 9567;
|D| = 88388;
support=9461/88388=10.7% confidence=9461/9567=98.9%

We can get a conclusion of the association rule of CO_L:
VI_M^FS_L^GQ_M^CLL_L^CS_H=>CO_L[support=10.7%, confidence=98.9%]

This rule tell us that the VI_M^FS_L^GQ_M^CLL_L^CS_H has a strong association with CO_ L because the computed confidence is 98.9%. Following the same computing steps, we can get a set of association rules listed as follows:

VI_M^FS_L^GQ_L^CLL_L^CS_H => CO_L [support=3.4%, confidence=100%]
VI_M^FS_L^GQ_M^CLL_L^CS_L => CO_L [support=4.0%, confidence=98.4%]
VI_M^FS_L^GQ_M^CLL_L^CS_M=>CO_L [support=7.0%, confidence=98.7%]

The results of association rule mining can be used to feature selection. If one feature is kept unchanging among all the rules, we do not need to select it into training data. If one feature is always appeared with another feature and its confidence is high, we can select one of them into our training data. Here association rule mining is used as feature selection or dimension reduction.

In our experiment, we use CO concentration data, wind speed data, light intensity data, humidity data, temperature data, traffic volume data, and vehicle speed data. We get CO concentration data, wind speed data, light intensity data, humidity data, and temperature data from environmental monitoring system, environmental monitoring system is mainly used to detect and control the tunnel environment. And we indirectly get traffic volume data and vehicle speed data from traffic control subsystem. Traffic control subsystem is mainly used to measure the total traffic volume and vehicle speed.

3.3 Situation Judgment

We need to find strongly correlated combination according to the mass-monitoring data, and then manually make a judgment for tunnel traffic situation. Traffic situation is divided into four categories: free, moving, congestion and heavy jam.

4 Realizing the Traffic Situation Awareness Based on Machine Learning

In this paper, we adopt machine learning methods and conduct state training through historical data, then realize state recognition. State identification will compare two methods of Markov chain and support vector machines. We make an analysis for applying them to traffic situation awareness, and then give the corresponding process.

4.1 Data Processing

(1) The experimental data are measured data of a tunnel, selecting part of the data to create the SVM model, and some of the data to predict.

(2) There are seven feature items, including CO concentration, wind speed, light intensity, humidity, temperature, traffic volume, and vehicle speed.

(3) Taking into account the different time, the range of these feature items are different, we establish the spring, summer, autumn, winter four training samples based on time and a test sample.

The part data of training sample is seen in Table 1.

Table 1. Part of training samples

CO Concentration	Wind Speed	Light Intensity	Humidity	Temperature	Traffic Volume	Vehicle Speed
0.72	1.54	10	60.42	11.5	1	78
0.72	2.32	10	60.5	11.5	3	78
0.72	1.71	10	60.62	11.5	1	77
0.72	1.54	10	62.42	10.78	10	77
0.72	1.34	10	62.58	10.82	11	78
0.72	1.19	10	62.62	10.82	16	77
0.72	1.36	10	62.58	10.82	15	78
0.72	2.07	10	62.82	10.86	21	77
0.72	1.66	10	62.9	10.86	16	78
0.72	1.15	10	62.9	10.86	14	77
0.72	1.32	10	62.9	10.86	19	77
1.08	2.89	10	53.29	10.78	13	78

The part data of test sample is seen in Table 2.

Table 2. Part of test samples

CO Concentration	Wind Speed	Light Intensity	Humidity	Temperature	Traffic Volume	Vehicle Speed
1.08	3.53	120.09	49.6	29.91	53	58
0.72	3.18	110.08	47.32	30.4	54	58
0.72	2.35	130.1	50.37	29.87	43	61
1.08	2.69	130.1	50.29	29.87	54	61
1.08	2.54	130.1	50.29	30	51	59
1.08	3.09	50.04	45.08	14.86	10	77
1.08	2.59	50.04	45	14.78	10	77
1.08	2.4	50.04	45	14.78	6	60
1.08	3.26	60.04	45	14.86	15	78
1.08	2.11	60.04	44.96	14.78	9	81
1.08	2.74	50.04	44.92	14.82	8	72
1.08	4.05	50.04	45	14.86	14	113

4.2 Hidden Markov Chain (HMM)

If the "future" of a process only depends on "now" not "the past", we can say that this process is Markov style, called as a Markov process. If both time and state are discrete, the Markov process that can be called Markov chains. Hidden Markov Chain (HMM) was based on the Markov chain [11]. HMM includes dual stochastic processes. One is Markov chain which describes Invisible state transition. The other is random process that describes the probability relationship both state and observation values.

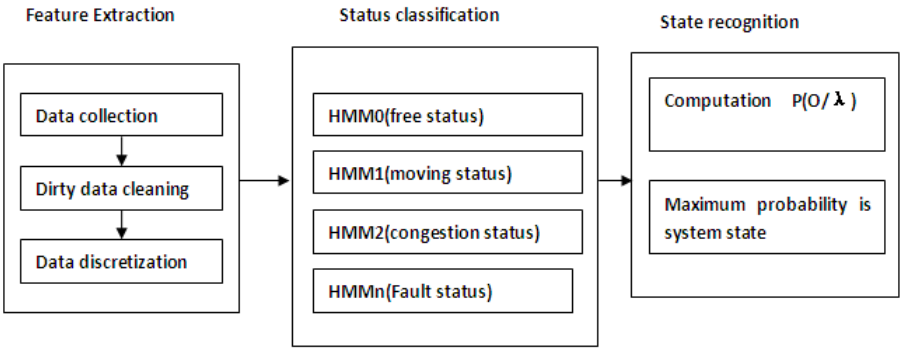


Fig. 3. The tunnel traffic situation awareness based on HMM

Tunnel traffic situational awareness based on hidden Markov consists of two parts, the state training and state identification. State training system is collecting output signal under certain state, then through data merging, dirty data cleaning, data discretization, HMM training, to get corresponding HMM. In order to achieve State recognition, state identification uses the feature extraction and state classification, as shown in Figure 3. Feature extraction similarly has these three steps: data merging, dirty data cleaning, data discretization. State classification is letting extracted features for a probabilistic reasoning, which utilizes HMM trained in various security status, so as to get the probability value of HMM corresponding various state. The state that the maximum probability value of HMM is that the state of system.

4.3 Support Vector Machine (SVM)

Support Vector Machine (SVM) is a supervised learning method, and can be used for data classification and regression [8-10]. For linear-inseparable or non-linear classification, we can introduce kernel functions to solve the problems. The idea is that through a kernel function, the space variables of non-linear classification are converted into the space of linear-separable. Finally we separate them by hyperplane.

SVM classification can classify system states according to system input, such as free, moving, congestion and heavily jam, which is called situation evaluation. SVM regression can make a prediction according to system inputs, and then use SVM classification to determine future system states, which is called trend prediction. Therefore support vector machines can be used to implement tunnel traffic situation awareness. Its details are shown in Figure 4.

The mechanism of SVM can be simply described as: Looking for an optimal hyperplane that meets the classification requirements, such that the hyperplane while ensuring the classification accuracy, it is possible to maximize the hyperplane on both sides of the blank area. In theory, support vector machine can achieve the optimal

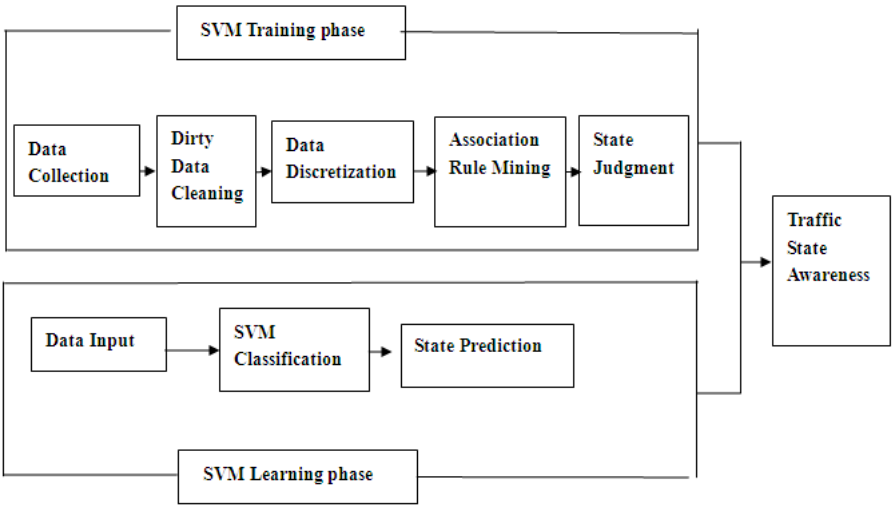


Fig. 4. The tunnel traffic situation awareness based on SVM

classification of linearly separable data. In order to solve the nonlinear problem, Vapnik et al introduce the kernel mapping method to transform the nonlinear problems into higher dimensional linear space separable problems to solve.

The schematic diagram of support vector machine is shown as Figure 5. Choose a different kernel function K (Xi, X) can generate different support vector machine.

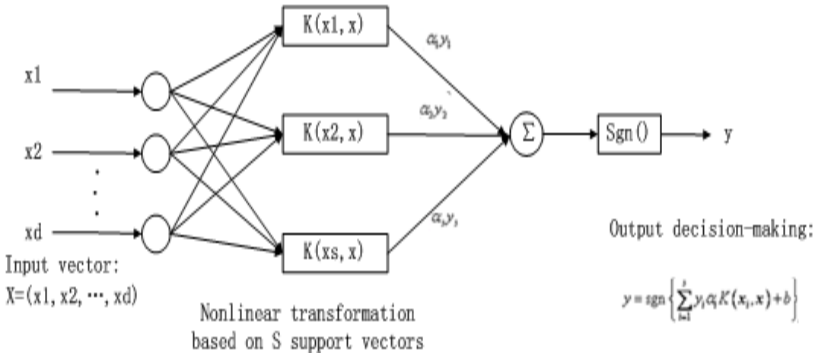


Fig. 5. The schematic diagram of support vector machine

SVM is essentially a binary classification, however, in practical applications, such as speech recognition, font recognition, face recognition, and etc. is using SVM to solve the multi-classification problems. For multi-classification recognition problem, SVM must be improved and generalized. There are several common multi-classification algorithms [12], in our experiment we use Multi-class SVM classification method based on the binary tree, for K-class training samples, training K-1 support vector machines. The schematic diagram of Multi-class SVM classification method based on binary tree is shown as Figure 6.

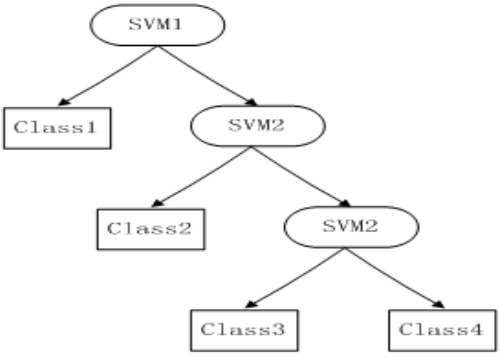


Fig. 6. Multi-class SVM classification method based on binary tree

5 Conclusions

This paper has deepen the algorithm application of data mining and machine learning, then do the job of mine city tunnel monitoring data and implement traffic situation awareness, finally provide the decision support for city information services. The feature of this paper is that we presented the method of using association rules to establish training data sets and solved the problems of practical applications which have a large number of multi-source data for machine learning under a complex environment. At last we give out the detailed processes of two machine learning methods, i.e., SVM and HMM.

Acknowledgements. This research was undertaken as part of Project 61003130 funded by National Natural Science Foundation of China and Project 2011CDB254 founded by Ministry of Education of China.

References

1. Endsley, M.R.: Design and evaluation for situation awareness enhancement. In: Human Factors Society 32nd Annual Meeting, Santa Monica, CA, pp. 97–101 (1988)
2. Wang, H., et al.: A review on Situation awareness of Internet. Computer Science 33(10), 5–10 (2006)
3. Stanton, N.A., Chambers, P.R.G., Piggott, J.: Situational awareness and safety. Safety Science (39), 189–204 (2001)
4. Endsley, M.R., Sollenberger, R., Stein, E.: Situation awareness: A comparison of measures. In: Proceedings of the Human Performance, Situation Awareness and Automation: User-Centered Design for the New Millennium (2000)
5. Endsley, M.R., Rodgers, M.D.: Distribution of attention, situation awareness, and workload in a passive air traffic control task: Implications for operational errors and automation. Air Traffic Control Quarterly 6(1), 21–44 (1998)
6. Lan, F., Chunlei, W., Guoqing, M.: A Framework for Network Security Situation Awareness Based on Knowledge Discovery. In: International Conference on Computer Engineering and Technology (ICCET), pp. 226–231 (2010)

7. Agrawal, R., Imieliński, T., Swami, A.: Mining association rules between sets of items in large databases. In: Proceedings of the 1993 ACM SIGMOD International Conference on Management of Data, pp. 207–216 (1993)
8. Gu, S., Guo, Y.: Learning SVM Classifiers with Indefinite Kernels. AAAI (2012)
9. Wassantachat, T., Li, Z., Chen, J., Wang, Y., Tan, E.: Traffic Density Estimation with On-line SVM Classifier. In: AVSS, pp.13–18 (2009)
10. Kumaraswamy, R., Prabhu, L.V., Suchithra, K., Pai, P.S.S.: SVM Based Classification of Traffic Signs for Realtime Embedded Platform. In: ACC, pp. 339–348 (2011)
11. Xie, Y., Tang, S., Huang, X.: A new model for generating burst traffic based on hierarchical HMM. In: FSKD, pp. 2212–2216 (2011)
12. Fang, H.: Support vector machine multi-classification algorithm 2009(4) fragments based on the SVM algorithm. In: BMEI 2011, pp. 1738–1742 (2011)

Arbitrary Waveform Generator and Total Distortion Evaluation

Jingyu Sun[1,2] and Puzhong Ouyang[2]

[1] BeiHang University, Beijing, 100191 China
[2] Changcheng Institute of Metrology & Measurement, Beijing, 100095 China

Abstract. The paper design a good performance arbitrary waveform generator with AD's digital-analog chip AD760 as the core. Thesis mainly research hardware and software design of computer-controlled arbitrary waveform generator for the corresponding. Hardware part of the arbitrary waveform generator take Cygnal's high-speed micro-controller C8051F310 micro-controller as the controller and AD760 circuit components of the design. We carried out anti-jamming design when hardware design, Including filtering, spare part handling of input pin, grounded, the power handling. The paper also describes a total distortion of periodic arbitrary waveform definition and evaluation process and methods, the energy is calculated using the time domain distortion, and the influence of the measurement system itself is compensated. This method can be used for arbitrary waveform generator arbitrary waveform distortion for accurate measurement and measurement calibration.

Keywords: Arbitrary Waveform Generator, AD760, direct digital synthesis, Distortion.

1 Introduction

Arbitrary waveform generator is capable of generating a large number of standard signals and user-defined signal, and to ensure accuracy, high stability, repeatability and ease of electronic instruments [1-4]. Arbitrary waveform generator having a continuous phase transformation, and frequency stability, etc., can simulate all kinds of complex signals, but also on the frequency, amplitude, phase shift, the waveform dynamic and timely control of, and to work with other instruments communications, an automatic test system, it is widely used in automatic control systems, vibration excitation, communications and instrumentation fields [5, 6].

Currently based arbitrary waveform generator is a direct digital synthesis technology that makes the memory lookup table stored in digital form through a waveform by digital / analog converter generates the required arbitrary waveforms [7-10]. In recent years, the international arbitrary waveform generator technology is mainly reflected in the following aspects:

1) Past, because of the low frequency range is relatively narrow application, the output waveform frequency increases, making the arbitrary waveform generator can be applied more widely in the field. Arbitrary waveform generator software development is making arbitrary waveform input more convenient and easy. Arbitrary waveform

Y. Yang, M. Ma, and B. Liu (Eds.): ICICA 2013, Part II, CCIS 392, pp. 95–104, 2013.

generator usually allows a series of point, line, and a function of fixing the waveform data in the memory segment. While using a very powerful input mathematical equations, complex waveforms can consist of several relatively simple formula compound as v = f (t) in the form of the wave equation mathematical expression yields. The rapid development of a variety of computer languages also contributed to an arbitrary waveform generator software technology. Can now take advantage of visual programming languages (such as Visual Basic, Visual C, etc.) to write arbitrary waveform generator soft front panel, which allows freehand input from a computer display arbitrary waveform to achieve the input waveform.

2) Combined with VXI resources. Currently, arbitrary waveform generator consisting primarily of independent desktop machines and cards for personal computers as well as newly developed VXI modules. Because VXI bus matures and the high demand for measuring instruments, used in many areas that need to measure and VXI systems generate complex waveforms, VXI system resources provide obvious advantages, but because of a long development cycle VXI modules and require specialized VXI chassis supporting the use of such arbitrary waveform generator VXI module is limited to aviation, military and defense and other large areas. In the civil context, VXI module is far more convenient than desktop machines.

The task of this paper is design a simple arbitrary waveform generator and application characteristics to use and combined with a new generation of high-performance chips, excellent performance, arbitrary waveform generator, arbitrary waveform generator that can produce sine, square, triangle wave other commonly used standard signal, but also according to the user's need to generate arbitrary waveforms, the entire system using micro-controller.

2 Direct Digital Synthesis Based on Phase Accumulator

Frame structure of direct digital synthesis based on phase accumulator is shown in Figure 1. This structure mainly composed of phase accumulator, data ROM, D / A converter component, which is a kind of new direct digital synthesis method.

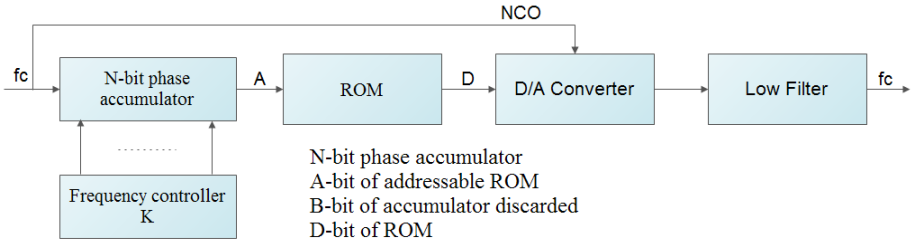


Fig. 1. Block diagram of a direct digital synthesis based on the phase accumulator

Working principle: the generated waveform data to be stored in the waveform memory, and the role of the reference clock, the frequency of the input data is accumulated, and the accumulator output waveform generator as part of the read address, the read-out waveform data by D / A converted to the corresponding voltage

signal, D/A converter output signal by a series of stepped voltage after the output of the low-pass filter is a smooth waveform signal [11-14]. Synthetic sine wave, for example, usually when we consider the habit of using a sine wave, sine wave amplitude - time expression is:

$$S(t) = A\sin(ax + \phi)$$

The amplitude of the sine function is nonlinear which generate any frequency sine wave amplitude is very difficult. However, we note sine phase is linear, DDS technology is the key to make full use of the sine wave phase linear variation of this feature, the DDS chip species, the core component is the phase accumulation and SIN function table.

3 Hardware Design of Arbitrary Waveform Generator

3.1 Hardware Design

Shown in Figure 2. Arbitrary waveform generator connect with the PC via RS232 interface,which is connected to the receiver that sent from the PC arbitrary waveform data. Wave form generator is divided into upper and lower hardware and application software of two parts [15-19]. In which the upper layer software provides interactive interface that Operator Console, select the waveform used to generate the waveform data, and serial communication control and man-machine communication; underlying hardware I / O controller, the CPU, DAC, counters, timers other modules (expansion module can be considered), the upper waveform is mainly used for data reception, storage, and can be used alone, the CPU required for transmitting waveform data to the DAC.

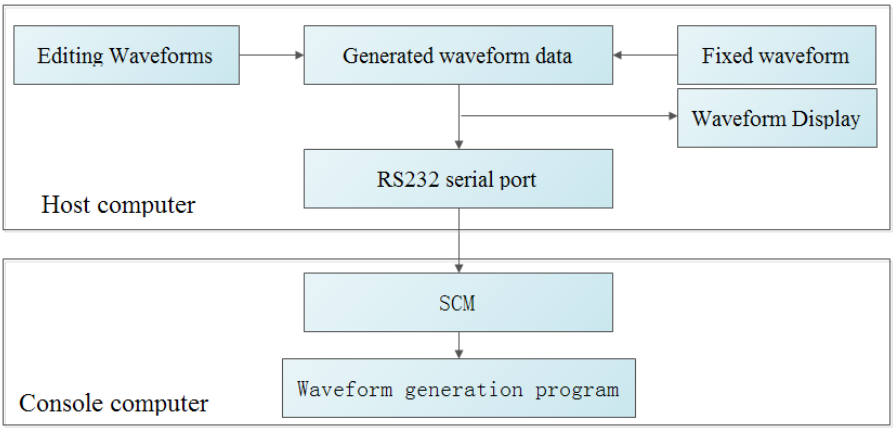


Fig. 2. Arbitrary waveform generator function

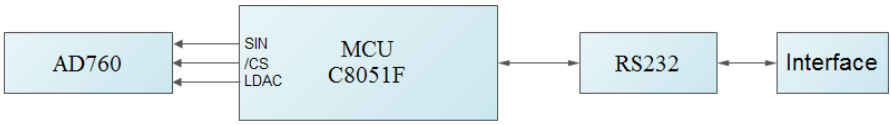


Fig. 3. System circuit diagram

3.2 The Flow Chart for Software

System control process, including the contents of two parts, part of the boot or the system itself after a power-on reset initialization, the other part is the interrupt respond to the host PC, the user convert the input information, and transmits control information to the AD760, that enable AD760 to work, the user needs to generate a waveform and output. Initialization part of the contents includes the following elements: the system self-test, the main source for the system clock and is used to store waveform data of the FLASH block [20-25]. And initialize the serial interface to respond to the host PC. Shown in Figure 4.

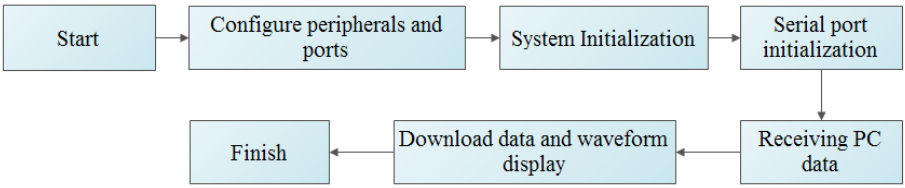


Fig. 4. Flow control of system

4 Error Analysis

DDS mathematical model can be attributed to [26-30]: At each clock period Tc, the frequency control code K and the N-bit phase accumulator cumulative time, and at the same time 2N modulo operation, the phase value obtained in binary code form of the RAM search address, and outputs the phase value corresponding to the number of amplitude values, and then completed by the DAC digital-analog converter discrete signal changes to a continuous signal, the final low-pass filter output signal can be obtained.

4.1 The Phase Noise of DDS

DDS is mainly determined by the phase noise of the reference clock source noise. DDS can be considered a division ratio of 2N / K divider, set the system clock phase noise of LC (fm), then the output signal of the phase noise is:

$$L_0(f_m) = L_c(f_m) - 20\lg(2^N / K)$$

4.2 The Spurious Analysis of DDS

In practice,we often take N = 32 or 48. If N bits are all used to address the RAM, requires a tremendous amount of storage, which is unrealistic. Thus, in the ROM can only be used when addressing a high W bit phase accumulator to address, which is set to introduce the phase error. Figure 5: Figure is a model DDS spurious sources.

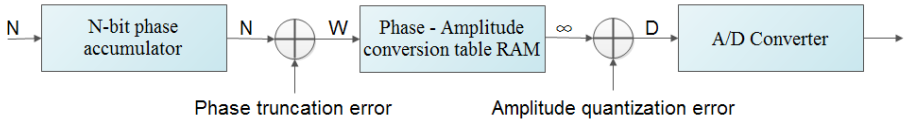


Fig. 5. DDS Model of spurious interference

Deserted phase error：

$$\xi_p(n) = F_r * n - 2^B * \mathrm{int}(F_r / 2^B * n)$$

Above formula int (x) expressed do not greater than x rounding operation. Visible, rounding B is reduced by a stray can be improved about 6db. Rounding fewer stray magnitude smaller; Conversely, the greater.

4.3 Quantization Error Introduced Spurious of ROM

SIN ROM table stores the value is not an exact value. At this time, 2L / (2L, Fr) for the period is a series of $\varepsilon_M(n)$：

$$\varepsilon_M(n) = \sin(2\pi * F_r * n / 2^L) - (1/2^D)\mathrm{int}\left[2^D \sin(2\pi 2^B / 2^L * \mathrm{int}\left[F_r n / 2^B\right])\right]$$

The signal processing knowledge shows εM (n) can be approximated in a cycle as zero mean and variance of the Gaussian white noise 2-2D/12.

4.4 Effect of DAC Devices on the Spectrum

In addition to the limited resolution DAC usually considered outside the median, DAC moment glitch, DAC nonlinearity of digital noise and clock feedthrough leakage exacerbating factors all contribute to the spectrum, which is the DDS output spectrum increased background noise and spurious. In many applications, DAC output spectral purity of major consideration indicators. Even an ideal N-bit DAC, the system should generate harmonics. Because the amplitude of these harmonics is strongly dependent on the output frequency of the frequency of the clock rate, it is because the DAC quantization noise spectral components, with this ratio changes.

5 The Results Analysis by Actual Test

Analog supply is 15v, DC digital power is 5v. Also, to make AD760 temperature stabilized measurements after 10 minutes. Test data: ad760 maximum output signal 9.9867v; minimum signal 0.18mv; intermediate signal 4.9633v; comes with the reference voltage is 9.9881v. Calculated output signal amplitude waveform precision can reach 12. Second test, and the first data obtained in a slight difference in the analysis may cause the reduction in power consumption and battery are. Noise measured value: the value of which has been given by subtracting the initial value of zeroing. Negative power 0:0.21-0.22 mv. The positive power supply 0:0.24 mv. AD760 25-foot (reference voltage input) is: 0.21mv-0.22mv. AD760 point 23 feet (signal output) is 0.25mv. Experimental data show that the output noise precision of less than 0.003%. A plurality of time periods after the test, the measured data and the previous one a slight difference, but as time increases, the change is relatively fast and analysis because the system is directly related to temperature rise, as a great influence on the accuracy of the signal degree from the DAC device temperature drift; either because of the reduced analog supply electricity. Figure 6 is one of several in the debugging process using an oscilloscope measured waveform.

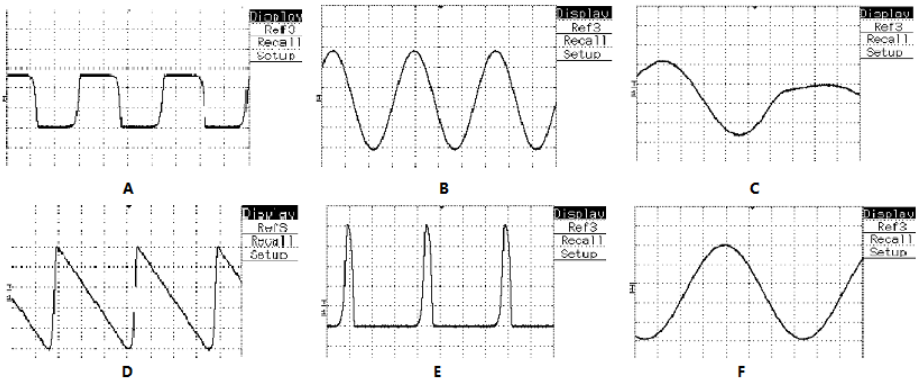


Fig. 6. A. square wave; B. wave; C. damping function; D. triangle wave; pulse E.; F. Gauss function

6 Conclusion

Arbitrary waveform generator is capable of generating a large number of standard signals and user-defined signal, and to ensure accuracy, high stability, repeatability and ease of electronic instruments. So it has broad application prospects.

With modern measurement technology of exploration and development, arbitrary waveform generator signal is generated as a common source, in the field of communication, satellite television, aerospace and transportation simulation test and measurement instrumentation, and other areas, is playing an increasingly important role. In some high-end and new arbitrary waveform generator, DDS technology

because of its unique advantages has been widely used. DDS, also known as direct frequency synthesis technology, is a common modern arbitrary waveform generator design technology. Direct frequency synthesis to precede to the phase, by way of frequency synthesis, the user directly generates the desired signal. It is the first signal to generate sample data temporarily stored in the signal storage device, and then follow the step frequency synthesis, storage structures addressing operation to produce a digital signal, so you can easily store a signal by changing the sample data to output the different types of signals. The phase and frequency synthesis step adder size is controlled to achieve precise phase synthesis, in order to achieve the purpose of the waveform adjustment control, therefore, the use of arbitrary waveform generator DDS technology has a very high frequency resolution, fast the conversion rate and the conversion can be done when the continuous phase. DDS technology -based arbitrary waveform generator, a signal is generated as a modern source, not only can generate a pulse waveform, sawtooth wave, square wave, pulse train signal generator can generate other conventional waveform, but also can achieve a variety of waveform modulation, such as frequency shift keying, phase shift keying, the sequence generation, frequency modulation, frequency scanning, can also be produced by way of hand-painted arbitrary shape signal. In some special applications, such as the output sweep signal, the request signal in the selected frequency band, the wave amplitude as flat, ideally close to the sweep waveform. However, due to the zero- order hold DAC unique characteristics, as well as analog channels of the device within the nominal value and the difference between the actual value and other reasons, resulting in the output signal amplitude flatness can not meet the requirements of the signal within the passband amplitude varying degrees frequency distortion. Domestic remedies for the DAC amplitude-frequency studies have been relatively mature, but for arbitrary waveform generator 's overall amplitude-frequency distortion correction scheme is studied and few. Therefore, an arbitrary waveform generator amplitude-frequency passband distortion analysis and study, and take appropriate corrective measures, so that the waveform amplitude-frequency characteristics as flat, is imminent. Foreign amplitude-frequency distortion for this study more, such as AD produced digital-analog converter chip, some types of chips built-in digital filter module, devoted to the zero-order hold by the DAC characteristic amplitude-frequency distortion caused. This digital filter module uses a multiply-accumulate the implementation structure, taking up fewer resources, is easier to achieve. As multiply-accumulate structure applies only to the case of low-speed handling, so only part of the DAC has this correction module. Applications in a digital radio, through digital quadrature amplitude modulator of the amplitude modulated waveform data into the DAC for the digital-analog conversion, inevitably introduces the zero-order hold characteristic of the amplitude- frequency distortion. To correct the distortion, Thu-jiLin and HenrySamueli a novel calibration method. They used a CMOS integrated chip in the form of internal integration of the multiply-accumulate structure based on digital filters, and the use of current popular CSD coding of filter coefficients can be simplified, using a pipelined manner, improve the processing of data speed. Using digital ASIC technology, design special-purpose chips, will be the future of mainstream digital signal processing solutions. It is obvious, however, this method requires a dedicated chip designers and chip production lines, the cost is relatively high, the design cycle is relatively long, not suitable for the cycle, the relatively high cost

requirements of the occasion. With DSP, ARM chip development and wide application, more and more digital signal processing functions to be implemented in its interior. HyungW.Paik and TimothyW.Grablander proposed a solution: the DSP to achieve 17 -order digital filter, go through the software calculates the new waveform data, and then convert the data into the DAC. This method is suitable for the large amount of data computing the occasion of use, if the operation was smaller, and then only for the implementation of digital filters and the use of DSP is even worth the candle. LC parallel network circuit is generally tuned amplifiers and oscillators for frequency selection. It has a band -pass filter characteristics. The resonance point to the signal amplification, while at other frequencies the signal points can be suppressed. Wang Wenliang in "100MSa / s arbitrary waveform generator hardware design," a paper presented a novel calibration hardware circuit LC Methods: LC resonant circuit in the case of non-ideal characteristics of the resonant wide bandwidth, making DDS output signal attenuation LC parallel band and a band corresponding to the amplification circuit. Since the DAC amplitude-frequency characteristic curve in the pass band is monotonic decline, after appropriate to select the appropriate L, C values, you can make the LC resonant circuit of the amplitude -frequency curve is exactly the opposite of the DAC. Such data is sent out through the DAC LC resonant circuit for subsequent correction, can play a significant effect, but increases the overall hardware circuit area. LC correction circuit on elliptic low -pass filter prior to the resistance, capacitance and inductance parasitic capacitance, parasitic inductance, and its value and the nominal value of the differences, the filter itself is not elliptic flatness, the actual measurement results and simulation results were different. Through appropriate modifications of the LC circuit component values, you can get better amplitude-frequency characteristics. For domestic research continues this phenomenon, such as automatic control of the National Defense University Graduate team Mr. Xu Xiangzhong in "DAC digital compensator design," a paper, we propose a zero- order hold the digital compensation formula approach, based on zero order to keep the transfer function, and zero- order hold and digital compensator overall transfer function of the digital compensator is derived transfer function and the remainder with the type used in the form of digital compensation formula for expression, and gives the simulation results. Beijing University of Aeronautics and Astronautics Zhu Yuhong, etc. In the "GPS signal simulator x / sin (x) digital filter design and implementation," a paper, we propose a x / sin (x) filters, amplitude-frequency characteristics of the DAC correction is similar. Combined with a real digital IF GPS signal simulator system target, describes the design using matlab narrowband x / sin (x) digital filter method and Xilinx's FPGA to implement the design of seven order FIR filter, the design the filter output delay of 8 clock cycles after the simulation is greater than the maximum speed of 120MHz. In summary, at home and abroad and the amplitude-frequency distortion correction method has a considerable amount of research, including through the design of specialized correction chip, through the embedded FIR correction filter, through a dedicated DSP chip programming operations and LC circuit design simulation through other ways. However, the domestic correction method is in its infancy, with only several calibration method is through the analog LC circuit or DSP operations to achieve, these programs or increase the hardware circuit of the area, or can not achieve real-time processing. To this end, by way of digital filtering of the amplitude-frequency characteristics of an arbitrary waveform generator for real-time

correction of research has a very important value. This paper analyzes the arbitrary waveform generator amplitude flatness above problems, compared to research status at home and abroad, introduced the arbitrary waveform generator works, from the structure analysis of the cause arbitrary waveform generator amplitude-frequency distortion causes - the zero-order hold DAC analog channel device characteristics and the nominal value and the actual value of the differences. Find out why, in the analysis and comparison of several correction methods at home and abroad, the paper presents a relatively simple and practical correction scheme - digital pre- correction method, and how to achieve a preliminary study. Designed to achieve a single -channel digital correction module based on multi-channel mode for arbitrary waveform generator, digital correction filter for a theoretical derivation and design the corresponding multi-channel correction module, and finally the simulation, and downloaded to the actual circuit is validated.

References

1. Zhou, X., Zheng, Y., Jin, Z., Wang, Y.-L.: FPGA-based digital capacitance detection system. Sensing Technology (April 2008)
2. Zeng, Z.-H., I has a soul, Xu, W.-S.: One way to reduce spectral leakage synchronization algorithm. Electrical Measurement & Instrumentation (November 2005)
3. Bian, X., Zhou, X., Quyou, A.: Based on virtual instrument measuring frequency phase compensation algorithm. Electronic Measurement Technology (September 2008)
4. Wei, H.N.: Distortion data acquisition device testing method. Electronic Measurement Technology (January 2009)
5. Liang, Z., Sun, J.Y.: Sinusoidal modeling measurement method and its application. Aviation Metrology & Measurement Technology (June 2001)
6. Yuan, X., Xu, H., Chen, S.: Jiang four thick, sinusoidal signal distortion measurement method Commentary. Aviation Metrology & Measurement Technology (January 2004)
7. Liang, Z., Sun, J.Y.: Pulse waveform duty cycle digital measurement and uncertainty analysis. Aviation Metrology & Measurement Technology (February 2004)
8. Wang, H., Liu, Z.-S.: Hilbert - Huang Transform signal total distortion Evaluation Huazhong University of Science and Technology (Natural Science Edition) (2008) -S1
9. Liang Z., Sun J.- Y.: Waveform measurements of the basic discussion. Measurement Technology (December 2004)
10. Liang, Z., Zhuji, J., Sun, J.-Y.: Sinusoidal waveform distortion signal source an accurate evaluation method. Measurement Technology (February 2003)
11. Shen, Y.: Doppler velocimetry system signal processing technology research. Zhejiang University (2011)
12. Qing, W.: Vibration calibration system to improve performance-related technical issues and its implementation. Zhejiang University (2003)
13. Li, H.: Power system dynamic simulation error of North China Electric Power University Beijing (2005)
14. Liu, H.-C.: Voltage capacitor DC partial discharge characteristics of Chinese Academy of Engineering Physics (2005)
15. Dong, Z.: Vehicle Tire Pressure Monitoring System. National Defense Science and Technology University (2005)
16. Jun, J.: Vibration calibration accuracy related technology research. Zhejiang University (2006)

17. Zhou, X.: Capacitive micromachined accelerometer processing circuit research. Zhejiang University (2008)
18. Liang, L.: VI-based fiber optic power transformer test system Algorithm and its Implementation. Huazhong University of Science (2007)
19. Liu: High speed and heavy load three DOF platform for closed-loop control design and research. Shanghai Jiaotong University (2009)
20. Xu, Y.: Precision capacitive micromachined accelerometer system research and design. Zhejiang University (2010)
21. Silver, Z.S., Tao, Z., Dragon: Improved data acquisition system performance testing Sine Fitting. Beijing University of Technology (June 2000)
22. raw whole paragraph, He, Z., Huang, D.: HHT method and its seismic signal processing applications. Chengdu University of Technology (Natural Science) (April 2005)
23. Liang, Z., Sun, J.Y.: Digital FM signal demodulation. Test and Measurement Technology (February 2005)
24. Liang, Z., Sun, J.Y.: Arbitrary waveform generator application and calibration Review. Test and Measurement Technology (February 2006)
25. Liu, J., Lin, H., Wang, X.: Based on EMD and ARMA model bridge vibration signal noise reduction processing method. Railway Science and Engineering (May 2006)
26. Liang, Z., Sun, J.Y.: Arbitrary function digitized waveform generation method. Electrical Measurement & Instrumentation (July 2004)
27. Huang, T., Floor, M.: HHT -based system identification of nonlinear structures. Earthquake Engineering and Engineering Vibration (March 2006)
28. Xiao, D., Xin, L., Jing, Z.: Damage detection using empirical mode decomposition method. Dalian University of Technology (March 2005)
29. Qian, Y., Huang, C.-J., Chen, C., Jiang, X.: Empirical mode decomposition based Denoising of Partial Discharge. Power Systems (December 2005)
30. Hu, J., Xi, Y.: HHT -based fault diagnosis of the rotor transverse crack. Power Engineering (2004)

Synchronous Analysis of Electroencephalogram Based on Nonlinear Independence

Bai Yifan

Yan Shan University, Qinhuangdao, Hebei 066004, China

Abstract. The brain has the characteristic of high non linearity. Analyzing the synchronous relation between electroencephalograms of each channel can achieve important information integrated, transmitted and processed in different areas of the brain. The paper uses phase-space reconstruction and nonlinear interdependence to explore the characteristic of general synchronization between two-line channel electroencephalograms. Synchronous analysis method based on nonlinear independence is applied to analyze epilepsy signals. From the analysis, we can see that the brain strengthens synchronously when epilepsy attacks, which not only can make us determine the areas of the brain which play a leading role in epilepsy, but also intuitively displays propagation characteristic of epilepsy.

Keywords: Synchronous, Electroencephalogram, Nonlinear Independence.

1 Introduction

Electroencephalogram is comprehensive performance of neuronal group action potentials in the brain tissue. Analyzing electroencephalograms can achieve large amount of information about human function and diseases. Checking electroencephalograms can provide information for the diseases including epilepsy, brain tumor, nervous system disease and intellectual disability.

2 Electroencephalograms and Synchronous Feature

The brain is the most complicated organ of functional structure of the human body. When a neuron in the brain receives the signal from other neurons, the neuron sends nerve impulses. And the final overlay of more neurons discharging can form electroencephalograms. Electroencephalograms are weak, susceptible to disturbance, random, no stationary, evidence frequency domain,coupled and nonlinear.

Synchronization is a funny natural phenomenon. Synchronization can be generally defined as the adjustment of weak interaction between oscillators on frequencies[1], which embodies interconnection between oscillators. When the systems are inconsistent, there is still synchronization, which is called generalized synchronization. Generalized synchronization can exist in the noise, which makes generalized

Y. Yang, M. Ma, and B. Liu (Eds.): ICICA 2013, Part II, CCIS 392, pp. 105–115, 2013.

synchronization achieve more and more applications to analyze neurophysiological signals.

3 Nonlinear Independence Algorithm

The methods of analyzing electroencephalograms includes linear method and nonlinear method. As the nature of neurons is nonlinear, and dual-channel nonlinear analysis method is appropriate for analyzing non-stationary and multi-noise features of signals, the paper uses the method of nonlinear interdependence to analyze synchronous relation between electroencephalograms of each channel, which can achieve important information integrated, transmitted and processed in different areas of the brain.

Nonlinear independence is the most reliable method that estimates the existence of generalized synchronization in time sequences. It depends on the reconstruction. The following is the definition of nonlinear independence [1]. Two columns of time sequences are achieved from the system X and the system Y. After phase-space reconstruction on two time sequences, we can get $x_n = (x_n, \mathrm{K}, x_{n-(m-1)\tau})$ and $y_n = (y_n, \mathrm{K}, y_{n-(m-1)\tau})$ in which $n = 1, \ldots N$, m is embedded dimension,and τ is time delay. $r_{n,j}$ and $s_{n,j}$, and $j = 1, \ldots, k$ represent k subscripts of the nearest neighbor of x_n and k .The definition of mean square euclidean distance of k subscripts of any x_n is

$$R_n^{(k)}(X) = \frac{1}{k}\sum_{j=1}^{k}(x_n - x_{r_{n,j}})^2 \tag{1}$$

Mean square euclidean distance of X under the condition of Y is defined as that of x_n and k point which replaces the subscript $r_{n,j}$ into $s_{n,j}$, that is,

$$R_n^{(k)}(X/Y) = \frac{1}{k}\sum_{j=1}^{k}(x_n - x_{s_{n,j}})^2 \tag{2}$$

If mean square radius of $\{x_n\}$ is $R(X) = \frac{1}{N}\sum_{n=1}^{N} R_n^{N-1}(X)$, when there is high correlation between the systems, $R_n^{(k)}(X/Y) \approx R_n^{(k)}(X) << R(X)$. If the systems are mutually independent, $R_n^{(k)}(X/Y) \approx R_n^{(k)}(X) >> R(X)$. And we can define

$$S^{(k)}(X/Y) = \frac{1}{N}\sum_{n=1}^{N}\frac{R_n^{(k)}(X)}{R_n^{(k)}(X/Y)} \tag{3}$$

As $R_n^{(k)}(X/Y) \geq R_n^{(k)}(X)$, $0 < S^{(k)}(X/Y) \leq 1$. The small value of $S^{(k)}(X/Y)$ indicates that X and Y are interdependent. Conversely, it indicates that X and Y are synchronized. According to literature [3], nonlinear interdependence can be defined as

$$H^{(k)}(X/Y) = \frac{1}{N}\sum_{n=1}^{N} \log \frac{R_n^{(k)}(X)}{R_n^{(k)}(X/Y)} \tag{4}$$

If X and Y are completely independent, $H^{(k)}(X/Y) = 0$. When the direction near Y is consistent that near X, $H^{(k)}(X/Y) > 0$, or $H^{(k)}(X/Y) < 0$. The research shows that

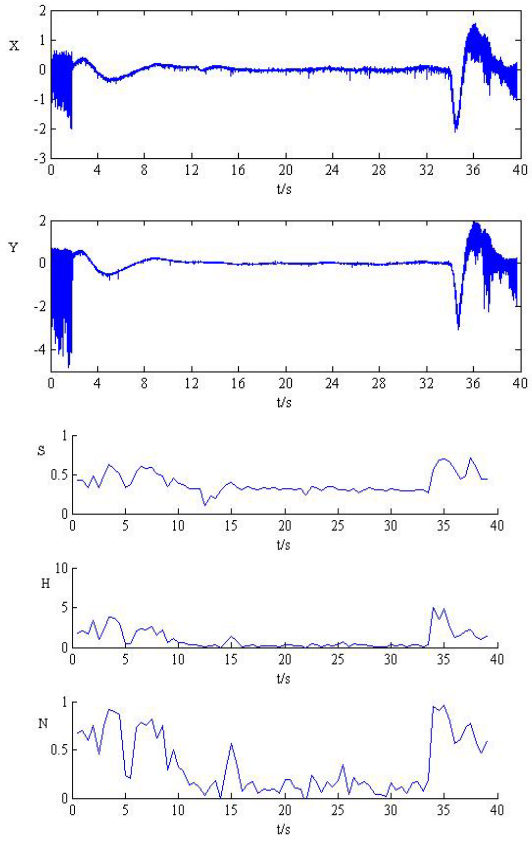
$$N^{(k)}(X/Y) = \frac{1}{N}\sum_{n=1}^{N} \frac{R_n(X) - R_n^{(k)}(X/Y)}{R_n^{(k)}(X)} \tag{5}$$

$R_n(X)$ is the average distance between vector x_n and the other vectors. Theoretically, the antinoise ability of N is stronger than that of S . The maximum of N is 1. The definitions of $S^{(k)}(Y/X)$, $H^{(k)}(Y/X)$ and $N^{(k)}(Y/X)$ can be analogized according to formula 3, 4 and 5. They are not equal to $S^{(k)}(Y/X)$, $H^{(k)}(Y/X)$ and $N^{(k)}(Y/X)$, which is the main advantage of nonlinear independence relative to other nonlinear methods.

The stimulation of signals can be achieved by the asymmetry.

4 Synchronous Analysis of Epilepsy Electroencephalograms

Using synchronous analysis of multichannel electroencephalograms can help us to know about complicated synchronous dynamic behavior formed by the coupling of neuronal group when epilepsy attacks, and understand transmission mechanism of epilepsy. In order to describe the intensity change of synchronization between two-channel field potentials more clearly, the width of time widow is setted to be 1 second, and the technology of sliding window is used to track the characteristics change of synchronization between two-channel field potentials, and it is overlapped for 0.5 second. The following figure is time domain and nonlinear dependence of dual-channel epilepsy electroencephalograms.When epilepsy attacks, neuron group of the brain discharges synchronously, which means that field potentials increase or change greatly [3]. From figure 1 (b) , we can observe that when epilepsy attacks, electroencephalograms strengthens synchronously.



a Time-domain waveform b Nonlinear independence

Fig. 1. Time-domain waveform graph and nonlinear independence of dual-channel epilepsy electroencephalograms

The following is an inspection on electroencephalograms of 9 channels. The time-domain waveform graph of channel 9 is shown in figure 2. Figure 3 shows the result of computing nonlinear interdependence between channel 1 and the other 8 channels. From figure 3, we can see that when epilepsy strokes, synchronism of electroencephalograms of each channel changes variously.Different stages of epilepsy are traced to analyze synchronizing property between two channels. As shown in Figure 2, four stages are intercepted, epileptic seizure prophase, epileptic seizure intermission, epileptic seizure 1 and epileptic seizure 2. 2 seconds are intercepted from each stage. Figure 4 is block charts of time-domain waveform and nonlinear independence of each stage. From figure 4, we can see that electroencephalograms of

seizure intermission and seizure prophase have no evident synchronization. In epileptic seizure, channel 2, channel 3, channel 4 and channel 5 begin to have synchronization and high-amplitude behavior, which can illustrate that the seizure areas are channel 2, channel 3, channel 4 and channel 5. We can see from figure 4（c） and 4（d） that the channels which have synchronization are changing, from wich we can know about the transmission state of epilepsy.

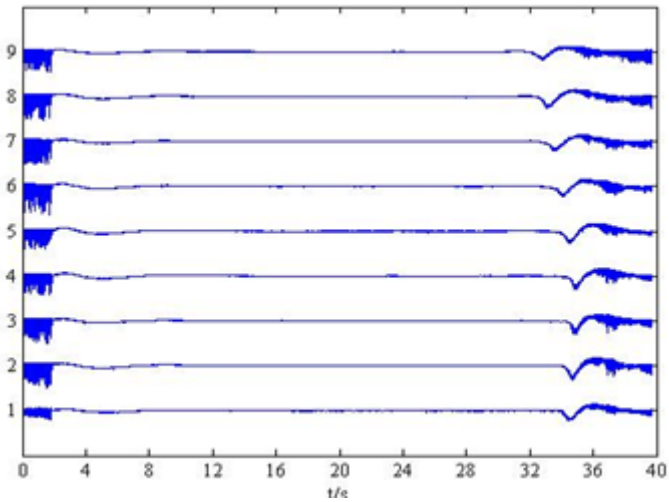


Fig. 2. Time-domain waveform of electroencephalograms of channel 9

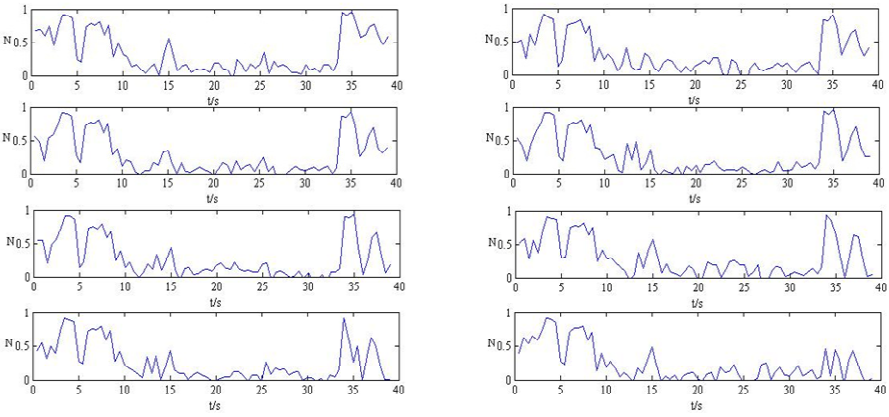
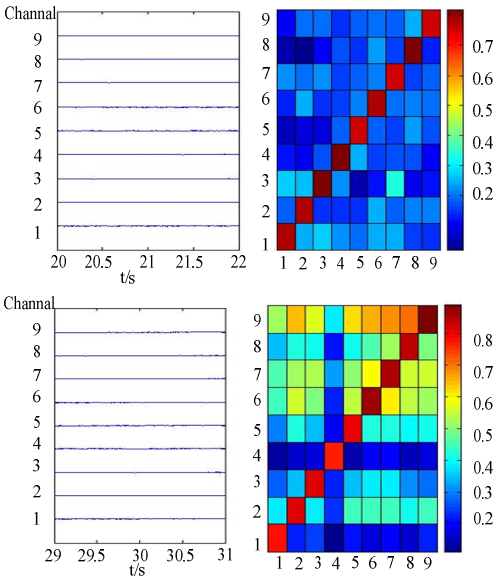
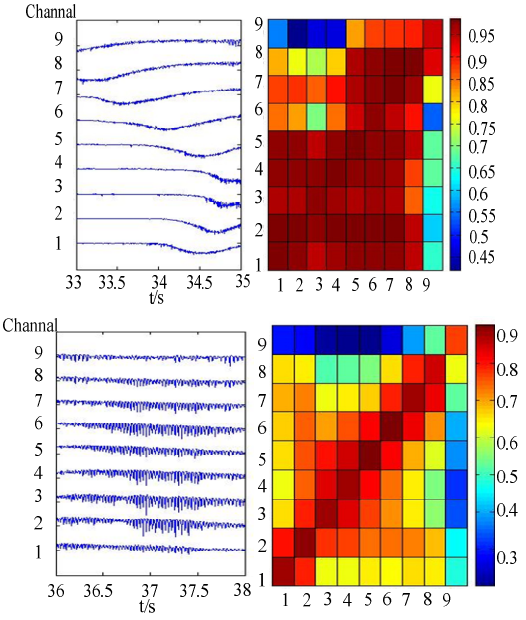


Fig. 3. Nonlinear independence between channel 1 and the other 8 channels



a Intermission b Seizure prophase



c Seizure 1 d Seizure 2

Fig. 4. Time domain graph and synchronicity intensity of epilepsy electroencephalograms in each stage

5 Conclusion

The paper makes synchronous analysis of epilepsy electroencephalograms based on nonlinear interdependence. After analyzing multichannels, we can see that when epilepsy attacks, the synchronism of electroencephalograms strengthens evidently. From the analysis on epilepsy electroencephalograms in different stages, we not only can get the brain areas which plays a leading role in epilepsy, but also can know about transmission mechanism of electroencephalograms.

EEG signal analysis method is applied to monitoring depth of anesthesia research has made some achievements, but its progress has not been happy to researchers, a variety of methods in clinical applications are still many limitations. Bispectral index for monitoring depth of anesthesia research is gaining importance, the index can more sensitively reflect the depth of anesthesia, but because it exists in different anesthetics, different anesthesia methods reflect the different shortcomings, so can not be a perfect the technology used in clinical anesthesia and independent monitoring. Anesthesia trend, artificial neural networks, complexity, wavelet analysis and nonlinear EEG analysis is developed in recent years has the potential to anesthesia monitoring methods, but also requires a lot of research to further demonstrate its validity and clinical applications feasible.

BIS (bispectral index, BIS) bispectral index is included time domain, frequency domain and higher order spectral variables (double spectral analysis) three characteristics of the electroencephalogram (EEG) quantitative analysis of indicators, which through a specific non- linear algorithm (Aspect Medical Systems, Natick, USA), the four different EEG parameter, the burst suppression ratio (burst suppression ratio, BSR), "QUAZI", β ratios (beta ratio) and the speed of the wave relative synchronization (synch fast slow), integrated into a non-dimensional number 100-0 [1], for the degree of inhibition of brain. 1997 approved by the FDA as a monitoring depth of anesthesia and sedation level indicators into clinical applications and research stage. As bispectral index includes the phase information, including high-level information so that analysis of the electroencephalogram bispectral index signal analysis has important value, thus BIS EEG monitoring anesthesia is a common method of sedation, so far BIS has been conducted on the application of a large depth of anesthesia assessment reports [2-4]. The study confirmed, BIS can be used to evaluate the anesthetic sedation [5,6] and noxious stimuli. With the deepening of the level of anesthesia or sedation, BIS values decreased. BIS specificity, sensitivity and accuracy of better, and little variability. Sebel et al found that, by BIS monitoring routine monitoring of patients than other patients in the ICU stay a short time. BIS monitoring in the induction and maintenance phases based on loss of consciousness value (about 60 or so), to adjust the amount of anesthetic induction and maintenance dose and reduce the level of uncertainty caused by anesthesia improper drug use prevalence, but also may be associated with reduced some of the negative role in enhancing the safety of patients undergoing for additional guidance anesthetic use and recovery in patients with early projections provide an effective assessment tools. Glass, etc. [3,7] studies have confirmed that the BIS -based continuous real depth of anesthesia monitoring methods to optimize each patient's anesthetic drug use, thereby reducing the anesthetic dose may

be insufficient and excessive situations. However, BIS as the depth of anesthesia monitoring technology has its obvious limitations, many factors will have an impact on the BIS. Detsch other studies [8-10] found, BIS EEG effect of significantly dependent on the use of anesthetics, such as BIS and the inspired concentration of sevoflurane good correlation with the concentration of isoflurane was no correlation, and, N2O and xenon anesthesia had no effect on the BIS values . Ortolani, etc. [11] study found that, BIS has ethnic differences.

Anesthesia trend (narcotrend, NT) NT is a new EEG for measuring depth of anesthesia method using multi-parameter Kugler statistics and computer processing, to form six stages EEG levels of quantitative indicators 14, i.e. A, B0 ~ 2, C0 ~ 2, D0 ~ 2, E0 ~ 1, F0 ~ 1, and also shows α, β, γ, δ wave power spectrum of the changes and trends. Phase A represents the waking state ; B is sedation (0, 1, 2); C is lightly anesthetized (0, 1, 2); D is the conventional general anesthesia (0, 1, 2); E is the depth of anesthesia (0, 1, 2); F stage (0, 1) is the disappearance of EEG activity. Kreuer, etc. [12] study confirmed the trend of anesthesia classification method can indeed detect desflurane anesthesia accident occurred when the concentration of desflurane decreased, while doing the typical depth of anesthesia classification, but also found narcotic trend analysis and bispectral analysis depth of anesthesia has a similar effect [12]. Schultz, etc. [13] conducted a feasibility study on Narcotrend found, Narcotrend depth of anesthesia and sedation levels for judgments, prediction probability PK is 0.90, the correlation coefficient γ is 0.90. Wilhelin, etc. [14] By 4630 clinical cases of intravenous anesthetic drug monitoring proved, Narcotrend is a new type of high credibility depth of anesthesia monitoring methods.

ANN (artificial neural networks, ANN) method of artificial neural network monitoring depth of anesthesia is the recent development of EEG analysis technique, which is the second after one kind BIS EEG power spectral data from all extracted out of a single variable. Usually four characteristic of EEG waveforms α, β, γ, δ, the average power as a characteristic parameter of EEG spectrum, coupled hemodynamic parameters such as blood pressure, heart rate, and MAC narcotic drug dose that data, using the AR model, cluster analysis and Bayes estimation theory, eventually forming a representative of the absolute values of the depth of anesthesia. Numerous studies have showed that neural networks can indeed successfully EEG data were graded according to the level of sedation [15,16]. Ortolani [17] and so on 200 patients were tested, of which 150 patient data used for training ANN, another 50 patients for testing ANN and data for statistical analysis. Extracted from the 14 EEG variables outside the BIS 13 ANN training variables as input value, the output target value is determined by both the depth of anesthesia anesthesiologist index. Instrument (Aspect A-1000) is given for each recording data signal quality index (signal quality index, SQI), and SQI <50% of the record data is removed from the training database. ANN MLP network performance analysis options (MLP), using standard back propagation (BP) algorithm for network training. First with Neural SIM software to optimize the network, and then start ANN training. When the mean absolute error < 4 and the rms error of < 5:00, stop training. Depth evaluation of the resulting neural network (network evaluated depth, NED) in the range of 0 to 100. In training, ANN calculates the target output and the corresponding correlation coefficients estimated NED, which shows the ANN training

is successful. To evaluate the linear correlation of the pre-treated and BIS EEG variables for multiple regression analysis. The test group BIS and NED data to calculate the Pearson correlation coefficient. Then the data according to Bland and Altman method, the method to make the average of the deviations between their maps to evaluate the consistency between the two. The results showed that, after multiple regression analysis to calculate the NED and the BIS Pearson correlation coefficient was 0.9411, by the Bland and Altman method to calculate the deviation of -0.199, consistency extremes of -10.19 (lower limit) and 9.79 (upper limit), indicating minimal deviation and good consistency. The experimental results show that the neural network by analyzing the preprocessed EEG data, can provide the depth of anesthesia index. Robert, etc. [18] also compared the neural network and BIS EEG analysis methods as depth of anesthesia monitoring results, the results also show that the neural network by analyzing the processed EEG data generated anesthesia index and a good correlation between BIS and neural network also able to handle the lack of phase information EEG data to assess the depth of anesthesia.

Complexity and wavelet analysis method (complexity and wavelet analysis) complexity in recent years, the emergence of a nonlinear dynamic analysis methods, because EEG is a non- stationary signal, so the complexity analysis is portrayed dynamics of EEG signals effective way. Chen et al [19] using the Lempel-Ziv algorithm complexity C (n) [20], after a study found, EEG sequences between the complexity and depth of anesthesia has a close relationship. At different depth of anesthesia, the complexity of the changes are more sensitive. More shallow depth of anesthesia, complexity, the greater the value ; deeper anesthesia, complexity, smaller value. As the complexity of algorithm is simple and easy to implement, fast speed, so the parameter for clinical anesthesia depth real-time monitoring possible. Zhang and Roy [21] used a combination of wavelet analysis, complex analysis and neural network system for monitoring depth of anesthesia. The first with a wavelet transform to the original EEG signal into components of different scales sixth row is calculated and the corresponding raw EEG complexity of each section, then the resulting complexity C (n) is applied to a four-layer neural network to predict the action of anesthesia. The results show that the system sensitivity of 88% and an accuracy of 92% and a specificity of 97 %, which is promising monitoring methods.

EEG nonlinear analysis (EEG-nonlinear) Learning in the use of nonlinear analysis and fractal theory chaotic nonlinear dynamics theory and methods to study and analyze the patient's perioperative changes in brain function status, has become a new research hotspot. Nonlinear Dynamic Analysis of EEG analysis method represents a future direction. Shen et al [22] using higher order spectral analysis techniques [23] to establish EEG nonlinear model of EEG third-order statistics information. Then establish a microcomputer as the core of EEG bispectral analysis system for clinical EEG nonlinear analysis. Zhang et al [24], the nonlinear dynamics of the Kolmogorov entropy applied to the analysis of rat EEG monitoring depth of anesthesia, the calculation and analysis of the rats anesthetized by intraperitoneal injection of sodium pentobarbital during dynamic Kolmogorov entropy of EEG curve, the results showed dynamic changes during anesthesia Kolmogorov entropy curve and the original EEG

brain regions, as reflected in the excited state suppression and trends are in good agreement, namely Kolmogorov entropy for clinical provides real-time monitoring depth of anesthesia a new approach. Bruhn, etc. [25] Application of approximate entropy of a series of studies, the results showed that: Approximate Entropy can accurately identify the occurrence of burst suppression ; With anesthetic (desflurane, propofol, etc.) concentration increased, the rules of EEG is also a corresponding increase ; compared with EEG spectral analysis, approximate entropy artifacts even less sensitive to intra-individual or inter-individual variability is relatively small. Vakkuri, etc. [26] proposed a time-frequency balanced spectral entropy concept, its clinical application results show that : In the propofol, sevoflurane and thiopental anesthesia, state entropy and response entropy can distinguish between conscious and unconscious states. Spectral entropy through (amount) EMG activity reflects the extent of sedation and pain response, reflecting the depth of anesthesia monitoring using a variety of methods for integrated analysis of trends (eg EEG analysis combined EMG), deserves further attention. EEG nonlinear dynamics analysis leads us up at the system level understanding of EEG, combining theory research tools, understand its contents, previously known information is difficult. We have every reason to believe that non-linear dynamic analysis in clinical applications continue to expand today, EEG nonlinear analysis will change and awareness during anesthesia EEG changes in cognitive function, as well as the depth of anesthesia changes provide a more effective means of research and more meaningful information.

An ideal depth of anesthesia monitor a number of conditions must be met : (1) how changes regardless of anesthetic concentration, can accurately monitor the depth of anesthesia and sedation and clinical manifestations of good correlation ; (2) the data is correct and reliable, have a higher time resolution ; (3) for all anesthetics can use a common grading scales anesthesia ; (4) can help clinicians provide conclusive ; (5) anti-electromagnetic interference capability ; (6) cost-effective easy to use. To achieve such a goal, further research is needed for the depth of anesthesia and improve detection methods, while also using a variety of relevant medical signals, to obtain a more comprehensive and integrated indicators.

References

1. Kraskov, A.: Synchronization and Interdependence Measures and their Applications to the Electroencephalogram of Epilepsy Patients and Clustering of Data. Germany. PhD thesis of University of Wuppertal, pp. 40-44 (2004)
2. Fine, A.S., Nicholls, D.P., Mogul, D.J.: Assessing instantaneous synchrony of nonlinear nonstationary oscillators in the brain. Journal of Neuroscience Methods 186(1), 42–51 (2010)
3. Quian Quiroga, R., Arnhold, J., Grassberger, P.: Learning driver-response relationships from synchronization patterns. Phys. Rev. E 61, 5142 (2000)
4. Pereda, E., Quiroga, R.Q., Vhattacharya, J.: Nonlinear multivariate analysis of neurophysiological signals. Progress in Neurobiology 77(1-2), 1–37 (2005)
5. Cui, D., Liu, X.Z., Wan, Y., et al.: Estimation of genuine and random synchronization in multivariate neural series. Neural Networks 23, 698–704 (2010)

6. Cantero, J.L., Atienza, M., Cruz-Vadell, A., et al.: Increased synchronization and decreased neural complexity underlie thalamocortical oscillatory dynamics in mild cognitive impairment. NeuroImage 46(4), 938–948 (2009)
7. Quian Quiroga, R., Kraskov, A., Kreuz, T., et al.: Performance of different synchronization measures in real data: a case study on electro-encephalographic signals. Phys. Rev. E 65, 041903 (1903)
8. Uhlhaas, P.J., Pipa, G., Lima, B., et al.: Neural synchrony in cortical networks: history concept and current status. Front Integr. Neurosci. 3, 17 (2009)
9. Wu, Q.: EEG source numerical solution of Hebei University of Technology (2001)
10. The Perspectives; different methods of psychological relaxation training on subjective sense of relaxation and self- physiological reaction of Beijing Sport University (2002)
11. Lu, Y.: Conscious rats neurotransmitters and EEG detection method combined mechanism of sedative drugs in the application. Zhejiang University (2004)
12. Ming, C.: EEG -based brain - computer interface research Tsinghua University (2004)
13. Wen, L.: Sibelium lit effect on epilepsy in rats and surgical treatment of intractable epilepsy clinical research. Shandong University (2004)
14. Wang, N.: Nonlinear dynamics methods in time series analysis application. Fudan University (2005)
15. Wang, J.: Selective head hypothermia new hypoxic-ischemic brain injury in pigs EEG physiological changes. Fudan University (2005)
16. Zhang, X.: Blind Source Separation and Its Application in EEG signal processing research. Dalian University of Technology (2006)
17. Li, B.: -Line handwritten signature verification technology research. Harbin Institute of Technology (2006)
18. Xu, Y.: Obstructive sleep apnea syndrome causes learning and memory decline biochemical mechanisms and treatment of clinical and basic research. Chinese Peking Union Medical College (2002)
19. Wang, T.: Epilepsy epileptic discharge and structural abnormalities comparative study. Postgraduate Medical School (2001)
20. Cai, C.: Chaotic system and its largest Lyapunov exponent is calculated in EEG data analysis application. Xiamen University (2001)
21. Zhang, D.: Based on wavelet and independent component analysis of EEG signal processing. Anhui University (2002)
22. He, J.: Hypothyroid rat auditory evoked potential postnatal development study. Jinan University (2003)
23. Chen, X.: Healthy Chinese idioms and sentence event-related potentials N_ (400) research and clinical application of First Military Medical University (2003)
24. States Korea Ru: Perfusion CT, EEG topographic diagnosis in patients with acute cerebral infarction application of Hebei Medical University (2003)
25. Yu, M.: Supplement malate on synchronized swimmers acting ability and EEG effects. Beijing Sport University (2003)
26. Praise: Based Alpha Waves Brain - Computer Interface Design. Tianjin University (2004)

MIDI Music Generator Based on STM32

Xinyuan Guo

Changsha University, Changsha, Hunan, China

Abstract. This paper describes a use of ARM processors STM32F407ZET6's and MIDI music generator hardware and software design and implementation, STM32F407ZET6 communicate with the host computer through serial ports, and the cpu can access and quickly identify the vocal baseband data, and finally convert it to a MIDI sound chip that can recognize MIDI message, so that the MIDI sound source and vocal fundamental frequency corresponding to the different sounds of electronic tones.

Keywords: ARM, STM32F407ZET6, MIDI, music generator.

1 Introductions

In recent years, computer technology has caused a series of interdisciplinary booming; computer music technology is one of the more active subjects [1-6]. It's traditional music with modern computer technology, the development of traditional music provides a broader space for development.

MIDI is an acronym for musical instrument digital interface, which is an electronic musical instrument digital interface, it is used in a music synthesizer (music synthesizers), musical instruments (musical instruments) and the exchange of musical information between computers a standard protocol [7-11]. MIDI files transmitted between the MIDI device MIDI message composition, the MIDI file is not sampled music samples, but rather a number of music data, a smaller storage space occupied. Therefore, the storage space is relatively sensitive embedded devices; MIDI music generator design is very practical significance [12-15].

Development board with independent design, the main chip use ST STMicroelectronics latest Cortex-M4 with FPU (floating point unit) features a high-performance MCU, its frequency operating at 168MHz, is controlled not only has a highly efficient operation, but also the floating point arithmetic operations are faster hardware arithmetic unit support [16-21]. Therefore, either the control or the audio and video processing is an ideal solution for high-end chip.

Circuit take STM32F407ZET6 chip, the chip is ARM32 bit CortexTM-M4F core with FPU, the maximum operating frequency at 168MHz, up 210DMIPS (1.25DMIPS/MHz), and support for DSP instructions [22-27]. Chip comes (192 +4) KB of RAM (more than usual 64KB STM32F103 has more memory) and 512KB of

Y. Yang, M. Ma, and B. Liu (Eds.): ICICA 2013, Part II, CCIS 392, pp. 116–125, 2013.

Flash (ART accelerator components with real-time, making the instruction is executed 0 wait cycles).

Development board complete with a variety of interfaces, supporting the program code, you can quickly get started ST chip various control and computing. Specific resources [28-30]: fully functional USART serial, USB HOST (full speed), USB DEVICE (full speed), Micro SD card, SPI Interface (SPI Flash), IIC Interface (24C08), CAN Interface, DCMI interfaces (OV7670, OV9650, OV9653, OV9655, supporting directly with Mini2440 development board CAM130 module is compatible), FSMC interfaces (LCD, ILI9320), TIMER Interface (buzzer), PIO (LEDx2 output, the key input), ADCx4, DACx1, PWMx4, ETH Interface (DM9161AEP, 100MHz), RTC, FPU and so on.

2 The Hardware Design of Embedded MIDI Music Data Generator

Block diagram of system hardware shown in figure 1.

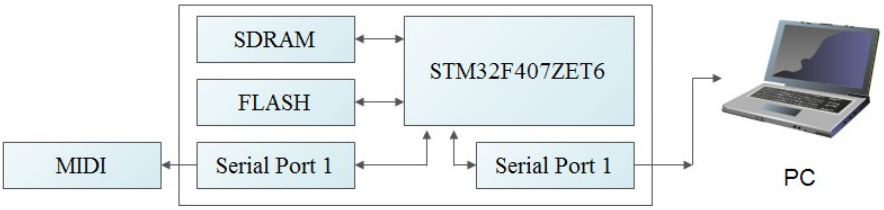


Fig. 1. The system hardware principle diagram

Figure 1, PC send MIDI notes information to STM32F407ZET6 through the RS232 serial port 0. PC will debug system, and terminal communication through the serial port 1.The STM32F407ZET6 received information is stored to note SDRAM, here is a group of people storing audio baseband data due MIDI sound only recognizes MIDI note number, so the baseband data to be converted to the corresponding note number, which corresponds to the relationship as shown in Table 1.

Table 1 cycle and MIDI note number (referred tone number) preformed in FLASH. Execute table program to find sound period corresponding MIDI note number data, and then send MIDI data to an audio chip via the serial port, resulting in a corresponding MIDI tone, but also to the human voice for each baseband data to specify a different tone sounds, and thus produce different forms of accompaniment. For example, the resulting baseband data 0x84, by the look-up table data corresponding MIDI note number 36, in accordance with the following MIDI commands, MIDI message format can be prepared as follows.

Table 1. The corresponding table of cycle、fixed do system、tone number

Cycle	Fixed Do System	Tone number	Cycle	Fixed Do System	Tone number	Cycle	Fixed Do System	Tone number
C4	F	29	5D	#f	42	2C	g1	55
BA	#F	30	58	g	43	2A	#g1	56
B0	G	31	53	#g	44	28	a1	57
A5	#G	32	4E	a	45	26	#a1	58
9C	A	33	4A	#a	46	24	b1	59
94	#A	34	46	b	47	22	c2	60
8C	B	35	42	c1	48	20	#c2	61
84	C	36	3E	#c1	49	1E	d2	62
7D	#C	37	3B	d1	50	1C	#d2	63
76	d	38	38	#d1	51	1B	e2	64
70	#d	39	35	e1	52	19	f2	65
69	e	40	32	f1	53	17	#f2	66
63	f	41	2F	#f1	54	15	g2	67

1) 0xC4 40; this command indicates the tone changed five channels for violin.

2) 0x94 36 68; this command indicates to open five channels, note 36, the key strength is 68.

3) 0xC5 56; this command indicates changing six-channel sound for the trumpet.

4) 0x94 36 68; this command indicates to open six channels of 36 notes, key strength is 68.

5) Delay some sound long.

6) 0x84 36 68; expressed concern over the command channel 5, note 36, the key strength is 68.

7) 0x85 36 68; expressed concern about the order 6-channel, 36 notes, key strength is 68.

Of course, you can also specify or add other different instrument sounds, enabling richer accompaniment

3 The Software Design of Embedded MIDI Music Data Generator

STM32-based MP3 player design methods, the system implements the SD card to store MP3 files decode and play, and in the TFT touch screen display playing a song title, song playback time remaining and display song lyrics counterparts through TFT touch screen graphical user interaction diagram circles, you can achieve the song fast

forward, rewind, pause and change circulation patterns, etc., can be achieved with a touch screen function MP3 player.

Decoding hardware solution using MP3, the direct use of audio decoder chip VS1003 decoding. Because MP3/WMA/MIDI VS1003 is a monolithic audio decoder and ADPCM encoder. It contains a high-performance, proprietary low- power DSP processor core VS_DSP4, working data memory for user applications 5KB instruction RAM and 0.5KB of data RAM. Serial control and data interfaces, four general-purpose I / O port, a UART, also has a high-quality variable sampling rate AD... (View More) C and stereo DAC, there is a headphone amplifier and ground buffer, powerful and cheap, VS1003 supports MP3, WAV, WMA, MIDI and many other audio formats. And using a touch screen instead of the traditional buttons, making input more convenient, GUI interface look more beautiful. This design can be seen on the market than the MP3 player has some advantages.

The system supports long file names and character display, if you use short file names, places ASCII encoding ; when the file name is longer than 8 bytes, the extension over three bytes, with regard to long file names in the form of storage, long file names character is encoded using UNICODE, each character occupies two bytes of space, its directory entry is 11 bytes offset that Attribute field is 0FH. In the long file name stored in 13 characters as a unit for cutting, each group occupying a directory entry, it may require more than one file directory entries, then the long file name each directory entry in the directory by descending the table, name to prevent confusion with other files.

The system uses the serial port as the input and output interfaces, via the PC HyperTerminal, enter the command to be controlled and displayed in the HyperTerminal output results, such as the path, directory listings, file name, and other related information. HyperTerminal to GB2312 Chinese character input and output encoding, input and output in ASCII encoded English characters, and long file names with UNICODE encoding. It needs to be converted to UNICODE encoding GB2312 or ASCII encoding and user interaction. Sort using a two-dimensional table here, there are more than 7,000 Chinese characters corresponding UNICODE encoding GB2312 coding. This search algorithm using binary search algorithm, greatly improving the search efficiency, which is critical for embedded applications.

VS1003 can be used as a slave microcontroller, the serial SPI interface to receive the input bit stream, the input bit stream is decoded, via a digital volume control reaches an 18-bit oversampled multi DAC. Such use of a processor with VS1003 chip STM32F103x, STM32 processor reads the SD card MP3 files sent to it through the SPI interface VS1003 chip play and then use STM32F10X processor GPIO port to control some of which can be achieved VS1003 an MP3 Player prototype designs.

3.1 U-BOOT Transplant

U-BOOT is universal boot loader and follows the terms of the GPL open source project. It supports MIPS, x86, ARM, NIOS, XScale, and many other popular series of processors. U-BOOT mainly used for system hardware initialization, the embedded

operating system and application software to provide a good working platform, which provides online upgrade system, debugging, and many other features for embedded systems development to provide a greatly facilitated. U-BOOT transplants are considering amending the following documents:

(1) <Target board>. H header files, such as include/con gs/RPxlite.h. U-Boot source code can be already in the target board header file, it can be newly named configuration header files; Most of the registers parameters are set in this file complete.

(2) <Target board>. C files, such as board / RPXlite/RPXlite.C. It is the SDRAM drivers, mainly to complete the SDRAM UPM table settings, power-on initialization.

(3) Flash drivers, such as board/RPXlite/flash.C or common/cfi-flash.C. In reference to an existing Flash-driven, based on the combination of target board Flash Data Sheet for appropriate changes;

(4) Serial driver, such as modifying cpu/mpc8xx/seria1.c serial transceiver chip enable parts.

3.2 System Software Design

Because the system does not exist multi task scheduling and therefore not transplant operating system, when the system power on initialization, the 0x0 boot U-BOOT.BIN, copy it to SDRAM 0x0c000000, executive system initialization program. Then, U-BOOT to 0x50000directly call system applications, the main program begins execution. Flow chart of the system is shown in figure 2.

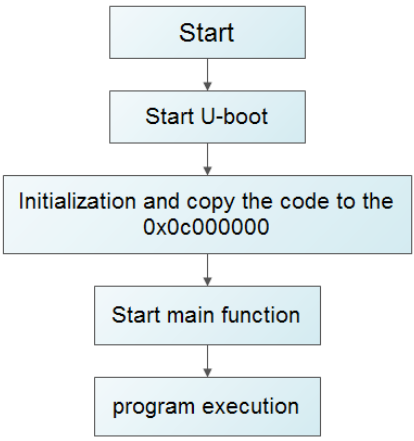


Fig. 2. System flow cha**rt**

The main program flow chart is shown in figure 3.

As can be seen from the chart, the main program is to eliminate invalid baseband data and check the MIDI note number data to find the effective data correspondence, and transformed into the corresponding MIDI message through the RS232 serial port,

send to MIDI audio chip. In order to improve lookup efficiency, the corresponding optimization algorithm, in order to improve the efficiency of the program. We can see the correct MIDI note number data obtained by AXD debugging software. In this example, figure 4 shows the acquisition cycle data sound, table 2 is the MIDI note number corresponding to the data search.

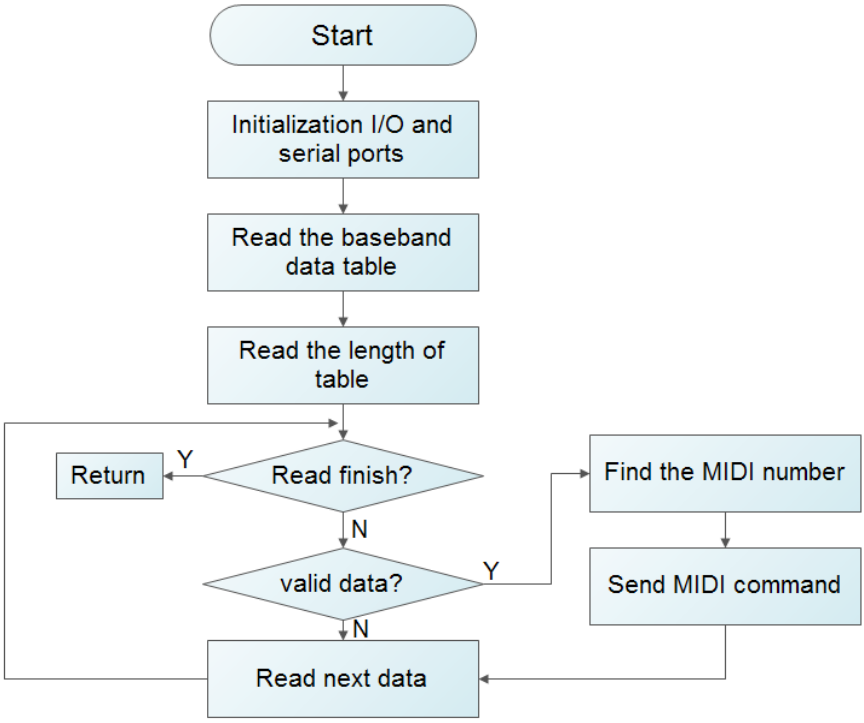


Fig. 3. The flow chart of main program

4 Programming and System Testing

4.1 U-BOOT Program

In this paper, using FlashPgm write Flash. FlashPgm is a GUI interface of the Flash programming tool, because of its friendly programming interface, programming speed and a high success rate which is in common use. When using the FlashPgm programming Flash, attention to the Wiggler the use of the JTAG interface board that is connected to the target board.

```
{0xc4,0xba,0xa50x9c,0x94,0x8c,0x84,0x7d,0x76,0x5d,0x58,0x53,
0x4e,0x46,0x3e,0x3b,0x38,0x32,0x2f,0x2c,0x2a,0x28,0x26,0x24,0x2
2,0x20,0x1e,0x1c,0x2b,0x19,0x17,0x17,0x15,0x15,0x17,0x19,0x1b,0
x1c,0x1e,0x20,0x22,0x24,0x26,0x28,0x2a,0x2c,0x2f,0x32,0x35,0x38,
0x3b,0x3e,0x42,0x46,0x4a,0x4e,0x53,0x58,0x5d,0x63,0x96,0x70,0x7
6,0x7d,0x84,0x8c,0x94,0x9c,0xa5,0xb0,0xba,0xc4,0xc4,0xba,0xff,0xf
e,0xb0,0xa5,0xf5,0xfc,0xf8,0xf3,0xe0,0xe2,0xc0,}
```

Fig. 4. The sound cycle data table

Table 2. MIDI note number table

Address	0	2	4	6	8	a	c	e
0X0C00D000	1E1D	201F	2221	2423	2625	2827	2A29	2C2B
0X0C00D010	2E2D	302F	3231	3433	3635	3837	3A39	3C3B
0X0C00D020	3E3D	403F	4241	4343	4142	3F40	3D3E	3B3C
0X0C00D030	392A	3738	3536	3334	3132	2F30	2D2E	2B2C
0X0C00D040	292A	2728	2526	2324	2122	1F20	1D1E	1E1D
0X0C00D050	201F	E7FF	E800	E800	0001	E7FF	E800	E800

4.2 Program Download

U-BOOT programming is successful, you can adopt U-BOOT to download the system main program to the target board of Flash, in this process we should pay attention to that: the main program is downloaded to the 0x50000 address of Flash, so that U-BOOT can successfully guide the system main program.

4.3 System Test

System testing if U-BOOT guides the main program, realizing the transformation frequency data and MIDI note number data, so as to realize the MIDI music player. The test current to write a set of fundamental data, write data form a one-dimensional array:

```
music_data[]={0x58,0x58,0x63,0x58,0x4a,0x42,
0x38,0x42,0x4a,0x42,0x38,0x32,0x38,0x42,
0x4a,0x4a,0x58,0x58,0x63,0x70,0x84,0x94,0x84,0x70,0x63,0x58,0x4a,0x42,0x4a
,0x58,0x58,0x4a,0x58,0x63,0x58,0x63,0x70,0x70,0x70,0x94,0x84,0x70,0x63,0x58,0
x4a,0x42,0x4a,0x58,0x58,0x42,0x38,0x42,0x4a,0x58,0x4a,0x4a,0x38,0x42,0x42,0x4
a,0x42,0x4a,0x42,0x4a,0x58,0x84,0x70,0x58,0x4a,0x58,0x63,0x63,0x63,0x58,0x58,
0x63,0x58,0x4a,0x70,0x84,0x94,0x84,0x84, 0x42, 0x38,
0x42, 0x4a, 0x42, 0x4a,
0x58,0x4a,0x42,0x94,0x84,0x70,0x63,0x58,0x4a,0x42,0x4a,0x58,0x58,
0x4a,0x58,0x63,0x58,0x63,0x70,0x70,0x70,0x70}
```

The array is according to the song "grass inch heart" music compilation, is the standard MIDI scale data of the song melody, send it to MIDI source can be drawn from the test results, the test can choose different timbre, using the speaker listener or oscillographic instrument detection. In a MIDI audio output end of the trumpet sound waveform and frequency diagram as shown in Figure 6, beautiful sound.

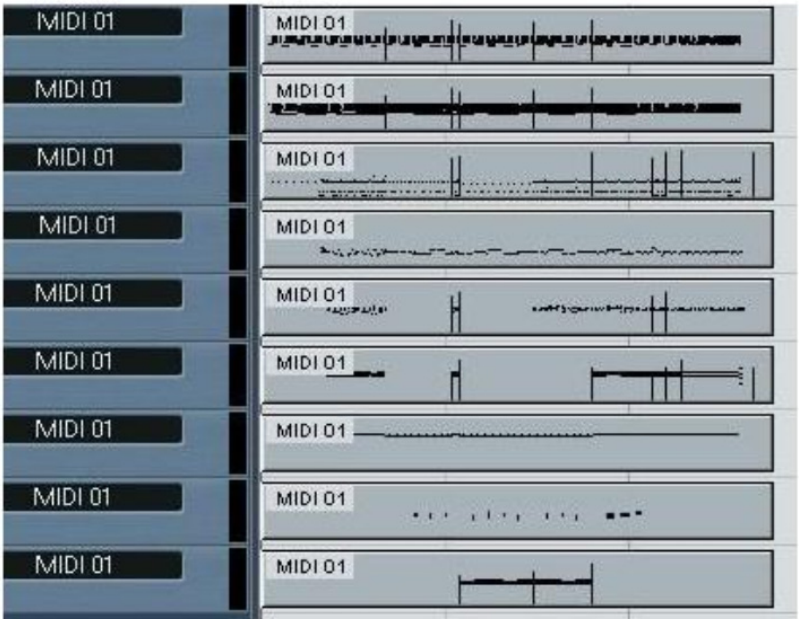


Fig. 5. Output of MIDI waveform

5 Conclusions

STM32-based MP3 player design methods, the system implements the SD card to store MP3 files decode and play, and in the TFT touch screen display playing a song title, song playback time remaining and display song lyrics counterparts through TFT touch screen graphical user interaction diagram circles, you can achieve the song fast forward, rewind, pause and change circulation patterns, etc., can be achieved with a touch screen function MP3 player. The STM processor has high performance, low power consumption, small volume and development, particularly suitable for portable MIDI music electronic products.STM32F407ZET6 successfully developed MIDI music generator based STM processor, is an attempt to music data, the results show that the STM microprocessor can real-time processing MIDI data, it has a high application value in this field. Combined with the MIDI music production software and music instruments of knowledge, by further improved, can be used in the development of sound control electronic musical instrument, for the pleasure of life.

Acknowledgement. The 379 provincial general subject : Practice and reform of the ethnic dance teaching in universities.

References

1. Cheng, Y., Wan, M., Liu, Y., Wu, S.: Programmable music generator design. Chongqing Institute of Technology (Natural Science Edition) (April 2008)
2. Cheng, Z.X., Shen, J.O.: From relational space to Euclidean space. Changzhou Institute of Technology (March 2005)
3. Zhong industry: PLD technology to develop new trends. Electronics World (August 2001)
4. Tao, J., Zhu, P.C.: Computer optical music recognition technology. Popular Science (August 2005)
5. Ling, X.: Editing Songs played Design. Electronic world (October 2003)
6. Jie, G., Ming, Y.: Weighted Hausdorff distance based on spatial face recognition algorithm. Hebei University of Technology (June 2004)
7. Bin, Z.: PIC16C5X Microcontroller with music theory and method. North China Institute of Technology (March 2002)
8. Ma, J., Chen, L.C., Zhang, Y.: Contour tracking and edge detection image recognition. North University of China (Natural Science Edition) (May 2006)
9. Wang S., Kangrun: Simple music generator with CPLD. Jiaozuo Institute of Technology (Natural Science Edition) (February 2001)
10. Mao, H., Ma, Y., Zhou, Y., Xiong, P.: Face Recognition Algorithms - Multiple orthogonal projection based on the image and the minimum distance classifier class. Computer Engineering and Applications 27 (2005)
11. Lao, L.: Thermostat based on CPLD chip design. Guangxi University of Technology (April 2002)
12. Qi, U.: CPLD-based pseudo-random sequence generator. Electronic Design & Application (November 2003)
13. Qi, U.: CPLD- based pseudo-random sequence generator. Modern Electronics Technique (March 2004)
14. Bo, Yi, D.S., Wang, D.: CPLD in the circuit board fault diagnosis. Microcomputer Information (March 2002)
15. Jiang, L.: CPLD in passive jamming equipment application. Optoelectronic confrontation with passive interference (January 2002)
16. Zhou Y., Yang L.: With CPLD digital system controller. North China Institute of Aerospace Engineering (2001) -S1
17. Liu, Y.A., Wu, H.: EDA experimental platform for the design and application. Electronics Engineers (December 2003)
18. Song, G.: Zhang T.H., Jiang H.: Digital circuits in VHDL synthesis technology. Modern Computer (March 2001)
19. Shen, H.: CPLD digital circuit experiment system design. China Educational Technology Equipment (June 2006)
20. Huang post macros, Ping, S.: CPLD-based EDA innovative experimental box design application. Laboratory Science (February 2011)
21. Chen, J.: EU introduced regulations music player limit. Chinese abroad Times (2009)
22. Liu, W.: Magnificent nine days, general manager. Vigorously develop the characteristics of domestic EDA tools. China Electronics News (2011)
23. Liang, H.: China EDA strategy seminar held in Beijing. China Electronics News (2009)
24. Yang, S.: Jinan ICC Development Center IC Division. Jinan: 12 in. line and EDA are two key. China Electronics News (2010)
25. High, J.: Altera out MAX II series CPLD. Electronic Information Times (2004)

26. Chen, S.: Music Player: from “talking “to” intelligent ”. Chinese Culture Daily (2010)
27. PLAYGAMER.COM; ’s most popular music player WINAMP China Computer Newspaper (2000)
28. Hunan leaf; Built-in hard drive, portable music player “iPod” newsletters (2002)
29. Degen, G.G.: AVR and CPLD development combining experimental platform. Newsletters (2003)
30. Li, M.: iPod iTunes model can domestic flowering. China Press reported (2009)

Operating Policy of Short-Term Power Dispatch for a Wind Farm

Cheng-Xiang Sun[1], Wenxia Pan[2], and Gengtao Guo[2]

[1] Lianyungang Normal College, 222006 Lianyungang, China
[2] Hohai University, 210098, Nanjing, China

Abstract. This paper investigates diesel units (DU)-battery energy storage systems (BESS) scheme to enable short-term dispatch commitment from a grid-connected wind turbine generating power station.The scheme is supplied by two distributed resources (DR) units, i.e., a synchronous machine and an electronically interfaced BESS. An operating policy to schedule short-term power dispatch for the scheme based on wind power forecast of the next interval,load demand and the charging/discharging characteristics of BESS is proposed. The studies show that an appropriate control strategy which by tracking active currents of BESS can ensure an acceptable level of dispatch security while maximizing wind energy harness.

Keywords: Battery energy storage systems (BESS), operating policy, PID neural network (PIDNN), power generation dispatch.

1 Introduction

During the past several decades, large amount of natural resources have been unlimitedly consumed, and our living environment has been severely destroyed. Global environmental protection concepts and concerns have been widely excited and several new forms of renewable resources such as photovoltaic systems and wind power generation systems to supplement fossil fuels have been developed in the whole world [1]. However, although a favourable wind power potential might exist, the constraints introduced by the time-variability of the harnessed wind energy result in relatively low wind penetration levels, typically up to 15%–20% of the grid system peak demand[2]. Large variations in the injected wind power could impact negatively on supply reliability and network security, with the result that wind power generation is often considered nondispatchable. It is even claimed that the renewable source will never be able to contribute significantly toward utility-scale supply or provide base-load power [3].

To overcome such limitations, introduction of energy storage is necessary [4],which counter-balance the inherent intermittency of wind power supply and permit its adaptation to the load profile. A hybrid power generation/energy storage system may combine available wind energy associated with available energy storage units. Design, simulation, and optimization of hybrid power systems have been the subject of several publications [5-13]. Reference [5] focus on a control strategy for wind farm dispatching

Y. Yang, M. Ma, and B. Liu (Eds.): ICICA 2013, Part II, CCIS 392, pp. 126–134, 2013.

has been developed in which a BESS has been used to smooth out the fluctuating wind power. In [6], where it is shown how suitably-designed ESS can bring additional economic benefit to the power generation endeavor is investigated. A dual-BESS scheme has been developed in [7] in which one BESS is to be charged by the wind, while concurrently, the other BESS is discharging power to the grid. In [8-13], the optimal design of a hybrid system is investigated, applying a variety of heuristic optimization techniques.

However, wind power dispatch has neither been fully understood nor investigated yet.In view of the above,the objectives of this paper are:

- To demonstrate the proposed hybrid power generation/energy storage system comprising DU and BESS.
- To illustrate that a properly controlled BESS unit which is interfaced to the grid with a power electronic converter can determine the short-term power dispatch strategy which will achieve an acceptable level of dispatch security while maximizing the wind energy harness?
- To highlight further areas for investigation.

2 Configuration of the Study System

The configuration of the proposed hybrid power generation/energy storage system is shown in Fig. 1. It includes a conventional autonomous power stations (APS) with diesel units (DU) supplying the load demand. A hybrid station is considered, consisting of wind generators, as the primary energy source, and a BESS system to balance the intermittent wind energy source supply.

The net power generation Pnet (i.e., the output power from the DU-BESS scheme)shown in Fig. 1 is determined by: 1) the output power from WTG, 2) the output power of DU, 3) the output power of dc–ac converter connected to BESS. The arrows indicate assumed direction of power flows. If power losses of converters are ignored, the expression for Pnet is given by

$$P_{net} = P_w + P_{DU} \pm P_{BESS} \tag{1}$$

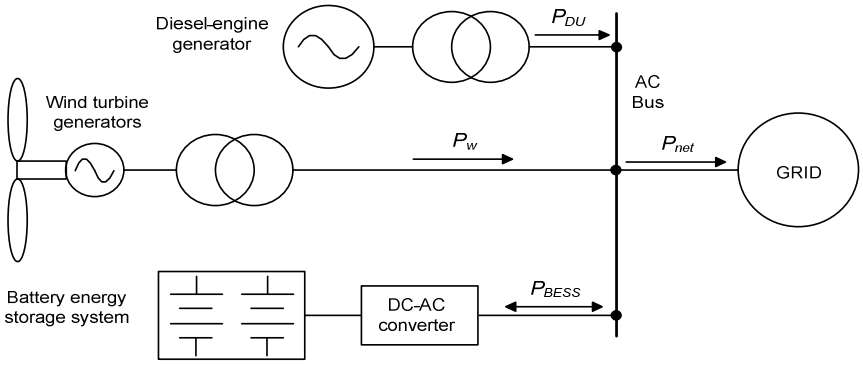


Fig. 1. Configuration of the hybrid power generation/energy storage system

3 Operating Policy

To conduct the proposed studies, This paper proposes scheduling power smoothing and peaking operating strategy,which relies on a basic decision making interval of 2 h. Unit commitment decisions are taken half an hour before each interval, to provide enough time for the startup of DU. The strategy based on wind power forecast of the next interval,load demand and the charging/discharging characteristics of BESS.

The following two basic operating modes are foreseen for the scheduling power smoothing and peaking operating strategy:

Mode 1: "Smooth Control" mode of operation where unit commitments are conducted to reduce the wind farm output power fluctuation. In this case, the batteries are sufficiently charged to undertake regulation of the system and provide the required reserve.

Mode 2: "Peak Load Regulating" mode, where BESS capacity (discharge power and energy) is available to meet the short-term power grid peaking requirement. In this case, if the BESS system has reached its lowest acceptable state of charge (SOC) and requires recharging, DUs are dispatched for the additional power grid peaking requirement, to operate in parallel to the BESS and wind turbine. Available wind power is fed to the load, while excess power charges the batteries.

After the decision making stage, the phase of actual operation follows. During this phase, the wind power absorption capability of the system is evaluated and set-points are dispatched to the wind power stations to curtail power, if necessary. Operation in each Mode is explained in the flowcharts presented in Fig. 2 for Mode 1, and in Fig. 3 for Mode 2 (which differs from each other only in the scheduling phase).

First,two parameters are cosidered, SOCmin, which is the limit under which the BESS is considered fully discharged and requires recharging.and SOCmax, the limit above which the BESS system is allowed to discharge, supplying the load.

In Mode 1 operation, the BESS system undertakes frequency and voltage regulation and provides the required reserve to compensate for wind power variations. The wind power absorption limit is given by:

In Mode 2 operation,

$$P_{GRID,\max}^{Mode1} = P_w + P_{BESS} + P_{DU} \tag{2}$$

$$P_{GRID,\max}^{Mode2} = P_w + P_{BESS} + P_{DU} \tag{3}$$

$$P_{GRID,\max}^{Mode2} = P_D \pm 0.1P_D \tag{4}$$

$$P_{GRID,\max}^{Mode2} = \max\left\{P_{MIN}, \min\left\{P_{GRID,\max}^{Mode2}, P_{MAX}\right\}\right\} \tag{5}$$

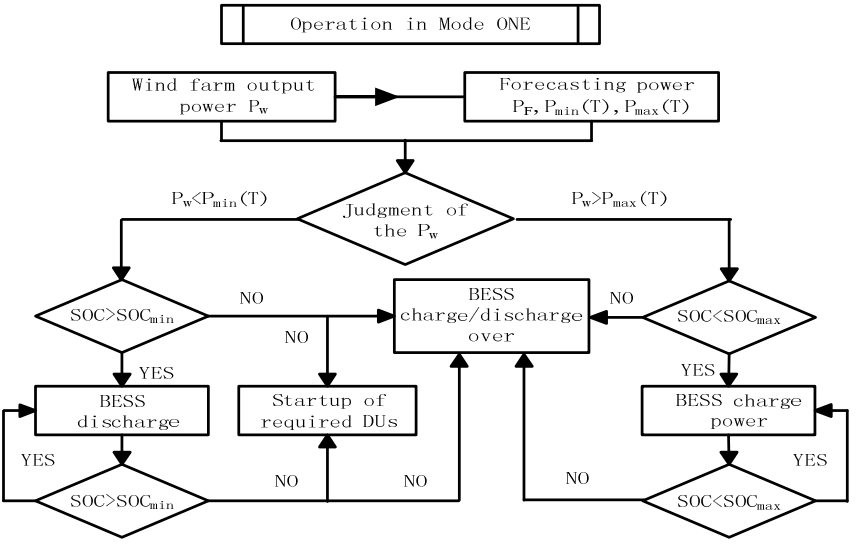


Fig. 2. Flowchart for operating "Smooth Control" mode

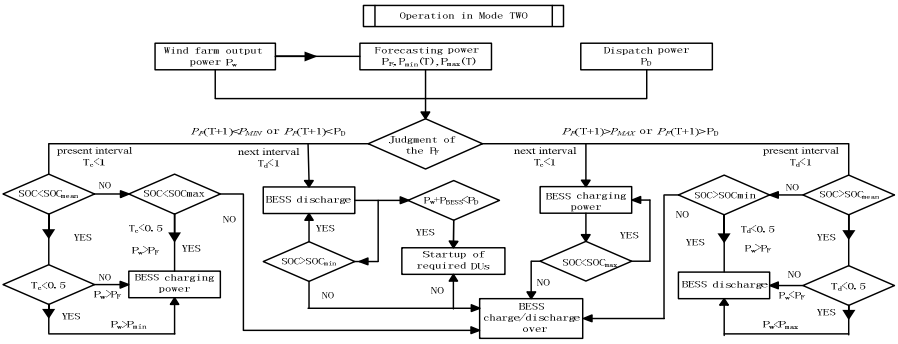


Fig. 3. Flowchart for operating "Peak Load Regulating" mode

4 Study Case

To determine the dispatch power schedule described in the previous section, several case studies are conducted to examine the DU-BESS scheme operation in power grid directly mode, power grid with BESS mode, scheduling power smoothing and peaking mode and in scheduling power smoothing and peaking with dispatch comand mode. The design of the dispatch is based on the actual wind speed shown in Fig. 4. A typical 2-MW WTG is adopted and the output power is shown in Fig. 5.

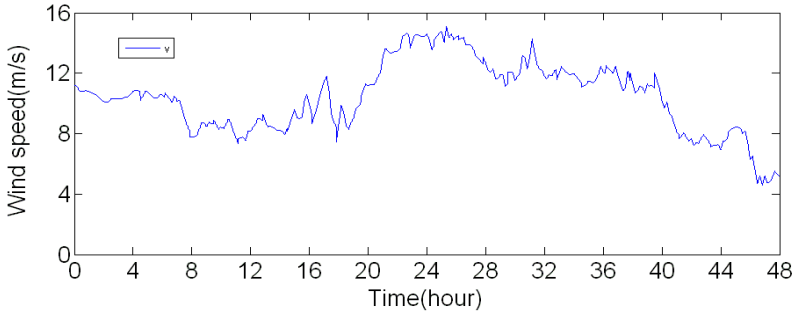


Fig. 4. Wind speed profile over a 48-hour period in a certain coast wind farm

Fig.4 shows the wind speed profile over a 48-hour period in a certain coast wind farm. The least wind speed was 5 meters per second ,the highest was 15 meters per second.

Case (1): Power Grid Directly Mode:

The objective of this case study is to illustrate how the dispatch can be affected without BESS.

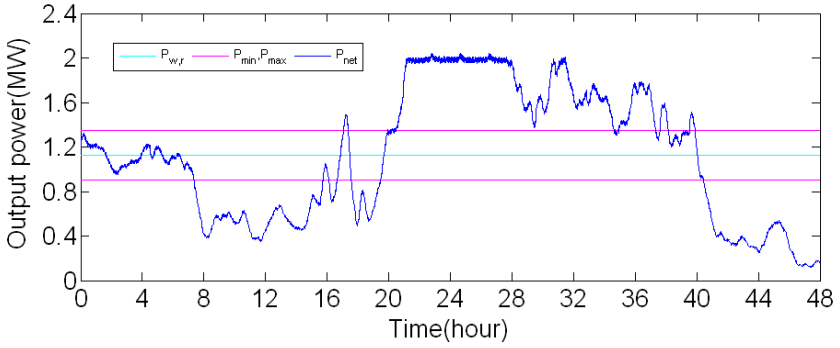


Fig. 5. A certain wind farm 2 MW wind generator's output power after the limitation within 48 hours

Fig.5 illustrates the impact of limitation on the 2MW wind generator's output power. In this study, power fluctuation rate ofλP = 20% is adopted, from the figure it is clear that,during 12-20 hour period and 44-48 hour period, wind generator's output power is so small that the possibility of power system islanding appears,however, a large amount of energy will be rejected during 24-44 hour period.

Case (2): Power Grid with BESS Mode:

To improve wind generator's power quality and efficiency, introduction of energy storage is necessary. Fig.6 (b) shows the battery SOC level. The BESS with 1KV voltage ,5000Ah capacity and 0.6MW charge -discharge power.The initial SOC is 0.5, i.e., where is half of the whole BESS energy.To prolong the battery lifetime, SOCmin is 0.1.

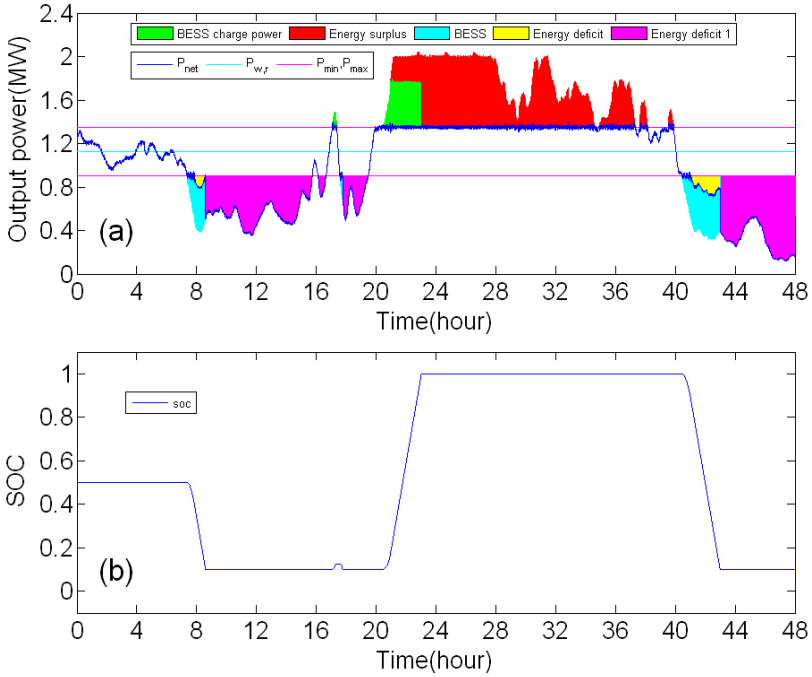


Fig. 6. System operating in power grid with BESS mode during a typical period. (a) Contribution to the grid (Concept of power and energy deviations). (b) SOC level.

In this section, simulation results are presented for the base case scenario shown in Fig.6, the operation of the system in power grid with BESS mode during a typical period is presented, including the contribution of all units to the grid, the SOC variation of the BESS.

Case (3): Scheduling Power Smoothing and Peaking Mode:

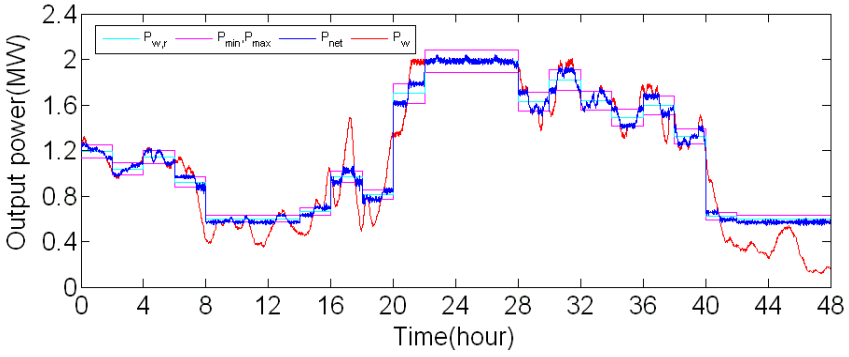


Fig. 7. Wind generator's output power by the adoption of scheduling power smoothing and peaking approach

In Fig.8(a), green areas represent the BESS charge power, Red areas represent the energy surplus,in this case, the batteries are practically fully charged, load is met by wind power, while excess wind generation is rejected. Cyan areas represent the energy deficit to the Pmin (lower limit of power smooth), which is covered by BESS, including the contribution of DU at around 12 o'clock and during 24-44 hour period.

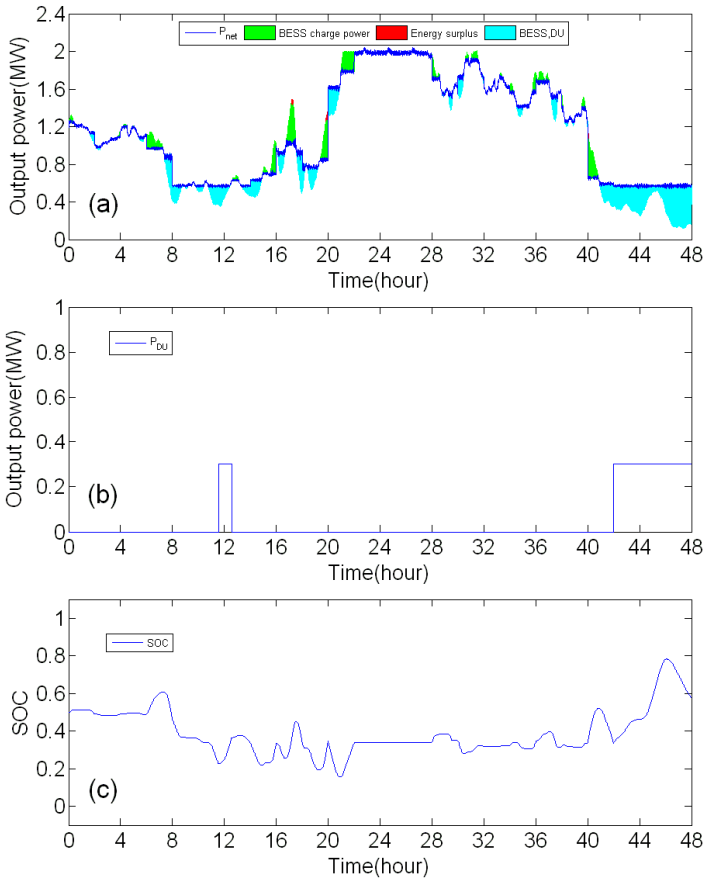


Fig. 8. System operating in the scheduling power smoothing and peaking mode. (a) Contribution to the grid(Concept of power and energy deviations). (b) DU output power (c) SOC level.

Secure load supply can be effected either by the BESS system, within the limits of its capacity, or by DUs. When wind generator output lower power, due to the limited BESS capacity, transfers "Smooth Control" mode to "Peak Load Regulating" mode of scheduling power smoothing and peaking strategy To cater for "Peak Load Regulating" mode requires the startup of DUs, choosing a 800KW diesel generators here. Fig.8(b) illustrate that steady 300kW power output last 1 hour at around 12 o'clock,and 6 hours during 42-48 hour period.In this case, BESS regulate power fluctuations,DUs output

constant power, additional demand, in excess of the BESS capacity, is covered by DUs. This mode favours both eases the control of DUs's output power and greatly improved the efficiency of DUs.

5 Simulation Results

In this section, various operating mode simulation results are presented for the hybrid system shown in Table 1.

Table 1. Summary Of Various Operating Mode Results

Contribution / Study Case	E_w (MWh)	E_{net} (MWh)	E_{BESS} (MWh)	E_{DU} (MWh)	E_R (MWh)	λ_P (%)
Case (1)	54	54	0	0	0	82%
Case (2)	54	47.8752	0.4	0	6.5248	20%
Case (3)	54	55.9200	-0.0764	2.100	0.1036	10.20%

Under lower wind speed condition, DU is adopted to improve the reliability of grid-connected wind power generation systems. DU is only available to PMIN, but wouldn't have favor to reduce rejected electrical energy much.

6 Conclusion

The application of hybrid DU-BESS scheme is designed to achieve an acceptable level of dispatch security has been investigated in the paper. The system is supplied by two distributed resources (DR) units, i.e., a synchronous machine and an electronically interfaced BESS. The former DG unit is equipped with excitation and governor controls and the latter BESS unit includes fast, independent real and reactive power controls. The simulation studies show that the latter BESS unit:

i) Can maintain stability of the system, mainly through reducing wind power fluctuation rate effectively,

ii) Can reduce the generation cost of the system and improve the wind electricity utilizing efficiency, through reducing BESS capacity and rejected electrical energy, respectively,

References

1. Lee, D.-J., Wang, L.: Small-Signal Stability Analysis of an Autonomous Hybrid Renewable Energy Power Generation/Energy Storage System Part I: Time-Domain Simulations. IEEE Trans.Energy Conver. 23(1), 311–319 (2008)
2. Vrettos, E.I., Papathanassiou, S.A.: Operating Policy and Optimal Sizing of a High Penetration RES-BESS System for Small Isolated Grids. IEEE Trans. Energy Conver. 26(3), 744–756 (2011)

3. Sovacool, B.K.: The intermittency of wind,solar,and renewable electricity generators:Technical barrier or rhetorical excuse? Utility Policy 17(3-4), 288–296 (2009)
4. Boyle, G.: Renewable Electricity and the Grid: Challenge of Variability. Earthscan, Sterling (2007)
5. Teleke, S., Baran, M.E., Huang, A.Q., Bhattacharya, S., Anderson, L.: Control Strategies for Battery Energy Storage for Wind Farm Dispatching. IEEE Trans. Energy Conver. 24(3), 725–732 (2009)
6. Voller, S., Al-Awaad, A.R., Verstege, J.F.: Benefits of energy storages for wind power trading. In: Proc. 2008 IEEE International Conference on Sustainable Energy Technologies, pp. 702–706 (2008)
7. Yao, D.L., Choi, S.S., Tseng, K.J., Lie, T.T.: Determination of Short-Term Power Dispatch Schedule for a Wind Farm Incorporated With Dual-Battery Energy Storage Scheme. IEEE Trans. Energy Conver. 24, 3(1), 74–83 (2012)
8. Wang, C., Nehrir, M.H.: Power management of a stand-alone wind/photovoltaic/fuel cell energy system. IEEE Trans. Energy Convers. 23(3), 957–967 (2008)
9. Fernandes, L., Figueiredo, J.: An economic analysis of a power production hybrid system. In: Proc. 5th Int. Conf. Eur. Electr. Market, Lisboa (May 2008)
10. Mitchell, K., Nagrial, M., Rizk, J.: Simulation and optimization of renewable energy systems. Int. J. Electr. Power Energy Syst. 27(3), 177–188 (2005)
11. Valenciaga, F., Puleston, P.F.: Supervisor control for a stand-alone hybrid generation system using wind and photovoltaic energy. IEEE Trans. Energy Convers. 20(2), 398–405 (2005)
12. Wies, R.W., Johnson, R.A., Agrawal, A.N., Chubb, T.J.: Simulink model for economic analysis and environmental impacts of a PV with diesel-battery system for remote villages. IEEE Trans. Power Syst. 20(2), 692–700 (2005)
13. Emadi, A., Williamson, S.S., Khaligh, A.: Power electronics intensive solutions for advanced electric, hybrid electric, and fuel cell vehicular power systems. IEEE Trans. Power Electron. 21(3), 567–577 (2006)

Mechanical Virtual Experiment Based on Flash Technology

Limei Wang

Siping Vocational College, Siping, Jilin, 136002 China

Abstract. Virtual experiment teaching platform can effectively solve the problem of insufficient funding experiments. Articles can be developed advanced complexity equipment through the virtual platform, to allow students a deep understanding of the image of the device body movement principle, to acquire knowledge itself; the other hand, the platform is a good solution to student learning in time and space on limitations. This paper analyzes the development of virtual experiment at home and abroad, through the comparison of various learning theories and studies comparing various virtual reality technologies, the characteristics for the mechanical discipline. The paper developed a mechanical virtual experimental teaching platform based on Flash technology, the device will integrated techniques of self-observation, self-disassembly, self-planning connection, 3D simulation, automatic calculation of the corresponding motion analysis in one, with a visual, interactive, scalable, and many other advantages that can be developed .

Keywords: experimental teaching, virtual experiment, Flash technology.

1 Introduction

The development of modern educational technology in teaching computer virtual experiment is widely applied [1-3], especially in some universities As the school resulted in the rapid expansion of the scale of the corresponding matching funds can not be put in place, there has been lack of practice teaching conditions and aging equipment update inconvenience and other problems face reduced funding for colleges and universities, to improve the quality of teaching and other policies, the traditional experimental teaching model has been a huge challenge, then the virtual experiment teaching has become a good choice. present our virtual experiment in teaching the study also a certain lack of systematic and good humanized design [4-7], and the development of virtual experiment experimental teachers in general there is a certain degree of difficulty in the virtual experiment progresses, the virtual instrument for the experimenter to correct all operations the response and the experimenter to get exactly the same manner as the actual experimental experience, is to build the core virtual experiments, even more important is the content according to the specific virtual experiments select the appropriate construction techniques [8-12]. article experimental teaching of mechanical prototype, using Flash technology for the design platform, combined with Pro/E, AutoCAD and other machinery commonly used software, the

Y. Yang, M. Ma, and B. Liu (Eds.): ICICA 2013, Part II, CCIS 392, pp. 135–144, 2013.

use of modern educational technology, design and develop a virtual experiment in which five machines.

2 Design and Development of Virtual Experiment

Article takes virtual simulation technology into teaching, mainly tries to develop the institutions and agencies of cognitive experiments [13, 14], Kinematic Diagram experiment, and four kinematic principle experiment, involutes gear experimental and statistical analysis of experimental machining errors, certain extent, solve the problem of insufficient laboratory equipment, but also to improve the student's positive initiative, and achieved good teaching effect. experimental system developed by the design of the general structure is shown in Figure 1, to help some of the major shows the experimental system operating environment and software use, the significance of the relevant interface and buttons, etc.; experimental guidance section describes the purpose of the experiment, principles, methods, procedures, experimental significance, and provide test report single download; experimental part is the main part of virtual experiments, according to experimental purposes and requirements of completion of the relevant experiments; data analysis part is a relatively large amount of data for some experiments designed to further experimental research or for teachers on student test report marking convenience.

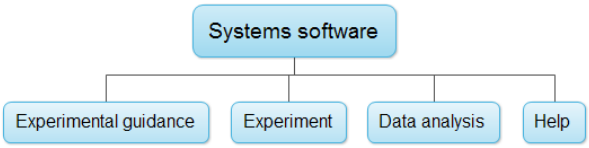


Fig. 1. The general structure of system software

2.1 Mechanism and the Mechanism of Cognitive Experiments

Institutions and cognition are commonly observed planar linkage, spatial linkage, cam, gear, gear train, intermittent motion mechanism and the combination type of institution and movement [15-18]. Bodies, machines, exercise deputy member has a certain understanding for us, and the basic elements of the machine have a preliminary understanding to students. Traditional experimental methods are generally visited mechanical teach showcase to complete, but the limitations of teaching and updating showcase growing problem show.

Experimental Center through Pro / E, UG, AutoCAD and other software on the part of institutions to establish a three-dimensional model, through Flash technology for the design platform, the use of XML, Web and other technologies, the development of institutions and agencies cognitive experimental system, which from the multi-angle shows the structural characteristics of institutions or movement characteristics, the system can be realized with a link to an external file system to facilitate the achievement of institutional repositories expand or update Figure 2 (left) selection screen for the organization of cognitive model experimental system, click on the selected picture the agency will enter a three-dimensional

display interface; Figure 2 (right) is the use of Pro/E assembly exploded diagram created in the system, which can see the assembly of the agencies teach animation, a clear understanding of body assembly process, enabling organizations consisting of cognitive, reaching experimental teaching purposes.

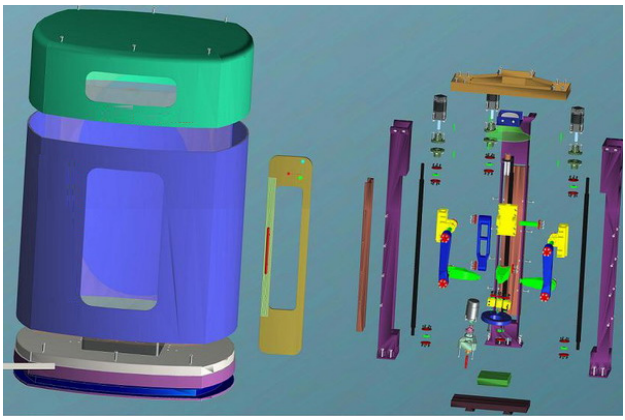


Fig. 2. Mechanism and experimental system

2.2 The Motion Diagram of Mechanism

Experiments require students to be familiar mechanism motion mapping methods, master mapping agencies for the actual mechanism motion map skills [19-23]. Experiment with how kinematic mechanical structure diagram and the difference between the actual and further deepen our understanding of body composition principle, Student can calculate the degree of freedom and their meanings.

Fig. 3. Mechanism of and mapping experiment system

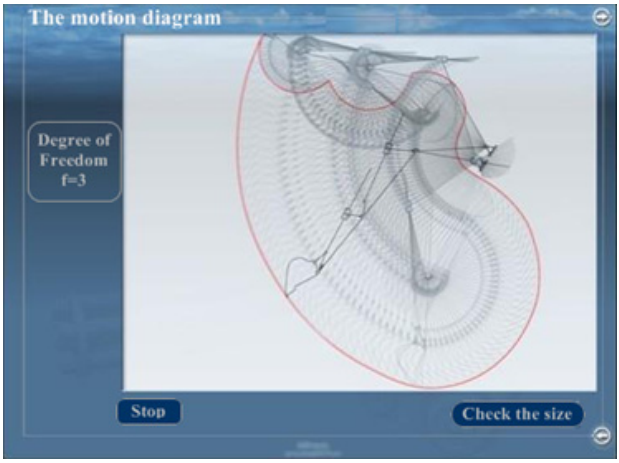


Fig. 4. The excavator arm movement

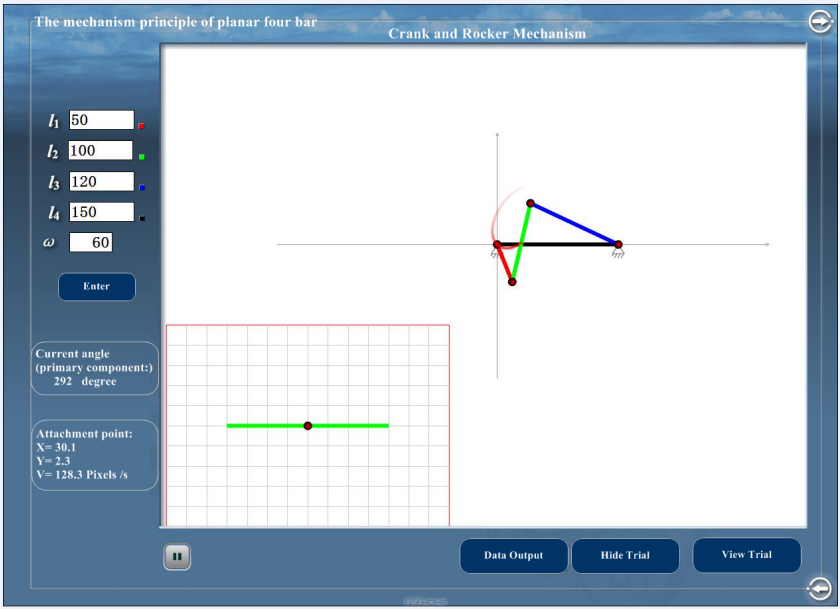


Fig. 5. Different design parameters into gear after the overall comparison

Developed mapping system kinematic diagrams, frameworks and institutions and agencies cognitive experimental system is similar, but we adds mapping and reference results through password access and many other features. Figure 3 shows the system during the measurement, the measured value will be displayed in the left panel; Figure 4 shows a schematic diagram of the excavator arm movement, body movement through dynamic diagram, play with the gradual control, clearly show the significance of the degree of freedom, and enhance the understanding of kinematic diagrams, to experimental teaching purposes.

2.3 Experiment of Four Bar Kinematic Principle

Four mechanism motion experiments are to enable students to master institution and its law of motion under different conditions of the formation for a long rod type. the development of mechanism motion principle experiment system achieve a body movement accurate simulation, according to different rod lengths automatically determine the four types of institutions according to their law of motion exercise, can dynamically adjust the speed and direction of movement of bodies (enter a negative movement in the opposite direction), you can display any link and its vicinity point trajectory, velocity and acceleration of the four institutions trajectory study has important significance.

2.4 Involutes Gear Experiment

Involutes gear experimental system can overcome the lack of the traditional mechanical meter range. We can energy intuitive, dynamic simulation of gear machining situation and to demonstrate the gear design parameters for involutes tooth profile curve, in order to better achieve the experimental teaching purposes. the experimental system can achieve different design parameters tooth profile comparison, either during the process or in the tooth profile of the comparison process, accessible by clicking the left mouse button, and with the middle mouse button wheel to achieve gear pan and zoom, easy to achieve timely content viewport observed more clearly show the shape of the various

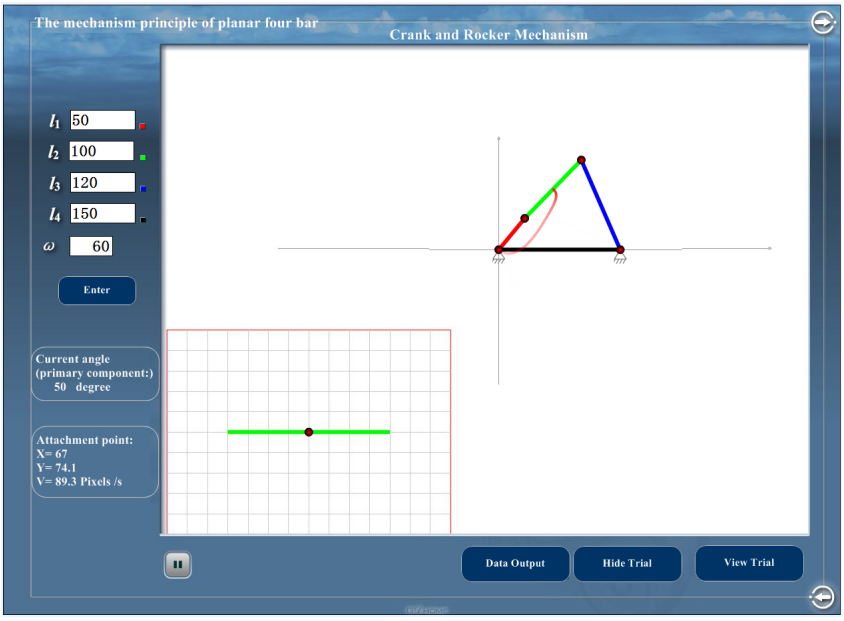


Fig. 6. A common geometric parameters of the gear

parts of the tooth profile Figure 5 range for different design parameters into gear after the overall comparison; Figure 6 is determined by visual mouth scaling function displays the shape of tooth profile can clearly see which undercut phenomenon, when the mouse pointer on the appropriate gear will show a common geometric parameters of the gear.

2.5 Statistical Analysis of Experimental Machining Error

Statistical analysis is done by a number of workpiece machining error manifestation to study the cause of error method. Machining error statistical analysis of experimental aims to consolidate the already learned the basic theory of statistical analysis, master the use of statistical analysis method steps. Through this experiment, students are required to master the actual size of the workpiece drawing maps of the method, and according to the distribution analysis of the nature of the error processing, computing process capability factor, rate of qualified products and scrap rates, etc., can put forward process improvement measures; master drawing XR dot pattern method, according to the process XR point of the stability diagram. Figure 7 sample data for the distribution obtained after the treatment, with a combined error statistics for the normal case; Figure 8 graphics for the point, when the points beyond the control line, that will be a special display, move the mouse to each point, the point will be able to display the relevant information.

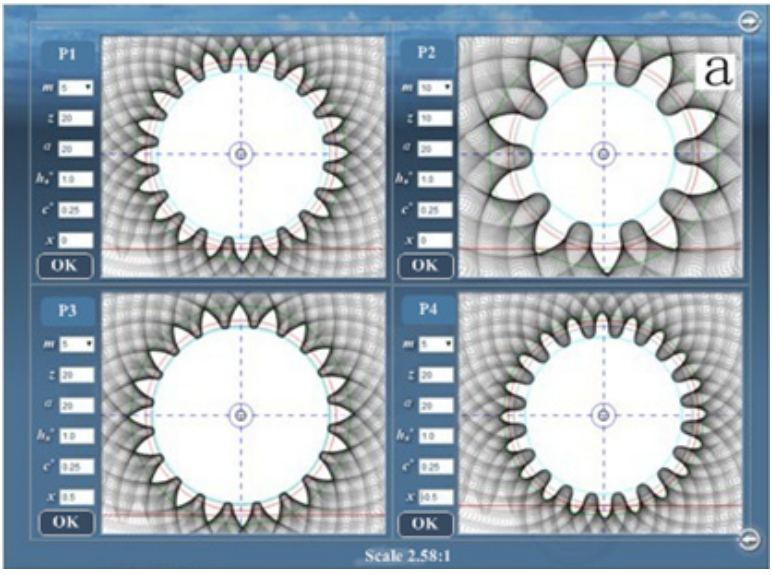


Fig. 7. Tooth profile Comparator Interface

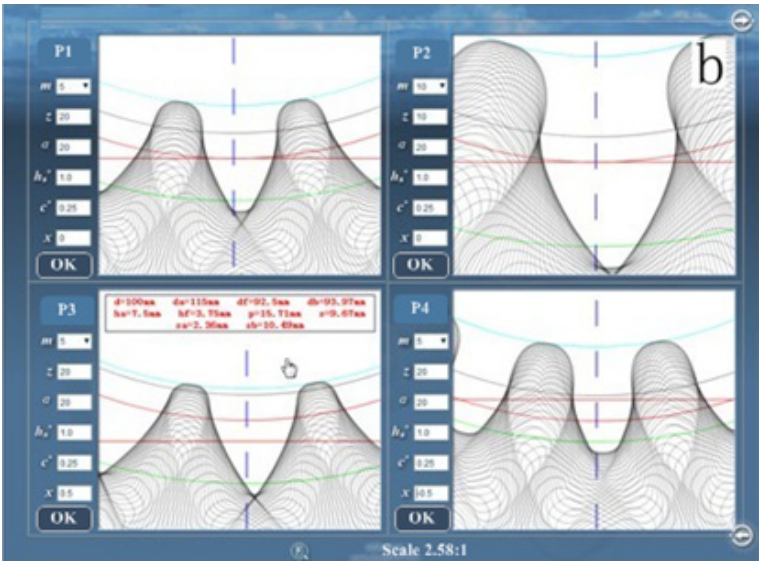


Fig. 7. (*Continued.*)

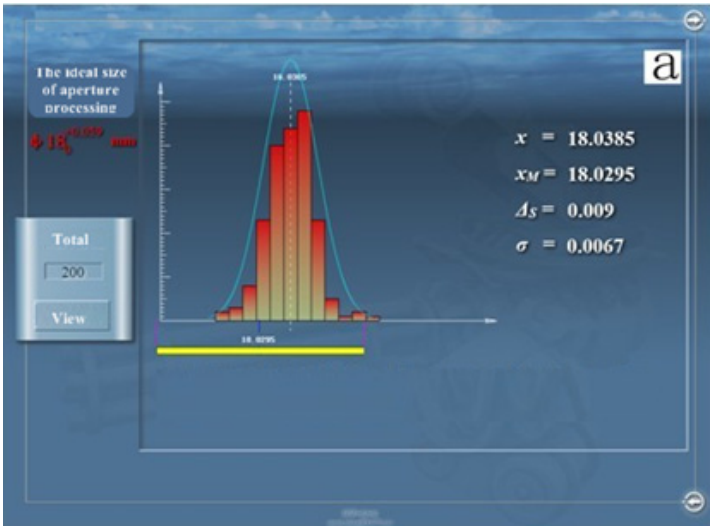


Fig. 8. Results of data processing

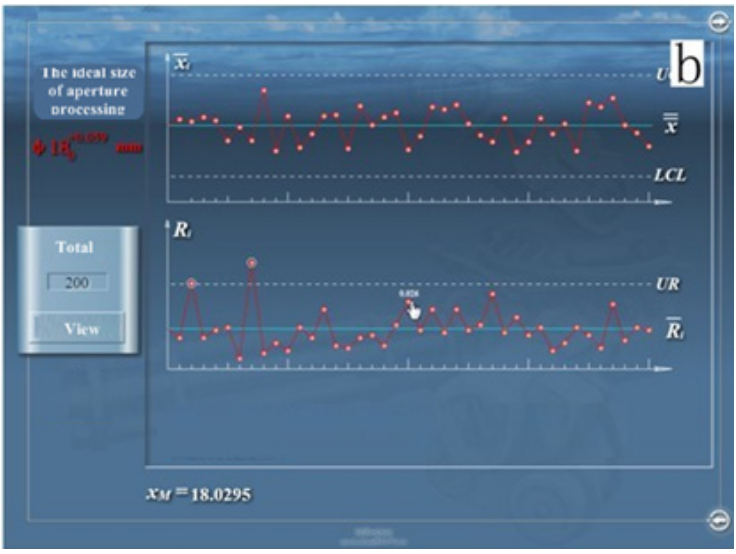


Fig. 8. (*Continued.*)

3 The Construction of Virtual Laboratory

Virtual Experiment Center create virtual experiments, carried out experimental teaching network, that can enrich the learning resources, alleviate the lack of equipment, but also reflects the personalized learning environment, improve teaching model, to improve teaching effectiveness. According to existing laboratories and experimental development project information, initial ideas mechanical engineering virtual laboratory structure is shown in Figure 9. Main function is to display institutions, virtual simulation and performance testing, along with the deepening of the reform of experimental teaching; experimental type and quantity will gradually increase, eventually forming the perfect virtual experiment teaching system.

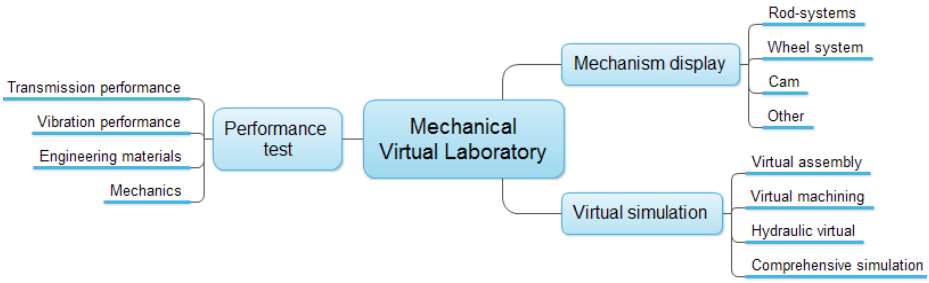


Fig. 9. The structure of Virtual Laboratory

4 Conclusions

Virtual experiments overcome the shortcomings of traditional test, which has become a powerful supplement traditional experimental teaching. Virtual experiment teaching enhances students' interest in experiment, reducing the cost of experiments, the overall teaching better. And virtual experiment software with small, high quality, fast speed, and easy operation, with a strong portability, interactivity and simulation is very suitable online learning based on flash technology.

References

1. Cai, W.: Experiment Teaching of Virtual Simulation Technology in Mechanical Engineering. Experimental Technology and Management 28(8), 76–78 (2011)
2. Ping, G.G., Song, N.: Virtual Experimental Teaching Status and Problems. Laboratory Science 13(5), 175–178 (2010)
3. Wei, Y.: Analysis and Research of Virtual Experiment. Science and Technology Information (35), 5–6 (2010)
4. Wei, G., Ke, J.: Design and development Kinematic Diagram of experimental Based on Flash Action Script. Computer and Digital Engineering 40(2), 105–107 (2012)
5. Agrawal, N., Prabhakaran, V., Wobber, T., Davis, J.D., Manasse, M.S., Panigrahy, R.: Design tradeoffs for ssd performance. In: Proceedings of the USENIX Annual Technical Conference, pp. 57–70 (June 2008)
6. Chang, Y.-H., Hsieh, J.-W., Kuo, T.-W.: Endurance enhancement of flash-memory storage systems: An efficient static wear leveling design. In: Proceedings of the 44th Annual Conference on Design Automation, pp. 212–217. ACM, New York (2007)
7. Chung, T.-S., Park, D.-J., Park, S., Lee, D.-H., Lee, S.-W., Song, H.-J.: System Software for Flash Memory: A Survey. In: Sha, E., Han, S.-K., Xu, C.-Z., Kim, M.-H., Yang, L.T., Xiao, B. (eds.) EUC 2006. LNCS, vol. 4096, pp. 394–404. Springer, Heidelberg (2006)
8. Gal, E., Toledo, S.: Algorithms and Data Structures for Flash Memories. ACM Computing Survey 37(2), 138–163 (2005)
9. Ganger, G., Worthington, B., Patt, Y.: The DiskSim Simulation Environment Version 3.0 Reference Manual
10. Page-level Address Mappings. In: Proceedings of the International Conference on Architectural Support for Programming Languages and Operating System (ASPLOS), pp. 229–240 (March 2009)
11. Kang, J., Jo, H., Kim, J., Lee, J.: A Superblock-based Flash Translation Layer for NAND Flash Memory. In: Proceedings of the International Conference on Embedded Software (EMSOFT), pp. 161–170 (October 2006) (Pubitemid 47168585)
12. Kim, J., Kim, J., Noh, S., Min, S., Cho, Y.: A Space-Efficient Flash Translation Layer for Compactflash Systems. IEEE Transactions on Consumer Electronics 48(2), 366–375 (2002)
13. Gupta, A., Kim, Y., Urgaonkar, B.: DFTL: A Flash Translation Layer Employing Demand-based Selective Caching of Page-level Address Mappings. In: Proceedings of the International Conference on Architectural Support for Programming Languages and Operating System (ASPLOS), pp. 229–240 (March 2009)
14. Kang, J., Jo, H., Kim, J., Lee, J.: A Superblock-based Flash Translation Layer for NAND Flash Memory. In: Proceedings of the International Conference on Embedded Software (EMSOFT), pp. 161–170 (October 2006) (Pubitemid 47168585)

15. Kim, Y., Gurumurthi, S., Sivasubramaniam, A.: Understanding the Performance-Temperature Interactions in Disk I/O of Server Workloads. In: Proceedings of the International Symposium on High- Performance Computer Architecture (HPCA) (Febuary 2006)
16. Kim, Y., Tauras, B., Gupta, A., Urgaonkar, B.: FlashSim: A Simulator for NAND Flash-based Solid-State Drives. In: Technical Report CSE 09-008, Penn State University (May 2009)
17. Lee, J., Byun, E., Park, H., Choi, J., Lee, D., Noh, S.H.: CPS-SIM: Configurable and accurate clock precision solid state drive simulator. In: Proceedings of the Annual ACM Symposium on Applied Computing (SAC), pp. 318–325 (March 2009)
18. Lee, S., Moon, B.: Design of Flash-based DBMS: An In-Page Logging Approach. In: Proceedings of the International Conference on Management of Data (SIGMOD), pp. 55–66 (August 2007)
19. Agrawal, N., Prabhakaran, V., Wobber, T., Davis, J.D., Manasse, M.S., Panigrahy, R.: Design tradeoffs for ssd performance. In: Proceedings of the USENIX Annual Technical Conference, pp. 57–70 (June 2008)
20. Chang, Y.-H., Hsieh, J.-W., Kuo, T.-W.: Endurance enhancement of flash-memory storage systems: An efficient static wear leveling design. In: Proceedings of the 44th Annual Conference on Design Automation, pp. 212–217. ACM, New York (2007)
21. Chung, T.-S., Park, D.-J., Park, S., Lee, D.-H., Lee, S.-W., Song, H.-J.: System Software for Flash Memory: A Survey. In: Sha, E., Han, S.-K., Xu, C.-Z., Kim, M.-H., Yang, L.T., Xiao, B. (eds.) EUC 2006. LNCS, vol. 4096, pp. 394–404. Springer, Heidelberg (2006)
22. Gal, E., Toledo, S.: Algorithms and Data Structures for Flash Memories. ACM Computing Survey 37(2), 138–163 (2005)
23. Ganger, G., Worthington, B., Patt, Y.: The DiskSim Simulation Environment Version 3.0 Reference Manual

Digital Library Based on Cloud Computing

Fu Jia[1], Yan Shi[2], and Li Tongwei[3]

[1] Library of Mudanjiang Medical University, Mudanjiang, Heilongjiang, China, 157011
[2] Information Technology Center of Mudanjiang Medical University, Mudanjiang, Heilongjiang, China, 157011
[3] Hongqi Hospital of Mudanjiang Medical College, Mudanjiang, Heilongjiang, China, 157011

Abstract. In order to improve the intelligent and personalized services, we implement the intelligent service of digital library by the strong cloud computing and background processing ability. First carry out detailed introduction on the adaptive systematic content and form of digital library under cloud environment, then expose the key modules in detail, finally using the example simulation to verify the performance of this method in digital library intelligent and personalized service.

Keywords: Cloud computing, digital library, adaptive system, FCM clustering, Apriori association rules.

1 Introduction

Digital Library is a library of using digital technology to process and storage various illustrated literature, and which is essentially a distributed information system of multimedia production. It stores the information resources of all sorts of different carrier and the geographic location by using digital technology, for the purpose of regional across, object-oriented network query and spread. It involves information processing, storage, retrieval, transmission and utilization of resources. In layman's terms, digital library is a library of virtual, and without walls, which system of establishing the sharing and extensive knowledge network system based on network environment, and is a knowledge center of large scale, distributed, easy to use, without time and space limit, and can realize seamless link libraries and intelligent information retrieval. Cloud computing is Internet based related services increased, use and delivery mode, usually involved providing extensive, virtual resources through the Internet. Cloud computing will change the internet technology base, even affect the entire industry landscape. The emergence of cloud computing brings new vitality for the construction of digital library.

Cloud computing has brought many convenient for digital library construction, such as improve the service efficiency, expand the service resources, etc., this paper mainly studies how to use cloud computing to provide users with more efficient intelligent digital library service, how to help users to get value and interested data in the vast ocean of knowledge, which is the key of applying the cloud computing into digital library, which means use the powerful calculation ability of cloud computing to

Y. Yang, M. Ma, and B. Liu (Eds.): ICICA 2013, Part II, CCIS 392, pp. 145–155, 2013.

analysis and process in the background, provide more personalized service for digital library users, set up the different adaptive systems, meet the individual needs of different users, and provide accurate intelligence service.

2 Digital Library under Cloud Computing Environment

There are many difficulties that we should overcome on the construction of digital library, including the intelligent active technology, the digital library services are no longer completely on passive mode for users, which means it will actively provide services according to users' historical requirements and interest preference, this largely increases the interaction between library and user. The researched adaptive system implement the active service of digital library, considering the differences between users, it implements adaptive push according to different users, actively send the users' habit related resources to users, figure 1 is the adaptive system structure of digital library.

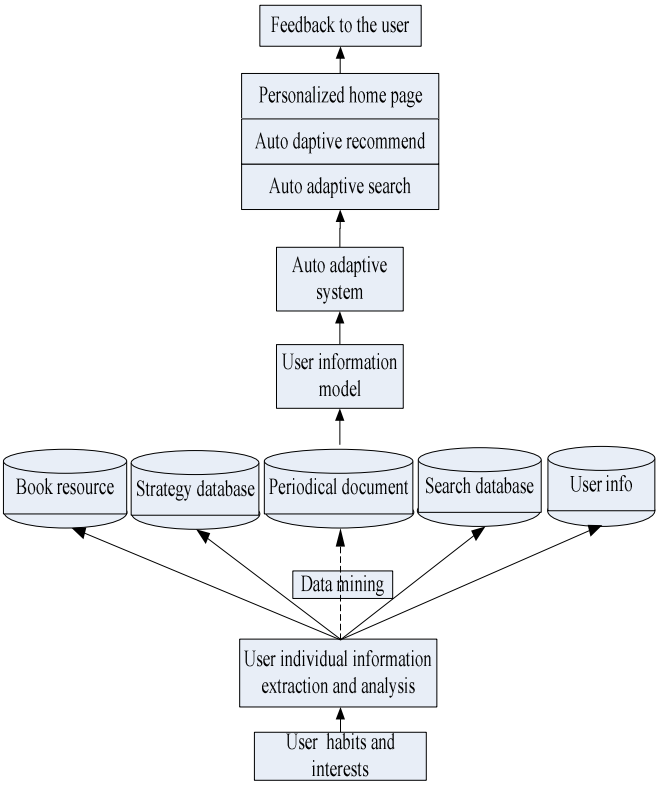


Fig. 1. Adaptive system structure of digital library

After extraction and analysis of users' habits and interests and combining the books repository, strategies data repository, periodical document repository, search repository, we use the data mining to analyze users' personal information model, launch adaptive system by combining user information model with the digital library resources, thus feed back the adaptive search, adaptive recommendations and personalized home page to users. The core part of whole structure is data mining, the clustering analysis and association rules processing according to all the users' operation after load on digital library is the most important part of adaptive system. How to effectively use the user irregular hit and download operation to analyze the user interested resources is also the research emphasis of the paper.

2.1 Adaptive System Effect

Adaptive system of digital library adds intelligent feedback on the basis of the traditional digital library and is responsible for the adaptive push, updates the reading contents, intelligent retrieval and personalized home page rendering. This adaptive system can better enhance the intelligent, personalized and interaction of digital library.

Other adaptive effects include personal home page custom made, personalized interface style, etc. Users' home page prompt the users' login time, users will pay attention to information update in the field, such as announcement of forthcoming books, resources preview, etc.

B. Main forms of self-adaption

B.A.Adaptive information retrieval

Information retrieval is the main method for users to find resources, most of the existing digital library does not take the user's difference into account, for all the users, the same input, the same feedback search structure, users need to search one by one in the ocean search results if they want to find their own needs, so the search efficiency is low. This method does not take user's knowledge background, interests, hobbies into account; adaptive search feed back different results according to different background and interests, the result that most close to the user professional will return on the previous page.

Adaptive retrieval list user retrieval conditions as a keyword list, combine the results by list keyword search with the user personalized information, finally to return search results to pages, the vector space and vector model can be used to node classification of users' information, and carry out result output after matches with keyword search results.

2.2 Adaptive Recommendations

Adaptive recommendations is the online personalized recommendation of user satisfied information based on user information, after users login on the library system, it carries out the user habit information online push, such as new books recommendation, the new literature, etc., which present no longer a unified login page; also for the offline recommendations, when there's user satisfied documents and materials, it will inform users timely by way of email and text message.

2.3 Key Module Implements of Adaptive Digital Library System

The key part of the adaptive system implementation of the digital library is user data mining, only a clearly classification of user data can realize adaptive system, and user data mining mainly needs to implement clustering and association [5]. In this paper, we use fuzzy clustering and Apriori algorithm.

2.4 User Data Standardization

Assuming the user information area $X = \{x_1, x_2, x_3, ..., x_n\}$ as the classified object, each object is measured by m's index, $x_i = \{x_{i1}, x_{i2}, x_{i3}, ..., x_{im}\}, i = 1, 2, ..., n,$ in which x1indicates the user, x2 indicates the password, x3 indicates the education background, x4 indicates the major, x5 indicates the CLC number that largely searched by users, x6 indicates the reference index of the library, x7 indicates the load record in different time of users, x8 indicates the material subject category that users loaded.

The original data matrix is

$$\begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1m} \\ x_{21} & x_{22} & \cdots & x_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ x_{n1} & x_{n2} & \cdots & x_{nm} \end{pmatrix} \tag{1}$$

After get the original data , we can carry out the data standard deviation transform :

$$x'_{ij} = \frac{x_{ij} - \overline{x}_j}{s_j} \quad (i = 1, 2, ..., n, j = 1, 2, ..., m) \tag{2}$$

In which $\overline{x}_j = \frac{1}{n}\sum_{i=1}^{n} x_{ij}$, $s_j = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_{ij} - \overline{x}_j)^2}$.

After the transform, all the variables unit dimension are removed, and are 0, standard error is 1, following is the range transformation:

$$x'_{ij} = \frac{x_{ij} - \min\{x_{ij} \mid 1 \le i \le n\}}{\max\{x_{ij} \mid 1 \le i \le n\} - \min\{x_{ij} \mid 1 \le i \le n\}} \tag{3}$$

After the transformation, all the variables values are between [0,1].

After the original matrix has been established, we can establish fuzzy similar matrix according to the original matrix, the fuzzy similar matrix indicates the similarity degree between various objects, $r_{ij} = R(x_i, x_j)$, there are mainly 6 method of calculating r_{ij}

Included angle cosine method to solve similarity factor

$$r_{ij} = \frac{\sum_{k=1}^{m} | x_{ik} - \overline{x}_i | | x_{jk} - \overline{x}_j |}{\sqrt{\sum_{k=1}^{m}(x_{ik} - \overline{x}_i)^2} \sqrt{\sum_{k=1}^{m}(x_{jk} - \overline{x}_j)^2}} \tag{4}$$

Correlation coefficient method to solve similarity factor

$$r_{ij} = \frac{\sum_{k=1}^{m} | x_{ik} - \overline{x}_i | | x_{jk} - \overline{x}_j |}{\sqrt{\sum_{k=1}^{m} (x_{ik} - \overline{x}_i)^2} \sqrt{\sum_{k=1}^{m} (x_{jk} - \overline{x}_j)^2}} \tag{5}$$

In which, $\overline{x_i} = \frac{1}{m}\sum_{k=1}^{m} x_{ik}, \overline{x_j} = \frac{1}{m}\sum_{k=1}^{m} x_{jk}$.

The maximum and minimum method

$$r_{ij} = \frac{\sum_{k=1}^{m} \min\{x_{ik}, x_{jk}\}}{\sum_{k=1}^{m} \max\{x_{ik}, x_{jk}\}}, i, j = 1, 2, \cdots, n \tag{6}$$

Arithmetic mean and minimum method

$$r_{ij} = \frac{\sum_{k=1}^{m} \min\{x_{ik}, x_{jk}\}}{\frac{1}{2}\sum_{k=1}^{m} \left(x_{ik} + x_{jk}\right)}, i, j = 1, 2, \cdots, n \tag{7}$$

Geometric mean and minimum method:

$$r_{ij} = \frac{\sum_{k=1}^{m} \min\{x_{ik}, x_{jk}\}}{\sum_{k=1}^{m} \sqrt{x_{ik} x_{jk}}}, i, j = 1, 2, \cdots, n \tag{8}$$

The fuzzy similarity matrix can be get after calculate rij, which is the similarity between different elements, such as we can now determine a user's major is for engineering, and computer specialty or communication major, which has a high similar degree between these two, but the confirmed specialty need to be sure by clustering.

2.5 FCM Clustering

Digital library clustering mainly realize three aspects of tasks: the library resource clustering, the user retrieval clustering, user recommend items clustering.

As for library resources cluster, we can get the requirement of different books by clustering analysis according to all the user's search condition, and the user's professional and academic, thus timely replenish supply corresponding books; Users retrieval clustering mainly record users' searching preferences to provide users with personalized service, clustering analysis is doing the classification according to the situation of users' multiple search and fuzzy matrix result, judging user retrieval classification belongs to which category; Users recommend clustering is to cluster

according to the user's access data, deciding to recommend users the relevant type of literature.

The paper uses fuzzy clustering C algorithm (FCM) to carry out users data clustering. First we need to confirm four basic definitions:

Definition 1 : uncompleted data sample set $X_A = \{x_1, x_2, ..., x_n\} \subset R^S$, $x_k = (x_{k1}, x_{k2}, ..., x_{ks})^T \in R^S$. Sample set X_A 's subset X_W and X_P , in which $X_W = \{x_k \mid x_k \in X_A \} 1 \le k \le n\}$, and x_k is a complete data sample, $X_P = \{x_i \mid x_i \in X_A \} 1 \le i \le n\}$, and x_i is a uncompleted data sample, $X_W \bigcap X_P = \varnothing$, $X_W \bigcup X_P = X_A$. Feature value set $X_M = \{x_{kj} \mid x_{kj} = ?1 \le j \le s, 1 \le k \le n\}$, $X_U = \{x_{kj} \mid x_{kj} = \text{confirm } 1 \le j \le s, 1 \le k \le n\}$, $X_M \bigcap X_U = \varnothing$, $|X_M| + |X_U| = |X_A|$,|| indicates the elements number in the set.

Definition 2 : data sample X_A 's complete ratio η :

$$\eta = \frac{n_u}{n_s} = \frac{|X_U|}{|X_A| \times s} \tag{9}$$

Definition 3: the affect factor α_k of data sample x_k on clustering analysis

$$\alpha_k = \begin{cases} 1 & , x_k \in X_W \\ 1 - \dfrac{|x_{kj}|}{|x_k|} & , x_k \in X_P, x_{kj} \in X_M, 1 \le j \le s \end{cases} \tag{10}$$

Definition 4: the similarity β_{ij} of data sample x_i and x_j :

$$\beta_{ij} = \| x_{ik} - x_{jk} \|, (1 \le k \le s) \wedge (x_i \in X_U) \wedge (x_j \in X_U) \tag{11}$$

Assuming the sample set $X = \{x_1, x_2, ..., x_n\} \subset R^S$ as a limited sample set of n's modal in the model space, in which $x_k = (x_{k1}, x_{k2}, ..., x_{ks})^T \in R^S$ as the character vector of data sample x_k , x_{kj} is the assigned value of dimension j of feature vector x_k [6]

Sample x_k and subset $x_i (1 \le i \le c)$'s relation is indicated by using $u_{ik} = u_{X_i}(x_k) (u_{ik} \in [0,1])$,in order to record the relation of multi subsets. We use matrix method to complete, marked as $U = [u_{ik}]_{c \times n}$.

X's fuzzy C dividing space M_{fc} :

$$\begin{aligned} M_{fc} &= \{U \in R^{cn} \mid u_{ik} \in [0,1], \forall i, k; \\ &\sum_{i=1}^{c} u_{ik} = 1, \forall k; 0 < \sum_{k=1}^{n} u_{ik} < n, \forall i\} \end{aligned} \tag{12}$$

In which clustering original model is $P=\{p_1, p_2, ..., p_c\}$, $p_i (i=1,2,...,c)$ indicates the center of class i , $p_i \in R^S$

The calculation process of FCM algorithm is to solve dividing matrix $U=[u_{ik}]_{c\times n}$ and clustering original model $P=\{p_1, p_2, ..., p_c\}$ [7-8] on guaranteeing objective function J_m minimum, Expression of J_m is :

$$J_m(U,P)=\sum_{k=1}^{n}\sum_{i=1}^{c}u_{ik}^m d(x_k, p_i) \tag{13}$$

In the formula, $\sum_{i=1}^{c}u_{ik}=1, u_{ik}\in[0,1], \forall k, d(x_k, p_i)=\| x_k, p_i \|$, m is for the weight index, $d(x_k, p_i)$ indicates the difference between x_k and p_i , use the distance between two vectors to measure.

Iteration regular formula is as follows :

$$p_i=\frac{\sum_{k=1}^{n}u_{ik}^m x_k}{\sum_{k=1}^{n}u_{ik}^m}, \text{(i=1,2,…,c)} \tag{14}$$

$$\begin{cases} u_{ik}=[\sum_{j=1}^{c}(\frac{d_{ik}}{d_{jk}})^{\frac{2}{m-1}}]^{-1} & ,\exists d_{ik}\neq 0, 1\le i\le n \\ u_{ik}=1 & ,\exists d_{ik}=0, k=1 \\ u_{ik}=0 & ,\exists d_{ik}\neq 0, k\neq i \end{cases} \tag{15}$$

2.6 Application of Apriori Association Rules

Assuming $I=\{i_1, i_2, ...i_m\}$ is a group of retrieval set(or item-sets), D is for the transactional databases, each affair can be represented as {TID,T}, in which $T=\{i_1, i_2, ...i_k\}, i_j \subseteq I(j=1,2,...,m)$, it indicates the involved item set in the affair (user-retrieval content). Three definition of association rules [3]:

Definition 1: relevancy of association rule:

Namely the ratio of affair set that contains X and Y with the entire set, marked as $support(X \Rightarrow Y)$. Which is

$$support(X \Rightarrow Y)=\frac{\left|\{T: X\bigcup Y\subseteq T, T\in D\}\right|}{\left|D\right|} \tag{16}$$

Definition 2: reliability of association rules: namely the ratio of affair containing X and Y with the affair only contain X, marked as $confidence(X \Rightarrow Y)$, which is :

$$confidence = (X \Rightarrow Y) = \frac{\left|\{T : X \bigcup Y \subseteq T, T \in D\}\right|}{\left|T : X \subseteq T, T \in D\right|} \tag{17}$$

Definition 3: when the item set relevancy support(X) is bigger than the threshold minsup, then we call the item is frequent item set.

Apriori algorithm is one of the most influential Boole association rules mining frequent item sets algorithm. Its core is two phase frequency set idea based recursion algorithm. We use this algorithm to obtain the maximum itemsets. Apriori algorithm flow chart is shown in figure 2.

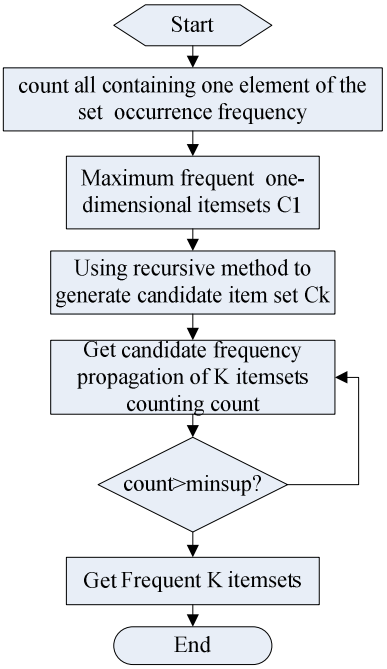


Fig. 2. Flow chart of Apriori algorithm

In which Minsup is for the minimum confidence, after using recursive algorithm to get K candidate priority, we get the frequent itemsets through comparison of minimum confidence.

Association rules can find users' tendency of downloading books resources and borrowing books, the library supply corresponding resources in time according to the needs, which also can get users preference browsing resources, and finally implementing adaptive recommendation and retrieval according to user's preferences.

3 Instance Simulations

In order to test the performance of adaptive system of digital library under cloud computing environment, first register10 accounts which respectively represent different professional users, then retrieval the professional knowledge according to the own profession, check and download. Simultaneously open the adaptive system.

Every reader can register an account on the service platform, organize their own personalized space according to their willingness. Readers can read all the given labels and the label related books in the personalized space; and you can also brows custom multimedia rules list, click one of rule, the system returns the corresponding multimedia push resources; You can browse bookmarks recorded in the book; and also the hot books selected by the platform.

Service platform provides three custom multimedia rules of book, image and calligraphy characters. In order to improve the user experience, we apply the AJAX technology to implement custom multimedia rules add and delete functionality, which makes the browser do not need to refresh the entire page when users add or remove a variety of media rules on the form. It can be creat directory in the rules list area, users can organize to add multiple rules, when users click on one of these rules, the browser will display the corresponding result on the right side area, which do not need to refresh the entire page.

Using the.net to construct web platform of digital library and name the library as delicious library. For example, select a literature student account to login on the system, click on recently recommended recent how book labels, then on the right side, the system will list some media related books and film and television works, which is the push services based on the user search and browsing habits, as shown in figure 3.

Fig. 3. Adaptive "recent recommended books" push services

Also, use a administrative management professional account to login digital library system, after entering the system, the system has been put some books in the home page in accordance with the reader habit for readers to choose before research, which is very intelligent and personalized. As shown in figure 4.

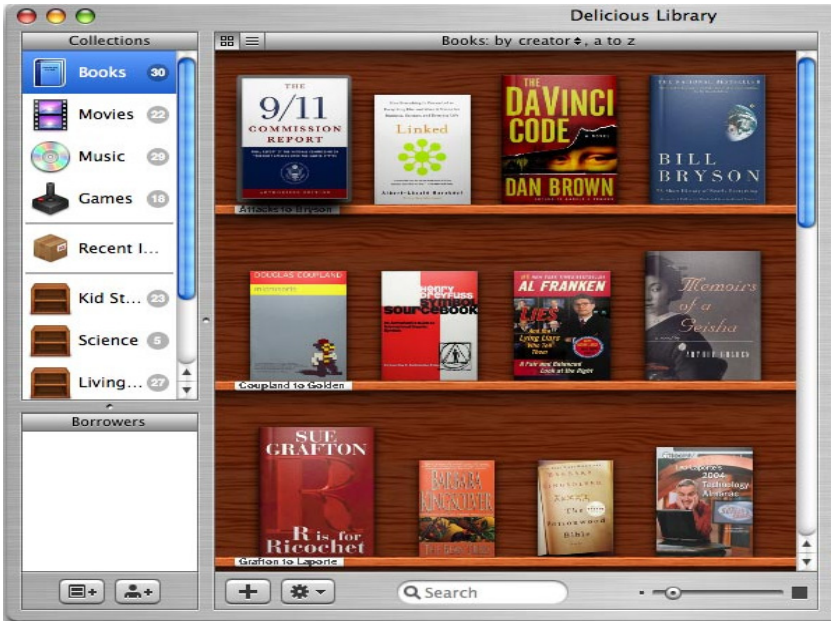


Fig. 4. Personalized home page

This kind of user behavior and search history based adaptive system need a large number of data mining and processing of backend server, which can not be satisfied by normal server, and it must rely on the help of powerful computing capacity of cloud computing to analyze the data processing, this adaptive system provides great help for the digital library construction.

4 Conclusions

Cloud Computing provides convenient for digital library construction, this paper analyzes the effect of cloud computing in digital library construction from intelligent and personalized service, and establishes digital library adaptive system by taking the advantage of powerful computing ability of the cloud computing, so as to provide users with personalized service, it has far-reaching significance for digital library construction, and which has the broad market application prospect.

Acknowledgement. Foundation Item. 2013 annual scientific project, Heilongjiang Province Office of education, Project number: 12532397.

References

1. Klein: CAD-Based off-line programming of painting robots, Robotics 5, 267–271 (1987)
2. Xu, R., Zhang, K., Nie, P., Qian, L.: A Formal Business Protocols with Composition Algebra for Cloud Computing 7(4), 1328–1335 (2011)
3. Guo, T., Wen, S., Yang, Y., Chen, J.: An Elastic Deployment Mechanism of Web Application Migrating to the Cloud Adopting IPv6 7(9), 3202–3208 (2011)
4. Li, J., Peng, J., Zhang, W., Han, F., Yuan, Q.: A Computer-supported Collaborative Learning Platform Based on Clouds 7(11), 3811–3818 (2011)
5. Li, X., Fan, Y.: Analyzing Compatibility and similarity of Web services processes. Chinese Journal of Computers 32(12), 2429–2437 (2009)
6. Li, Y., Huai, J., Guo, H., Duo, Z.: A Hierarchical Service Composition Framework Based on Service Overlay Networks. Journal of Software 18(12), 2967–2979 (2007)

Design of Second Language Acquisition Mobile Learning Resources

Zhang Peng[1] and Li Jing[2]

Beijing China[1], Tangshan China[2]
[1] International Communication Institute
Beijing International Studies University
Beijing, China
swanpeng@126.com

Abstract. The process and characteristics of the instructional design of mobile learning resources are studied in this paper. First we analyzed the characteristics of mobile learning resources, and studied each step in the designing process, considering the characteristics of learners' mobility. Then we referred to international digital learning resources standards and criteria, considering the characteristics of mobile equipments, learners and second language acquisition. We developed a resources designing model suitable for mobile learning. In the end, we also gave some advices to the resource-designers.

Keywords: Mobile learning, Learning resources, Second language acquisition.

1 Characteristics of Mobile Learning Resources

Mobile learning resources is one type of digital learning resources [1, 2]. Because of the uniqueness of their own mobile terminals characteristics, mobile learning resources differ from the general digital learning resources in the following aspects:

Simply presented and summarized [3]. Learners access to, browse and visit mobile learning resources through mobile terminals, such as smart phone, iphone, ipad, and tablet pc, and so on. Because those mobile devices' screens are generally small, the learning content pages are relatively small, and menus and navigations are also relatively simple. Content is always manifested through texts and vectorgraphs.

Learning resources organized by modules [4]. Mobile learning is flexible and random, so the information learners need is comparatively random and concise. Therefore the content of mobile learning resources is concise and abstracted. According to this, we should organize the learning content by the idea of "module", and cut them into little "information pieces", and put these little pieces into individual themes according to learners' special needs. As shown in Figure 1.

Besides of the characteristics of mobile learning resources, adult mobile learning resources must be suitable for adult psychology and learning pattern, because most second language learners are adults, and should satisfy the special need of adult

Y. Yang, M. Ma, and B. Liu (Eds.): ICICA 2013, Part II, CCIS 392, pp. 156–164, 2013.

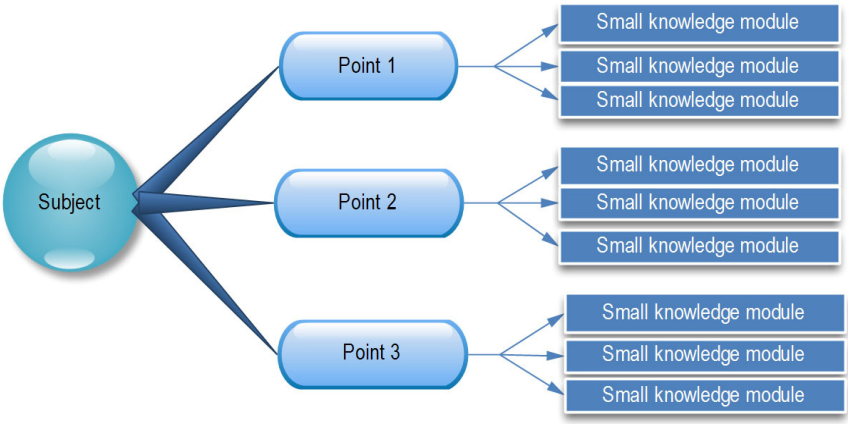


Fig. 1. Component Divisions of Mobile Learning Resources

mobile learning [5-7]. For instance, resources organizing are aimed to solve practical problems, and resources content organizing is aimed to improve practical abilities, and so on.

2 Instructive Design of Adult Mobile Learning Resources Process

The designing process of mobile learning resources includes analysis, plan, explore, and evaluation, the same as traditional education designing [8]. However, each part has its own characteristics compared with the latter one.

2.1 Demand Assessment

Demand assessment aims to locate the gap between ideal and practical situations, to make sure of the assignments and their order, and to find out required Knowledge, Skill, and Attitude (KSA) [9]. After the assessment, we do not have to carry those problems to the next step, which are caused by ideal and practical situation differences which could not be amended by mobile learning methods.

2.2 Analysis

The main tasks of Analysis include the analysis of learners and learning objects. Although Analysis is the indispensable step in both traditional and mobile education, there are a lot of differences between them.

2.2.1 Learners Analysis

The characteristic analysis of second language learning always start with psychological analysis, such as, learning attitude, learning motivation, and learning

capability, and so on. The basic characteristics of second language learning are very important in instructing the mobile learning resources designers and keeping them close to the adult learners, in order to design resources model really suitable for learners.

2.2.2 Learning Object Analysis

During this process, specific and particular teaching details should be paid attention. Teaching content should correspond with objectives, and each objective should guide learners to a higher level by creating a sequence of ascending difficulty. Every section of education is based on knowledge from the last section. It is necessary to consider the difficulty span between different sections, which should be accessible for the majority of learners and not too easy for them at the same time. This sense of achievement could motivate learners. And the proper amount of exercises corresponding with the designing and education objectives could bring learners closer to those objectives through practicing and feedbacks, and create easy access for systematic operation and evaluation.

2.2.3 Learning Content Analysis

Learning content should be organized based on learning component, which, therefore, should be decided in the first place. According to the actual need, learning components could be chapters, sections, topics, and units, and specific learning materials.

Upon completion of all of the analysis, it is necessary to write an analysis paper. The results of the analysis are the basis of resource-design, and could ensure that the resources are rightly placed in accordance with the due objectives.

2.3 Design

It includes objective design, resource construction design, learning strategy design, and learning interface design. The specific tasks are:

- Designing resources components.
- Determining necessary text content.
- Determining necessary voice content.
- Determining visual content illustrated by graphics, icons, pictures, animation or video images.
- Collecting or selecting resourceful materials according to teaching objectives and content.
- Designing resource structure and interface configuration according to the results of the learnng content analysis.
- Selecting and placing different resources components into the corresponding resource structures and interface configurations in according with different logical structure.

- Designing mutual relation between different resources components.
- Designing practice approaches for learners during their learning process.
- Designing learner-controlled self-testing approaches.
- Designing learner-controlled interaction and feedback approaches.

Resource design standards should be set up to enable massive produce, so the whole development team could refer to the same rule to make sure the consistency of the entire learning resources design. After confirmation of all the designing tasks, the operating team should start with a small unit or subject as an example to from an embryonic pattern of resource development. This enables both the designers and users to understand the resource structure and appearance, and leads to large-scaled development.

2.4 Development

2.4.1 Script Design

Script writing and media material production is the first task. Script should include: screen layout presentation, media special effects, narration, sound and music effects, and video pictures, and also should specify the serial number of each screen, coincide with the nearby serial number. Names of text documents, graphic files, audio files, animation files, and video files should also be annotated

2.4.2 System Integration

Upon the completion of the script, the department or individual using those resources should be asked to make sure the correctness of the content and the integrity of the resources. The process of media production of learning resources can only be conducted after this confirmation, which includes handling different assignment to various professionals based on the previously finished design concept, and integrating those productions into a complete software structure. Upon the completion of resources components development, each component must be labeled with descriptive meta-data information in order to facilitate its retrieval and reuse.

2.5 Test

Because the unique hardware environmental requirement of mobile learning resources, the finished software must run tests in a simulated environment, such as on a simulator, before proceed to the actual environment. Main test items are:

- Correctness test;
- Fault tolerance (robustness) test;
- Integrity (safety) test;
- Inter-module coordination and communication test;

- Efficiency test;
- Maneuverability test;
- Portability test.

After finishing these tests, a specific test report should be sent to the staff to make amendments, which should also be documented at the end of the report.

2.6 Evaluation

This step could be conducted in a small-scaled but real mobile environment to test the quality of learning content design, structural framework design, resources component design, interface design, and media design, and so on. Evaluation could be conducted through questionnaire survey. The objects of this survey contain representatives of learners and instructional designers, content experts, and so on, in order to get different views to understand the resource quality. There are the evaluation items of adult mobile learning resource.

2.6.1 Learning Content Design

- The scientific tenability of learning content.
- Learning content and learning objectives.
- Learning contentand learners.
- structure of learning content.

The Evaluation Details include: Learning content presented is correct in order to ensure that no scientific error occurs; whether the designed learning contents correspond with the learning objectives; whether the designed resources components correspond with the learning objectives; whether the learning contents suit the needs of adult learners; the continuity, sequence and integration of learning content structure are reasonably designed or not; whether the relation between learning contents is correct.

2.6.2 Structural Framework Design

- The Flexibility of the Structural Framework.
- Structure of Framework.

The Evaluation Details include: Whether there is a flexible and effective guidance in the structural framework; whether the structural relation between concepts is clear; whether the framework allows adult learners to reorganize the learning content; whether the content is provided to learners for comparison.

2.6.3 Resources Component Design

- The scientific nature of Resource Component.
- The Reusability of Resources Component .
- The Aesthetic Design of Resources Component.

The Evaluation Details include: Whether the resources components are scientifically and rationally divided; whether the resources components match their content and structure; resources components information is clearly indicated; resources components are restructuring; the resources components, including their composition, figure, and color, are aesthetically designed; whether the animation, video and other dynamic components are designed by the aesthetic standard.

2.6.4 User Interface Design

- Interactive design of user interface
- User interface designing style.

The Evaluation Details include: Whether feedback guidance and assistance instruments are effective; whether the convenient tools, such as Link, Skip, Backward, Guide, and Help, can facilitate the use of adult learners; whether the overall interface style are unified and coordinated; whether the text and background colors are unified, coordinated.

After amendments to the designing process according to the results of the evaluation, completed products can be promoted into massive use. As the wide use of the product, the application situations and users' feedback should be paid close attention to, in order to provide necessary follow-up services and technical support for users within the life cycle of resources.

3 Instructional Design Model of Adult Mobile Learning Resources

Both domestic and foreign experts and organizations raised a lot of expectations for the digital learning platform system and the corresponding instructional design of learning resources, and made a lot of instructive development. For instance, the Shareable Content Object Reference Model (referred as SCORM) aims to ensure the on-line access to exercises and learning resources of high quality, and improves the real-time, interoperable, and durable performance, which coincides with the mobile learning resources designing concept. Based on the above study of every step analysis during the resources adult mobile instructional design, according to the SCORM 2004, this study develops the following study instructional design model of adult mobile resources. As shown in Figure 2.

Compared with ordinary learners, adult learners have more social experience, social resources, and more complex interpersonal relations, which make the analysis of adult learners' initial behavior more complicated. Therefore this model analyzes basic characteristics of adult learning resources separately from both the internal and external dimensions. Detailed learning resources design is base on the stage analysis results, in order to guarantee the consistency, reusability, and flexibility of the entire learning resources design. It is necessary to develop specifications to make sure the entire development team stick to the same guidance. Upon the completion of resources components development, each component must be labeled with descriptive

meta-data information in order to facilitate its retrieval and reuse. Because the unique hardware environmental requirement of mobile learning resources, the finished software must run tests in a simulated environment, such as on a simulator, before proceed to the actual environment. This model is designed to apply on the development of learning resources products, therefore the evaluation objects are products, teaching quality.

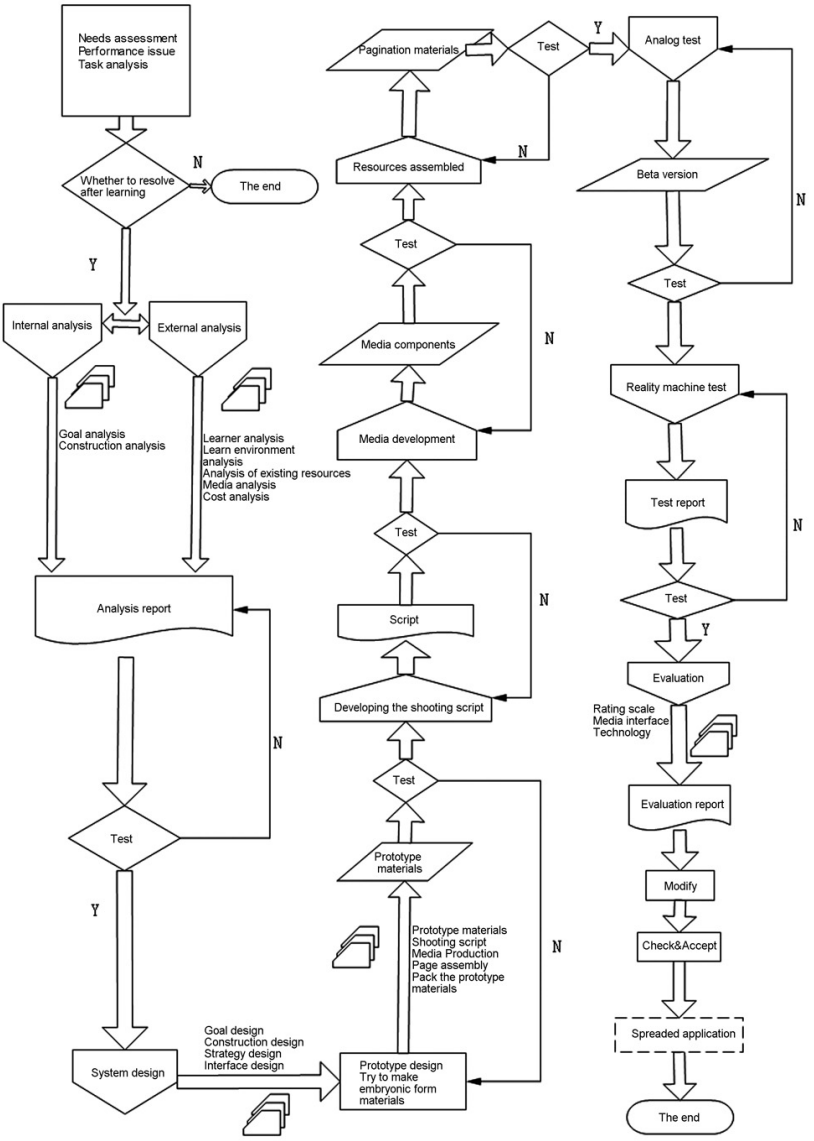


Fig. 2. Second Language Acquisition Mobile Learning Resources Design

4 Some Advices for Adult Mobile Learning Resources Designers

During the practical application of this model, the instructional designers should design and analysis of resources and curriculum units by components, then clarify the logical relationship between them, and chain the curriculum modules together.

During the actual design and development of mobile learning resources, some details should be paid attention to. Here are several advices for the learning resources designers.

4.1 Analysis

- Make sure of suitable contents as mobile learning materials.
- Make sure whether there is a potential secondary users of these resources.
- Decide on which resources will be shared by different learners.
- Decide on the mode in which storage resources components and their information are use.
- How to get easy access to the component-base in search for reusable learning resources components.

4.2 Design

- Cutting learning resources components appropriately for future use.
- Determine the manner in which to organize content to meet the learning needs.
- Determine sharing components of textbook and considering the experience of adult learners.
- Design basic media material to facilitate the repeated use.

4.3 Development

Comprehensive and detailed documents are required for those resources probably being re-used, including complete script documents, detailed notes for the program, specific meta-data information resources components, and so on.

References

1. Shuzhi, B.: ComparativeAdultEducation. Beijing Normal University Press (1995)
2. Knowles, M.S.: Modern Practice of Adult Education. People's Education Press (1989)
3. Xu, F., Ye, C.: Mobile learning and its theoretical foundation", Open Education Research, pp.23–26 (March 2004)

4. Xiong, Z.: Research on mobile learning and resource design Master's thesis. EastChinaNormalUniversity (May 2005)
5. Houle, C.O.: The Design of Education. Jossey-Bass (1996)
6. Chengtian, O.: Study on Mobile Education Platform Supporting English Words Intelligent Ubiquitous Learning System. Capital Normal University (May 2004)
7. Shen, S.: Resources are Extension of The Media. China Educational Technology, pp.13–16 (June 2000)
8. Liu, C.: Network Communication and modern education. New Century Essays on the development of network communication, pp. 485–489. Beijing Broadcasting Institute Press (2001)
9. Ji, D.: AdultEducational Psychology, 2nd edn. Shanxi People's Publishing House (May 2006)

Secondary Developments Based on BIM Virtual Technology

Siwen Bai

China Construction Steel Structure Corp.Ltd, Beijing, 100089 China

Abstract. The article introduces and summarizes the existing BIM technology and virtual construction technology, starting from the construction status and the actual demand, to analyze the application of BIM in construction engineering; and practice of virtual construction technology is studied through the professional modeling software and software development platform combining, realize visual management construction process; and combined with the two development achieved acquisition of specific information in the BIM model and component collision detection function, and provide powerful technical support for the further construction of buildings.

Keywords: virtual construction technology, BIM concept, secondary development.

1 Introduction

The rapid development in the globalization of information technology today, information technology has become the impetus to push the development of human society, especially the application and popularization of Internet technology, information technology has penetrated into all fields of social life and economic [1-4]. With the deepening of the development of information technology, there is a hitherto unknown revolutionary change to the traditional management of construction industry. In the construction of large and complex, in order to ensure the normal operation of the construction work order, in all aspects of the work can accurately collaboration and accurate construction plans [5, 6]. However, in the construction of complex, only the calculation tool, the traditional method has been to meet the information processing the big many, there will be no more construction scheme optimization [7-10]. Therefore, the application of BIM technology in building design and construction process is complex, to ensure the reliability, safety and economy of the project construction [11-15].

2 The Concept of BIM

Virtual construction technology of BIM is a new subject [16-20]. Technically, virtual construction technology can realize the comprehensive transformation of the traditional construction method, the construction enterprise organization, and management and production aspects of change. But at this stage because of various subjective and

Y. Yang, M. Ma, and B. Liu (Eds.): ICICA 2013, Part II, CCIS 392, pp. 165–173, 2013.

objective reason shave failed to develop the universal, still need the power to actively promote and support, and constantly improve its system of theory and technology, in order to promote the progress and development of construction technology [21-26].

BIM is the abbreviation of Building Information Modeling. This is the numerous domestic and foreign designers called "revolutionary" technology, is the first by the United States Georgia Institute of Technology (Georgia Tech College) architecture and computer professional chuck Eastman (Chuck Eastman) is a concept proposed by Dr.: "building information model integrated all the geometry information, functional requirements and component performance, a construction project in the whole life cycle of all the information into a separate building model, but also includes the process of information construction schedule , construction, maintenance and management ". As shown in figure 1.

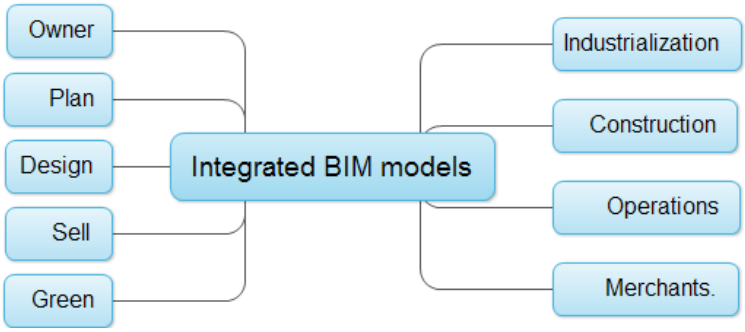


Fig. 1. BIM Integration Function Chart

BIM is an integrated process that is used to support construction workflow, generate coherent, reliable 3D format design information [27-29]. It uses a digital method for virtual simulation of the building, in order to understand its performance, and to use a set of coordinated and complete information as a design platform. First, it establishes a platform for architects, engineers and construction side can truly understand their design to be achieved; Second, it not only provides an integration of all platform collaborators and change processes, as well as the ability to obtain a more in-depth design insight, grasp the sustainability of the design, production and installation of the gripping member, and also supports the construction process. In fact, the future of building information modeling will become an essential tool, a collaboration and implementation of projects for the construction of platforms, it will become architects, structural engineers and construction side in the future work of the indispensable platform.

BIM carrier: Through 3-D digital technology, integrated project-related information construction engineering data model. The model can be provided for the design and construction coordination, information and analysis for computing. The model can in that environment constantly updated information and provide access to, and the parties to the project can be obtained from the model, they need information, you can more clearly and more comprehensive understanding of the project.

BIM features: Because BIM is needed to support construction of an integrated life cycle management environment, so is the structure of the building information model contains the data and behavior of the composite structure of the model. It not only contains data such as data relating to geometry model also includes the management of the behavioral model of a combination of both for the data given through the associated meaning, which can be used to simulate the behavior of the real world.

BIM process: BIM flow chart is shown in figure 2.

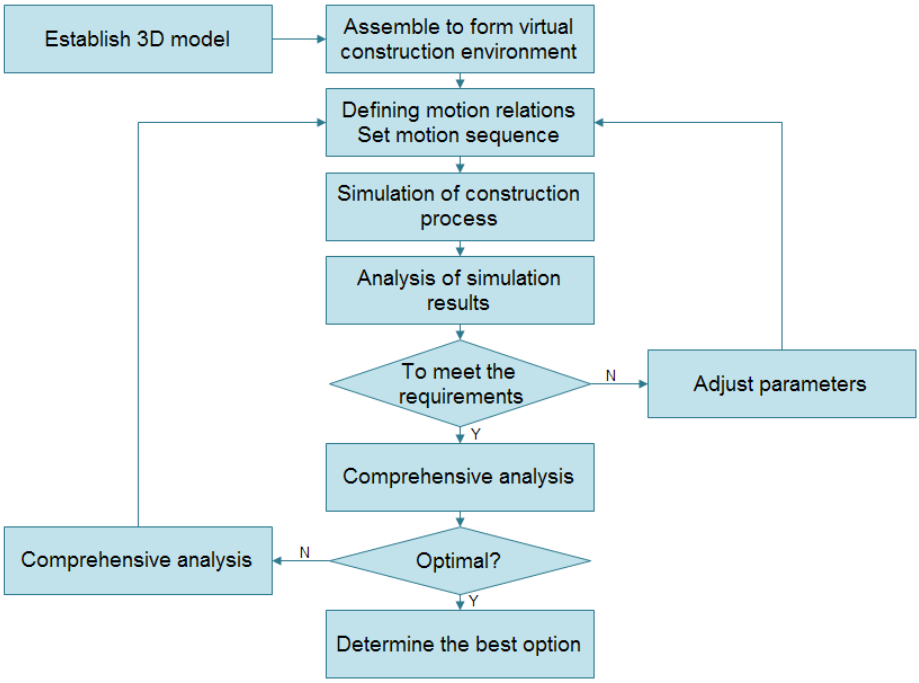


Fig. 2. BIM flow chart

3 Virtual Technology Application Based on BIM

3.1 3-D Modeling for Building

Began to emerge of the BIM eventually replace the traditional CAD design method, it set off a building industry's second revolution. In order to realize the BIM technology requires the help of BIM software, the software BIM concept for building entity components based on digital and as a design element, can automatically calculates and displays the content space design elements and functional relations, in order to give full play to the designer's imagination provides great space.

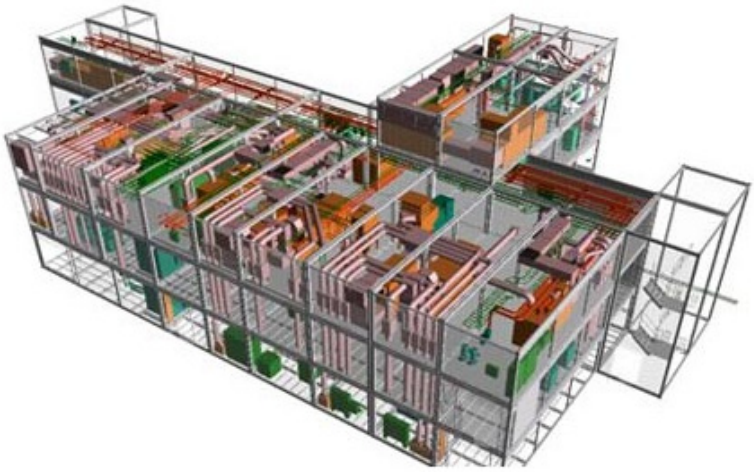
In the construction field, you can create a BIM software contains all the features of the actual construction of the building three-dimensional model, so that the parties involved in the construction project on the senses to establish three-dimensional effect, so as to effectively guide the engineering design and construction. BIM software applications can be efficiently and flexibly paper documents into electronic documents, three-dimensional rendering and detailed construction drawings and other functions. BIM effective use of information is brought to the construction industry is one of the main values, when the model-based three-dimensional and two-dimensional technology and information, when combined, it means that the architect has a richer and more efficient design process , which reduces the construction risk, improved quality control, to ensure the design intent, but also makes communication more smooth. In addition to drawing drawings, create schedules, document generation and other lower-level design tasks have been automated, but also for architectural design have also been modified to achieve automatic correlation and real-time updates. In short, through integration of BIM technology, prompting architects were able to play to maximize the utility of the computer.

3.2 Virtual Construction Management

Compared to traditional manufacturing industries, the construction industry can't match in production efficiency. According to incomplete statistics, in a project that approximately 30% of the construction process requires rework, 60% of the labor resources are wasted, and 10% of the waste material loss phenomenon. We can not difficult to calculate, tens trillions of dollars and capital loss in the huge construction industry every year.

BIM technology created Architectural 3D model contains all the comprehensive information in construction project. 3D models will generally be the time, cost elements combining 5D construction management, through the application of computer technology, to simulate the construction process of construction projects, before construction can actually be able to determine the feasibility of the construction program. By 5D construction system can quickly create a three-dimensional construction model is designed to reduce or avoid the most errors exist in the document, thus saving costs. Through this model can also be very easy to analyze the reasonableness of the construction process, and then generate the corresponding procurement plan and a list of financial analysis costs, efficient optimize the construction program. For example, in 2005 it released a set of Graphisoft's Virtual Construction software application that can be used 5D construction management.

In the 5D construction management system, the design, cost, schedule of three parts related to each other, can be updated in real time, thereby reducing the construction project evaluation takes time budget, the budget significantly improved accuracy, and enhance the controllability of the project construction sex. Construction simulation can be found in advance through the design and construction of the problem, design, budget, schedule, and other attributes can be updated, and to ensure access to data consistency and accuracy of the information shown in Figure 3.



(a)

	A	B	C	D	E	F	G	H
1	pes	ry in	story name	layer	element type	fill	volume	surface (top)
2								
3	1	-2	PARKING_B2	_Slab	Slab	_S_und_found	4 228.83	45 671.21
4	1	-1	PARKING_B1	_Slab	Slab	_S_und_cei	1 536.35	55 311.02
5	17	0	LEVEL_1	_Slab	Slab	_F_s_conc	801.67	31 017.52
6	6	0	LEVEL_1	_Slab	Slab	_S_betw	17.55	1 137.50
7	1	0	LEVEL_1	_Slab	Slab	_S_roof	1.97	106.32
8	3	0	LEVEL_1	_Slab	Slab	_S_und_cei	1 560.77	56 028.10
9	3	0	LEVEL_1	_Slab	Slab	_S_und_found	4 364.77	47 668.40
10	20	1	LEVEL_2	_Slab	Slab	_F_s_conc	1 240.99	32 661.60
11	7	1	LEVEL_2	_Slab	Slab	_S_betw	1 912.46	80 246.16
12	8	1	LEVEL_2	_Slab	Slab	_S_roof	704.60	20 498.61
13	2	1	LEVEL_2	_Slab	Slab	_S_und_found	179.54	2 644.19
14	2	2	LEVEL_3	_Slab	Slab	_R_steel	129.67	7 012.79
15	1	2	LEVEL_3	_Slab	Slab	_S_betw	271.99	7 340.69
16	1	3	LEVEL_4	_Slab	Slab	_S_betw	271.44	7 328.90
17	1	3	LEVEL_4	_Slab	Slab	_S_roof	153.90	4 156.35
18	1	4	ROOF	_Slab	Slab	_R_steel	291.85	7 879.98
19								

(b)

Fig. 3. Virtual construction model and engineering quantity statistics table by virtual construction software

3.3 Construction Process Management

In recent years, with the popularization of BIM technology, the realization of the theory and practice of BLM application provides powerful support. The main direction of development of the construction industry information technology research and application of the BLM will become the future; also will become the main task facing the development of BIM technology.

As everyone knows, in the construction process of each stage of the problems to be solved and the conditions are different, so the architectural design is a process of gradual and continuous feedback. Especially with the construction process of in-depth, the previous stage results in the condition known as later stage. In each stage the architects need to repeatedly modify the design scheme to obtain the best design. By establishing the relation between specific in each stage, and finally integrate them into the design process of a covering the whole building life cycle.

3.4 Storage and Access Data in BIM Model

For huge information building information model of the life cycle, how to store and access to relevant data information in the BIM model has become the key problem of the underlying data information management architecture in the whole life cycle support. But because of the complexity of the BIM model, involving many professional personnel, diversity, will undoubtedly hinder the information integration and management of BIM based on the process. In recent years, part of the Business Companies and scientific research units have been developed and developed based on the BIM data access tool, it has realized the data information in the BIM model of access to a certain extent, but there are still many problems: such as commercial confidentiality, toolbox source code is not open to the outside world, cannot on its function deeply extension and customization, but in general these access tools only support a specific programming language and platform, processing only to the specific version of the IFC model and data, while ignoring the problem of semi-structured and unstructured data, we also need to continue in-depth study.

4 Secondary Development of BIM

According to the concept of BIM is in the.NET programming and Visual technology, based on Studio 2010 platform, through the API software function in Revit using C# language development. The Revit software for detailed process function develops mainly in the following steps, as shown in figure 4.

For example, the dialog box displays a hello world, compile and run in Figure 5.

```
Public class Helloworld : IExternalCommand
{
Public Result Execute(ExternalCommandData commandData, ref
string messages, ElementSet elements)
{
MessageBox.Show("Hello world");
return Result.Succeeded;
}
}
```

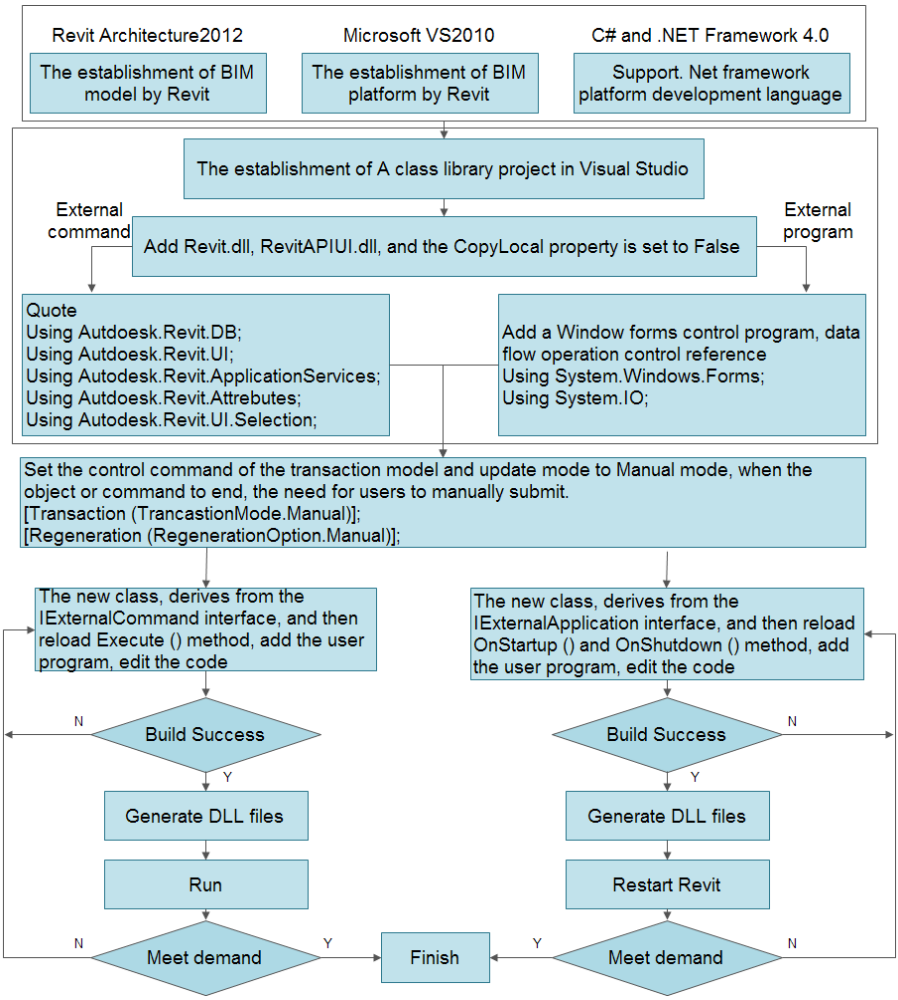


Fig. 4. Program Development Process

Fig. 5. External command runs results based on Revit

5 Conclusions

Virtual construction technology of BIM is a new subject. Technically, virtual construction technology can realize the comprehensive transformation of the traditional construction method, the construction enterprise organization, and management and production aspects of change. But at this stage because of various subjective and objective reason shave failed to develop the universal, still need the power to actively promote and support, and constantly improve its system of theory and technology, in order to promote the progress and development of construction technology.

In conclusion, the application of virtual construction technology in the construction field will be an inevitable trend, application prospect in building design and construction in the future broad. Believe that with the development of virtual construction technology and perfect, will promote the construction industry in China entered a new era of engineering.

References

1. Autodesk Software (China) Co., Ltd.. BIM technology and force the construction industry to promote change in the industry. Survey and Design (2), 44–47 (2011)
2. Building Information usher layout opportunity [EB]. Chinese BIM portal (2011), http://www.chinabim.com/news/construction/2011-12-27/3010.html
3. Zhu, J.: Design Study of BIM railway. Railway Engineering Society (10), 104–108 (2010)
4. He, G.P., Li, G.: BIM applications will also bring any architectural changes. Information Technology (01), 10–17 (2010)
5. Bo, R.: From 2-dimensional to 3-dimensional computer-aided architectural design. Chong qing University (2010)
6. Adriaanse, A., Voordijk, H., Dewulf, G.: The use of interorganisational ICT in United States construction projects. Automation in Construction 19(1), 73–83 (2010)
7. Aish, R.: Collaborative design using long transactions and “change merge”. In: Donath, D. (ed.) Promise and Reality: State of the Art versus State of Practice in Computing for the Design and Planning Process – 18th eCAADe Conference Proceedings, Weimar, Germany, pp. 107–111 (2000)
8. Akkermans, H., Gordijn, J.: What is this science called requirements engineering? In: 14th IEEE International Requirements Engineering Conference (RE 2006), pp. 273–278. IEEE, Minneapolis (2006)
9. Alshawi, M., Ingirige, B.: Web-enabled project management: an emerging paradigm in construction. Automation in Construction 12(4), 349–364 (2003)
10. Amor, R., Augenbroe, G., Hosking, J., Rombouts, W., Grundy, J.: Directions in modelling environments. Automation in Construction 4(3), 173–187 (1995)
11. Amor, R., Bloomfield, D., Cerovšek, T., Finne, C., Groosman, M., Gudnasson, G., Hutchison, A., Hyvärinen, J., Ólafsson, G., Oliveri, E., Rio, O., Turk, Ž.: Information Services to Enable European Construction Enterprises Building Research Establishment, Watford, UK (2001)
12. Amor, R., Jiang, Y., Chen, X.: BIM in 2007 – are we there yet? (2007)
13. Rebolj, D. (ed.): CIB W78, University of Maribor, Maribor, pp. 159–163 (2007)
14. Arayici, Y., Coates, P., Koskela, L., Kagioglou, M., Usher, C., O’Reilly, K.: BIM implementation for an architectural practice

15. Dikbas, A., Ergen, E., Giritli, H. (eds.): Managing It in Construction/Managing Construction for Tomorrow, vol. 26, pp. 45–51. CRC Press, Taylor & Francis Group, Istanbul, Turkey (2009)
16. Ball, A., Ding, L., Patel, M.: An approach to accessing product data across system and software revisions. Advanced Engineering Informatics 22(2), 222–235 (2008)
17. Barak, R., Jeong, Y.S., Sacks, R., Eastman Unique, C.M.: requirements of building information modeling for cast-in-place reinforced concrete. Journal of Computing in Civil Engineering 23(2), 64–74 (2009)
18. Bazjanac, V.: Implementation o semi-automated energy performance simulation: building geometry
19. Dikbas, A., Ergen, E., Giritli, H. (eds.): Managing It in Construction/Managing Construction for Tomorrow, vol. 26, pp. 595–602. CRC Press, Taylor & Francis Group, Istanbul, Turkey (2009)
20. Bazjanac, V., Kiviniemi, A.: Reduction, simplification, translation and interpretation in the exchange of model data. In: Rebolj, D. (ed.) CIB W78, University of Maribor, Maribor, pp. 163–171 (2007)
21. Beetz, J., de Vries, B., van Leeuwen, J.: RDF-based distributed functional part specifications for the facilitation of service-based architectures. In: Rebolj, D. (ed.) CIB W78, University of Maribor, Maribor (2007)
22. Bjork, B.-C.: Requirements and information structures for building product data models
23. VTT Building Technology, VTT, Espoo, Finland (1995)
24. Bjork, B.-C., Penttila, H.: A scenario for the development and implementation of a building product model standard. Advances in Engineering Software 11(4), 176–187 (1978, 1989)
25. Björk, B.-C., Laakso, M.: CAD standardisation in the construction industry – a process view. Automation in Construction (in press)
26. Boddy, S., Rezgui, Y., Wetherill, M., Cooper, G.: Knowledge informed decision making in the building lifecycle: An application to the design of a water drainage system. Automation in Construction 16(5), 596–606 (2007)
27. Bogen, J.: Theory and observation in science. In: Zalta, E.N. (ed.) The Stanford Encyclopedia of Philosophy, Stanford University, Stanford (2009)
28. Borrmann, A., Ji, Y., Wu, I.C., Obergrießer, M., Rank, E., Klubert, C., Guenthner, W.A.: ForBAU – the virtual construction site project
29. Dikbas, A., Ergen, E., Giritli, H. (eds.): Managing It in Construction/Managing Construction for Tomorrow, vol. 26, pp. 528–534. CRC Press, Taylor & Francis Group, Istanbul, Turkey (2009)

Three-Dimensional Modeling Technology in Virtual Reality

Heng Chen

Sichuan Tourism University, Chengdu, Sichuan Province, 610100

Abstract. With the rapid development of our country' economy, related technology of computer has obtained quick development, among which computer hardware and software have a good performance used in every industry and every field. During recent years, computer technology has entered towards the era of 3D, namely entering into the age of using 3D technology of computers. More and more enterprises have completed various scientific research using computer three-dimensional models. Virtual reality, visualization of mechanical calculation and computer animation has already become three major means of 3D computing. In virtual reality, modeling technique is the most critical method. In this paper, we would explain the construction method of virtual reality with the method of 3D modeling method of engineering survey.

Keywords: virtual reality, 3D modeling, signal reception, optimization design.

1 Introduction

Virtual reality refers to artificial media space constructed by using computer, but it is virtual with realistic feel which makes people enter a virtual environment through multimedia sensor interface device, producing the feeling of being personally on the scene so as to achieve the purpose of technology exploration.3D modeling technology which is widely used in surveying and mapping engineering has its very big feature in operation [1-8]. Its features such as quick speed of observation, 24-hour available for use and no regional restriction have brought many benefits to engineering construction. We used to establish multiple control points on the surface of the earth level and use measuring method of multiple stations to increase the workload obviously. By using 3D space geodetic control network in virtual reality, observers are able to record the parameters of the control points within the area without the restriction of time and space in a short time and quickly complete the analysis and integration process [9-13].

Three-dimensional modeling techniques widely used in all areas of life, including in the construction, medical image, heritage, 3D animation games, movies special effects, virtual reality plays an important role.

In the construction sector, if the method of making three-dimensional model used to make three-dimensional models of buildings, it can be more intuitive to observe all sides of the building, and the entire structure, and even the internal structure, which is designed for both the architect viewing results, or for the customer view is very convenient.

Y. Yang, M. Ma, and B. Liu (Eds.): ICICA 2013, Part II, CCIS 392, pp. 174–183, 2013.

In medicine, since 100 years ago since Roentgen discovered X-rays, medical image processing technology has come a long way. Obtain three-dimensional human anatomy has been one goal. Three-dimensional modeling techniques in medical applications of virtual surgery [14-17].

In conservation areas, three-dimensional model in heritage conservation plays an important role. Using three-dimensional modeling techniques will relics and ancient buildings through image acquisition, digital processing, data compression techniques into three-dimensional image, which is one way to preserve and study artifacts. In recent years, China has begun to attach importance to this work, such as the Forbidden City Digital Museum under construction in the positive, the Hall of Supreme Harmony and the surrounding three-dimensional model of the scene had been by the Treasury will be produced Toppan completed to achieve a scene roaming, with considerable realism [18-22].

Now, the rapid development of computer games today, the pursuit of the game's true and gorgeous screen effect has been the ownership of consensus. Character to the scene from the game, all using the three-dimensional model, depicting fine, have a good visual effects and impact.

Special effects in film productions, 3D modeling technology are also widely used.

In the field of virtual reality, 3D modeling technology plays a vital role. Currently, virtual environment scene biggest difficulty lies in roaming model fidelity and real-time rendering of contradictions. Because of this roaming seen near the scene from the observer, requires very realistic rendering. Therefore, the model structure should be fine, but they also consume a lot of time. Similarly, the computer performance constraints, the construction of more complex models out of the number of faces, the more you draw the more difficult to achieve real-time effects, if real-time bad, it will make the viewer can not accept.

Three-dimensional model from the current technology research situation, there are many issues need to be resolved. Three-dimensional model of the development of technology will promote the development of the entire virtual reality system for the entire virtual environment and lay a good foundation. Thus, in the construction of three-dimensional virtual environment model of virtual simulation technology has an extremely important significance.

From the above, three-dimensional model based on virtual reality technology is the technology of virtual reality system application model should be as simple geometry model, you should ensure that the quality of the model must be done case the amount of data as small as possible to ensure that the virtual reality application efficiency of the system design of a virtual reality system in addition to the general conditions of individuals can not customize the hardware, but can give full play to the initiative of individuals in the system can only be a soft aspect efforts [23-27].

Before the establishment of the campus model, we should be based on real-life scenarios building location of each entity to determine the entire campus all the entities in the virtual scene structure of the model distribution. Divided on the scene hierarchy, you can easily block the scene models, physical models of division of labor and organization and management structure of the model explicitly target, thereby greatly reducing the workload of the model. In modeling, even the simplest model we also need to adjust the model hierarchy, to achieve the purpose of optimization models.

In a virtual environment are the most tedious virtual environment model problems. Three-dimensional scene model in the entire virtual environment session workload, need to invest a lot of manpower and material resources. In the modeling process, should be treated differently according to different buildings. Can the campus buildings in the 3D model embodied and manifested.

We produce three-dimensional models, the use of today's most popular modeling software 3ds Max models, model mainly simple model based, as far as possible in the production process of removing excess surface, overlapping faces and a few extra lines and points, in order to reduce the amount of data the model number of faces in order to improve the software run faster. In the process of making models for some buildings, do not be too demanding details of the performance, as far as possible using the "simulated" geometry to represent the basic shape of the object.

Three-dimensional modeling techniques and environmental art model is between the plane between drawings and the actual three-dimensional space, which is the organic link between the two together, is a three-dimensional three-dimensional model, architectural model helps design creative scrutiny, you can visually reflect the design intent, to make drawings in the form of performance inadequacies. It is part of the process designer, and also belongs to the design of a form of expression, which is now widely used in urban construction, real estate development, real estate sales, design tender and Cooperation and other aspects [28-30].

Three-dimensional model design, does not require a lot of space for practitioners cultural level, age, gender, etc. are not many conditions are relatively limited. This is a no pollution of all kinds, spending fewer returns and more urban industries, the promotion of employment, social and economic development play a great role. Experts believe that the national norms new career, to develop professional standards of the profession of architecture model designed to measure staff qualifications and ability to provide the basis for a new career in favor of the scientific, standardized and modern management, fundamentally promote employees improve the overall quality.

In a variety of exhibitions at that magnificent estate model always so many people from wandering. These vivid architectural models, so that buyers feel very intuitive type design drawings, community planning, in reality, in the end is what kind. The production of these architectural models, the building designer's creativity on display in sandbox professionals, it is two years ago, the Ministry of Labor and Social Security announced the second batch of new jobs in the building model design staff.

2 Research on the Construction Features of 3D Space Model in Virtual Reality

In the computer surveying and mapping engineering of the past, engineers would select a number of control points in the measurement areas, and these control points selected are with good visibility and favorable terrain and in this control points, engineers can simply construct the measuring angle of the place to measure side, so that an integral control network can be formed, and we can use this method to construct a measurement model. However in the field operation, it is difficult to achieve this ideal stage, and using the construction of GPS three dimensional space geodetic control network based

on the original construction basis simplifies operation method. The basic idea of the two measurements is almost the same. In addition, within the control network of the measurement, 3D renderings will increase the precision of measurement, the number of characters of the data will show a rising trend, and the increase of the accuracy of each point which can not be achieved with traditional measurement can be realized in this method.

We will simply introduce the basic principle of 3D graphics modeling in virtual reality. Modeling technology is to get the final results after processing the data with mathematical model based on the data processing of the controlled network and means of gauss projection, combine the features of the geodetic control network and specific requirements, and re-integrate 3D data to form data parameters with higher precision, laying a solid foundation for later-period construction. Compared with the measurement and selecting places of the past, there are some requirements for 3D geodetic survey and the principle is to select the relatively average spots with high accuracy and high grade in the whole area, providing convenience for processing data and coordinate transformation, and collecting and storing data in relatively definite spots. The operation process of the proposed 3D visualized modeling and simulation for construction operation as shown in Fig.1.

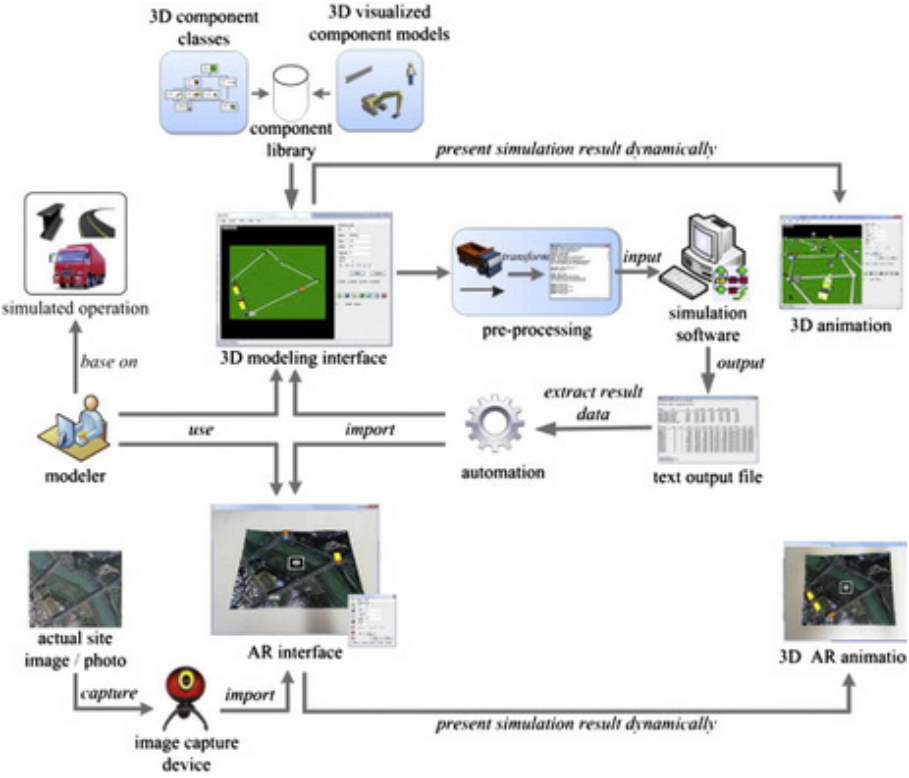


Fig. 1. The operation process of the proposed 3D visualized modeling and simulation

3 Analysis of Constructing 3D Model in Virtual Reality

Engineers would use 3D system, conduct level Ⅱ encryption of measured regions, select fixed parameters in 3D selected spots, project requirements of controlling modeling at a certain elevation, show the coordinates of control points on the horizontal size on computers, and have a good control of the proportion of the length and the width measured. In selecting height and side length measured, we should choose them as short as possible, so that data shown later will be more precise. There are following requirements for modeling:

3.1 Control Side Length

While selecting the side for measurement, long side plays an important role in controlling the model size in geodetic control. If the length of long side is too long, the whole 3D model would lose frames. If the length of long side is too short, the whole range of control network will decrease accordingly, increasing the difficulty of

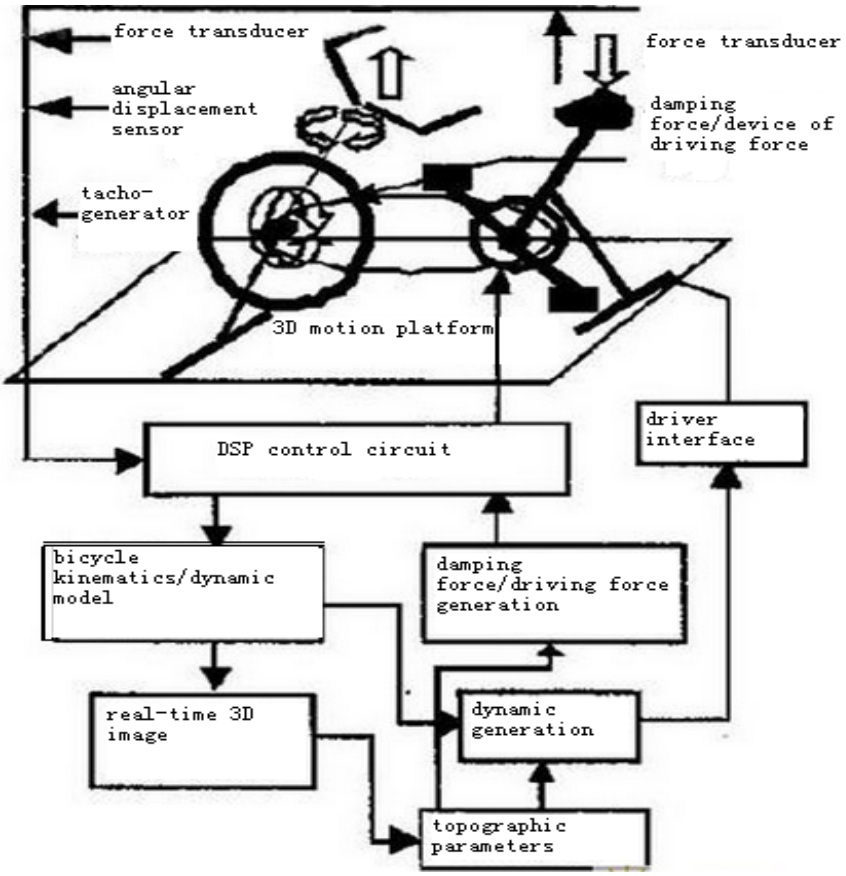


Fig. 2. 3D image generation model

measuring at later period. In selecting receiving frequency, using two-way frequency receiver can reduce adverse results of measurement due to uneven length of the long side and short side. The 3D image generation model is shown in fig.2.

3.2 Increase the Times of Observation

While measuring specification, one control point requires at least two control measurement record, and within 1000 m range, error rate of two times should below 0.5%. When measuring long distance, we choose multiple times of measurement to reduce the error probability, and for the same control point, in 3D measurement, we would repeat the measurement method.

3.3 Control Signal

In the construction of 3D model in virtual reality, we should choose the reasonable signal receiving band, and collect data when the satellite's signal is sound to ensure the reliability and practical value of data. We should stop data collection timely once find abnormal signal phenomenon.

4 Guidance on Construction of 3D Model in Virtual Reality

A lot of hard tasks have been accomplished. In the construction of virtual reality 3D model, but a lot of problems happens in the use of it. Especially in the construction of models, we can not use geodetic control, and can not show the advantages of this technology. We will discuss the in the common problems in the construction of geodetic control network in 3D space.

4.1 Reduce the Probability of Errors

In a lot of data errors, some are due to human factors. Being not familiar with machines and negligence of operation would to a certain extent bring more trouble in constructing multiple models cross. Frequent study is needed for network equipment configuration and we should set more reasonable configuration to ensure the security of network system. To prevent more operation errors, we should select system login system; users enter the system after identification, and after collecting data, there is re-identification sigh in confirming the final data. System has some relevant data interpretation of servers and routers, and we should record the operation time, and make a backup in a timely manner.

4.2 Control Access

Access control is a kind of security protection measures for illegal operation of 3D system network and it gives certain permissions to users and user groups in constructing geodetic control network system, restricting users and user groups' scan and alteration of directories, subdirectories, files, printers, and other shared resources.

Network operation server when unavailable can give "no reply" operational instruction, immediately shut down the current improper interface, speed up the system operation speed, and have a real-time monitoring of daily files to resolve the problems in a timely manner.

4.3 Open Automatic Backup System

Sudden collapse of the system would appear in constructing 3D model in virtual reality, and it is most likely to go back to the original interface after the system deals with a lot of data. In order to ensure that the system data is not lost, we should make a backup of the system. We should select fixed disk to save the data and export data regularly to ensure the effective operation of the program. When the error occurs in the system, we can recover the data to the initial manner as soon as possible to save more time for measurement task. We need to reduce the interference around the signal source, regularly update the database in the system to maintain the stability of data acquisition, store data of 3D space nodes on relevant technical drawings and in this way we could know the measurement.

5 Optimization Techniques of 3D Modeling

An important step in the process of 3D modeling is optimization of the technique and the optimization result will directly restrict display speed and operation efficiency of the system. 3D visual observation points is shown in fig.3.

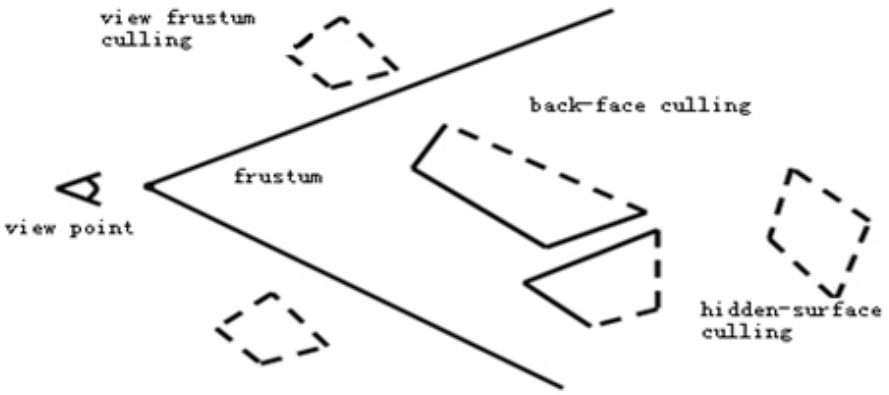


Fig. 3. 3D visual observation points

Steps of 3D modeling optimization process are as follows:

Step 1: structure optimization.

As for initializing the virtual scene, in accordance with the principle of scene block (or model segmentation), we should establish a hierarchy structure. As for the structure

dealt with in the last step, we should adjust structure in accordance with the principle of adjusting level for modeling.

Step 2: texture optimization.

We should deal with the texture preprocessed by data according to the actual demand of modeling, such as processing it into simple component texture, optimizing texture format and texture mapping

Step 3: model optimization.

Models created should be processed by using visible blanking technology, texture mapping technology, instantiating technology and LOD technology to simplify the number of the polygons and optimize models.

Step 4: optimization of the scene.

We can process created model system (within and out of the system) by using texture mapping technology, instantiating technology, LOD technology and external technology to carry on the optimization of the whole scene.

6 Conclusion

This paper expounds the construction process of each entity model, has a system research on the integration, management and schedules of scenario, summary the problems occurring in the process of modeling, and put forward the solution. This paper completes the regional measurement by using the key technology of modeling and related technology of modeling optimization to realize the purpose of authenticity, real-time and interactivity, which is of great reference value to realize the construction of a large-scale virtual reality scene. We produce three-dimensional models, the use of today's most popular modeling software 3ds Max models, model mainly simple model based, as far as possible in the production process of removing excess surface, overlapping faces and a few extra lines and points, in order to reduce the amount of data the model number of faces in order to improve the software run faster. In the process of making models for some buildings, do not be too demanding details of the performance, as far as possible using the " simulated" geometry to represent the basic shape of the object.

This paper is the research on the application of 3D modeling technology in virtual reality. The field of applying virtual reality technology is broadening, for example, it has been applied to the military, education training, engineering design, business, medicine, film and television, art, entertainment and many other fields, and there is higher demand for modeling of large-scale complex scenarios. 3D modeling technology is the basis of building 3D virtual scene, so it has a certain degree of practical significance to 3D modeling technology.

References

1. Hou, M., Zhen, L., Sun, W.F.: Based Skyline three-dimensional digital campus. Beijing Institute of Architectural Engineering (April 2008)
2. Yang, Y., Chen, J., Gang, K.: Characteristic pattern in the fuzzy comprehensive evaluation of the results expressed in the membership application, Beijing Water (January 2010)

3. Tao, Tian, H., Cen learn, Zhang, C.-Y.: Dimensional virtual reality research campus. Beijing Institute of Petrochemical Technology (February 2010)
4. ARCHIVES OF Yao, H., Ting, Y., Fan, Q.-Q., Lei, Y., Liang, L.: MultiGenn Creator & VRP based virtual reality platform to build campus Theory and Practice. Surveying and Mapping (January 2010)
5. Gang, M., Zang, L.M., Fan, W.: Dimensional virtual campus design and implementation. Geomatics (January 2008)
6. He Y., Rong, Z., Jing, Z.: Shaw construction. Qinghai Plateau snow Risk Assessment Model and risk zoning study. Grassland (November 2010)
7. INVENTORYShui deposit, Xie: Virtual reality technology research. Surveying & Spatial Information Technology (February 2009)
8. Yi, Y., Lu, Z., Wu, Q.-Y.: IMAGIS based three-dimensional digital campus design and implementation. Surveying & Spatial Information Technology (March 2009)
9. Zhang, J., Jun, T.: JAVA-based interactive virtual reality system design. Missiles and Guidance (January 2006)
10. Jie, W., Wu, T.: Based on the three-dimensional virtual reality simulation system of rivers. Computer Knowledge and Technology 35 (2005)
11. Huang, X.: Based virtual learning applications of VR technology research. Normal University (2005)
12. Shenhua: Environment based on Windows video capture technology research and application of Dalian Maritime University (2005)
13. Li, Y.: Virtual classroom distance education platform as the core of the research and implementation. Chongqing University (2006)
14. Tang, L.: Based on Virtools real-time three-dimensional scene roaming systems research and development. Daqing Petroleum Institute (2007)
15. Lu, Y.: Complex virtual scene modeling and model simplification method. Huazhong University of Science (2006)
16. Fan, C.D.: VRML-based virtual campus design and implementation of Soochow University (2008)
17. Zhen, L.: Virtual Campus System Research and Implementation of Southwest Jiaotong University (2008)
18. Liu, Y.: VRML-based visualization library system design and implementation of Qufu Normal University (2009)
19. Feng, Z.: Virtual campus roaming system research and practice. Chang'an University (2007)
20. Wei, X.: Virtual reality technology in campus network management application research. Chongqing Normal University (2009)
21. Liu, T., Wei, C., He, J.: Sichuan University of Arts and Modeling virtual campus scenes Research and Implementation. Digital Technologies and Applications (August 2011)
22. Liu, Y., Qin, J., Hung, C. wind: Dimensional virtual campus Model Construction. Digital Technologies and Applications (June 2011)
23. Zhao, H.: Based on ArcGIS and 3D Max virtual campus system establishment. Shanxi Architecture 19 (2011)
24. Liu, Q.: Vega-based virtual campus weather effects visual simulation. Information Technology (July 2011)
25. Meng increase, Ping, M., Liu R.: Creation of three-dimensional and three-dimensional virtual scene play. China Educational Technology (August 2011)
26. Liu, Y., Kun: 3DS Max in the virtual campus dimensional map production application. Popular Science (July 2011)

27. Wei, Y.: Interactive VRML-based virtual campus roaming system design and implementation. Science and Technology Information 20 (2011)
28. Ling, C.J.: Virtual campus and the Comparison and Application. Computer Knowledge and Technology 17 (2011)
29. Li, J.: Virtual reality technology and the virtual campus research and practice - to Weifang University Virtual Campus Case Study. Journal of Engineering Graphics (March 2011)
30. Yang H., Straw, Zhong, Y. as : Based on Virtual Reality Digital Campus Roaming System Design. Guilin University of Electronic Technology (April 2011)

Scene Matching Techniques: Modeling and Analysis

Li Xiang[1,2]

[1] College of Aerospace Science and Engineering, National University of Defense Technology, 410073, Changsha, China
[2] Science and Technology on Aerospace Flight Dynamics Laboratory, 100094, Beijing, China

Abstract. To guide the studies in scene matching, the problem of scene matching is modeled and its influencing factors are analyzed. According to sources, influencing factors are cataloged into three types: 1) changes of scenes; 2) changes of imaging conditions; 3) changes of sensors. For each factor, its mechanism is discussed. After simplification, three important factors, the illumination, projection and multimodal, are considered to be main constraints of current algorithms.

Keywords: scene matching, image registration, feature space, similarity measure, search space and strategy, remote sensing image.

1 Introduction

The human vision and memory system can sense and store the scene. And when met the same scene again, even with different illuminations and viewpoints, human can still recognize it [1-4]. In computer vision, the problem was named Image Matching/Registration [5]. In vision navigation of aircraft, it was called Scene Matching. Scene matching is fundamental to vision navigation missions, but has many problems, say [6-10]: 1) how to improve the correct rate, accuracy and efficiency of matching under different viewpoints, illuminations, and times; 2) how to expand the adaptability of the algorithm to area with less salient objects; 3) how to match images from different sensors.

The performance of a scene matching algorithm includes the adaptability, reliability, efficiency and accuracy. And the studies of scene matching algorithm consist of the followings [11-15]:

(1) Selection of scene matching area;
(2) Feature space;
(3) Similarity metric;
(4) Search Space and Strategy.

Many surveys [16-19] have discussed the state-of-art of current scene matching methods, but seldom addressed the model and influencing factors. To answer these questions, the scene matching problem was modeled. Then, its influencing factors were analyzed.

2 Definition and Modeling

Scene Matching algorithm geometrically aligns the sensed images and the reference images of the same scene, which were taken at different times, from different

Y. Yang, M. Ma, and B. Liu (Eds.): ICICA 2013, Part II, CCIS 392, pp. 184–193, 2013.

viewpoints, and/or by different sensors, according to similarity measurements [20-25]. Since the reference images are calibrated, the coordinate of targets in the sensed images could be known.

An image is a 2D function of grayscale on the coordinates (x, y), the reference image could be denoted as

$$\mathbf{F}_{realtime} = \mathrm{M}(\mathbf{F}_{reference}) \tag{1}$$

The sensed image could be referred as

$$\mathbf{F}_{realtime} = \mathrm{M}(\mathbf{F}_{reference}) . \tag{2}$$

If nothing has changed, the $F_{reference}$ and $F_{realtime}$ should be the same. But differences in viewpoints, times, sensors and illuminations, create the problem in scene matching. These influences are functionals since its independent and dependent variables are 2D functions.

$$\mathbf{F}_{realtime} = \mathrm{M}(\mathbf{F}_{reference}) \tag{3}$$

According to sources, the influences could be cataloged into: 1) changes of scenes; 2) changes of imaging conditions; 3) changes of sensors.

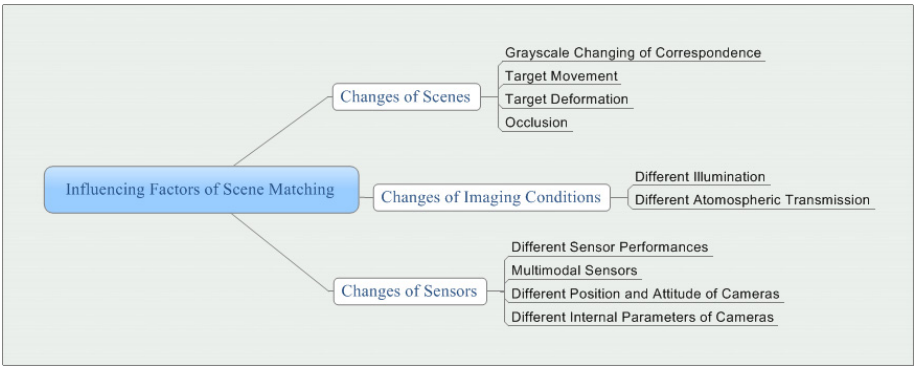


Fig. 1. Influencing Factors of Scene Matching

2.1 Changes of Scenes

2.1.1 Grayscale Changing of Correspondence

Due to the different times at which the reference and sensed images were acquired, the albedo and radiance of the same point may change [26-30]. This kind of changes are stochastic, but are decided by some hidden mechanism. The green grassland in summer may change into yellow when autumn comes. The influence of grayscale changing can be modeled by

$$\mathbf{F}_{featurechange} = M_{featurechange}(\mathbf{F}_{reference}, m) . \tag{4}$$

The generalized materials m denote but not limited to the materials of target points (x, y).

2.1.2 Target Movement, Deformation and Occlusion

The target movement and deformation could also cause differences between the reference and sensed images. The grayscale of the same point didn't change but its position changed. This can be modeled by

$$\mathbf{F}_{featuremovement} = M_{featuremovement}(\mathbf{F}_{reference}) = \mathrm{f}_{reference}(x(u,v), y(u,v)) \quad (5)$$

$$\mathbf{F}_{featuredeformation} = M_{featuredeformation}(\mathbf{F}_{reference}) = \mathrm{f}_{reference}(x(u,v), y(u,v)) . \quad (6)$$

x(u, v), y(u, v) are the movement of correspondence points. If it could be recognized, the influences functionals are modeled.

Occlusion is another influencing factor, which could be considered as grayscale changing.

$$\mathbf{F}_{\text{mask}} = M_{\text{mask}}(\mathbf{F}_{reference}, m) \quad (7)$$

2.2 Changes of Imaging Conditions

In scene matching, the major light source is the sun. The illumination in a sunny day could be modeled as parallel light $\mathbf{L}(l_x, l_y, l_z)$ whose magnitude denotes the light intensity and the direction represents the direction of light. The scene could also be modeled by a 3×1 vector $\mathrm{S}(S_x, S_y, S_z)$ whose magnitude denotes the albedo and the direction represents the surface normal[10].

$$\mathbf{F}_{illumination} = M_{illumination}(\mathbf{F}_{reference}, \vec{L}, \vec{S}) \quad (8)$$

(2) Different Atomospheric Transmission

Atmosphere may dissipate and absorb the signal, causing changes in brightness and contrast of images.

$$\mathbf{F}_{dissipation} = M_{dissipation}(\mathbf{F}_{reference}, d) = d \times \mathbf{F}_{reference} \quad (9)$$

$d \in [0,1)$ denotes the dissipation factor in transmission.

2.3 Changes of Sensors

2.3.1 Different Sensor Performances

Even in scene matching between the same modal of images, the different sensor performances might also influence the images.

$$\mathbf{F}_{sensorperformance} = M_{sensorperformance}(\mathbf{F}_{reference}, para) \quad (10)$$

The para describes the signal-to-noise ratio (SNR), the sensitivity and the resolution and so on.

2.3.2 Multimodal Sensors

To the same scene, the optical, infrared and synthetic aperture radar(SAR) image sensed different features and images.

$$\mathbf{F}_{sensordifference} = M_{sensordifference}(\mathbf{F}_{reference}, m) \tag{11}$$

The mechanism of multimodal images is similar to the grayscale changing of correspondence, but the differences are greater. It has a strong connection with the materials m.

2.3.3 Different Position and Attitude of Cameras

The image difference caused by cameras position and attitude are modeled by

$$\mathbf{F}_{positionattitude} = M_{positionattitude}(\mathbf{F}_{reference}, \mathbf{R}, \mathbf{T}) = \mathrm{f}_{reference}(x(u,v), y(u,v)) . \tag{12}$$

The position changes of correspondence points x(u, v), y(u, v) follow the perspective transformation which is the dependent variables of the translation and rotation.

2.3.4 Different Internal Parameters of Cameras

Focal length difference, lens distortion and other factors are the reason of image distortion. x(u, v), y(u, v) can also model the difference between.

$$\mathbf{F}_{innerpara} = M_{innerpara}(\mathbf{F}_{reference}) = \mathrm{f}_{reference}(x(u,v), y(u,v)) \tag{13}$$

To conclude Eq. (1~13), we get

$$\begin{aligned}\mathbf{F}_{realtime} = (&M_{innerpara} \bullet M_{positionattitude} \bullet M_{sensordifference} \bullet M_{sensorperformance} \\ &\bullet M_{dissipation} \bullet M_{illumination} \bullet M_{\mathrm{mask}} \bullet M_{featuredeformation} \bullet M_{featuremovement} \cdot \\ &\bullet M_{featurechange})(\mathbf{F}_{reference})\end{aligned} \tag{14}$$

Considering the following reasons, Eq. (14) could be simplified.

1) The occurring probability of some influencing factors is low, or can be eliminated through the calibration, say the target movement, deformation, occlusion and the lens distortion.

2) Some factors share similar models and could be merged. The grayscale changing of correspondence is a slight version of multimodal.

3) Some problems have been solved by current methods, such as the difference of brightness and contrast.

$$\mathbf{F}_{realtime} = (M_{sensordifference} \bullet M_{projection} \bullet M_{graymapping} \bullet M_{illumination})(\mathbf{F}_{reference}) \tag{15}$$

In Eq. (15), $M_{illumination}$ denotes the functional of illumination difference; $M_{projection}$ describes the functional of perspective projection; $M_{sensordifference}$ is the functional of multimodal; $M_{graymapping}$ means is the functional of linear and non-linear grayscale change between images. Since $M_{graymapping}$ could be solved by gray histogram transformation, $M_{illumination}$, $M_{projection}$ and $M_{sensordifference}$ are three main problems in current scene matching.

3 Discuss

High-tech is currently the worldwide rapid development of a wide range of high-tech applications, are profoundly changing the world the face of economic and military struggle, triggering a series of revolutionary changes in the military field, showing weaponry information, intelligence, integration trend, firing accuracy than ever before. Former U.S. Defense Secretary William - Perry in March 1996 to the President and Congress an annual report for the first time proposed a " military revolution " concept, refers to the use of new technologies with innovative military systems operational concepts and organizational adaptation combined, to fundamentally change the characteristics of military operations and processes. The report also predicts changes in future wars, the first is the long-range precision strike, and the second is information warfare. Also put forward depth precision strike likely to dominate future wars. Modern warfare must have a high-performance, multi- range precision attack missiles. Such missiles are implemented in future space warfare, the primary means of depth to combat the world are stepping up the development and deployment.

March 24, 1999, the US -led NATO against Yugoslavia brazenly taken a military air strike (hereinafter referred to as " Operation Allied Force "), and on May 8 savagely bombed my embassy in Yugoslavia. In this act of aggression, the U.S. military used a lot of advanced precision-guided weapons, precision strike weapons used by major cruise bombs, which hit precision and power in the world would concern. Tomahawk cruise missiles used by the U.S. military is mainly Block 3 type, which is used in the Gulf War of BGM-109C / D developed based on, using INS + match + GPS + terrain scene matching guidance, a maximum range of 1,667 km (ship-launched) / 1127 km (submarine), with an accuracy from 3 to 6 m (theory), cruising altitude of 15 to 150 meters, cruising speed of Mach 0.5 to 0.75, the price is about $ 1.4 million. In the " Operation Allied Force " in the use of precision-guided bombs mainly Joint Direct Attack bombs (JDAM) and so on. Improvements made by the conventional JDAM bombs, using GPS INS plus composite guidance, accuracy of 10 meters, the price of $ 12,000, the U.S. Navy and Air Force fighter F/A-18C/D B-1 \ B-2 bomber can carry.

Type Tomahawk missiles increased primary fuel, combustion efficiency and reliability can be improved, increasing the body processes than type 2 556km, reached 1852km. It uses global positioning system and inertial navigation system (GPS / INS)

for the guidance, with improved digital scene matching area correlator and auxiliary terrain matching for terminal guidance. Improved digital scene matching area correlator en route to the missile to the target display signs digital images, reducing the seasonal and diurnal changes in guidance accuracy, increasing the effective picture area can be accurately into a space the size of football gate. If the target changes (such as a target to be destroyed), playing on the computer can still recognize some scenes, accurately calculate the navigation flight attitude and correction data. Through the GPS system can readily determine the three-dimensional spatial position and missile flight speed, so that the missile does not deviate from the course, control the missile target from different directions, while significantly reducing mission planning time. These measures have greatly improved the accuracy of the missile, when it reaches the range of 2000km, circular error probability less than 10m. 3 type Tomahawk also improve the ability to accurately reach the target time in order to facilitate the theater with other missiles, aircraft and boats generally work together, the " arrival time " control software makes far apart from the different time zones fired several missiles almost Meanwhile the two countries hit land targets.

American strategists believe that any launch Tomahawk missiles will not endanger the safety of personnel, is a suitable "zero death " weapons used in the war. America is currently developing more advanced stealth cruise weak, the radar reflection area 0.01m2 below, and the use of low-altitude subsonic flight to avoid enemy radar detection and tracking, but also control cable of Mach 8 hypersonic cruise possibility missile. One of the most striking is to be deployed in 2005 on the use of Tomahawk 4 type GPS receiver, the receiver aircraft and satellite synthetic aperture radar images generated in the guidance, the use of infrared imaging for terminal guidance. And type 3 Tomahawk digital scene matching area correlation terminal guidance compared radar images from weather conditions can be achieved by less than 5m CEP accuracy.

Scene matching terminal guidance is precision-guided weapons hit the target accurately the key reason why cruise missiles can accurately destroy targets, mainly due to its homing phase using a digital scene matching area correlator (Digital Sceme Macthing Area Correlator, DSMAC), ie next, as the scene matching system. When the cruise missile flew close to the target late, the next scene matching system through the visual image matching technique to accurately determine air missile position and the target position deviation is calculated and sent to the navigation computer, and guide the missile accurately hit the target.

The so-called matching, means that two different sensor positions from the same scene admitted down the two images be aligned in space, to determine the relative displacement between the two processes. Will be pre- scene photo shoot to the ground, according to the pixel size digitized maps made in the development of the missile flight path before performing a task, select the range of response area of the scene as a benchmark figure matches the stored onboard computer. If the reference image scene features obvious, easy to locate missile exact match, it is called matching area. When the missile flew to a predetermined position, the playing ground below the video camera image, the image points according to size, flying altitude and field size and other parameters to generate a real-time graph, but also to match the computer. In matching computers, real-time map and the relevant comparative reference map to

find the location of the two. Since the geographic coordinates of the reference image (or the relative position of the target) is known in advance. Accordingly, it is with real-time graph of the registration location, the missile can be determined relative to the target position, which is the course of action scene matching system.

In the complete scene matching process need to use selection criteria to determine the matching area (such as the gray -related criteria, the criteria for space-frequency domain features, geometric criteria, etc.), scene preprocessing (image enhancement, histogram equalization, image thinning, image to noise, image geometric distortion correction, etc.), scene feature extraction (extraction of regional relevance scene, scene edge tracking, frequency domain parameter extraction scene, scene extraction of morphological parameters, etc.), scene analysis (by means of a existing set of scene selection criteria extracted from the scene image to compare the parameters, the synthesis process), scene evaluation (for comparison, the Composite after a given reference map available in the scene to evaluate and output), coarse matching location algorithm, fine positioning algorithm and match results matching evaluation algorithm.

Our scene matching research started relatively late, at present there is a certain distance from the world level, according to publicly available information, is still applied to the actual shape of weapons useless scene matching system. National Laboratory of Pattern Recognition scene matching related technologies are being studied, hope that through close cooperation with the military, Integration Laboratory has been in image processing, pattern recognition formed the professional advantages for the modernization of national defense and make new greater contributions.

4 Conclusions

To register image acquired at different times, from different viewpoints, and by different sensors, and improve the adaptability, reliability, efficiency and accuracy of scene matching algorithm, the problems of illuminations, projection and multimodal must be comprehensively investigated. Under the influence of various factors, finding the invariance of features is the key to scene matching.

The so-called matching, means that two different sensor positions from the same scene admitted down the two images be aligned in space, to determine the relative displacement between the two processes. Will be pre- scene photo shoot to the ground, according to the pixel size digitized maps made in the development of the missile flight path before performing a task, select the range of response area of the scene as a benchmark figure matches the stored onboard computer. If the reference image scene features obvious, easy to locate missile exact match, it is called matching area. When the missile flew to a predetermined position, the playing ground below the video camera image, the image points according to size, flying altitude and field size and other parameters to generate a real-time graph, but also to match the computer. In matching computers, real-time map and the relevant comparative reference map to find the location of the two. Since the geographic coordinates of the reference image (or the relative position of the target) is known in advance. Accordingly, it is with real-time

graph of the registration location, the missile can be determined relative to the target position, which is the course of action scene matching system.

Implement this guidance, be pre- selected in the vicinity of the target area as distinct geomorphological features scene matching area, obtained through reconnaissance optical image of the scene matching area, and to be divided into several square scene matching small units. According to the average light intensity per unit, converted into corresponding values constitute scene matching area reflects the light intensity of each unit digital scene maps stored in the missile 's computer. When the missiles fly through the scene matching area, playing on a television camera captured scene image has been digitized, with the pre- stored digital scene maps were compared to determine whether the missile deviated from the predetermined route. If the deviation occurs, the resulting error correction control signal missile route.

In the complete scene matching process need to use selection criteria to determine the matching area (such as the gray -related criteria, the criteria for space-frequency domain features, geometric criteria, etc.), scene preprocessing (image enhancement, histogram equalization, image thinning, image to noise, image geometric distortion correction, etc.), scene feature extraction (extraction of regional relevance scene, scene edge tracking, frequency domain parameter extraction scene, scene extraction of morphological parameters, etc.), scene analysis (by means of a existing set of scene selection criteria extracted from the scene image to compare the parameters, the synthesis process), scene evaluation (for comparison, the Composite after a given reference map available in the scene to evaluate and output), coarse matching location algorithm, fine positioning algorithm and match results matching evaluation algorithm.

Scene matching technology after decades of development, is widely used, such as digital image processing, environmental studies and so on based on it, the breadth of its application produced many different matching algorithms, each algorithm has its specific applicable scope of this paper is to study a major work in a changing environment, anti-rotation change, the resolution of inconsistencies, gray change, shift changes and noise matching algorithm, this paper describes in detail the scene matching principle relevant knowledge, areas for further study algorithm provides a theoretical basis and reference indicators 7 scene matching algorithm divided into four categories according to their characteristics, this paper is combined with a domain -based approach to change in scene matching, image rotation angle of resolution inconsistent, contrast changes, translation and noise is caused by changes in the mismatching of the main reasons for the algorithm used in this paper to try to overcome the above several distortion interference in order to accurately use a method that first removed the largest inscribed circle image vulnerable to loss of external rotation or additions from the pixel, and then the image is logarithmic polar coordinate transformation, logarithmic polar coordinate transformation Cartesian coordinate system can be rotated and sized images mapped to a logarithmic polar coordinate system horizontal and vertical shift change, the final use of the image fourier transform amplitude correlation matching translational invariance which search strategy to use method, measurement methods used to mean normalized product correlation algorithms.

In this paper, based on the probability of false positives based on minimum selecting scene matching method currently used evaluation criteria for selecting scene matching

methods for evaluation, for the existing methods influenced by the image quality, low immunity, low accuracy and computationally intensive, and many other issues of image content integration selecting scene matching method this paper, especially for the content of the fused image selecting scene matching method, the key factors (landscape scene content and the relationship between adaptation and simulation scene matching probability confidence) conducted intensive studies, given a different degree of adaptation of scene content value, the text that matches the simulation model used in high reliability of the paper selection method also feasibility and effectiveness of other aspects of the demonstration. this paper proposes two specific content fusion scene selecting scene matching method : content distribution based on GIS and landscape scene matching area selection method ; edge density based on GIS and Image Matching District selection method. These two methods presented in this paper based on the minimum probability of miscarriage of justice under the meaning of the evaluation criteria have higher matching zone selection accuracy and efficiency to meet selecting scene matching the requirements of the experiment shows both methods have a high reliability and good matching area selection feasibility.

Acknowledgement. This work is supported by the Open Research Foundation of Science and Technology on Aerospace Flight Dynamics Laboratory(Grant No. 2012afdl041).

References

1. Zitova, B., Flusser, J.: Image registration methods: a survey. Image and Vision Computing 21, 977–1000 (2003)
2. Wyawahare, M.V., et al.: Image Registration Techniques: An overview. International Journal of Signal Processing 2(3), 11–28 (2009)
3. Brown Gottesfeld, L.: A survey of image Registration Techniques. ACM Computing Surveys 24, 325–376 (1992)
4. Deshmukh, M., Bhosle, U.: A Survey of Image Registration. International Journal of Image Processing 5(3), 245–269 (2011)
5. Dawn, S., Saxena, V., Sharma, B.: Remote Sensing Image Registration Techniques: A Survey. In: The International Conference on Image and Signal Processing, pp. 103–112 (2010)
6. Wang, H., Zhang, K., Li, Y.: Research Progress on Image Matching. Computer Engineering And Applications 40(19), 42–44 (2004)
7. Tan, Z.-Y., Sun, J.-Y., Wang, Z., Zhang, C.-X.: Status and Prospect of Algorithm for Scene Matching Systems. Computer Technology And Development 22(9), 66–70 (2012)
8. Guo, Q.: Scene matching technology development outline. Infrared and Laser Engineering 36(z2), 57–61 (2007)
9. Guo, Q.: Modeling and simulation of the scene matching and orientation system. Infrared and Laser Engineering 35(z1), 295–300 (2006)
10. Oren, M., Nayar, S.K.: Generalization of Lambert's Reflectance Model. In: SIGGRAPH, pp. 239–246 (1994)
11. Liang, X., Wu, H.-Z., Liu, T.C.: Digital scene matching image simulation modeling and simulation generated. Ordnance Technology (January 2005)

12. Xing, S., Zhang catalog, Tao, L., Wang D.-B.: Based on Landsat ETM data larch forest biomass models. Fujian College of Forestry (February 2004)
13. Sun, G., Yu, G.: Remote sensing technology in highway engineering geological exploration balance date of application. Communication Science (January 2004)
14. Gao, Y., Qiang, Z.M., Zhu, J., Shuai: Jiangxi red soil area residents TM imagery to automatic information extraction Expert Mode. Sensing (April 2002)
15. Delger, Wang, Q.-C., Zhou L., Sun A., Xu Restoration,Ying, Q., Liu B.: 1997-1999 Yellow River ecological effects of artificial rainfall areas Maqu test. Plateau Meteorology (March 2005)
16. Qiang, Z.M., Gao, Y.: Surveying. Fuzhou area of soil erosion TM imagery expert mode of automatic information extraction East China Geological Institute (March 2002)
17. Cheng, J., Zhou, C.-P., Ding, M.-Y., Chao, C.: Different weather conditions, outdoor scene image generation method. Infrared and Laser Engineering (February 2001)
18. State wins, Zhao, H., Zhang Z.: Terrain digital landscape plan simulation involves several issues. Ocean Mapping (February 2003)
19. Zhao,Y., Lei, Mao, X.: Under visual scene simulation of real-time image distortion attitude. Computer Simulation (October 2006)
20. Chen, E.K., Yu, C.-Z., Wang, S., Tan, M.: One kind of vision-based real-time obstacle avoidance ROBOFISH integrated approach. Control and Decision (April 2004)
21. Wang, D.: Multi-modality image registration techniques and large research. Chinese Academy of Sciences (Institute of Electronics) (2002)
22. Zhao, B. love: Based on wavelet analysis method of high resolution radar target recognition. Harbin Engineering University (2003)
23. Wang, Y.: SAR image and optical image data fusion algorithm. Electronic Science and Technology (2003)
24. Shi, Y.: Digital information pattern recognition theory and its application. Shandong University of Technology (2003)
25. Zhang, J.: INS / GPS / SAR navigation system key problems of Northwestern Polytechnical University (2003)
26. Cheng, Q.M.: Content-based remote sensing image database search key technology research. Chinese Academy of Sciences (Institute of Remote Sensing Applications) (2004)
27. Liang, X.: Based on variational PDE and multi-fractal image modeling theory, algorithms and applications. Nanjing University (2004)
28. Mei, J.: Based on support vector machine High Resolution Image Target Detection. Wuhan University (2004)
29. Lu, H.: Formation based on satellite remote sensing images on the positioning algorithm. National Defense Science and Technology University (2004)
30. Shan, D.: Morphology in Remote Sensing Image Processing Application Research. Wuhan University (2004)

Networked Intelligent Retrieval System Based on Semantic

Meng Zhang, Chanle Wu, and Gang Ye

Computer School of Wuhan University, Wuhan, Hubei, China, 430072

Abstract. This paper makes full use of the advantages of the semantic ontology based on framework of semantic intelligent retrieval system, improved the traditional retrieval system that only query contains keywords,which lack limits between semantic relation conceptions, so as to enhance the relevance and associative retrieval results, improve system reusability,reliability, standardization and retrieval rate.

Keywords: semantic, ontology, Intelligent Retrieval System.

1 Introductions

On the Internet, dedicated to provide query service website is search engine, the use of network search software will be collected on the Internet web page for processing, and then establish the database, in response to user queries, and provide the information [1-5]. For the contradictions between Web information retrieval and the mass of information, Google, Baidu search engine play a role in mitigation in a certain extent. but the search engine is more of a keyword search under normal circumstances, keywords from the user's retrieval demand does not match, not a lot of webpage results and user demand were also retrieved, seriously affect the retrieval efficiency and quality [6-9]. And if the semantic Web search engine, you can better understand and judge the user demand, make the search more accurate.

2 Intelligent Retrieval System

Intelligent retrieval system is different from traditional information retrieval system model, in the information retrieval process, introducing the concept of semantic information processing of query objects, and query expansion and semantic analysis processing operations through the intelligent, improve the precision and recall of information retrieval algorithm, not only improves the performance of information retrieval system, but also improve the the user experience [10].

Intelligent retrieval system should have the ability to understand the semantic. This understanding of the semantics of the function should include synonyms reasoning user retrieval words and word retrieval analysis of correlation function [11-16]. When users enter search item, the system should be able to make a reasonable segmentation for search input, and analysis of the search term do inference relations between word and

Y. Yang, M. Ma, and B. Liu (Eds.): ICICA 2013, Part II, CCIS 392, pp. 194–203, 2013.

word association of each word, so as to understand the meaning of the most close to the user demand, and returns the best results and related results [17-20]. This paper presents the key technology of semantic intelligent retrieval system contains: Topic Crawler Based on domain ontology construction, construction technology

Construction technology and construction of the three pieces of topic search engine main parts technology (as shown in Figure 1).

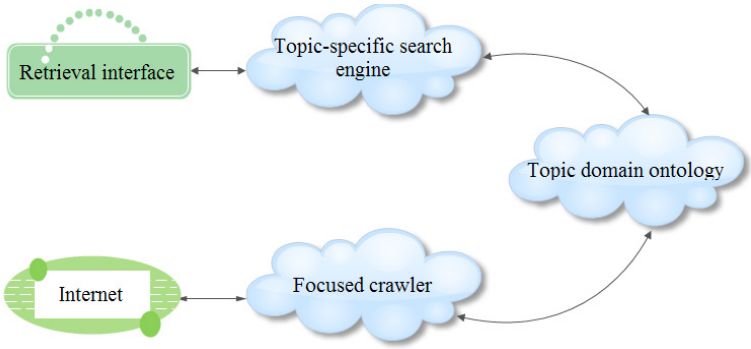


Fig. 1. The cloud picture of intelligent retrieval system based on semantic

3 The Design of Semantic Retrieval System Based on the Intelligent Network

3.1 System Model

The intelligent retrieval system should be composed of three models: Subject reptiles model, domain ontology model, themes search engine model. We will integrate these three models, you can get a complete semantic-based intelligent retrieval system model (Figure 2). The role of the functional modules is:

(1)Function module of subject reptiles crawled relevant pages related sites on the Internet. Crawl pages saved to disk, these web document data preprocessing by the file handler completes, and added to the database.

(2) Ontology functional module is the participation of experts in the field together to develop SWRL rule set to complete the construction of ontology and knowledge base, and build complete ontology file XML file parsing and conversion to form a triple characteristic data sheet.

(3) Search engine function module for each data table in the database to complete the creation of the index files and the associated index files are stored, and provides a user interface to be queried. Retrieved by the search user interface input request by the search engine function module to complete the search term word processing and indexing retrieval, and ultimately get back to return the search result set.

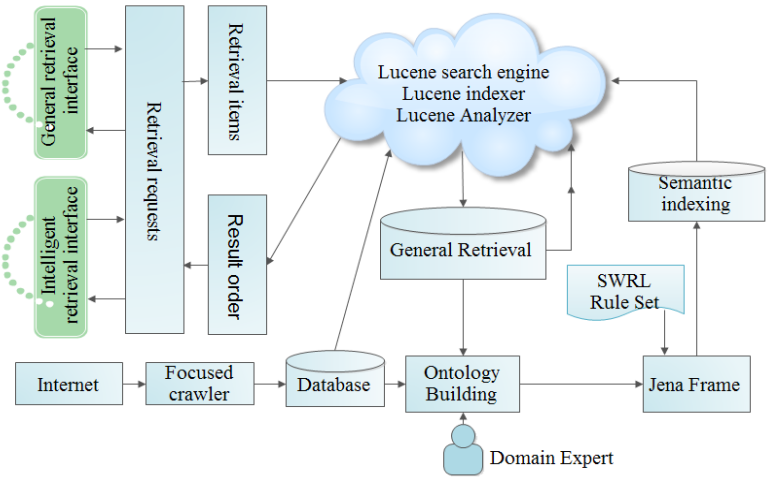


Fig. 2. Model of retrieval system

3.2 Query Expansion of Semantic

Document Retrieval Based on Semantic Web technologies to achieve not only functional of semantic retrieval system, but also to achieve a semantic extension query. For users, the function of the main manifestations: Enter keywords; users can not only get information with keywords related resources, but also can be obtained with the words synonymous resource information. Such as user input, "Lu Xun article" The key words in the search results you get "" Scream "", "" Weeds "" and other related resource information. The search results show that the system truly synonymous for keyword queries and keyword-based retrieval is impossible to achieve semantic query expansion. Able to get the concept of upper and lower relationship queries: for example, Lu Xun is a famous thinker and writer, then the body system, "Lu Xun" and "thinkers, writers," between the relationship is a bit up and down, if the user in the system to retrieve "Lu Xun," the keyword information in the system, enter "thinkers, writers," the same access to that realization concept hyponymy query.

These two functions are implemented to ensure retrieval recall ratio. Technical semantic retrieval system also can effectively guarantee retrieval precision, because in this system, using ontology concepts defined attributes and concepts described information resources, so the system can be entered by the user concepts and concepts attribute information resources to achieve accurate positioning. Example of the "modern writers" the specific described as: Users need to use the search system queries the "modern writer" relevant information, when the input "Zhejiang", "thinker" or "writer" These words, the system will accurately locate in "Lu Xun" and his work information for display. For the user, this results in one step, does not display other irrelevant information, thus ensuring the precision of search results.

```
<rdf:ID="modern writers">
<Website:regionOf Rdf:resource="# Zhejiang"/>
<Website:relatedFruit Rdf:resource="# Literary authors"/>
<Website:fullDesc Rdf:datatype="Luxun">
```

3.3 The Realization of Semantic Networks

Ontology be able to express accurately between domain concepts and relationships concepts, it is ontology-based technology to achieve system functions. System can achieve recall and precision. In a certain sense, the recall rate and precision are mutually contradictory, this system by extending synonymous and hyponymy to alleviate this problem, so that the system of recall and precision rate. This semantic network model is constructed based on the field of computers, is a semantic network consists of nodes and directed arcs semantic network consisting of a directed graph. Where nodes represent concepts, while these concepts are represented semantic relationships that exist between. Solid and hollow point and non-point represent keywords keywords. The concept of different levels of abstraction demonstrated the degree varies, the higher the level, it contains the generic concept will be more and embracing the stronger will. A group of the generic concept of combined together to form a broader concept, usually, the next bit of a broader concept of an abstract representation of the concept; while the generic concept of the concept is usually a host further refine and complement the unique properties of their own the same description also has the properties of a broader concept. From this perspective, the semantic model is like a tree, its first layer is the top of the classification tree, which means that a separate topic, and then the implementation of the subject under layers gradually refined. Based on a tree structure based on the added horizontal relations is to be relatively independent of each concept together, for example, "Lu Xun" and "Zhou Shuren" are synonymous relationship between, and the "Lu Xun" and "social drama" between the upper and lower part relationship, and ultimately the formation of semantic networks.

4 The Design Theme of Search Engine Architecture

Different from the traditional search engine model, we use the fully open source Lucene text search tool kit. Article presents a semantic-based search engine retrieval model, a key part of its implementation is to establish semantic indexing. On the one hand, carries a semantic technology semantic indexing spatial reasoning ability, it can be inferred that the implication of each ordinary relationship between the index. On the other hand, the semantic index associated with a common index, makes a lot of common underlying index because of association do not need to re-retrieved retrieved, thus greatly improving the retrieval efficiency. Lucene software package by the indexer, retrieval and memory components. Related classes with our design shown in Figure 6, which are interpreted as follows::

(1) BuildVocabulary.java: This class implements a dictionary to build, according to the literature database dictionary file titles and keywords built vacabulary.dic.

(2) SousuoAbout.java: This class is used to retrieve data stored in the database literature major fields, such as title, content, images Url, Url original address and so on.

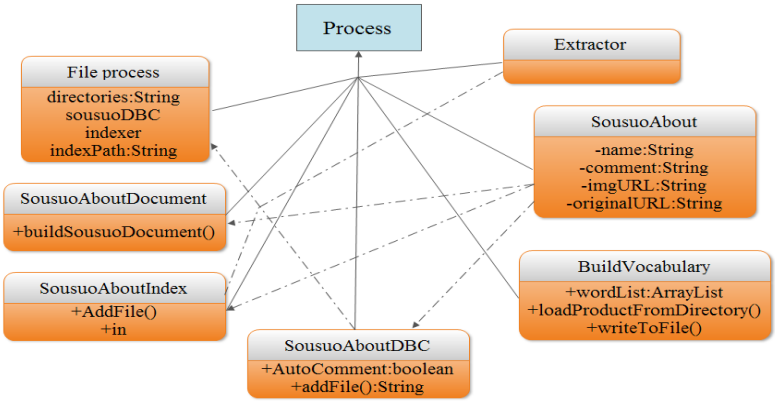


Fig. 3. Related classes diagrams of Lucene Class

(3) SousuoAboutJDBC.java: This class is mainly to complete the database links and data write operation.

(4) FileProcess.java: Completion of the main topics such as crawling reptiles network Chu literature page processing, and then stored in a database. For example, the contents of the file stored on the crawl, pictures Url Url address the storage and the original storage.

(5) SousuoIndex.java: This class is mainly to complete the index build operation. The key is to add the word.

(6) SousuoAboutDocument.java: This class is completed the construction of the literature indexed documents.

5 Search Example of Intelligent Retrieval System Site

We use the server platform operating system is Microsoft Windows Server2003, database MySql 5. A front-end interface to retrieve dynamic pages using JSP design by Web services tools Tomcat7. 0 release. For example, enter search term "Lu Xun", the jump results page shown in Figure 4.

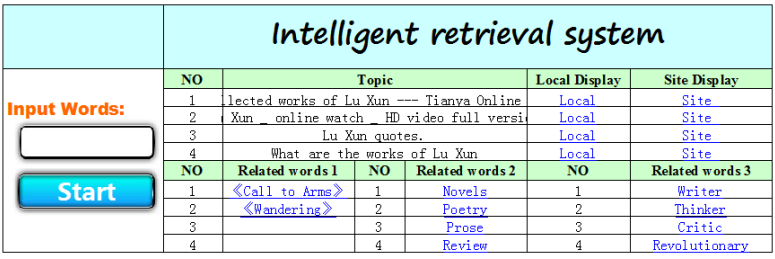


Fig. 4. "Lu xun" search result

6 Discuss

With the development of Internet, online information resources exponentially. If breathing sea in ho Internet information, the user should find useful information becomes very difficult. This is because the information needs of users of the limited nature of the opposite sex and inter- specific online information resource distribution and dispersion and infinite nature of contradictions ; addition, existing search engine does not provide good personalized service. Therefore, how to analyze the user's interest to create a user preference model, thus to provide users with excellent personalized service to become a hot research topic. User model is personalized service system key technologies. Personalized service system first need to create a user model, and then to different interests for different users prefer to provide personalized service. Traditional user models are mostly used words to characterize user interest, without considering between words intrinsically linked to the information source itself does not use domain knowledge for the user interest model building services. In order to provide better personalized service, this paper presents a recommendation based on semantic Web technology systems, with an emphasis on user model which has been studied. The research work mainly includes the following aspects: (1) based on the analysis of the user model, we propose a new user model. The model is based on the established interest in classification ontology, ontology classification by instantiating interest to get the body of each user profile. This paper draws on information retrieval commonly used spreading activation model to complete the user model is updated. (2) In this paper, the hierarchy of user interest representation, the traditional content-based recommendation method has been improved. To recommend to the user when this full account of the user's interest in the hierarchy of the concepts of interest in the role of values , by calculating the value of each class of interest in the concept of adding a user interest vector. This can better reflect the user's interest. (3) In this paper, the theoretical basis of the study, the realization of a Semantic Web-based recommender system technologies to help government departments to complete the work of public opinion analysis. In the framework of the system, including three modules: news and information collection and processing module, the user model module and the recommended policy module. In the news and information collection and processing module, through the news from news sites crawled information to complete its review information collection, and the comments under each news information in a simple clustering process. User model module uses the proposed ontology-based user model and update method. We first created news areas of interest classification body, and then to get the user by instantiating description file, and use spreading activation model to update the user model. These recommendation strategies used to improve the content -based recommendation method. A new generation WWW will also be able to provide data processed by a computer, which will make a lot of intelligence services possible " ; Semantic Web research goal is to" develop a series of computer-readable semantic information and processing of language and expression technology to support extensive and effective network environment automated reasoning. " semantic Web founder Tim Berners-Lee on the semantic Web is defined as follows : " semantic Web is a network that contains the document or part of a document that describes between things the significant relationship, and contains the semantic information, to facilitate the automatic processing machine. " despite the understanding and description of the

Semantic Web is different from but still be seen in the description and understanding of some basic features of the Semantic Web : (1) semantic Web Unlike now WWW, it is the expansion and extension of the existing WWW ; (2) the existing WWW is a document-oriented and document-oriented semantic Web is data represented ; (3) semantic Web will be more conducive to the computer "understanding and treatment." and has a certain judgment, reasoning ability.

Semantic Web is dependent on three key technologies: XML, RDF, and Ontology. In Section 3 This article will discuss them. Although the Semantic Web to show us the bright future of the WWW and hence bring the Internet revolution, but the realization of the Semantic Web is still facing enormous challenges: (1) the content of accessibility, which is based on Ontology and Semantic built Web pages are still rare ; (2) the development and evolution of the body, including the core ontology for all areas of the development process of developing methods and technical support, annotation and ontology evolution and versioning issues ; (3) content scalability, ie with the Semantic Web content after how scalable way to manage it, including how to organize, store, and search, etc. ; (4) multi- language support ; (5) body language standardization.

The first layer : Unicode and URI. Unicode is a character set, the character set, all characters are represented by two bytes, can represent 65,536 characters, basically the world including characters in all languages . Benefits of using Unicode data format is that it supports the entire world's major languages mix, and can be retrieved simultaneously. URI (Uniform Resource Identifier), or Uniform Resource Locator, which uniquely identifies a concept or on a network resource. In the Semantic Web architecture, the layer is the basis for the Semantic Web, which is responsible for processing resources Unicode encoding, URI identifier of the resource is responsible.

The second layer: XML + NS + xmlschema. XML is a simplified SGML, which combines the rich features of SGML and HTML, ease of use, which allows users to add arbitrary structure of the document, without having to explain the implications of these structures. NS (Name Space) that namespace identified by the URI index, the purpose is to avoid different applications use the same character describe different things. XML Schema is DTD (Document Data Type) alternatives, which itself uses XML syntax, but more flexible than the DTD provides more data types, can better serve as a valid XML documents and provide data verification mechanisms. Because XML is a flexible structured, indexed by the URI NS brought certainty and XML Schema data can be provided by a variety of data types and inspection mechanism, making it the Semantic Web architecture important part. This layer is responsible for representing data from the grammatical content and structure of language through the use of standard forms of network information, data structures and content separation.

Third layer: RDF + rdfschema. RDF is a description of the information resources on the WWW a language whose goal is to create a variety of metadata standards for coexistence frameworks. The framework can fully exploit the advantages of various metadata, Web-based data exchange and re-use. RDF to solve is how to use the XML standard syntax unambiguously describe the resource object problems, make the resources described in the metadata information to become machine-understandable information. If you look at the XML metadata as a standardized syntax specification, then the RDF can be seen as a kind of standardized metadata semantic description specification. Rdfschema use of a machine system to be understood that the resource definitions describe terms, the aim is to provide a mechanism or framework embedded

words, the framework may be integrated together to achieve the various expressions of the description of Web resources.

The fourth layer: Ontology vocabulary. This layer is in RDF (S) defined on the basis of an abstract description of concepts and their relationships, used to describe the application of knowledge in the field, describing various types of resources and relationships between resources, to achieve the expansion of vocabulary. In this layer, users can not only define the concept and to define the relationship between the concept of wealth.

Fifth to seven: Logic, Proof, Trust. Logic is responsible for providing axioms and inference rules, and Logic, once established, it can be a resource through logical reasoning, the relationship between resources and the inference results have proved its effectiveness. Through Proof exchange and digital signatures, establish a certain trust relationship, thus proving the reliability of the Semantic Web output as well as its compliance with the requirements of users.

Semantic Web architecture under construction, within the scope of this current international architecture research has not yet formed a satisfactory alternative to strict logical description and theoretical system architecture of our scholars only in the research done on the basis of foreign brief introduction, have not formed a systematic exposition.

Semantic Web implementation requires the support of three key technologies : XML, RDF, and Ontology. XML (eXtensible Marked Language, you can Extensible Markup Language) allows information providers as needed, self- defined tags and attribute names, so that the structure of XML documents can be complicated to any extent. It has good data storage format and scalability, and ease of highly structured network transmission, etc., coupled with its unique mechanism of the NS XML Schema supports multiple data types and checking mechanism, making the Semantic Web one of the key technologies. Currently on the Semantic Web discussion of key technologies focused on RDF and Ontology body.

7 Conclusion

For traditional search engines is inaccurate, this paper presents a new semantic-based literature search engine, to achieve the exact keyword search. In the semantic network environment literature retrieval system design process, we discussed the function of system requirements, design, implementation, and the function of literature retrieval system flowchart. The system basically realizes the semantic retrieval system capabilities.

In addition to the Semantic Web country scholars conducted a systematic introduction, is also on the Semantic Web and its key technology to do some useful experimental and applied research, from the technical point of view with which they can be broadly divided into three categories: Semantic Web architecture and its implementation, RDF technology, Ontology and construction applications. In outlining the basic Semantic Web technologies, based on the design of a business-oriented Semantic Web portal architecture SPortal; put forward the concept and role of proxy page and function; designed the Semantic Web query language CDQL. This paper SPortal the overall structure and the functional modules made a detailed introduction, is

done in the framework of the Semantic Web a very useful first attempt, they made the pages agents, Semantic Web query language designed for domestic semantic CDQL Web development has important reference value.

In the RDF technology applications, respectively, from the XML / RDF features start discussions based on XML / RDF for MARC and DC metadata description techniques and gives a detailed description of examples ;

Our research on the Semantic Web has the following characteristics:

(1) In recent years, research on the Semantic Web surge in the number of papers described the Semantic Web as a future trend of development of the Internet has caused great attention of scholars.

(2) The references of the papers are all essentially foreign literature, indicates that China's Semantic Web research in its infancy.

(3) Research on the Semantic Web has a certain hierarchy, from a basic introduction to the theoretical discussion, from the theoretical discussion to practical test and application layers and progressive, reflecting a certain hierarchy.

(4) On the Semantic Web research is focused on the discussion on the body, including the body 's basic meaning of formal representation, ontology, ontology languages and ontology-based application of various theoretical and experimental study. This is because the Semantic Web architecture in the fifth to the seventh layer (Logic, Proof, and Trust) is also being built into the other, that Ontology is the core of the Semantic Web, Semantic Web research is currently the crux. It is not only the formation of the resources, knowledge and so on the basis of common understanding, but also support the inference mechanism of the computer may " understanding and treatment " basis.

Our research on the Semantic Web has made some progress, but there are still some shortcomings and deficiencies : (1) lack of theoretical studies, mostly in brief Semantic Web and its related technologies on the basis of further discussion, while the semantic Web architecture, methodological principles, building strategy and other expositions less ; (2) references from the papers of view, research on the semantic Web has not yet formed their own professional features and system ; (3) for the preparation of the body and build just try and do some preliminary testing, not to mention the preparation of the corresponding norms and standards ; (4) the latest theories of foreign research results and practical application tracking is not enough.

Therefore, the main future research directions and key issues include: (1) Semantic Web basic theoretical research, including Semantic Web architecture, guidelines, construction methods, etc. ; (2) the practical application of the Semantic Web, Semantic Web vision and establishing a WWW page or site ; (3) the preparation of norms and standards body for guiding Chinese network ontology ; (4) the practical application of research body, the establishment has a certain practical value vocabulary (Vocabulary) and inference mechanisms ; (5) for Logic, Proof, Trust research layers ; (6) for the latest foreign theoretical research and practical application of timely information and discussions.

References

1. Lu, X., Zheng, Q.: Based on user behavior analysis of search engine model. North China University of Technology (March 2004)
2. Feng, W., Bing, D.: Domestic Chinese Search Engine [J]; Computer Knowledge and Technology (Academic Exchange) (August 2007)

3. Wu, Q.-M., Yong, Zhou, C.: Knowledge-based engineering of complex product development process model. Journal of Engineering Design (May 2008)
4. Chen, J., Deng, F., Liu, F.: Intelligent search engine analysis and discussion. Guangzhou University (Natural Science) (March 2002)
5. van Harmelen, F.: Landsat, leading experts. Secret Semantic Web. Programmers (November 2007)
6. Mei, W.-K., Lu, Z., Sun, X., Li, R.: Semantic Search Research. Computer Science (May 2008)
7. Xiangyang, Min, Ma Q.: Jena -based ontology construction method. Computer Engineering, 14 (2007)
8. Hu, Y., Li, S., Guo, M.: ontology-based product knowledge representation. Computer Aided Design & Computer Graphics (December 2003)
9. Pay, X., Shanping, G.M.: Product data model ontology knowledge representation. Computer Aided Design & Computer Graphics (March 2005)
10. Peng, G., Lin, L., Ming, C., Dong, J.: Product configuration based ontology mapping model automatically obtain. Computer Integrated Manufacturing Systems -CIMS (September 2003)
11. Li, Y., Lu, Z., Sun, X., Li, R.: Privacy protection based on the semantic data integration. Huazhong University of Science and Technology (Natural Science Edition) (2005)-S1
12. Nie, Left, X.-R., Chen, D.-L.: Ontology mapping an improved concept similarity computing approach. Computer Applications (June 2008)
13. Cattle for collar, Discussion WordNet trends and application. Technology Rich Wizard 26 (2010)
14. Wei, H., Minxi: Ontology-based intelligent retrieval system modeling books and Application Research. Changchun University of Technology (Natural Science) (February 2006)
15. Jiang, Y., Li, B., Wu, Z.: semantic integration of distributed information systems architecture SP4IDIS. Computer Engineering and Design (March 2008)
16. Xu, S., Cai, Y.: Before, books based on ontology model of intelligent retrieval system. Sichuan Institute of Technology (Natural Science) (May 2009)
17. Chen, Lu, Z., Wang, Y.: Ontology- based XML Schema multiplexing method. Computer Engineering and Science (November 2005)
18. Zhang, S.Q.: Qiang, Ping, Chen X.: Ontology based manufacturing grid construction Mechanism. Application Research of Computers (August 2008)
19. Xian, L., Hessler, Fan, L.: In the ocean. Ontology-based natural language information extraction in software testing application. Central University for Nationalities (Natural Science Edition) (March 2008)
20. Jiang, L.: Semantic Web-based courses Retrieval System. Nanchang College of Education (May 2011)

Design and Implementation of Computer Immune System

Zhao Tao

Yunyang Teachers'College, Shiyan Hubei, China, 442000

Abstract. Designing a novel computer security system based on the basic principles and ideas of Artificial Immune System (AIS) is a new direction in computer security technology research. Based on artificial immune principles, the paper makes deep exploratory research on computer immune technology based on the host. Starting from the system call, the paper makes exploratory research on computer immune response mechanism, which has important practical significance for studying computer security technology.

Keywords: computer security, biological immune system, computer immune system.

1 Design and Implementation of System Calling Sensor

1.1 Linux System Call Mechanism

Under Linux operation system, the process can't access the kernel, which means that it not only can't access the memory space of kernel, but also couldn't call kernel functions [1-4]. CPU hardware determines the protected mode. Linux operation system includes kernel space and user space which mean the memory reserved by the kernel and user process. As the memory spaces are independent of each other, the user process can't directly access kernel space, and the kernel can access user space only by put_user and get_user macro and the similar macro. System call is the interface between operating systems of processes, so system call needs to interact with user space frequently [5-9].

In Linux system, the process of Linux executing system call is as shown in Figure 1. It includes four steps [10-13]. (1)When the user requests system call, a soft interruption of Int0x80 is firstly produced to tell the kernel and transmit parameter information. (2) After the interruption is responded, it enters kernel mode. In interrupt service routine, the names of interrupt request are used to search interrupt service entrance address table sys_call_table and find the corresponding service program entrance address. (3) By using the addressed pointer, the kernel jumps to the system call entry to implement genuine system call functions. (4) Restore_all ends system call and returns the user process.

Sys_call_table which is defined in arch/i386/kernel/entry.S records the entry address of system calling service functions. Open system call needs to be implemented. SYMBOL_NAME (sys_open) in sys_call_table is the entry address of service function sys_open () of the kernel. Ret_from_sys_cal is a part of system_call. It can be directly

Y. Yang, M. Ma, and B. Liu (Eds.): ICICA 2013, Part II, CCIS 392, pp. 204–213, 2013.

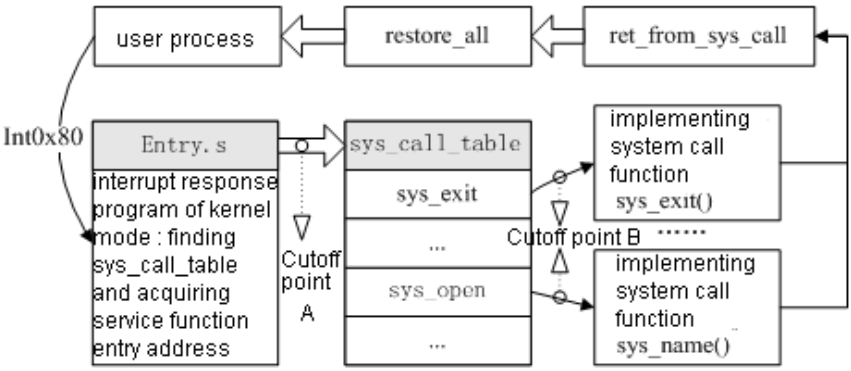


Fig. 1. Implementation process of Linux system call

called by C and can skip from other parts of system_cal. When a system call or interrupt is completed, the function is called to process outstanding work. Before system_cal returns, restore_all is used to make pull operation on the values in the register, and returns the user mode.

1.2 Linux System Call Interception Mechanism

In Linux system, strace and ps order the program to automatically track system call of target process [14-18]. Strace is used to trace the process and debug. It uses ptrace system call to monitor application programs called by other system calls. Strace program is easy to be used, but strace program can make key security procedures like Sendmial collapse. And when strace program works, it makes the implementation efficiency of the monitored program reduce by 50% or more. After these are ready, a data collection thread is generated to collect the data collected by data collection threads together. And the collected data is formatted, and receives data match. If it is mismatching, it needs to report to the checking platform in the upper layer, and the activities in the process need to be stopped temporarily until it receives the detection reply of the upper layer. The method is similar to unobtrusive scripting. And the advantage is that the programming is easy and the debug is convenient. And it is generally used for system debugging.

2) Writing LKM

LKM program is loaded into the kernel of operation system as a part of kernel. So its source file is not like common C language program which has main function. The kernel of the whole operation system operates as a main function, so LKM program has no main functions. Interception point of LKM method is point B in Figure 1, the reason for which is that it only uses the variables in kernel symbol table of modifying the system to replace the original system call.

Typical LKM program mainly includes two parts, init_module and cleanupe_module. Init_module is the called function for LKM program initialization, and cleanup_module is the called function for LKM program unloads.

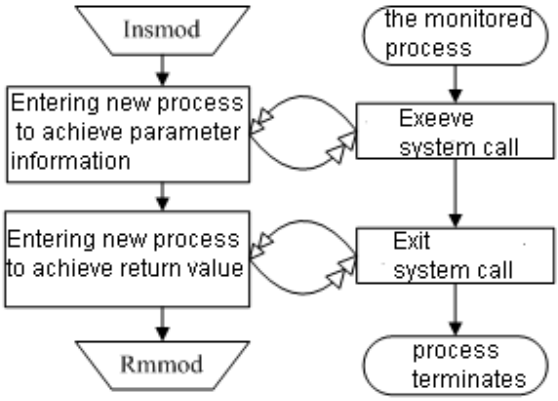


Fig. 2. Process of designing LKM sensor

Table 1. System calls which can't be intercepted by LKM

System call number	System call name
11	sys_ execve
22	sys_oldumount
52	sys_umount
113	sys_vm86old
114	sys_wait4
119	sys_sigreturn
127	sys_create_module
128	sys_init_module
129	sys_delete_module
149	sys_sysctl
162	sys_nanosleep
167	sys_query_module

1.3 Performance Requirements and Evaluation of System Call Sensor

Linux system call mechanism is very complicated, which has a higher requirement to acquire Linux system call information. On one hand, system call sensor needs to intercept system call data in kernel space. On the other hand, it needs to establish user space interface of data to make other working parts in user space process the data effectively.

The requirements on system call sensor include:

(1)The core part should be small to reduce the impact on the system.

System call sensor works in system core and intercepts system call process. When the system calls operate, they need to receive information filter of sensors firstly, and then skip to the interception point, which completes the operation of the original system call. When the frequency of system call is higher, the core part of sensor has an

influence on the implementation efficiency of the monitored process. Therefore, the core part of system call sensor should be small and only can complete basic data interception. The other work can be completed in user space, which can reduce the impact on implementation efficiency of system service process.

(2)User space should be rapid, continuous and real-time.

The user space of system call sensor is the interface of other working parts accessing system call information. It realizes data buffer and completes preliminary data process such as format conversion. User space must process large amount of original data timely and rapidly, and the work should be continuous.

2 Immune Cell Generation Module

In the intrusion detection system based on immune mechanism, the detector is a system call sequence with the length of K. In order to make negative selection, the system firstly generates two sequence libraries. One is to generate system call sequence randomly, and the other is to generate system call sequence when the process implements normally under confinement. The latter uses rough set theory to solve the problem that priori knowledge of immune system is deficient and to construct perfect vaccine based on prior knowledge. The design of the former is to take account of the randomness of system call, which makes the detector detect abnormal situations as much as possible. 164 system calls received random process. K system calls were selected randomly every time to compose a detector of system calling short sequence with the length of K. And the program is as follows.

```
For( i=0;i<7;i++)
{
    Randomize();
    A[i]=random(164);
    Printf("%d\n",a[i];
}
```

3 Negative Selection

Before lymphocytes of an organism become mature and are used as detector cells, they need to go through the process called immune tolerance, which means not respond to autoantigen. Antibody generated randomly can combine with self-generation autoimmunity, so if a mature lymphocyte combines with itself successfully, it will be cleared. And lymphocytes without combining with itself will become mature detector, and the process is negative selection.

The key of negative selection introduced in the paper is to match and compare the system call sequence which is used as quasi- detector and the system call sequence in normal-behavior pattern library established after rough set analysis.

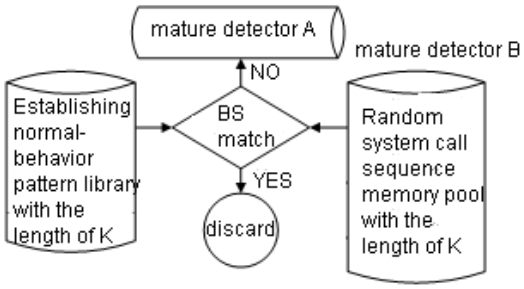


Fig. 3. Negative selection algorithms

4 Dynamic Selection

4.1 Generation of Memory Detector

Memory detector has the characteristic of longevity. But the number is limited, and the idle time of rule in life cycle is greater than the threshold, so the rule in memory detector will degenerate into mature detector. Memory detector applies misuse detection method for pattern matching to find a self-nonself model, which will make immune alarm.

For the given sequences, j and j, the maximum hamming distance of them is calculated, which means non-matching degree of the system call sequence.

dmax(i,j)=max{d(i,j),j is a system call sequence used as detector.

The method of solving d(i,j) is introduced in Hamming algorithm in Chapter 4.3. And the concrete algorithm is shown in literature [40].

From the above comparison, we can know time complexity of the whole matching process. In order to determine is a new sequence is mismatching, no more than k-1 times of comparison are needed. The system call sequences in detector library are generally stored with the form of number, so every root node corresponds to a system call. If the short sequence has never appeared in detector, its Hamming distance needs to be calculated. As dmax(i,j) is the maximum Hamming distance between sequence i and all sequences in detector library, the sequence i must be compared with all sequences in detector library, which needs to compare for N(k-1) times (N is the number of short sequences in detector library).

When dmax (i,j) is less than the threshold, the detector should be recorded, and the detector should be transferred from mature detector set into memory detector set. After dynamic selection for a while, the detectors which are not activated in life cycle will be abandoned.

4.2 Update of Mature Detector

If the sequence set CA matched with mature detector A doesn't include all non-selves, the attacker can find and use the leaks to avoid the detection of the detector. The longer

the time that the set CA matched with the detector keeps invariant is, the greater the possibility that the attacker finds the leaks is. Once the leaks are found, the attacker can use it to attack the network. One scheme of reducing the leaks is to increase the number of detectors, so that the set CA includes all non-self series. But the scheme is not feasible because of the resource. The other scheme is that the set CA changes with the time. Like lymphocytes which have life cycle, it can provide a certain life cycle for each detector. When the life ends, the new detector is used to replace. If all detectors have the same life cycle, all detectors may die simultaneously at some time. When age is greater than the life cycle, the detector dies with the probability P. If the detector detects one match, the age reduces some time t, and the age should be ensured to be greater or equal to 0. The updating algorithm of the age can ensure that the most mature detector die eventually.

4.3 Advantages of Dynamic Selection

If the series of non-self set has never been detected by memory detector, they will be not detected by memory detector in the future. Therefore, mature detector is necessary. It makes non-self set matched with the detector changes continuously, which can detect the non-self which has never been detected. Memory detector ensures that non-self series which has been detected can be detected again, which accelerating the speed of detecting non-self. Overall, memory detector and mature detector are combined mutually, which not only improves the detection efficiency and deletes permanent detection leaks, but also improves the speed of detecting non-self which has been detected. It makes the detection system have the advantage of misuse detection and abnormal detection method.

5 Life Cycle of Detector

As the network stage changes continuously, the detector must make timely response. According to the principles of Biological Immunology and the objectives that we should achieve, the dynamic life cycle of the model detector is summarized, as follows.

1) Immature Detector. It is the system call sequences which are generated by the system randomly, and there is no negative selection. It corresponds to immature T cell which has not went through thymus negative selection in Biological Immunology system.

2) Mature Detector. It corresponds to immature B cell.

3) Memory Detector. It corresponds to memory B cell.

The life cycle of the model detector can be divided into three stages, as follows.

1) The system generates immature detector set randomly and applies immature detector to normal mode library or self set for matching by negative selection. And the immature detector which matches with self becomes mature detector. The process is continuous in the system and keeps updating like immune cells.

2) Mature detectors monitor system network. When the mature detector detects once matching, the matching times of the detector increase by 1. After the appointed time,

mature detector detects if its matching time exceeds the pre-set activation threshold and if the life of the mature detector achieves the pre-set life value. If the matching count of a mature detector exceeds the activation threshold, and the detector detects once intrusion again, the detector becomes memory detector. If the matching count doesn't exceed the activation threshold and achieves the pre-set life, the detector will be abandoned.

3) Memory detector uses misuse detection technology to detect system call sequence of privileged process timely. If it detects the intrusion, the alarm should be made immediately for immune response and the detection time of memory detector should be updated. The idle time of memory detector should be detected regularly. If it is greater than the threshold, it will degrade into mature detector.

6 Computer Immunology Research Applied to Network Security

At present, the international community launched the "Computer Immunology" study focused on defense, military, security sector applications, more representative are:

① University of New Mexico FORREST American research team with an in-depth analysis of the biological immune mechanism is proposed based on a computer model of the immune system, and the corresponding algorithm - negative selection algorithm. The experimental results show that this method can easily find unknown virus infection, to further improve the security of the computer system.

② IBM Research Center Kephart and others by simulating biological immune system functional components and identify foreign antigens, analysis and removal process, designed a computer model and the immune system, which is designed primarily " bait " to catch the virus samples extracted virus signatures, virus removal and design appropriate procedures.

③ Spafford Purdue University and Wright-Pafferson AFB Air Force Institute of Technology Marmelstein have in-depth analysis of computer virus research significance, research methods and security requirements, and the corresponding computer model of the immune system.

④ Japan Ishida Toyohashi University of Science and Technology is also based on the immune system of a computer virus defense techniques conducted in-depth study and application of multi- Agent technology and implementation, carried out in a computer network monitoring and removal of computer viruses work, and gives against networks will be characterized by the virus infected files and system repair method.

National immunization research in the computer just started. Wuhan University proposed multi-agent based model of computer security and the immune system detects the structure and evolution on Self collection methods, and in the "Self", "Nonself" recognition rules in conducting research with the evolution of the mining method to extract rules based on system call is established based on the bit string identifier, learn some of the characteristics of the food chain, the establishment of a multi- recognizer collaborative recognition model ; Wuhan University and the Northern Jiaotong University, proposed host-based security scan computer immune system detects ; Northern Jiaotong University presents a intrusion detection model based on immune

and immune stochastic processes introduction of computer studies ; Nanjing University of Aeronautics and Astronautics on the use of antiviral immune mechanism techniques were studied ; Beijing Institute of Technology Department of Automatic Control from the control theory point of view on a computer immunological and biological immune similarity between the field of computer anti-virus application of multi- agent control technology to build a computer simulation of biological immune system feasibility and practicality.

Network security is a three-dimensional depth, multi-layered defense integrated system for anomaly intrusion, viruses, and so can be obtained from the natural immune system a lot of inspiration. Future research directions will concern the following areas:

① distributed feedback control : the use of natural immune system highly distributed nature and mechanism of parallel processing, in the framework of a computer intrusion detection system on a distributed architecture that detects when the parties jointly to obtain data for analysis, and to take joint action defensive measures, and efficient respond to a variety of complex attacks.

② The hybrid intrusion detection : with the traditional design ideas in a simple system architectures, hybrid intrusion detection is a multi -level, based on the structure of the system host and network-based hybrid architecture ; detection algorithm in use anomaly detection and misuse combined with a detection mode ; detection method is applied in real-time detection and interval detection based mixed complementary strategy.

③ Multi Protection System features : many of today's information security systems are borrowed from some of the characteristics of natural immune system, but it has all the characteristics of information security systems has not yet appeared, efforts in this direction, research a robust, distributed, adaptive information security systems has important practical significance.

④ genes Computer: Immune-based computer system has a stronger genetic identification and protection, it is through the detection of genetic code to determine the legitimacy of data, only the genetic code coincides with the end of the transceiver can manipulate data, genetic code is automatically generated and can not be artificially intervene, with good safety performance.

Immune computer technology has developed rapidly, more and more people at home and abroad engaged in theoretical research and design, along with natural immune subjects, intelligent simulation technology, will be able to build a more powerful computer immune system, to solve more and more serious computer virus problem has become the mainstream of network security technology.

7 Conclusions

The paper analyzes operation mechanism of Linux system call in detail, introduces interception methods of system call under Linux, expounds using kernel patch method and LKM technology to implement workflow and programming method of system call sensor, and puts forward the requirements on system call sensor. The paper also introduces computer immune response mechanism based on system call, and the paper

introduces and analyzes security policies including forcing to end illegal process, analyzing interception information and delaying implementation process.

Computer security has been explored in the field of computer hot topic, a variety of new technologies and products are endless, but the practical application is far not meet the requirements, however, vigorously strengthen computer security issues in the information society talks about the process of building more obvious and complicated. In recent years, a method of biological immune mechanism is applied to the field of computer security technology, namely the emergence of computer immune system but also for the resolution of this issue provides a new way. Mimic biological immune mechanism, we have designed a computer system security model GECISM, the model constituted by a plurality of agents, each agent mimic different immune cell function and mechanism, through mutual cooperation to protect the security of the host. This article focuses on GESCIM classes MC Agent Agent tectonic studies rely on the operating system environment is Linux. GESCIM MC Agent class is the interface with the external environment, is mainly responsible for the anomaly detection, the detection is based on the generated run-time system call sequence. Collection, the rule base and the detector are the main class of MC Agent component. Class MC Agent collector in the Linux kernel-level implementation, the article describes in detail the class MC Agent on system call system call sequence acquisition and will be divided into a short sequence of methods, and to analyze system call sequence mode. This article gives a short sequence set in the system calls the application of data mining techniques to generate rule-based approach, and details of each application and the CART algorithm C4.5 algorithm to generate the rules of procedure and experimental results. Class MC Agent rules in the rule base is a set of IF... THEN... statement that sets the rules can be judged through a short sequence of system calls are normal. The article also details the class MC Agent detector used to generate rules for abnormal process detection methods, and the experimental results and analysis.

This article learn the basic principles of artificial immune technology and ideas to build a new type of computer protection system, the immune system is the computer technology research field of computer security a new direction. Based on artificial immune principle, the paper -based host computer technology for more in-depth immune exploratory study: This thesis work includes the following parties where: Artificial Immune Systems and Computer Immune technical introduction. This paper describes the basic concepts of artificial immune system and the development history, discussed several typical algorithms, including those based on the basic mechanisms of the immune system, immune algorithm, negative selection, dynamic selection algorithm, immune lifetime Computer construction of the immune system. Reference biological immune mechanism, this paper discusses the immune system to build a computer general principles. On this basis, the biological immune principles to computer security system design and implementation in order to improve existing methods inadequate. This article from the perspective of system calls, the computer immune response mechanism of exploratory research, the immune response of a computer model that can well reflect the initial response of biological immune response and secondary response mechanism, to achieve a computer immune response mechanisms of learning and memory.

References

1. Xie, X., Wu, F.: Foreign overview of the development on computer security. In: Fourth National Computer Security Seminar Proceedings (1989)
2. Liu, D.: How we work to strengthen computer security monitoring. In: Tenth National Computer Security Symposium (1995)
3. Hui, H.: Computer security policy banks in one instance. In: Eleventh National Computer Security Seminar Proceedings (1996)
4. Editor's Note. First National Computer Security Seminar Proceedings (1986)
5. Pei, L.: Foreign legislation on computer security and the domestic situation. In: First National Computer Security Seminar Proceedings (1986)
6. Nakai, D.: Computer security management status of personnel needed safety training. In: Ninth National Computer Security Symposium (1994)
7. Zhang, J.: Computer security software environment for building some of Thinking In: Eighth National Computer Security Symposium (1993)
8. Peng, H.: Based on rough set theory Intrusion Detection Method. University of Electronic Technology (January 2006)
9. Zhang, X., Lin, J.-J.: Registry- based Distributed Intrusion Detection System. East China University of Technology (June 2004)
10. Zhang, X., Gu, C.-H., Lam, D.: Based on support vector machine Windows host intrusion detection system. East China University of Technology (Natural Science) (March 2006)
11. Lu, H.: Wei K., Kong H.: An Improved KMP efficient pattern matching algorithm. Huazhong University of Science and Technology (Natural Science) (October 2006)
12. Li, Y., Li, W.: For intrusion detection pattern matching algorithm. Computer Engineering and Applications (June 2003)
13. Bin, W., Fang, B.-X., Yun, X.: Integrity checking file-based intrusion detection and recovery technology research. Computer Engineering and Applications, 31 (2003)
14. Fu, Y., Tang, G., De, L.: Procedures based on system call sequence abnormal behavior detection. Computer Engineering and Applications (June 2006)
15. Chen, X., Yang, S.-B.: Intrusion detection system file integrity check of Computer Science (September 2002)
16. Zhang, Y., bright, Xiao S.-P., Guoyu: Based on Rough Set Attribute Reduction Support Vector anomaly intrusion detection method. Computer Science (June 2006)
17. Clear, Zhang, Y., Xiao, S.-P., Guoyu: Based on Rough Set Theory and process system call sequence anomaly detection method. Computer Science (August 2006)
18. Xingwang, Ming, C.: Immune -based Intrusion Detection System Framework. Technological Development of Enterprise 19 (2010)

Incremental Updating Algorithm of Weighted Negative Association Rules

He Jiang and Wenqing Lei

School of Information, Shandong Polytechnic University, Jinan, China,
Jianghea09@126.com
School of Information Engineering,
Shandong Youth University of Political Science, Jinan, China

Abstract. Incremental updating algorithm for mining negative association rules is different from positive association rules mining. With the continued increase of the data records in the database, incremental updating association rules technique represent an important class of knowledge that can be discovered from data warehouses. Incremental updating algorithm is important for mining infrequent item sets in dynamic databases. In this paper, we proposed an incremental updating algorithm of weighted negative association rules (WNARI). By comparing weighted itemsets and unweighted case, the number of incremental updating negative association rules on weighted condition is less than the unweighted.

Keywords: Negative Association Rule, Weight, Incremental Updating, Infrequent Item Set.

1 Introduction

The association rule is an important topic in the data mining and knowledge discovery research field, was originally pointed out by R. Agrawal in 1993 [1]. Has attracted many scholars' wide attention and a number of algorithms have been proposed to generate frequent item sets. The algorithm is built on the basis of the database is static. And Items in the database are equally important. In the real world, the database is often dynamic. Such as add or delete some database records in the original database when necessary. And the importance of items is often different. The model of weighted association rule and corresponding algorithm are proposed in [2]. An Incremental updating algorithm of negative association rules were proposed in [3].

All the traditional association rule mining algorithms were developed to find positive associations between items. By positive associations we refer to positive correlation between items existing in transactions. But few consider the negative correlation between items. The forms $A \Rightarrow \neg B$, $\neg A \Rightarrow B$ and $\neg A \Rightarrow \neg B$ mean negative correlation between them. Correspondingly, together with the rules of the forms $\neg A \Rightarrow B$ and $\neg A \Rightarrow \neg B$, are called negative association rules (NARs) and the rules of the form $A \Rightarrow B$ positive association rules (PARs). Positive association rules exist only

Y. Yang, M. Ma, and B. Liu (Eds.): ICICA 2013, Part II, CCIS 392, pp. 214–222, 2013.

in frequent item sets, while negative association rules exist not only in frequent item sets, but more exist in infrequent item sets.

The negative rules are mined from frequent item sets in [4][5]. An algorithm WNRIF, mining weighted negative association rules from infrequent items is proposed in [6]. Paper [7], an algorithm for mining weighted negative association rules from infrequent item sets based on multiple supports (WNAIIMS) is proposed. In [8], three algorithms are proposed to prune weighted negative association rules. It ignored the dynamic changes of the database. To solve this problem, an algorithm which mining positive and negative association rules from dynamic database was proposed [3]. But the algorithm did not consider the importance of the items in transaction database and negative association rules from infrequent item sets.

In this paper, we comprehensively considerate the problems of all previous researches, improve them, propose an weighted negative association rules of incremental updating algorithm(WNARI) to mine the weighted negative rules from incremental updating database. The algorithm simultaneity considerate the problems of dynamic changeable of transaction database, the importance of different items and negative rules from infrequent item sets The original database and new added database are scanned only once generating infrequent item sets in the algorithm. The efficiency of generating is improved and the space of search is shortened. The concept of correlation is introduced to the algorithm made the rules are more practical significance. Experimental results show that the algorithm is effective.

2 Problem Statements and Related Work

2.1 Negative Association Rule

A negative association rule is an implication, such as $A \Rightarrow \neg B$ (or $\neg A \Rightarrow B$), where A, $B \subset T$, and $A \cap B = \Phi$. Given the support s and confidence c, if D has $(100 \times s)\%$ affairs which contain A while do not contain B, so, the support of the negative association rule $A \Rightarrow \neg B$ is the frequency of occurrence of transactions with item set A in the absence of item set B, remembered as $supp(A\neg B)=s$; If the affairs that contain A has $(100 \times c)\%$ affairs which do not contain B, then the confidence of the negative association rule $A \Rightarrow \neg B$ is c, remembered as $conf(A\neg B)=c$. We can also define the support and confidence of the negative association rules $\neg A \Rightarrow B$. Mining negative association rules is that selecting all the negative association rules $A \Rightarrow \neg B$ (or $\neg A \Rightarrow B$) which can satisfy the users' minimal support-minsup and the minimal confidence-minconf in *D*, that is to say, the support and confidence of the negative association rules are not less than the minimal support and minimal confidence respectively, and both *A* and *B* are frequent item sets.

2.2 Incremental Updating Algorithm for Mining Association Rules

With the passage of time, the size of the database will continue to expand or need to remove part of records, or need to adjust the minimum support, so gradually to gather

the frequent item sets we are interested in. So how to update the already derived association rules effectively has very important application value from the database that data has changed. This is the so-called incremental updating question for the association rules.

The generalized updating question is that, mining new association rules in the new database DB+, which after added(or subtracting) data set db in the base of original database DB. The incremental updating question for the association rules includes three new questions:

In the condition of minimal support and minimal confidence remain unchanged, how to generate the association rules in the $db \cup DB$ when a new data set db adds to database DB; In the condition of given minimal support and minimal confidence, how to generate the association rules in the DB-db when the database DB subtracts a data set db; How to generate the association rules in the database DB when the minimal support and minimal confidence have changed.

Agrawal R and Srikant R proposed the FUP updating algorithm [9], which solves the first question. In paper [10], the FUP2 algorithm which can solve the first and the second questions was proposed. The IUA and PIUA algorithms [11] were proposed, which can solve the third question. The study on the incremental updating algorithm for mining negative association rules is few compared with the incremental updating for mining positive association rules [12-16].

In fact, the incremental updating algorithm for mining negative association rules is similar to the incremental updating for mining association rules. That is, mining negative rules in the updating database. However, they also have some different performances: The positive association rules only exist in the frequent item sets, but the negative association rules not only exist in the frequent item sets, but more exist in the infrequent item sets [17-21]; The positive association rules only have one form A=>B, while the negative association rules have three forms: $\neg A \Rightarrow B$, $A \Rightarrow \neg B$ and $\neg A \Rightarrow \neg B$; When solving the incremental updating question for mining positive association rules, we just need to obtain all the frequent item sets in the updating database, then produce the positive association rules using the formulas; while solving the incremental updating question for mining negative association rules, we need to obtain all the frequent and infrequent itemsets, then mining the negative association rules by using the mining algorithm.

3 Incremental Updating of Weighted Negative Association Rules

Mining association rules of transaction database can be seen as a collection of item sets. The collection is divided into two mutually exclusive parts: frequent item sets and infrequent item sets. We use DB and db to respectively represent original database and new added database. T represents the total number of the transaction database DB, and t denoted db. DB.F, db.f and DB+.F respectively represents the frequent item sets of DB, db and DB+. The infrequent item sets of DB, db and DB+ denoted by DB.NF, db.nf, DB+.NF.

When DB ∩ db=∅, clearly, any subset of DB.NF and db.nf is all infrequent itemset

When DB∩db≠∅, itemsets divided into disjion part and intersection part. Disjoint part may arise a collection of frequent and infrequent itemsets: L=DB.F–db.f, L'' = db.f–DB.F, NL= DB.NF – db.nf, NL' = db.nf – DB.NF. Intersection part can be divided into four subsets L1, L2, L3, L4.

L1=DB.F∩db.f, each itemset of L1 is frequent itemset in both DB and db.

L2=DB.NF∩db.nf, each itemset of L_2 is infrequent itemset in both DB and db.

L3=DB.F∩db.nf, each itemset of L3 is frequent itemset in DB and infrequent itemset in db.

L4=DB.NF∩db.f, each itemset of L_4 is infrequent itemset in DB and frequent itemset in db.

3.1 Algorithm Design

The main idea of WNARI algorithm

Input: the original database DB, additional database db, wminsupp, wminconf.
Output: weighted frequent and infrequent item sets (WFI, WIFI) of DB, db, DB+

Weighted positive and negative association rules (WNARs) of DB, db ,DB+

Steps:

Determine DB ∩ db=∅ or DB ∩ db ≠∅

If DB ∩ db=∅ do

//weighted frequent itemsets and infrequent itemsets were Generated when DB ∩ db=∅

L_1= find_frequent_1_itemsets (DB); NL1=C_1-L_1;

C_k=apriori-gen (L_{k-1}, wminsupp);

For (i=2; i<=size; i++)

{Ci=join (C_{i-1}); Y= C_{ij}

For (j=1; j<= C_i.count;i++)

For (k=i+1; k<=size; k++)

$$W[Y,k]=\sum_{i_j\in Y} w_j+\sum_{j=1}^{k-q} w_{r_j}$$

B(Y, k)=[(wminsupp*T)/ w[Y, k];

If SC(Y)>=B(Y, k) // prune step using k-support expectation

L_k=L_k ∪Y; WFI =WFI∪L_k,

Else

NL_k=NL_k∪Y; WIFI= WIFI∪ NL_k

L_k={c∈C_k |c.count•wminsupp}; NLk= C_k–L_k

//With the same mothed generate all possible *k*-item sets in *db*.

//mining weighted positive and negative association rules in Frequent item sets when DB ∩ db=∅

For any itemset A∪B=X and A ∩ B=∅

If wsupp(A∪B)-wsupp(A)(B)>0

{If wconf(A⇒B)>=minconf

Then WPARs=WPARs∪ {A⇒B}

If wconf(¬A⇒¬B)>=minconf

Then WNARs=WNARs∪ {¬A⇒¬B}

Else

{If wsupp(A ∪B)-wsupp(A)(B)<0

{If wconf(A⇒¬B)>=minconf

Then WNARs=WNARs∪ {A⇒¬B}

If wconf(¬A⇒B)>=minconf

Then WNARs=WNARs∪ {¬A⇒B}

If DB ∩ db≠∅, infrequent itemsets are Generated when DB ∩ db ≠∅ For DB ∪db, each itemset of L1 is certainly frequent itemset and L2 is certainly infrequent itemset. Of L3 and L_4

For all t ∈ db

For all itemset A ∈L_3

If A∈ t then L_3.A.count++;

For all itemset A ∈L_3

Prune step using k-support expectation

If L3.A.count •WS* (T + t) then L_3= L_3∪ {A};

Else NL_3=NL_3∪ {A};

With the same method finds L_4 and NL_4

L''=L_1∪L_3∪L_4, NL''= NL_2∪NL_3∪NL_4

// mining negative association rules in inFrequent itemsets when DB ∩ db≠∅

For any itemset A ∪B=X and A ∩ B=∅

If wsupp(A∪B)-wsupp(A)(B)>0

{If wconf(A⇒B)>=wminconf

Then WPARs=WPARs∪ {A⇒B}

If wconf(¬A⇒¬B)>=wminconf

```
Then WNARs=WNARs∪ {¬A⇒¬B}
Else
{If wsupp (A∪B)-wsupp(A)(B)<0
{If wconf( A⇒¬B)>=wminconf
Then WNARs=WNARs∪ {A⇒¬B}
If wconf(¬A⇒B)>=wminconf
Then WNARs=WNARs∪ {¬A⇒B}
```

3.2 Analysis Experiment Results

We conducted our experiment on a man-made dataset to study the algorithm. The dataset has 300+50 transactions, DB=300, db=50, when eight largest categories were kept.

When the items are not weighted, the minimum support (minsupp=0.05) is equal, but the minimum confidence (minconf) is set different values (20%, 40%, 60%, 80%). the result as follows in table 1 and table 2

Table 1. The number of NARs on the condition of the same minsupp from dataset *DB*

minconf	PARs from FI	NARs from FI			NARs from IFI		
	A⇒B	-A⇒-B	A⇒-B	-A⇒B	-A⇒-B	A⇒B	-A⇒B
20%	127	33	46	38	13	59	33
40%	70	33	46	28	13	59	12
60%	29	15	27	1	5	56	0
80%	1	5	13	0	3	50	0

Table 2. The number of NARs on the condition of the same minsupp from dataset DB+db

minconf	PARs from FI	NARs from FI			NARs from IFI		
	A⇒B	-A⇒-B	A⇒-B	-A⇒B	-A⇒-B	A⇒-B	-A⇒B
20%	136	54	11	7	32	56	33
40%	88	54	11	5	32	56	23
60%	30	44	11	0	25	51	7
80%	1	29	4	0	22	50	3

From the experimental results, when the minimum support is certain, minimum confidence increases gradually. The numbers of positive and negative association rules reduce gradually.

Set the weight of each item based on the user's interest.

A: 0.1, B: 0.2, C: 0.9, D: 0.4, E: 0.6, F: 0.3, G: 0.5, H: 0.8.

When added weight values in item, the weighted minimum support (wminsupp=0.05) is equal, but the weighted minimum confidence (wminconf) is set different values. Using the same data, the result is as follows in table 3 and table 4.

Table 3. The number of NARs on the condition of the same *wminsupp* from dataset *DB*

wminconf	WPARs from WFI	WNARs from WFI			WNARs from WIFI		
	A⇒B	-A⇒-B	A⇒-B	-A⇒B	-A⇒-B	A⇒-B	-A⇒B
20%	104	37	0	0	68	5	3
40%	66	37	0	0	68	5	0
60%	30	37	0	0	68	5	0
80%	12	24	0	0	55	2	0

Table 4. The number of NARs on the condition of the same *wminsupp* from dataset *DB+db*

wminconf	WPARs from WFI	WNARs from WFI			WNARs from WIFI		
	A⇒B	-A⇒-B	A⇒-B	-A⇒B	-A⇒-B	A⇒-B	-A⇒B
20%	58	33	0	0	51	8	6
40%	36	33	0	0	51	8	1
60%	20	33	0	0	51	8	0
80%	8	24	0	0	44	8	0

By comparing the two groups experiment, after the items have been weighted, the number of rules is less than the unweighted case.

4 Conclusions

WNARI algorithm mines frequent and infrequent itemsets from the DB and db which makes full use of the improved Apriori algorithm. WNARI algorithm uses the nature of set to divide DB and db into intersection and non-intersection. When $DB \cap db = \Phi$, some frequent itemsets in DB and db will become infrequent after adding db to DB. So the count of infrequent itemsets in the updating database will increase; When $DB \cap db \neq \Phi$, the infrequent itemsets are constituted by the infrequent itemset from the intersection of DB and db, DB.NF and db.f, DB.F and db.nf intersect respectively. The advantage of WNARI algorithm is that only scanning database DB and db a few times, which can short the scanning time. So the algorithm is effective.

Acknowledgments. This work was supported by Shandong Province Nature Science Foundation (No.Y2008G26) and A Project of Shandong Province Higher Educational Science and Technology Program (No.J09LG32).

References

1. Agrawal, R., Imielinski, T., Swami, A.: Mining Association Rules between Sets of Items in Large Databases. In: Proceedings of the 1993 ACM SIGMOD International Conference on Management of Data, vol. 11(3), pp. 207–216. ACM Press, New York (1993)
2. Cai, C.H., Fu, A.W.-C., Cheng, C.H., et al.: Mining Association Rules with Weighted Items. In: Proceedings of the International Database Engineering and Application Symposium, Cardiff, UK, vol. 13(2), pp. 68–77 (1998)
3. Sun, B.Y., Jiang, H.: Incremental updating algorithm for mining association rules. Computer Engineering 5(6), 676–677 (2009)
4. Zhao, Y.Y., Jiang, H.: Research of mining positive and negative weighted association rules based on Chi-squared analysis. In: Proceeding of The Second International Conference on Information and Computing Science Manchester, vol. 14(11), pp. 234–238 (2009)
5. Jiang, H., Zhao, Y.Y., Dong, X.J.: Mining both Positive and Negative Weighted Association Rules with Multiple Minimum Supports. In: Proceeding of 2008 International Conference on Computer Science and Software Engineering, China, vol. 24(22), pp. 407–410 (2008)
6. Zhao, Y.Y., Jiang, H.: Mining Weighted Negative Association Rules Based on Correlation from Infrequent Items. In: Proceeding of 2009 International Conference on Advanced Computer Control, Singapore, vol. 15(5), pp. 270–273 (2009)
7. Jiang, H., Luan, X.M., Liu, G.L., Dong, X.J.: Mining Negative Weighted Association Rules from Infrequent Itemsets Based on Multiple Supports. American Journal of Engineering and Technology Research 11(9), 3848–3854 (2011)
8. Jiang, H., Dong, W.J.: Algorithms for Pruning Weighted Negative Association Rules. In: Proceeding of 2011 4th IEEE International Conference on Computer Science and Information Technology, China, vol. 53(9), pp. 336–340 (2011)
9. Agrawal, R., Srikant, R.: Fast algorithms for mining association rules. In: Proceedings of the 20th Int'l Conference on VLDB, vol. 24(6), pp. 487–499 (1994)
10. Brin, S., Montwani, R., Silverstein, C.: Beyond Market Baskets: Generalizing Association Rules to Correlations. In: Proceedings for the ACM SIGMOD Conference, vol. 24(10), pp. 265–276 (1997)
11. Feng, Y.-C., Feng, J.-L.: Incremental Updating Algorithm for Mining Association Rules. Journal of Software 9(4), 301–306 (1998)
12. Agrawal, R., Srikant, R.: Fast Algorithms for Mining Association Rules in Large Databases. In: Proceedings of the 20th International Conference on Very Large Data Bases, September 12-15, pp. 487–499 (1994)
13. Berzal, F., Cubero, J.-C., Marín, N., Serrano, J.-M.: TBAR: An efficient method for association rule mining in relational databases. Data & Knowledge Engineering 37(1), 47–64 (2001)
14. Brin, S., Motwani, R., Ullman, J.D., Tsur, S.: Dynamic itemset counting and implication rules for market basket data. In: Proceedings of the 1997 ACM SIGMOD International Conference on Management of Data, Tucson, Arizona, United States, May 11-15, pp. 255–264 (1997)
15. Calders, T., Goethals, B.: Mining all non-derivable frequent itemsets. In: Elomaa, T., Mannila, H., Toivonen, H. (eds.) PKDD 2002. LNCS (LNAI), vol. 2431, pp. 74–85. Springer, Heidelberg (2002)
16. Han, J., Fu, Y.: Discovery of Multiple-Level Association Rules from Large Databases. In: Proceedings of the 21st International Conference on Very Large Data Bases, September 11-15, pp. 420–431 (1995)

17. Liu, B., Hsu, W., Ma, Y.: Mining association rules with multiple minimum supports. In: Proceedings of the fifth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Diego, California, United States, August 15-18, pp. 337–341 (1999)
18. Pasquier, N., Bastide, Y., Taouil, R., Lakhal, L.: Efficient mining of association rules using closed itemset lattices. Information Systems 24(1), 25–46 (1999)
19. Ramkumar, G.D., Ranka, S., Tsur, S.: Weighted Association Rules: Model and Algorithm. In: KDD 1998 (1998)
20. Tao, F.: Mining Binary Relationships from Transaction Data in Weighted Settings. PhD Thesis, School of Computer Science, Queen's University Belfast, UK (2003)
21. Wang, W., Yang, J., Yu, P.S.: Efficient mining of weighted association rules (WAR). In: Proceedings of the Sixth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Boston, Massachusetts, United States, August 20-23, pp. 270–274 (2000)

Research of Intelligent Evacuation System

Weijie He[1], Zhouxiong Wu[2], and Qingrui Liu[3]

[1] Institute of Electrical information, Hebei University of Science and Technology, Shijiazhuang, 050018
heweijie@guigu.org
[2] Shenyang Fire Research Institute, Shenyang, 110034
[3] Institute of Electrical Information, Hebei University of Science and Technology, Shijiazhuang, 050018

Abstract. Design a system that combines the professional intelligent evacuation system and GIS (Geographic Information Systems), which can dynamically guide people to leave the fire site in the shortest time through a safe route according to instantaneity situation.

Keywords: Intelligent evacuation dynamically guide GIS RS485.

1 Research Background

With rapid development of social economy, there are more and more constructions of high-rise buildings, large scale underground mall, metro, and highway tunnel in recent years. These places usually have a capacity of hundreds of or even thousands of people, once a fire breakouts, due to people's fear, hurry and confusion, as well as the complicated situation of passage way in the building and overmuch floors, serious consequence of heavy casualties is likely to take place and increase the difficulty of personnel evacuation [1-3].

Currently, evacuation instructions widely installed in the building are light emission type of low luminance. These instructions are used to indicate evacuation channel. However, the indicating directions of these instructions are fixed, once the road ahead is blocked by the fire, the instructions serve nothing to guide people towards safety place but misleading them to dangerous situation and lost the chances to escape [4-7]. At the same time, evacuation instructions of this kind are usually constant light while operating, the visible distance is apparently shorten when smoke and smoke layer are top-down spreading. In addition, people in emergency status may lost normal judgment and comprehensive ability due to over stressed and panic. Therefore, indicating affects of traditional evacuation instructions system is unsatisfactory in smoke condition, people may neglect or even do not pay attention to them, it happens occasionally that people died only a few meters away from emergency exit.

2 Design Philosophy

Considering the significant deficiencies of traditional evacuation instructions system, the brand new conception of "intelligent evacuation" is put forwarded. Core ideology

Y. Yang, M. Ma, and B. Liu (Eds.): ICICA 2013, Part II, CCIS 392, pp. 223–232, 2013.

of intelligent evacuation is according to exact fire scene, by introducing voice exit guider for exit of high location, evacuative light for exit of lower location and two-way, adjustable, continuous indicating luminous flux at both floor and wall in order to actively, quickly, and accurately lead people to emergency exit far away from the fire. In this case, "evacuate through the nearest fire exit" has changed to the safety evacuate manner of "stay away from fire as precondition and evacuate through the nearest fire as principle", which has dramatically reduced evacuation duration, avoid blindly escape.

3 System Configuration

This system is formed by fire detector, controller of intelligent evacuation (host), intelligent fire emergency lighting luminature and light with intelligent fire emergency indicating luminaire, and other communication (controlling) equipments. Make great modifications to traditional controller of fire emergency luminare based on the conception of "intelligent", by adopting advanced computer technique. It has realized many functions such as windows operating, building plane figure edit and display, and non-fire emergency evacuating. Schematic diagram of system configuration is shown as follows:

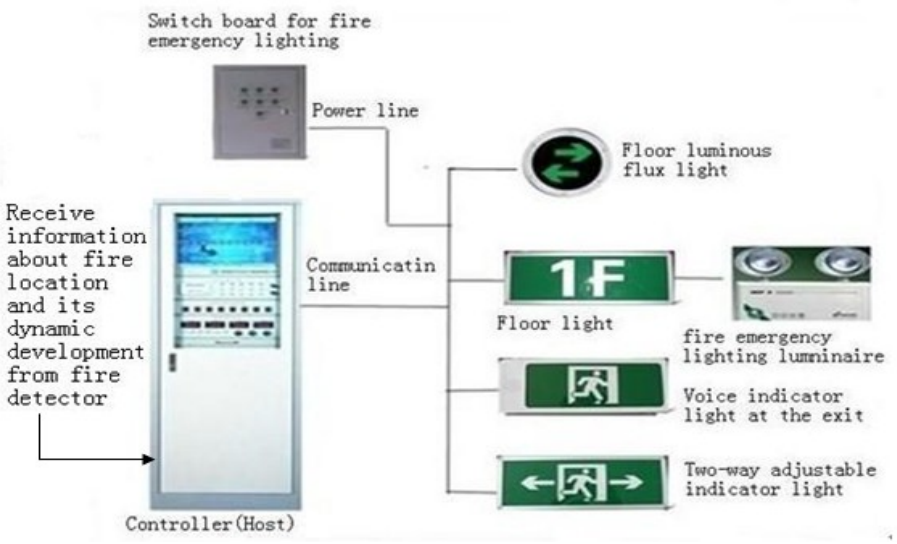


Fig. 1. System Stucture

When a fire breakout, intelligent evacuation controller (host) receive message from fire detector, Professional intelligent evacuation GIS system immediately generates the best evacuation channel, and activate fire emergency instruction light along the channel at the same time. These emergency lights make up luminous flux and guide people to the safety exit by flashing instructions Emergency lights. People are able to see the luminous flux clearly, and evacuate safely according to the guider.

4 Software Design of Intelligent Evacuation Controller (Host)

Intelligent evacuation controller adopt GIS as platform to realize the visible function of building plane figure and best evacuation route by combining professional intelligent evacuation system, which able to guide people evacuate from the fire site safely, accurately, and quickly according to the fire condition.

Geographic Information Systems, GIS, also known as geographical information science and information system, is a kind of spatial system supported by computer hardware and software technology, which use theories and methods of system engineering and information science to comprehensively, and dynamically acquire, store, transfer, manager, analysis and utilize geographic information. The feature of GIS is that it can visualize show all the useful information in form of graphics, or even human-machine interactive mode, so that people can use them conveniently. One important sign which makes GIS different from other information system is that GIS can acquire, store, analysis and display spatial data. With the development of GIS technology, the essential job of GIS is to resolve decision of geographic space by using application model.

However, traditional GIS focus more on data collection, storage, analysis and display, their ability of spatial knowledge discovering, spatial decision and geographical model imitation are relatively weak, they are inadequate in figure out complicated geographic information problem [8-12].

When generating the best evacuation route, most existing intelligent evacuation systems do not take dynamic development trend of the fire, people distribution in the building, and the status of evacuation exit of the building into account [13-16]. Considering so many problems bring in by dynamic changes in reality evacuation, it is difficult for traditional GIS to make accurate spatial decision. Therefore, we can introduce professional intelligent evacuation system, and combine it with GIS, use it to improve inferential analysis and intelligent decision making function. At the same time, the knowledge required in professional intelligent evacuation system is implicated in GIS data base [5].

Professional intelligent evacuation system contains a lot of evacuation knowledge from many fire prevention specialists. By taking their year' of experience and knowledge as reference, the computer software will automatically inference the accident, so as to imitate the process how the expert resolve the problem [6].

Professional intelligent evacuation system is combined with GIS system, the general systemic construction are shown as figure 2:

Storage space decision of professional evacuation system knowledge base, which means disaster reduction, fire prevention, specialized knowledge and experience for evacuation, those aspects are needed in the seeking process of intelligent evacuation approach. Inference engine of professional evacuation system can use GIS spatial decision and graphic display and Knowledge base of professional evacuation system, through imitate professional evacuation inference engine by using computer program, according to fire location and predict spread trend of the fire, take comprehensive consideration of evacuation channel and the positions of emergency light at the exit, and get the best evacuation plan.

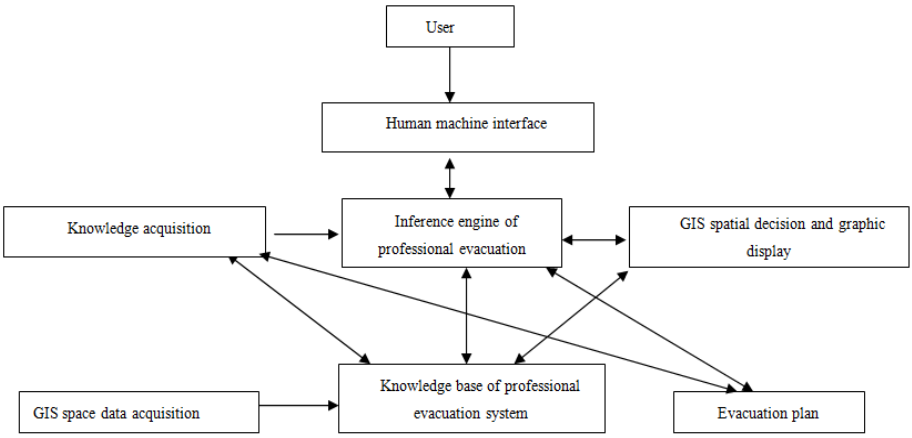


Fig. 2. Upper Computer System Software Structure

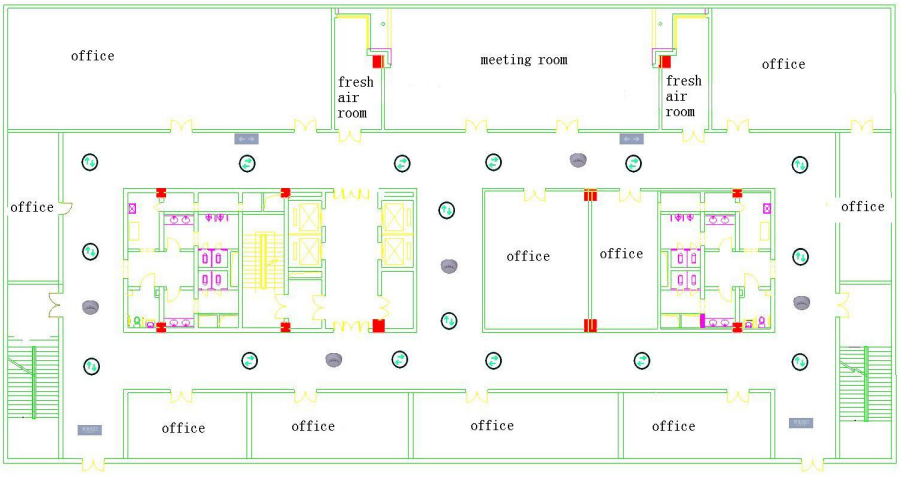


Fig. 3. The Route For evacuating To The Nearest Exits (In General Cases)

Besides, Fire alarm system and various kinds of fire linkage equipments should be installed in modern buildings, like fireproof rolling gate, smoke vent and air supply device, automatic fire extinguishing control and fire emergency broadcast. Once the fire occurs, all fire linkage equipments should realize the linkage operation.

The present system firstly obtains the spatial data and the linkage information of the fire compartment and fire alarm controllers through GIS and then the intelligent evacuation expert system implements the special decision and seeks the optimum evacuation route. The types of the spatial decision and evacuation plans implemented by the system are generally divided into the following kinds: 1) evacuation to the nearest exit when there is no fire disaster; 2) safe evacuation in fire disaster; 3) evacuation according to fixed routes

(like indicating the spectators in theaters or stadiums to the entrances or exits in order) in emergency (or special) events

4.1 Evacuation to the Nearest Exit When There Is No Fire Disaster

When there is no disaster, the system implements the default evacuation plan. That is turning off the emergency lights, keeping the emergency exits bright and making the emergency marker lamps point to the nearest emergency exit so as to help people accurately find the exit of the building. It is shown in Fig. 3.

4.2 Safe Evacuation in the Event of Fire Disaster

The implementation of spatial decision and evacuation plan if the present system is applied to general buildings: when a fire disaster occurs, the intelligent evacuation expert, GIS system, will generate the optimum evacuation route promptly between the compartment on fire and its nearby fire compartment based on the position information of the fire points, the linkage information and the evacuation exits situation (including the action message of the fireproof rolling gate). Then it will rapidly make the fire emergency indicating luminaire on the optimum evacuation route form a light stream by flashing one by one along the direction of emergency exits so as to help the escaping people clearly see the light stream along which they can evacuate safely. As for the fire emergency luminaire in non-firing compartments, the system will implement emergency order of evacuating to the nearest exit. That is making the fire emergency indicating luminaire form indicating light stream according to default direction, making the voice exit luminaire flash and make voice to remind the escaping people of the emergency exits and turning on the fire emergency luminaire . What is shown in Fig.4 and Fig.5 is the evacuation situations for two different kinds of fire points.

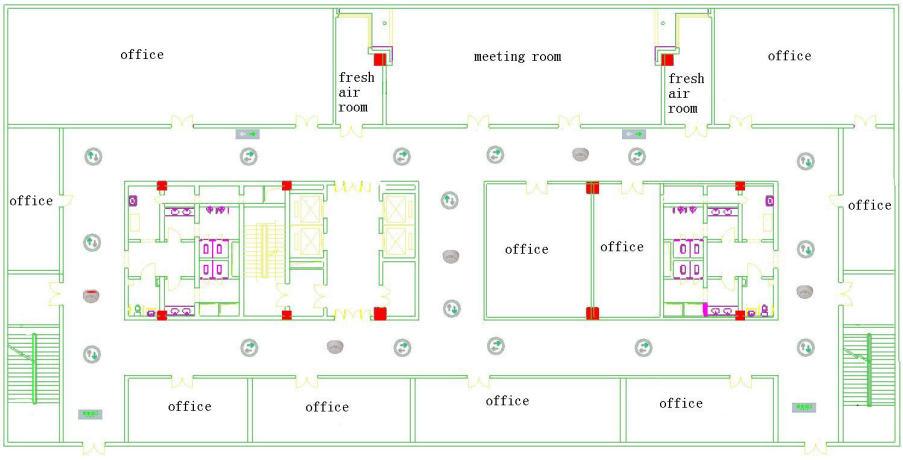


Fig. 4. Route for safe evacuation (in fire disaster)

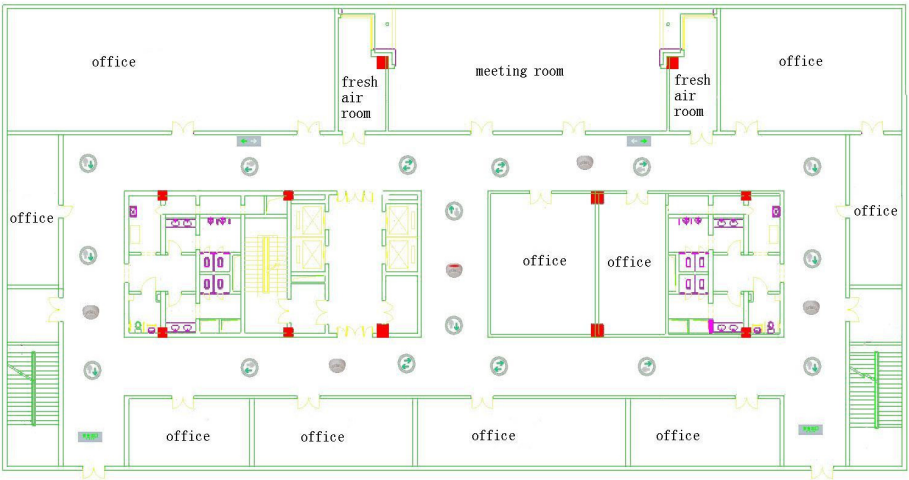


Fig. 5. Route For Safe Evacuation (In Fire Disaster)

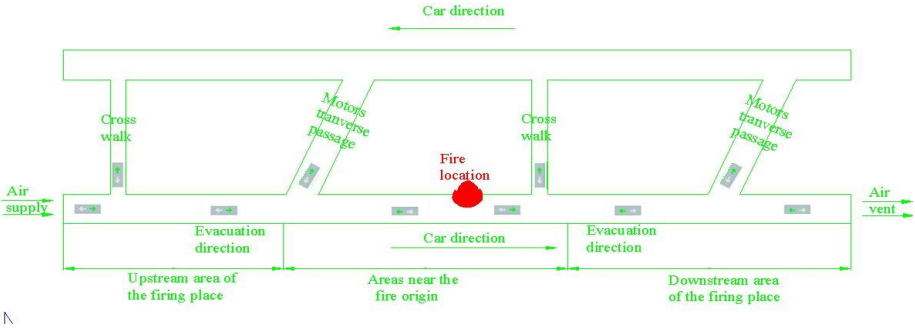


Fig. 6. The evacuation map in the event of a fire disaster in a tunnel

According to domestic and foreign statistics, most injury or death are caused by overwhelming smoke and fume, toxication and suffocation when fire disasters occur in subway tunnels or other highway tunnels. [7] Therefore, when the system is applied to subway tunnels or other highway tunnels, we should consider their special characteristics and implement the evacuation way of “giving primary consideration to staying away from the fire and obeying the principle of windward evacuation and evacuation to the nearest exit”. The intelligent evacuation controller should combine the direction of smoke vent and air supply of the smoke exhaust system and the position information of the evacuation facilities (including the contact platform and lateral evacuation platform of the subway tunnels and the pedestrian and traffic transverse tubes of the highway tunnels) in implementing spatial decision and evacuation plan. In the paper, one part of a long common highway tunnel in China is taken as an example to simply explain the evacuation situation in a fire disaster. It is shown in Fig.6:

The tunnel is an up-down separation tunnel with twin tubes between which pedestrian and traffic transverse tubes are built for evacuation and refuge in the event of fire disaster or other emergency cases. [10] Once fire disaster occurs in the tunnel, the automatic fire alarm system (FAS) will coordinately control smoke protection and smoke vent systems to supply air and vent smoke. The direction of air supply and vent of this part is shown in the figure. As for the upstream area of the firing place: evacuation to the nearest exit should be adopted because the wind blows toward the upstream direction. The vehicles in the upstream area should continue to drive along the original direction and enter into safe areas from the nearest pedestrian and traffic transverse tubes; as for the areas near the fire origin: the vehicles should keep away from the fire origin and enter into safe areas from the nearest pedestrian and traffic transverse tubes; as for the downstream areas: the air moves toward the downstream areas and the smog and smoke strengthen along with the fire, so the smoke and smog will concentrate heavily in the downstream air inlet. Thus, the vehicles should adopt windward evacuation and enter into safe areas from the nearest pedestrian and traffic transverse tubes. The intelligent evacuation controller makes all fire emergency lighting luminaire on the optimum evacuation route flash one by one to guide people and turns on all fire emergency lumiaire in the transverse tubes. It also informs us the situation about the fire disaster and the evacuation situation through fire emergency broadcast.

4.3 Evacuation According to Fixed Routes in Emergency (or Special) Events

Besides helping spectators evacuate safely in a fire disaster, the system also guides them to the entrances and exits in order according to fixed routes when it is applied to stadiums and theaters, etc. The stadium is a place of heavy crowd. Especially during international or national key sports events, there may be thousands of people. In terms of the safety problem of the stadium, we should not only consider the evacuation in emergency cases, but also consider how to maintain the order when people enter into and go out the stadium. What is shown in Fig.7 is the plan of the basketball hall of one stadium. The intelligent evacuation controller makes a plan in advance and then implements it when people start entering into the hall. Making all fire emergency indicating luminaire along the entry route flash one by one to form a light stream, the controller helps people clearly see the light stream along which they enter into the hall in order. The major entry route for the spectators is shown in Fig. 7.

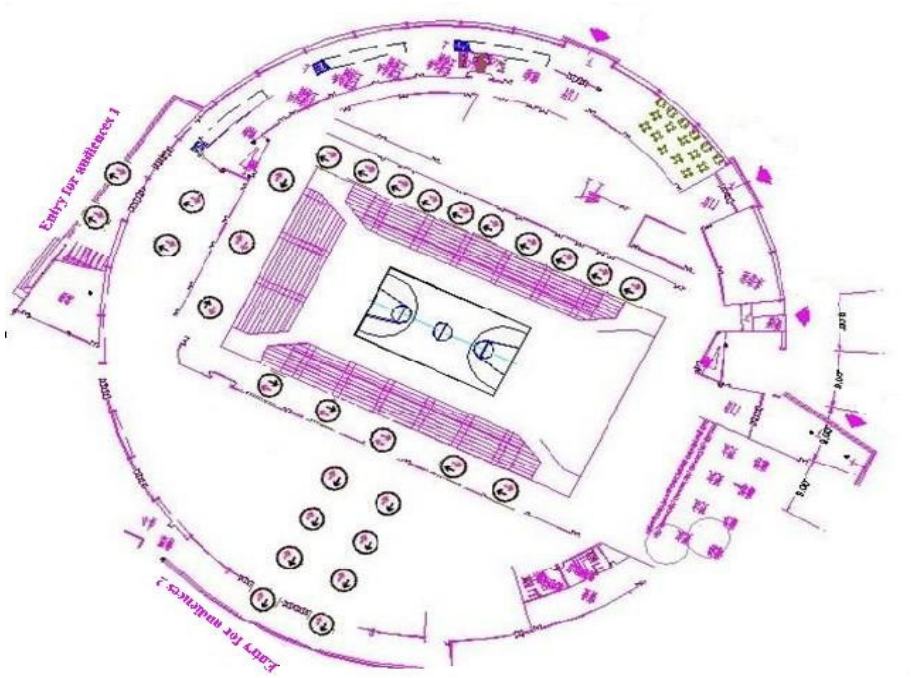


Fig. 7. Help People Enter Into The Stadium According To Fixed Route

5 The Design of Intelligent Fire Emergency Luminaire (Lower Computer)

The intelligent fire emergency luminaire consists of emergency exit (voice) sign, two way sign, geographical indications, floor number sign and fire emergency lighting luminaie. The fire emergency luminaire realizes its intellectualization by taking AVR microcontroller as its core and contacts with the intelligent evacuation controller (host) by adopting RS485 bus communication. As a kind of electric code of multi-node and differential data transmission, RS-485 bus has already been one of the most widely used standard communication interfaces. Its noise abatement capability, data transmission rate, cable length and its reliability are beyond of comparison. It is adopted as the data transmission link in many different fields.

Emergency exit (voice) sign—controlled by the host and equipped with voice, stroboscopic and lighting-off function. It is mainly set in the emergency exits.

Two way sign—controlled by the host and equipped with stroboscopic and direction indicating control function.

Geographical indication—controlled by the host and equipped with stroboscopic and direction indicating control function. It works under safe voltage.

Light stream sign—numerous two way signs or geographical indications form a group of luminaire by the way of stroboflash, shaping a indicating light stream which flows forward steadily at a certain frequency during evacuation.

Fire emergency lighting luminaire—providing illumination for evacuation and fire operation when a fire disaster occurs.

The intelligent fire emergency luminaire differs from common independent fire emergency luminaire with another characteristic: it possesses self-fault detection function. The intelligent evacuation controller keeps monitoring every lamp's state of the whole system. When there is a short and open circuit in the light path of the lamps, an under voltage problem of the battery, a short and open circuit in the charge return circuit and any faults of the links between the controller and the lamps, the intelligent evacuation controller will give an alarm by light or voice and indicate the position of the fault lamp on the controller to inform the maintainers to maintain. Thus, a large number of manpower and material resources can be saved and the delay of evacuation due to late detection of the problems of common fire emergency luminaries' be eliminated.

6 Conclusions

The present system adheres to the concept of "intelligent evacuation" and takes "evacuation expert system and GIS" as its basis. Replacing the previous "evacuation to the nearest exits or windward evacuation" with the evacuation way of "giving primary consideration to staying away from the fire and obeying the principle of windward evacuation and evacuation to the nearest exit", the system reduces the evacuation time for people and avoids blind escape, gaining precious time and safe escape for people when they are escaping in emergency events.

The system applies for: theaters, subway tunnels and other highway tunnels, shopping malls, hospitals, office buildings, hotel, airports and other large-scale modern buildings.

References

1. Chen, N.: Design of Monitoring System for Fire Alarm in Intelligent Buildings, pp. 5–6. Qinghua University Press (2001)
2. Zhao, Y.: The Current Status and Developing Trend of Domestic Fire Emergency Lighting System. Fire Technique and Products Information (11) (2004)
3. Liang, Q.: The Applications of Artificial Immune Systems and GIS Spatial Analysis. Wuhan University Press (January 2011)
4. Wang, D.: A Study on Intelligent Emergency Evacuation Indicating System. Shenyang Institute of Aeronautical Engineering, Master Thesis (February 2009)
5. He, G., Cai, H.: Expert System and Its Application in Fire Fighting Rescue (June 2010)
6. Li, X., Liu, X., Li, S.: Intelligent GIS and Spatial Optimization. Science Press (February 2010)
7. GB 50157-2003 Code for Metro Design. China Planning Press, Beijing (2003)
8. Wang, J.: Discussion of Subway Fire Fighting and Emergency Rescue in China. Fire Safety Science (January 2011)
9. GB 17945-2011 Fire Emergency Lighting and Evacuation Indicating System. China Standards Publishing House, Beijing (2011)

10. JTG/TD71-2004 Code for Traffic Design of Highway Tunnel. China Communications Press, Beijing (2004)
11. Di Febbraro, A., Giglio, D.: On Adopting a Petri Net-Based Switching Modeling System to Represent and Control Urban Areas. In: Proc. 8th IEEE Intl. Conf. Intelligent Transportation Systems (ITSC 2005), pp. 185–190 (2005)
12. Bertini, R.L.: The German Autobahn: An ITS Test Bed for Examining Dynamic Traffic Flow Phenomena. In: Proc. 8th IEEE Intl. Conf. Intelligent Transportation Systems (ITSC 2005), pp. 502–507 (2005)
13. Tango, F.: Towards a New Approach in Supporting Drivers Function: Specifications of the SASPENCE System. In: Proc. 8th IEEE Intl. Conf. Intelligent Transportation Systems (ITSC 2005), pp. 614–620 (2005)
14. Fuerstenberg, K.C.: A New European Approach for Intersection Safety - The EC-Project INTERSAFE. In: Proc. 8th IEEE Intl. Conf. Intelligent Transportation Systems (ITSC 2005), pp. 343–347 (2005)
15. Spielhofer, R.: Roncalli: Position Dependent In-Vehicle Real-Time Information for Road Users - An Overview. In: Proc. 8th IEEE Intl. Conf. Intelligent Transportation Systems (ITSC 2005), pp. 1012–1016 (2005)
16. Rabel, M., Schmeiser, A., Grossmann, H.P.: Ad-Hoc In-Car Networking Concept. In: Proc. 8th IEEE Intl. Conf. Intelligent Transportation Systems (ITSC 2005), pp. 143–148 (2005)

System Model of University Finance Based on Quantitative Decision-Making Process and Chain Interaction

Taohua Zeng and Feng Zhu

Financial Section, Southeast University, 210096, Nanjing, China
bulsan@seu.edu.cn
School of Continuing Education, Southeast University, 210096
Nanjing, China
pigazf@126.com

Abstract. This paper analyzes the features of university financial management under the new situation of economy age, on the basis of the new situation, new problems have been introduced, also put forward new measures to solve the problems of university financial management. These measures can meet the basic requirements of knowledge economy age on the financial management. In twenty-first Century, along with the economic globalization and the knowledge-based economy and the rapid development of electronic commerce, financial management theory and method has been greatly changed. Therefore, the practical significance of research on university financial management in the knowledge economy age lies in perfecting the financial management system and guarding against financial risks.

Keywords: Expert system, Quantitative decision-making process, Chain interaction, Comprehensive utility value, Frame structure.

1 Introduction

With the arrival of knowledge economy times, economy appeared hitherto unknown development, education globalization and continuously update of informatization means make financial system management become more and more complex. With the development of education industrialization, the content of university financial management expands increasingly, which has brought new challenges and problems [1-3]. In order to strengthen the development of internationalization and socialization, financial management must carry on the reform and innovation to deal with the new challenge and trend.

2 The New Features of Financial Management in Universities in Knowledge Economy Age

Knowledge economy also known as intelligent economy, which refers to the economy build based on the production of knowledge and information, allocate and use the

Y. Yang, M. Ma, and B. Liu (Eds.): ICICA 2013, Part II, CCIS 392, pp. 233–241, 2013.

economy [4-7]. Its a relative concept with agricultural economy, industry economy, science and technology, is a key of science and technology. Therefore, as a core, knowledge is the most important part of economic formation. In knowledge economy age, as the key resource support, high-tech is the pillar of intellectual economy. Injected with indispensable energy and power, this is different from any other early economic formation. The production speed of knowledge economy age depends on the rate of high-tech development, including multimedia, network and computer [8-10]. The application of these tools will greatly enhance the speed of knowledge production. Universities are representatives in the high tech domain, we should use these tools to improve the financial management system, so that universities can blend in the knowledge economy age, stand in the forefront of the development of the age. Accordingly, accelerate the development and help to prevent financial risks in universities, the deepening of financial management system reform has a profound meaning.

The financial management in universities appeared many new characteristics under the background of knowledge economy age [11-13]. The first is the impact of economic globalization on university financial management.With the development of international exchanges and cooperation in all areas, there are more and more financial investment and trade have been put into the universities, education will gradually towards internationalization and marketization. University financial management must make the corresponding reform of globalization, seize market opportunities, and improve the efficiency of capital operation and benefits. Second, with the integration development of China's economic, universities facing how to improve the internationalization level of financial management under the new situation. In the actual university financial reimbursement policies, some standard repayment and financial compensation does not adapt to new forms of economic growth, and severely suppress school development [14-17]. According to the method of comparative analysis, we need to present a valid suggestions for the differences between regions, we need to put forward a effective suggestion in order to reform the economic compensation standards which does not meet the current economic development, provide approaches to improve the current university financial reimbursement policy, to make it keeps pace with the times. At present, the use efficiency of university budget capital is low, capital mainly reflect in interim payment amount of settlement claims, often temporary emergency fund, however, financial appropriation budget payments is not able to meet the cost. Therefore, it is a prerequisite for enhancing the university settlement and the overall reform of regulations or policies, make full use of efficient capital budget. Since 1920s, university financial management mode changed from a highly centralized planned economy to a system integrated with the concentration and classification model of market economy. With the development of China 's market economy and the reform of university education system, the education funding sources gradually developed from simple financial allocation to the diverse ways. Balance of payments also

changed from simple units of government input and expenditure to more complex business accounting. Economic calculation is also changed greatly, the university financial activities presents new characteristic and diversification of financing method. In addition to the the basic requirements of economic activities, the development of teaching and technology, the independent accounting enterprises and commercial capital appear more acerb conflict. Under the dependence on the knowledge, speed of information dissemination such a kind of economic formation, the application of multimedia technology is the most important means of financial management [18-20] .The University financial management system should gather important information and resources; the most important is to use high-tech means and related software to deal with the financial information and data. So the universities are facing new condition about how to strengthen financial management and prevention of financial risk. University financial department should devote sufficient financial resources, material resources, to strengthen the scientific and professional management, to adapt to the change of times.

3 The Main Defects of Financial Management in Universities

3.1 The Defects of University Financial Capital Sources

At present, the main capital source in universities is the national financial investment. In addition, fee income, research fund, school-run industries and other investment of social all circles and donations are also important sources of funds. However, the national finance investment is still the main source. With the development of socialist market economy and the higher education needs, we are in urgent need of funds from a plurality of sources. Therefore, it is necessary to reform the current university financial management goal to make the sources of funds to adiversified development.

3.2 The Defects of Financial Management System in Universities

At present, a prominent character of university is "no master". Universities often become the net assets; asset does not require any reimbursement or any economic benefits. So cost audit does not participate in the gains and losses of university financial management, the distribution of surplus without calculating make the financial management quite flexible, the financial management of universities is easy to abuse authority to get profit. The basic theory of financial management is very important for enterprise to get benefits by the coordination of financial activities in different disciplines. At the same time, the capital investment of country doesn't have any financial supervision, resulting in a lack of effective resource allocation and management.

3.3 The Lack of Scientific Economy Decision-Making

The knowledge economy promotes management departments to realize the resource allocation efficiency, rationalization. For the distribution of capital, the financial management in universities are usually not proceed from the university and the development strategy of allocating capital as a whole, not even for the purpose of plan and phase distribution. In addition, the use of capital structure can not achieve the best state; all kinds of information economic decisions always from a subjective perspective, people always ignore the importance of decision, ignore the economic benefits of the use of capital, resulting in the waste of capital to some extent. Financial management should focus on solving capital allocation more scientific and flexible [4]. The economic decision-making is an intentional program in the overall consideration of higher subjective department, which belongs to the whole macro management and development plan. As a particular university, the applying for a important project must be based on the project feasibility study of decision-making process. Review the above problems; we must establish a systematic model and a sound financial management system in universities, to realize the financial management efficiently, prevent financial risks.

4 The Construction of University Financial Management System Model and Countermeasures to Solve the Financial Management Problems

4.1 Integral Model Method

The "3 bases and 3 levels" of the subject structure of Chinese management sciences and the main development point in"3 areas" [5] is presented by Mr. Cheng siwei. In order to further the development of university financial management, we can divide financial management in universities into 3 levels. The first level is regarding the financial management of universities as a whole system; second, regard each project system of financial management in universities as the second level. For instance, the capital budget, capital operation, the university financial performance evaluation are called the topic, also the main topic of the financial management in universities. Third, regard net cash flow, capital cost measurement, net present value decision and the bottom subsystem of each topic as the third level.

4.1.1 The Theoretical Framework of Overall Model

The presence of correlated characteristics of branch system in university financial management, in which the overall model contains 4 major key elements Respectively are mathematical calculation method, mathematical model, management information collection, scientific knowledge and experience of experts, also contains intelligent inference and decision-making quantification in management process. n subsystems

composed the complex university financial management system. If S refers to the special system, the *st* (*t*=1,., *n*) represent the *T* subsystem, you can use the *N* =*S*{*s*1, *s*2, *si*}represent the university financial management system. The model is composed of four elements, respectively, the mathematical models of *em*, relevant information *ei*, calculation method *ec*, also contains the expert knowledge and experience *ek*[6]. The whole model structure of university financial management system can be represented as follows:

$$F(N)=f(ei,em,ec,ek) \tag{1}$$

From the analysis on the internal subsystem, the t subsystem is composed of four parts, respectively, relevant information *eit*, mathematical model *emt*, calculation method *ec*, expert knowledge and experience *ek* 4 elements(*t*=1,⋯,*n*), therefore:

$$F(nt)=ft(eit,emt,ect,ekt),\ t = 1, \cdots, n \tag{2}$$

From the existed chain interaction between n subsystem, so:

$$F(N)=f(s1,s2,\cdots,sn) \tag{3}$$

From above we can see that, there are 2 main levels of relationship between composition of university financial management system and the overall framework of the model. First, regard the four elements as the first layer, respectively, the inside information of the system, mathematical model, calculation method, expert knowledge and experience. This layer mainly embodies the two processes, respectively, reflection of quantitative decision-making process, the intelligent reasoning process. Second, regard the relationship between systems as a second layer, through this hierarchy can reflect matching degree of the system coordination, as well as the effect of this level to the whole subject system.

4.1.2 The Model Framework of University Financial Performance Evaluation System

First of all, before the analysis of the thematic system structure, we need do a lot of research on university financial performance evaluation project; secondly, to find out the four subsystems involved in the topic system, mainly the exchange decision-making system, evaluation index system, weight coefficient decision subsystem and the evaluation criteria system. Finally, to analysis their inner relationship [7]:

① Input decision subsystem information;

②According to the first step input evaluation index system;

③According to the above relevant information input power factor decision-making subsystem;

④From the evaluation index we get in the second step and weight coefficient input evaluation standard system from step three;

⑤The evaluation result output evaluation standard system.

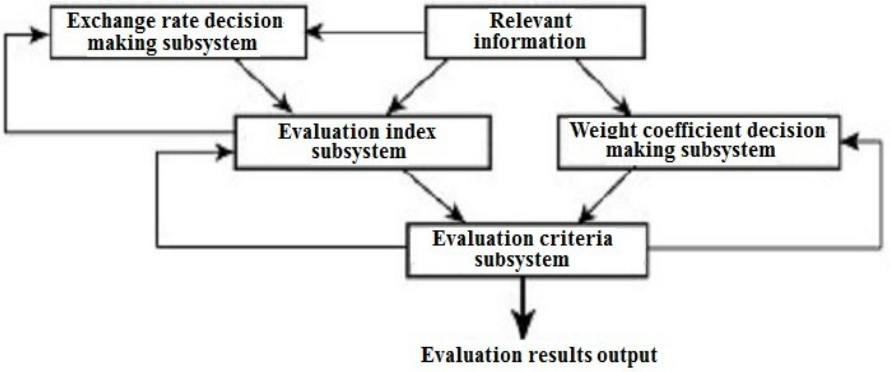


Fig. 1. The university financial performance evaluation system model framework

4.1.3 The System Model of University Financial Performance Evaluation Subsystem

In order to elaborate university financial performance evaluation system model, we take the evaluation criteria system model and the weight coefficient of decision-making system model as an example.

In order to evaluate the university financial performance of each part, first, we must carry on the evaluation index system of subsystem, which mainly according to the exchange rate decision subsystem to select the appropriate exchange rate portfolio. But many divisions are different in the two index value, including 13 financial indexes and 11 non-financial indexes [8-10]. Through these indicators is very difficult to directly evaluate the division's financial performance. Therefore, it is necessary to construct a unified nondimensional utility values as the standard comparative financial performance.

The evaluation standard of comprehensive utility value in t subsector can be represent by $Ut = \sum_{i=1}^{n} \alpha i Ui(oi)$. The index weight of i can be represent by αi; the utility values of index i can be represent by $Ui(oi)$; the comprehensive utility values of n are indicated by Ut. Figure 2 is the quantitative decision process of evaluation standard subsystem.

①The utility function of index i can be performed as $Ui(ox)$ by university financial management experts according to the evaluation index background and meaning .

②In order to solve the problem of dimension unification. We will have the first i indicators of the actual data Oi into the index of the utility function $Ui(ox)$, it can be calculated that the index dimensionless practical utility value $Ui(oi)$.

③The comprehensive utility value calculation of $Ut = \sum_{i=1}^{n} \alpha i Ui(oi)$ in t (t=1,., n) we need to know αi and $Ui(oi)$, each index corresponding weight coefficient αi (i=1,., 24) can be determined according to the weight coefficient of decision-making subsystem, the actual utility values $Ui(oi)$ (i= 1,., 24) have been calculated from the above.

④In order to evaluate the university financial performance, we will first sort the n division of the comprehensive utility value Ut (t=1,., n). Then according to the principle that the greater the comprehensive utility value, the better financial performance evaluation [11-14].

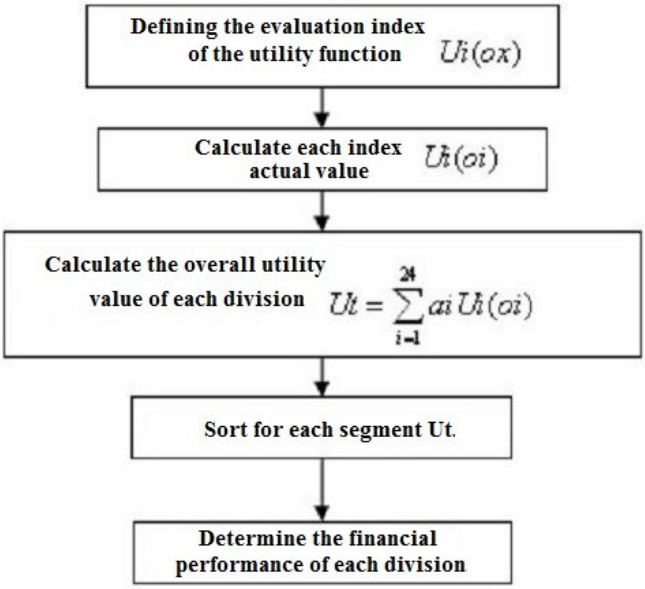


Fig. 2. The flow diagram of evaluation standard subsystem

4.2 The Make the Logistics Operation Based on the Rule of Market Economy

Universities can learn foreign universities' logistics mode. For example, in a French university logistics service has a high level of socialization, the student canteen, dormitory, library, and even some recreation and sports center or center of obtain employment guidance are not part of the university. Only in this way can we realize fair competition, to improve the quality and efficiency of service. So universities can better concentrate on teaching and research, logistics according to market economy rule execution services, they can provide mutual aid and perform their task to promote the long-term development of university.

4.3 Control Liabilities Scale

State funding is the main capital source for universities, the majority of universities used to free use of capital is not repay. Some universities have a lack of bank loans, as well as the lack of capital cost and financial risk. So the university must learn to use the bank loans risk consciousness, reasonably determine the loan scale and time. Higher income minus the cost to reach the current annual cash surplus amount.

Considering the factors of cash flow, 50% cash surplus of the annual funding, which is the maximum credit amount? Load size determines the university development needs, will be the largest average annual credit equal to the period of the loan. According to the principle of prudent requirements, university should be estimating costs even more income to ensure the security of university financial.

4.4 The Use of High-Tech Means to Reform Financial Management Level in Universities

With the raising the level of computer technology and financial information management, financial management software has become a more important accounting auditing and management tools. Good financial software can greatly improve the working conditions and efficiency; make financial personnel optimize resource management, so as to realize the standardized and scientific management, which has an important sense to improve the university financial management.

5 Conclusion

Based on knowledge economy age, this paper introduces the characteristic and deficiencies of financial management under the new situation, accordingly proposed the corresponding countermeasure, and established the whole model of financial management system, which laid a solid foundation for compiling computer program for further development and the development of university finance management of the intelligent system project. So as to achieve the electronic network, improve the work efficiency of the financial management in universities.Under the new situation of of knowledge economy age, theapplication of multimedia, network and computer greatly speed the knowledge production. Universities, as the representative of the national science and technology development, should use these tools to improve the financial management system, make universities blend in the knowledge economy age, to avoid financial risks, and walk in the forefront of the times.

References

1. Lui, B.: Study of Innovation for University Financial Management. Liao Ning Economic Management Cadre University (3), 32 (2011)
2. Chen, B.: University Innovative Fmancial Management Philosophy and System. Financial Accounting Communcation (4), 65 (2010)
3. Lui, W.: Further Strengthen Ways of University Financial Management in Our Country. Financial Accounting Study (7) 47 (2007)
4. Cheng, S.: The Present Situation and Prospect of Management Science. Journal of Management Sciences in China 1(1), 8–14 (1998)
5. Cheng, S.: China Management Science Subject Structure and Development Key selection. Journal of Management Sciences in China (1), 1–6 (2010)

6. Xu, Y.: Discussion on Intelligent Accounting Expert Decision System. Accounting Research (2), 33–36 (2008)
7. Ramshidi, J.: System Modeling and Control, vol. (5), p. 61. Science Press, Beijing (2008)
8. Bo, H.: Research on Multi-National Corporation Financial Performance Evaluation of Intelligent System Modeling. Journal of Shanghai Second Polytechnic University 17(2), 12–19 (2010)
9. Chen, Y.: Enterprise Performance Evaluation System. Foreign Economics & Management 22(4), 7–10 (2010)
10. Chu, X., Nadiminti, K., Jin, C., Venugopal, S., Buyya, R.: Aneka: Next-Generation Enterprise Grid Platform for e-Science and e-Business Applications. In: Proceedings of the 3th IEEE International Conference on e-Science and Grid Computing (e-Science 2007), Bangalore, India (December 2007)
11. Fu, Y., Chase, J., Chun, B., Schwab, S., Vahdat, A.: SHARP: an architecture for secure resource peering. ACM SIGOPS Operating Systems Review 37(5), 133–148 (2003)
12. Lai, K., Rasmusson, L., Adar, E., Zhang, L., Huberman, B.A.: Tycoon: An implementation of a distributed, market-based resource allocation system. Multiagent and Grid Systems 1(3), 169–182 (2005)
13. Au Young, A., Chun, B., Snoeren, A., Vahdat, A.: Resource allocation in federated distributed computing infrastructures. In: Proceedings of the 1st Workshop on Operating System and Architectural Support for the Ondemand IT Infrastructure (OASIS 2004), Boston, USA (October 2004)
14. Irwin, D.E., Chase, J.S., Grit, L.E., Yumerefendi, A.R., Becker, D., Yocum, K.: Sharing networked resources with brokered leases. In: Proceedings of the 2006 USENIX Annual Technical Conference (USENIX 2006), Boston, USA (June 2006)
15. Wang, S.-J., Xu, Y.-S., Yang, S., Chen, E.-Q.: Phase Behavior of a Hydrogen-Bonded Polymer with Lamella-to-Cylinder Transition: Complex of Poly(4-vinylpyridine) and Small Dendritic Benzoic Acid Derivative. Macromolecules 45(21), 8760–8769 (2012)
16. Gopinadhan, M., Beach, E.S., Anastas, P.T., Osuji, C.O.: Smectic Demixing in the Phase Behavior and Self-Assembly of a Hydrogen-Bonded Polymer with Mesogenic Side Chains. Macromolecules 43(16), 6646–6654 (2010)
17. Zhu, Q.-Y., Han, Q.-H., Shao, M.-Y., Gu, J., Shi, Z., Dai, J.: Supramolecular and Redox Chemistry of Tetrathiafulvalene Monocarboxylic Acid with Hydrogen-Bonded Pyridine and Bipyridine Molecules. The Journal of Physical Chemistry B 116(14), 4239–4247 (2012)
18. Bai, P., Kim, M.I., Xu, T.: Thermally Controlled Morphologies in a Block Copolymer Supramolecule via Nonreversible Order–Order Transitions. Macromolecules 46(14), 5531–5537 (2013)
19. Seto, H., Ogata, Y., Murakami, T., Hoshino, Y., Miura, Y.: Selective Protein Separation Using Siliceous Materials with a Trimethoxysilane-Containing Glycopolymer. ACS Applied Materials & Interfaces 4(1), 411–417 (2012)
20. Yu, A., Shang, J., Cheng, F., Paik, B.A., Kaplan, J.M., Andrade, R.B., Ratner, D.M.: Biofunctional Paper via the Covalent Modification of Cellulose. Langmuir 28(30), 11265–11273 (2012)

Analytical Approximation Solution for Differential Equations with Piecewise Constant Arguments of Alternately Advanced and Retarded Type

Qi Wang[1,*], Cui Guo[2], and Ruixiang Zeng[1]

[1] School of Applied Mathematics, Guangdong University of Technology, Guangzhou, 510006, China
qiwang@gdut.edu.cn
[2] College of Science, Harbin Engineering University, Harbin, 510006, China

Abstract. In this work, the variational iteration method is used for analytic treatment of differential equations with piecewise constant arguments of alternately advanced and retarded type. In order to prove the precision of the results, some comparisons are also made between the exact solutions and the results of the numerical method and the variational iteration method. The obtained results reveal that the method is very effective and convenient for constructing differential equations with piecewise constant arguments.

Keywords: Variational iteration method, Piecewise constant arguments, Analytical approximation solution, Lagrange multiplier.

1 Introduction

Differential equations with piecewise constant arguments (EPCA) have received extensive investigations [1-5]. In EPCA, the derivatives of the unknown functions depend on not just the time t at which they are determined, but on constant values of the unknown functions in certain intervals of the time t before t. These equations have the structure of continuous dynamical systems in intervals of unit length. Continuity of a solution at a point joining any two consecutive intervals implies a recursion relation for the values of the solution at such points. Therefore, they combine the properties of differential equations and difference equations.

EPCA has been under intensive investigation for the last twenty years. The theory of EPCA was initiated in [6, 7] and developed by many authors. The general theory and basic results for EPCA have been thoroughly investigated in the book of Wiener [8]. For more detailed information about analytical solution and numerical treatment of EPCA, the reader is referred to [9-15] and the references therein.

In this paper, we will apply the analytical approximation technique: the variational iteration method to the following EPCA:

* Corresponding author.

Y. Yang, M. Ma, and B. Liu (Eds.): ICICA 2013, Part II, CCIS 392, pp. 242–251, 2013.

$$y'(t)+ay(t)+by\left(\left[t+\frac{1}{2}\right]\right)=0,\quad y(0)=y_0, \tag{1}$$

where $a,b,y_0\in\mathbf{R}$ and $[\cdot]$ denotes the greatest integer function. Since the argument deviation $t-[t+1/2]$ is negative in $[n+1/2,n+1)$ and positive in $[n,n+1/2)$, (1) is said to be of alternately advanced and retarded type. The main purpose of this paper is to extend the variational iteration method to find the approximate solution of (1).

Many different methods have recently introduced to solve nonlinear problems, such as the homotopy analysis method (HAM) [16, 17], the variational iteration method (VIM) [18, 19], the Adomian's decomposition method (ADM) [20], and homotopy perturbation method (HPM) [21]. The VIM is strongly and simply capable of solving a large class of linear or nonlinear differential equations without linearization or small perturbation and also it reduces the size of calculations. The variational iteration method, which proposed by He [22-24], was successfully applied to autonomous ordinary and partial differential equations [25-27]. Recently, it has been used to solve effectively, easily and accurately a large class of linear and nonlinear differential equations.

2 Preliminaries

In this section, we will introduce some definitions and results which will be used later.

2.1 EPCA of Alternately Advanced and Retarded Type

Definition 1 (see [8]). A solution of (1) on $[0,\infty)$ is a function $y(t)$ that satisfies the conditions

(i) $y(t)$ is continuous on $[0,\infty)$;

(ii) The derivative $y'(t)$ exists at each point $t\in[0,\infty)$, with the possible exception of the point $t=n+1/2$, $n=0,1,\cdots$, where one-sided derivatives exist;

(iii) Eq. (1) is satisfied on $[0,1/2)$ and each interval $[n-1/2,n+1/2)$.

Theorem 1 (see [8]). If $b\neq a/(e^{a/2}-1)$, then (1) *has on* $[0,\infty)$ *a unique solution*

$$y(t)=m(T(t))\lambda^{[t+1/2]}y_0,$$

where

$$m(t)=e^{at}+\left(e^{at}-1\right)a^{-1}b, T(t)=t-\left[t+\frac{1}{2}\right], \lambda=\frac{m(1/2)}{m(-1/2)}$$

Theorem 2 (see [8]). The zero solution of (1) *is asymptotically stable* for any given y_0 if and only if

$$-\frac{a\left(e^{a}+1\right)}{\left(e^{a/2}-1\right)^{2}}<b<-a\text{, for } a>0,$$

$$b<-a \text{ or } b>-\frac{a\left(e^{a}+1\right)}{\left(e^{a/2}-1\right)^{2}}\text{, for } a<0,$$

$$b<0 \text{ for } a=0.$$

2.2 Variational Iteration Method

The VIM is the general Lagrange method, in which an extremely accurate approximation at some special point can be obtained. Next, we will present the essential steps for using the VIM and the determination of the Lagrange multipliers.

Consider the following differential equation

$$Ly+Ny=g(t), \tag{2}$$

where L and N are linear and nonlinear operator, respectively, and $g(t)$ is the inhomogeneous term. According to VIM, we can write a correction functional as:

$$y_{n+1}(t)=y_n(t)+\int_0^t \lambda(s)[Ly_n(s)+N\tilde{y}_n(s)-g(s)]ds, \tag{3}$$

where λ is a general Lagrangian multiplier which can be identified optimally via integration by parts and the variational theory, and $\tilde{y}_n$ as a restricted variation which means $\delta\tilde{y}_n=0$. Having λ determined, an iteration formula, without restricted variation, should be used for the determination of the successive approximations $y_{n+1}(t)$ of the solution $y(t)$. The zero-th approximation y_0 can be any selective function. Consequently, the solution is given by

$$y(t)=\lim_{n\to\infty} y_n(t)\cdot \tag{4}$$

3 Applications

In this section, the VIM is successfully applied for solving a linear EPCA of alternately advanced and retarded type.

For (1), the correction functional reads

$$y_{n+1}(t)=y_n(t)+\int_0^t \lambda(s)\left(y_n'(s)+ay_n(s)+b\tilde{y}_n\left(\left[s+\frac{1}{2}\right]\right)\right)ds. \tag{5}$$

Taking variational on both sides of (5), we have

$$\delta y_{n+1}(t)=\delta y_n(t)+\delta\int_0^t \lambda(s)\left(y_n'(s)+ay_n(s)\right)ds$$

$$= (1+\lambda(t))\delta y_n(t) + \int_0^t \left(a\lambda(s) - \lambda'(s)\right)\delta y_n(s)ds\,,$$

this yields the stationary conditions:

$$\begin{cases} 1+\lambda(t) = 0, \\ \lambda'(s) - a\lambda(s) = 0. \end{cases} \tag{6}$$

Thus

$$\lambda(s) = -e^{a(s-t)}\,, \tag{7}$$

so we obtain the following iteration formula

$$y_{n+1}(t) = y_n(t) - \int_0^t e^{a(s-t)}\left(y_n'(s) + ay_n(s) + by_n\left(\left[s+\frac{1}{2}\right]\right)\right)ds\,, \tag{8}$$

and the following initial approximation is chosen

$$y_0(t) = y(0) = y_0,$$

so we have

$$\begin{aligned} &y_0(t) = y_0, \\ &y_1(t) = y_0(t) - \int_0^t e^{a(s-t)}\left(y_0'(s) + ay_0(s) + by_0\left(\left[s+\frac{1}{2}\right]\right)\right)ds\,, \\ &y_2(t) = y_1(t) - \int_0^t e^{a(s-t)}\left(y_1'(s) + ay_1(s) + by_1\left(\left[s+\frac{1}{2}\right]\right)\right)ds\,, \\ &y_3(t) = y_2(t) - \int_0^t e^{a(s-t)}\left(y_2'(s) + ay_2(s) + by_2\left(\left[s+\frac{1}{2}\right]\right)\right)ds\,, \\ &\qquad\cdots\cdots\cdots\cdots \\ &y(t) = \lim_{n\to\infty} y_n(t)\,. \end{aligned} \tag{9}$$

During the process of computation, we find that the greatest integer function $[\cdot]$ brings us much trouble. To overcome it, we introduce a method: consider the above iteration formula in a series of intervals:

$$[0,1/2),[1/2,3/2),[3/2,5/2),\cdots,[(2n-1)/2,(2n+1)/2)\,,\ n=1,2,\cdots$$

Following this way, each integral in iteration formulas can be easily computed. Therefore, the following theorem is obtained.

Theorem 3. The VIM solution of (1) can be determined by (9) with the iteration (8).

Thus, the discussion is as follows.

When $t \in [0,1/2)$,

$y_{1,0}(t) = y_0,$

$$y_{1,1}(t)=y_{1,0}(t)-\int_0^t e^{a(s-t)}\left(y_{1,0}^{'}(s)+ay_{1,0}(s)+by_{1,0}\left(\left[s+\frac{1}{2}\right]\right)\right)ds=-\frac{b}{a}y_0+\frac{a+b}{a}y_0e^{-at},$$

$$y_{1,2}(t)=y_{1,1}(t)-\int_0^t e^{a(s-t)}\left(y_{1,1}^{'}(s)+ay_{1,1}(s)+by_{1,1}\left(\left[s+\frac{1}{2}\right]\right)\right)ds=-\frac{b}{a}y_0+\frac{a+b}{a}y_0e^{-at}.$$

Follow this way, we can obtain

$$y_{1,n}(t)=-\frac{b}{a}y_0+\frac{a+b}{a}y_0e^{-at},\quad n\geq 1.$$

When $t\in[1/2,3/2)$,

$$y_{2,0}(t)=y_{1,1}(\frac{1}{2})=-\frac{b}{a}y_0+\frac{a+b}{a}y_0e^{-at},$$

$$y_{2,0}(t)=y_{1,1}(\frac{1}{2})=-\frac{b}{a}y_0+\frac{a+b}{a}y_0e^{-a/2},$$

$$y_{2,1}(t)=y_{2,0}(t)-\int_0^t e^{a(s-t)}\left(y_{2,0}^{'}(s)+ay_{2,0}(s)+by_{2,0}\left(\left[s+\frac{1}{2}\right]\right)\right)ds$$

$$=-\frac{b}{a}(1-e^{-at})\left(-\frac{b}{a}y_0+\frac{a+b}{a}y_0e^{-a/2}\right),$$

$$y_{2,2}(t)=y_{2,1}(t)-\int_0^t e^{a(s-t)}\left(y_{2,1}^{'}(s)+ay_{2,1}(s)+by_{2,1}\left(\left[s+\frac{1}{2}\right]\right)\right)ds$$

$$=\left(\frac{b}{a}\right)^2(1-e^{-a})(1-e^{a/2-at})\left(-\frac{b}{a}y_0+\frac{a+b}{a}y_0e^{-a/2}\right),$$

$$y_{2,3}(t)=y_{2,2}(t)-\int_0^t e^{a(s-t)}\left(y_{2,2}^{'}(s)+ay_{2,2}(s)+by_{2,2}\left(\left[s+\frac{1}{2}\right]\right)\right)ds$$

$$=\left(\frac{b}{a}\right)^2(1-e^{-a})\left(-\frac{b}{a}y_0+\frac{a+b}{a}y_0e^{-a/2}\right)(\frac{b}{a}e^{-a/2}-\frac{b}{a}-\frac{b}{a}e^{-a/2-at}+\frac{a+b}{a}e^{-at}-e^{a/2-at}),$$

……

When $t\in[3/2,5/2)$,

$$y_{3,0}(t)=y_{2,1}(\frac{3}{2})=-\frac{b}{a}\left(-\frac{b}{a}y_0+\frac{a+b}{a}y_0e^{-a/2}\right)(1-e^{-3a/2}),$$

$$y_{3,1}(t)=y_{3,0}(t)-\int_0^t e^{a(s-t)}\left(y_{3,0}^{'}(s)+ay_{3,0}(s)+by_{3,0}\left(\left[s+\frac{1}{2}\right]\right)\right)ds$$

$$=\frac{b}{a^2}\left(-\frac{b}{a}y_0+\frac{a+b}{a}y_0e^{-a/2}\right)(1-e^{-3a/2})(b-(a+b)e^{-at}),$$

$$y_{3,2}(t)=y_{3,1}(t)-\int_0^t e^{a(s-t)}\left(y_{3,1}^{'}(s)+ay_{3,1}(s)+by_{3,1}\left(\left[s+\frac{1}{2}\right]\right)\right)ds$$

$$= \frac{b(a+b)}{a^2}\left(-\frac{b}{a}y_0 + \frac{a+b}{a}y_0 e^{-a/2}\right)(e^{-at} - e^{-3a/2-at})(\frac{b}{a}(e^{a/2} - 2e^{-a/2} + e^{at-2a}) - 1)$$

$$-\frac{b^2}{a^2}(1-e^{-3a/2})(e^{-at} + \frac{b}{a} - \frac{a+b}{a}e^{a/2-at})\left(-\frac{b}{a}y_0 + \frac{a+b}{a}y_0 e^{-a/2}\right),$$

…….

When $t \in [5/2, 7/2)$

$$y_{4,0}(t) = y_{3,1}(\frac{5}{2}) = \frac{b}{a^2}(-\frac{b}{a}y_0 + \frac{a+b}{a}y_0 e^{-a/2})(1-e^{-3a/2})(b-(a+b)e^{-5a/2}),$$

$$y_{4,1}(t) = y_{4,0}(t) - \int_0^t e^{a(s-t)}(y_{4,0}^{'}(s) + ay_{4,0} + by_{4,0}([s+\frac{1}{2}]))ds$$

$$= \frac{b}{a^2}(-\frac{b}{a}y_0 + \frac{a+b}{a}y_0 e^{-a/2})(1-e^{-3a/2})(b-(a+b)e^{-5a/2})(1-\frac{a+b}{a}(1-e^{-at})),$$

$$y_{4,2}(t) = y_{4,1}(t) - \int_0^t e^{a(s-t)}(y_{4,1}^{'}(s) + ay_{4,1}(s) + by_{4,1}([s+\frac{1}{2}]))ds$$

$$= \left\{1 - \frac{a+b+1}{a}(1-e^{-at}) + \frac{a+b}{a}(e^{-at} - e^{-2at}) + a(a+b)e^{-at} - (a+b)[1-b(1-e^{3a})]t\right\}$$

$$\frac{b}{a^2}(-\frac{b}{a}y_0 + \frac{a+b}{a}y_0 e^{-a/2})(1-e^{-3a/2})(b-(a+b)e^{-5a/2}),$$

…….

When $t \in [7/2, 9/2)$

$$y_{5,0}(t) = y_{4,1}(\frac{7}{2}) = \frac{b}{a^2}(-\frac{b}{a}y_0 + \frac{a+b}{a}y_0 e^{-a/2})(1-e^{-3a/2})(b-(a+b)e^{-5a/2})(1-\frac{a+b}{a}(1-e^{-7/2})),$$

$$y_{5,1}(t) = y_{5,0}(t) - \int_0^t e^{a(s-t)}(y_{5,0}^{'}(s) + ay_{5,0}(s) + by_{5,0}([s+\frac{1}{2}]))ds$$

$$= \frac{b}{a^2}(-\frac{b}{a}y_0 + \frac{a+b}{a}y_0 e^{-a/2})(1-e^{-3a/2})(b-(a+b)e^{-5a/2})(1-\frac{a+b}{a}(1-e^{-7a/2}))(1-\frac{a+b}{a}(1-e^{-at})),$$

$$y_{5,2}(t) = y_{5,1}(t) - \int_0^t e^{a(s-t)}(y_{5,1}^{'}(s) + ay_{5,1}(s) + by_{5,1}([s+\frac{1}{2}]))ds$$

$$= \left\{1 - (1-e^{-at})[1 + \frac{b}{a}(1 - \frac{a+b}{a}(1-e^{-4a})] + (a+b)te^{-at}\right\}\frac{b}{a^2}$$

$$(-\frac{b}{a}y_0 + \frac{a+b}{a}y_0 e^{-a/2})(1-e^{-3a/2})(b-(a+b)e^{-5a/2})(1-\frac{a+b}{a}(1-e^{-7a/2})),$$

…….

When $t \in [9/2, 11/2)$

$$y_{6,0}(t) = y_{5,1}(\frac{9}{2}) = \frac{b}{a^2}(-\frac{b}{a}y_0 + \frac{a+b}{a}y_0 e^{-a/2})(1-e^{-3a/2})(b-(a+b)e^{-5a/2})$$

$$(1-\frac{a+b}{a}(1-e^{-7a/2}))(1-\frac{a+b}{a}(1-e^{-9a/2})),$$

$$y_{6,1}(t) = y_{6,0}(t) - \int_0^t e^{a(s-t)}(y_{6,0}^{'}(s) + ay_{6,0} + by_{6,0}([s+\frac{1}{2}]))ds$$

$$= \frac{b}{a^2}(-\frac{b}{a}y_0 + \frac{a+b}{a}y_0 e^{-a/2})(1-e^{-3a/2})(b-(a+b)e^{-5a/2})(1-\frac{a+b}{a}(1-e^{-7a/2}))$$

$$(1-\frac{a+b}{a}(1-e^{-9a/2}))(1-\frac{a+b}{a}(1-e^{-at})),$$

$$y_{6,2}(t) = y_{6,1}(t) - \int_0^t e^{a(s-t)}(y_{6,1}^{'}(s) + ay_{6,1}(s) + by_{6,1}([s+\frac{1}{2}]))ds$$

$$= [1-\frac{a+b}{a}(1-e^{-at}) + \frac{b(a+b)}{a^2}(1-e^{-5a})(1-e^{-at})]\frac{b}{a^2}(-\frac{b}{a}y_0 + \frac{a+b}{a}y_0 e^{-a/2})$$

$$(1-e^{-3a/2})(b-(a+b)e^{-5a/2})(1-\frac{a+b}{a}(1-e^{-7a/2}))(1-\frac{a+b}{a}(1-e^{-9a/2})),$$

…….

In a word, in the interval $[(2n-1)/2,(2n+1)/2)$, $n=1,2,\cdots$ we have the following iteration formulas:

$$y_{n+1,m}(t) = y_{n+1,m-1}(t) - \int_0^t e^{a(s-t)}\left(y_{n+1,m-1}^{'}(s) + ay_{n+1,m-1}(s) + by_{n+1,m-1}\left(\left[s+\frac{1}{2}\right]\right)\right)ds$$

$$= y_{n+1,m-1}(t) - \int_0^t e^{a(s-t)}\left(y_{n+1,m-1}^{'}(s) + ay_{n+1,m-1}(s)\right)ds - \int_0^{\frac{1}{2}} e^{a(s-t)}by_{n+1,m-1}(0)ds$$

$$- \int_{\frac{1}{2}}^{\frac{3}{2}} e^{a(s-t)}by_{n+1,m-1}(1)ds - \cdots - \int_{\frac{2n-1}{2}}^{t} e^{a(s-t)}by_{n+1,m-1}(n)ds.$$

In view of (9), we can obtain the analytical approximation solution. Usually, the $m+1^{\text{th}}$ approximation is used for numerical purposes.

4 Numerical Simulation

In this part, we will present some examples to test the effectiveness of VIM and the correctness of our conclusions. The software we use is Matlab R2012a. All the figures are produced on it.

Let $a=1, b=-2$ and $y_0=1$ in (1). In Fig. 1 we compare the 7th VIM solution with the numerical solution of the θ-method [14] using $h=0.05$ and $\theta=0.6$. The graphs of the true solution and 8th approximate solution are shown in Fig. 2. Moreover, we also plot the 9th approximate solution and the numerical solution of the θ-method with $h=0.04$ and $\theta=0.7$ in Fig.3, the 10th approximate solution and the numerical solution of the θ-method with $h=0.02$ and $\theta=0.85$ in Fig.4. From these figures we can see that the higher approximation has better property than the lower approximation. Therefore, the VIM is useful for seeking the approximation solution of EPCA.

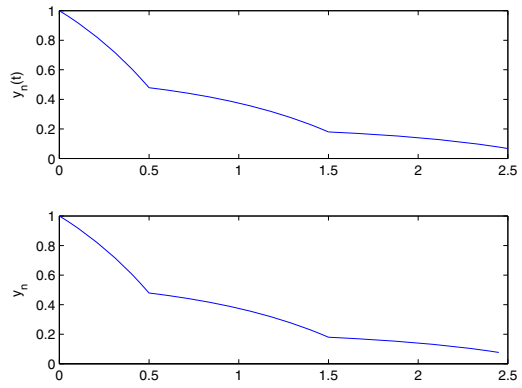


Fig. 1. A comparison between the 7th VIM solution (upper) and the numerical solution (lower)

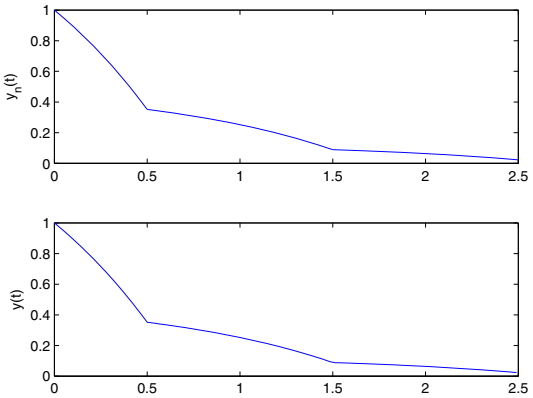


Fig. 2. A comparison between the 8th VIM solution (upper) and the true solution (lower)

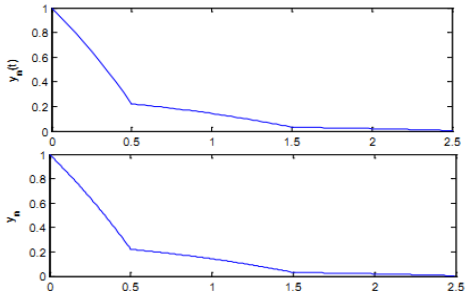


Fig. 3. A comparison between the 9th VIM solution (upper) and the true solution (lower)

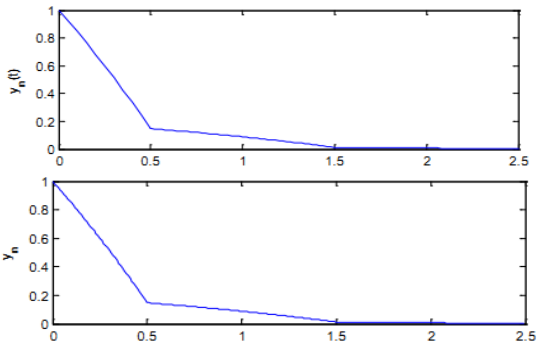


Fig. 4. A comparison between the 10^{th} VIM solution (upper) and the true solution (lower)

5 Conclusions

The VIM has been successfully applied for solving linear EPCA of alternately advanced and retarded type. An illustrative example is solved exactly. The results reveal that the VIM is very effective and simple.

Acknowledgements. The authors are very grateful to the reviewers for their helpful suggestions. This work is supported by National Natural Science Foundation of China (No. 11201084).

References

1. Wiener, J., Debnath, L.: Partial Differential Equations with Piecewise Constant Delay. J. Comput. Int. J. Math. Math. Sci. 14, 485–496 (1991)
2. Nieto, J.J., Rodríguez-López, R.: Study of Solutions to Some Functional Differential Equations with Piecewise Constant Arguments. Abstr. Appl. Anal. 2012, 1–25 (2012)
3. Akhmet, M.U.: On the Reduction Principle for Differential Equations with Piecewise Constant Argument of Generalized Type. J. Math. Anal. Appl. 336, 646–663 (2007)
4. Seifert, G.: Second Order Neutral Delay Differential Equations with Piecewise Constant Time Dependence. J. Math. Anal. Appl. 281, 1–9 (2003)
5. Gurcan, F., Bozkurt, F.: Global Stability in a Population Model with Piecewise Constant Arguments. J. Math. Anal. Appl. 360, 334–342 (2009)
6. Shah, S.M., Wiener, J.: Advanced Differential Equations with Piecewise Constant Argument Deviations. Int. J. Math. Math. Sci. 6, 671–703 (1983)
7. Cooke, K.L., Wiener, J.: Retarded Differential Equations with Piecewise Constant Delays. J. Math. Anal. Appl. 99, 265–297 (1984)
8. Wiener, J.: Generalized Solutions of Functional Differential Equations. World Scientific, Singapore (1993)
9. Ivanov, A.F.: Global Dynamics of a Differential Equation with Piecewise Constant Argument. Nonlinear Anal. 71, e2384–e2389 (2009)
10. Akhmet, M.U., Yılmaz, E.: Impulsive Hopfield-Type Neural Network System with Piecewise Constant Argument. Nonlinear Anal.: RWA 11, 2584–2593 (2010)

11. Yuan, R.: On Favard's Theorems. J. Differential Equations 249, 1884–1916 (2010)
12. Lv, W.J., Yang, Z.W., Liu, M.Z.: Stability of Runge-Kutta Methods for the Alternately Advanced and Retarded Differential Equations with Piecewise Continuous Arguments. Comput. Math. Appl. 54, 326–335 (2007)
13. Liu, M.Z., Ma, S.F., Yang, Z.W.: Stability Analysis of Runge-Kutta Methods for Unbounded Retarded Differential Equations with Piecewise Continuous Arguments. Appl. Math. Comput. 191, 57–66 (2007)
14. Wang, Q., Zhu, Q.Y., Liu, M.Z.: Stability and Oscillations of Numerical Solutions of Differential Equations with Piecewise Continuous Arguments of Alternately Advanced and Retarded Type. J. Comput. Appl. Math. 235, 1542–1552 (2011)
15. Wang, Q., Zhu, Q.Y.: Stability Analysis of Runge-Kutta Methods for Differential Equations with Piecewise Continuous Arguments of Mixed Type. Int. J. Comput. Math. 88, 1052–1066 (2011)
16. Allan, F.M.: Construction of Analytic Solution to Chaotic Dynamical Systems Using the Homotopy Analysis Method. Chaos Solitons and Fractals 39, 1744–1752 (2009)
17. Liao, S.J.: Beyond Perturbation: Introduction to the Homotopy Analysis Method. CRC Press, Boca Raton (2004)
18. He, J.H., Wu, X.H.: Variational Iteration Method: New Development and Applications. Comput. Math. Appl. 54, 881–894 (2007)
19. He, J.H., Wu, G.C., Austin, F.: The Variational Iteration Method which Should be Followed. Nonlinear Sci. Lett. A 1, 1–30 (2010)
20. Al-Sawalha, M.M., Noorani, M.S.F., Hashim, I.: On Accuracy of Adomian Decomposition Method for Hyperchaotic Rössler System. Chaos Solitons and Fractals 40, 1801–1807 (2009)
21. He, J.H.: Homotopy Perturbation Method for Bifurcation of Nonlinear Problems. Int. J. Nonlinear Sci. Numer. Simu. 6, 207–208 (2005)
22. He, J.H.: Approximate Analytical Solution for Seepage Flow with Fractional Derivatives Inporous Media. Comput. Method Appl. Mech. Eng. 167, 57–68 (1998)
23. He, J.H.: Variational Iteration Method–A Kind of Non-linear Analytical Technique: Some Examples. Int. J. Nonlinear Mech. 34, 699–708 (1999)
24. He, J.H.: Variational Iteration Method for Autonomous Ordinary Differential Systems. Appl. Math. Comput. 114, 115–123 (2000)
25. He, J.H.: Variational Iteration Method-Some Recent Results and New Interpretations. J. Comput. Appl. Math. 207, 3–17 (2007)
26. Wazwaz, A.M.: The Variational Iteration Method for Analytic Treatment for Linear and Nonlinear ODEs. Appl. Math. Comput. 212, 120–134 (2009)
27. Ali, A.H.A., Raslan, K.R.: Variational Iteration Method for Solving Partial Differential Equations with Variable Coefficients. Chaos Solitons and Fractals 40, 1520–1529 (2009)

Study on Safety and Reliability of Logistics Network System

Xiaoying Che

Qinzhou University, Qinzhou, Guangxi, 535000 China

Abstract. This paper proposed the construction of logistics requirements of the safety system in warehousing, transportation, information and transaction 4 features and technical equipment, process specifications, and personnel management 3 level,and the entire system is divided into the prevention, protection and emergency three sub-systems. Finally, according to the supply chain thinking, proposes construction of safety measures based guidelines for safety chain logistics system.

Keywords: logistics network, security system, reliability.

1 Introduction

Logistics security problem is not cause for concern, while also defining the scope of logistics security is no uniform standard. Some experts believe that the logistics security, including information security, transport safety, process safety and storage security four aspects. LUO Yixin [1] is defined the logistics security as logistics operation process, due to human error or technical defect or failure caused by damage to the goods, logistics damaged facilities and logistics information distortion and other security issues [2-5]. Enumerate the former security risk may exist logistics chain, logistics and security on the latter issue only limited to goods and logistics carrier itself damage. In fact, in order to provide safe and reliable logistics services, logistics and security for the issue under consideration needs to be more extensive. Therefore, this paper will further expand on the basis of the above definition [6].

Practice shows that the logistics and security not only to the people's property and life issues [7-10], but also the logistics enterprises to increase efficiency, improve service level guarantee. This article provides a general framework for logistics and security systems, the development of modern logistics to provide a reference.

In this paper, to improve computer information management in the enterprise logistics application level as a starting point [11-14], through the third-party logistics expertise Analysis on China's enterprises of the logistics management information system for the system planning, system planning framework is given, and to logistics management information system security recommendations.

For the development of enterprises, the logistics increasingly become the bottleneck. At present, China's public companies have no one to make convincing third-party logistics (referred TPC) business, so many companies have come up with a lot of

Y. Yang, M. Ma, and B. Liu (Eds.): ICICA 2013, Part II, CCIS 392, pp. 252–261, 2013.

money to do their own logistics. Therefore, third-party logistics enterprises should seize this opportunity to improve their logistics management. Modern logistics is a computer information management and communications technology as the core of the industry, but the current is still in the traditional enterprise logistics delivery, delivery phase. Therefore, third-party logistics enterprises have strict and scientific management system to achieve transaction information, information processing electronic ability to fully utilize computers and computer networks to process information in order to improve their competitiveness [15-18].

To achieve this, the key is to step up the development of logistics management information system construction. Logistics management information system design and development of third-party logistics do big, strong, showing their corporate image, from technology, the development of logistics management information system should focus on the overall situation, the overall consideration. First, managers and technical staff is fully co-ordinate planning, developing their own in-house information management system, establish their own supply and demand sites, followed by integration and optimization of the existing rules and regulations, and to develop new content. System Design Mode This system is available from third-party logistics enterprise's market position and aspects to be considered for development [19-21]. Third-party logistics in the market mode.

Thus, third-party logistics management information system consists of two parts : the external and internal information exchange online information processing. External information exchange customer management subsystem: accept online orders module; Query Module. Financial settlement subsystems: basic cost accounting module; special fee settlement module; inquiry fee settlement modules. Internal information processing warehouse management subsystems: Warehouse Management module; inventory management module; Query Module. Transportation management subsystems: vehicles, personnel change management module; transport task scheduling management module; Query Module. Financial settlement subsystems: the cost of clearing module; Query Module. Management Subsystem: Permission setting module; Query Module. Establishment of the enterprise logistics enterprises to build their own web site Internet site, the site will be in the mode of operation enterprises, distribution situation daily in online publishing. Through the use of modern e-commerce tools, and online distribution of information dissemination and collection, convenient for customers to save logistics costs and improve logistics effectiveness, so as to bring more business profits. Logistics management information system security according to the system of external and internal two modules, system security protection problem can be done in two ways. External information exchange security in the exchange of information with other companies, accept corporate orders, feedback and other important information on the Web [22].

Insecurity for the network (for example : network hacker attacks, network virus, network systems inherently insecure, etc.) may take appropriate measures, such as : data encryption, digital signatures, message authentication technology, access control technology, key management, authentication technology. Internal information processing security safe operating environment is the basic requirement of enterprise logistics information to form a highly conscious, law-abiding technology workforce computer information system security is the most important part of the work. To grasp this aspect of safety, can proceed in two ways : one is from a technical aspect. That the development and application of good business logistics management information

system permissions settings module, touching this information management system staff to set different access rights, set a password for identification, the user's regular, fixed terminal landing with the machine safety implement strict access control technology, strictly and effectively restrict unauthorized access to the computer to prevent unauthorized users invasion. The other is to start by management. Any enterprise development should be people-oriented, third-party logistics companies are no exception, companies can ideological quality, ethics, rules and regulations and so do some work. The development of modern logistics industry is the inevitable trend of development of the market, third-party logistics information management, design and development of enterprises will bring spillover effects, to achieve a comprehensive computerized management planning logistics management information system would link some of the more backward optimization at the same time, to speed up the modernization of the logistics industry, information technology, industrialization, the group, the professional development process, thus contributing to expand the market, to expand living space, improving the overall economic benefits, help logistics enterprises to achieve sustainable development.

This paper argues that the logistics security is through effective safety management measures to prevent and eliminate all lead logistics carriers and logistics operation devastating losses occurring factors, as well as the surrounding logistics and human health, the environment and the devastating social factor. This defines two types of logistics security issues.

2 The Problems Logistics Safety Management in China

At present, China first class logistics and security issues is very significant, serious impact on the logistics enterprise efficiency. As shown in Table 1, cargo damage rate compared with foreign enterprises, logistics operations cargo damage rate is still high. In addition, agricultural picking, transportation, storage and other logistics on the part of the loss rate of 20% to 30%, or about 1/4 of the agricultural products in the logistics chain is being consumed. Fruit loss rate in developed countries is below 5%, the United States is only 1% to 2%. Therefore, the strengthening of China's logistics security system is not only a safety measure, but also reduces logistics loss, improve economic means.

Table 1. Logistics enterprises damage rate (2004)

	Producers	Commercial enterprises	Fruits and vegetables	Fruits and vegetables (US)
Cargo damage rate	2%～5%	2%～5%	20%～30%	1%～2%

The second category Logistics is high frequency of security issues, greater harm to society, to the country and the people bring great losses. According to statistics, a logistics and transport of dangerous goods, 2000 accident year 514, the death of 785 people; to 2003 rose to 621, the death of 960 people. Since 2005, more than occurred all

over the country and leakage of hazardous chemicals major accident. Therefore, only improve our logistics and security system to ensure safety of people and property and lives.

Logistics enterprises pay enough attention to safety, security measures are still in place. Most logistics enterprises in the assessment of logistics performance is only concerned with two types of logistics costs and customer satisfaction indicators, ignoring the logistics safety management indicators, and in order to further reduce logistics costs at the expense of the necessary basic safety measures. Most logistics companies do not set up a full-time job security management, security audit jobs, security risk assessment and accident investigation jobs. In addition, in a survey, there are still 28% of those surveyed did not take simple physical security measures. Although the practice has proved that about 2/3 of the cargo theft is theft collusion with internal staff conducted. But in 2005 the staff review of measures has been taken only 53%. Thus corporate emphasis on security and logistics are still far from enough.

At the macro level, the government and industry associations to monitor the logistics, there are big safety deficiencies, mainly reflected in the management of fragmentation, and those policies, departmentalized; there is no unified planning, unified security and certification system. Connect supply logistics, manufacturing, wholesale and retail supply chain processes involved during the transportation, handling, warehousing, customs clearance and other links. Customs, public security, public health, transportation, railway, industry and commerce, commodity inspection, Highways, industry associations and other departments were only a single aspect of the logistics responsible for the safety, the lack of the entire logistics chain safety comprehensive monitoring, can not form a coherent security system, can not be inevitably appear on the regulatory vacuum, making accidents very often.

3 The Establishment of Logistics Security System

For the current security problems of logistics to make some countermeasures, namely: First, from the logistics function and system-level perspective constitute two proposed framework and logistics security systems specific instructions for each part; Second, according to the logistics industry characteristics, the supply chain management and value chain analysis of the two theories is proposed based on the preliminary safety chain logistics thinking.

3.1 Logistics System

Warehousing and logistics security systems can be divided into security, transportation security, and information security and transaction security four aspects. In each function, and then based on safety management level, divided into technical equipment, process specifications, and personnel management three levels, as specified in Table 2.

Table 2. Functions and levels of logistics safety system

	Storage Security	Transportation Security	Information Security	Transaction Security
Technical equipment	Security monitoring; Access control systems; Alarm system;	Vehicle safety monitoring Overload monitoring Driver monitor Items Security	Data Backup Communication Encryption System Recovery Intrusion Detection	Credit Management System Verification of the information system
Process specification	Database management; Obsolescence management Valuables Management	Vehicle Scheduling Cargo transfer Accident Unloading; loading	Permission settings Information Sharing	Customer credit investigation Information Review Guarantee transactions, Control system
Personnel management	Employee credit investigation; recruitment; dismissal; safety training; safety awareness; safety culture			

3.1.1 Storage Security

Warehouse items occur theft, fire, flooding, damage and deterioration phenomena, it will cause substantial damage to property. Storage should be done to strengthen safety management following two points.

1) From a technical point of view the device, you must install CCTV in the warehouse, on the key areas to monitor, understand the dynamics of the warehouse and the subsequent investigation of cases to help; warehouse entrance, window rails, cargo handling areas and valuables storage is the most important area, to use multiple cameras. In addition, access control systems, intrusion alarm systems and other physiological characteristics such as fingerprints scanning measurement techniques can be used to restrict access to warehouse personnel.

2) Well-designed goods and people out of storage process specifications, and make inventory items scrapped report and approval procedures, the procedures to eliminate all undesirable elements may be exploited. Safekeeping of valuables, you need to set aside a dedicated area in the reservoir area and itemized registration and strict management.

3.1.2 Transportation Safety

From the point of view adopted by means of technical equipment, transport safety measures include vehicle safety, cargo safety, driver safety and prevent overloading of four aspects, see Table 3.

Table 3. Technologies and facilities for transportation securities

Function	Equipment
Vehicle safety monitoring	Tachograph ; GPS satellite positioning ; GIS map guide ; Vehicle theft
Overload safety monitoring	Weighing; pressure sensitive sensor; infrared detection
Driver monitor	IC card technology; alcohol testing; fatigue driving monitoring
Cargo security monitoring	Compartment reinforcement; electronic tags; X-ray detector

But also attach importance to strengthen all aspects of transportation safety process management. Trucking operation process can be divided into delivery loading, transport and unloading destination 3 parts. The focus of loading and unloading cargo to discover and report exceptions, transportation focuses on preventing theft and carjacking, in which processes are very important emergency situations, based on actual experience, a lot of unexpected situations, such as anchoring, traffic accidents, which may be a these premeditated crime.

3.1.3 Information Safety

Information systems have become the important tools of the logistics enterprises to improve efficiency and enhance the competitiveness, so information systems collapse would hinder the normal operation of the logistics, and such as customer information, price information, driving directions and warehouse data and other information leaks, if competitors or criminals access, logistics enterprises will bring immeasurable loss.

1) Select the stable operating system
2) regularly to patch
3) Turn off the system default shared resources
4) timely backup the operating system
5) Install anti-virus software
6) caution unsolicited CD, USB
7) Use caution programs and files on the network
8) to prevent network attacks and intrusions

3.1.4 Transaction Security

As the domestic legal system is not perfect, commercial credit is not high, enhance transaction security to do the following.

1) To carry out customers and trading partners, credit investigation, credit building customer transaction records and files, create logistics enterprise credit management system is to improve the efficiency of credit investigation a good way.

2) the use of information technology, logistics information system documents automated matching feature, you can reduce the workload and improve the efficiency and accuracy of documents checked.

3) for the implementation of the credit guarantee system for new customers, such as stored deposits.

4) To strengthen the internal control system in order to prevent collusion fraud.

3.2 Logistics and Security Management Subsystem

Logistics security system is divided into prevention systems, security systems and emergency systems (Figure 1). Which are routine safety prevention systems design part of the emergency system is in an exceptional case of emergency safety measures designed to protect the system will provide logistics security system daily maintenance needs.

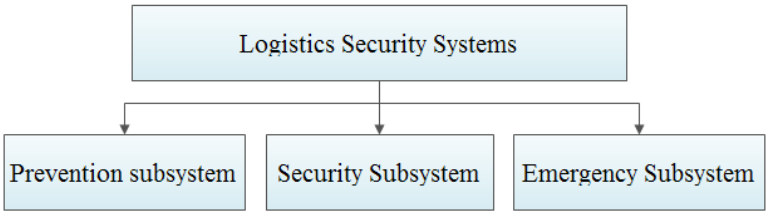


Fig. 1. Sub-modules of logistics safety system

3.3 The Construction of Logistics Safety Chain

With the theory of supply chain development, people gradually realized that security is not only the logistics associated with individual enterprises, but also with the entire logistics chain for all business-related nodes. Shown in Figure 2, the logistics permeates from the manufacturer to the customer's entire supply chain, each node in the enterprise, the completion of cargo agents, storage and processing, multimodal transport, customs clearance and urban distribution and other functions, during and after a freight forwarding company, warehousing enterprises, transport companies, customs brokers and distribution companies and other units. Meanwhile manufacturers reverse logistics chain through recycling waste products to their customers, which constitute the circulatory system of the logistics supply chain.

1) Logistics chain security systems including logistics enterprises, government agencies, industry associations and intermediary organizations. Shown in Figure 2, the government agency responsible for safety oversight policies and monitoring; logistics enterprises directly responsible for safety precautions, security and emergency work; industry associations and intermediary organizations to provide enterprises with security consulting and safety certification, and coordinate parties to develop industry safety standards while providing relevant government agencies to provide policy advice and logistics industry safety assessment feedback.

2) based on safety chain logistics ideology, government regulators to do: ①policy to the entire logistics chain must be safety-oriented and eliminate fragmented, not only to take into account policies are responsible for their part of the impact, but also take into account policy on the downstream part of the logistics chain effect; ②in order to timely and accurately grasp the logistics chain security situation, different government departments should strengthen communication and shared information systems.

Logistics companies need to do: ①upstream and downstream enterprises transfer responsibility for security, and retain the transfer record transfer process may eliminate loopholes; ②between upstream and downstream enterprises, through information systems interconnection,to realize information sharing, security risks promptly notify various enterprises on the chain, so that preventive measures ahead of other companies; ③by the logistics chain's core business lead, held on a regular basis to keep up with the downstream business safety seminars, joint research security problems and countermeasures.

3) Safety chain logistics and security as ideology requires an enterprise to create value for the work. Past, companies lack of security awareness, not fully aware of the importance of logistics security, without taking into account the economic value of security, safety and economic benefits of the opposition, is considered safe investment burden on enterprises, did not put the safety measures put in place.

Shown in Figure 2, the paper will be divided into basic logistics logistics activities and supporting activities. Basic activities including transportation, warehousing, value-added processing, marketing and customer service unit. Ancillary activities including logistics facilities and equipment, personnel management, technology development and management, outsourcing of procurement or inter-enterprise collaboration unit. The main idea is to analyze the system's various units, through the development of appropriate corrective measures and programs to achieve maximum security value of each unit. Then through the whole system analysis and calculations, to achieve the maximum value of the overall system security. Security through the logistics value chain analysis, we can understand the value of each part of the activities for security contribution helps to grasp the safety management priorities, establish safety can create effective concept.

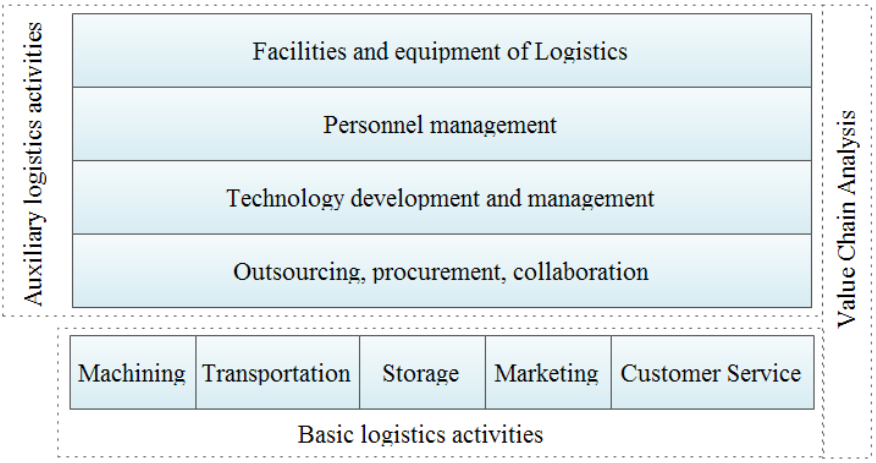


Fig. 2. Analysis of logistics based on safely value chain

4 Conclusion

With the logistics supply chain management continues to evolve, the complexity of various logistics information, companies urgently requires logistics information technology, computer network technology and the popularity of logistics information provided gave technical support. Therefore, the logistics information system in the enterprise took root, and the enterprise has brought greater efficiency. Business is based on the following background was vigorously develop the logistics information system. In today's world, are basically a buyer's market, purchased by the consumer to choose which company's products, they basically have complete freedom of decision. Producing the same product on the market plethora of companies, businesses in order to win the competition, we must continue to introduce new lower cost quickly meet consumers always changing consumer demand, which require rapid response logistics system. For a quick response, feedback must be timely, which necessarily requires companies to establish their own logistics information system. Modern competition among enterprises is largely manifested as the competition between supply chains, while in the entire supply chain, more links, information, relatively speaking, more complex, enterprise communication between them is much more difficult. Each link in order to communicate freely, to share information, build supply chain logistics information system is imperative. The rapid development of computer technology, network wide extension of the whole society into the information age. In this Internet age, only the information society, it may have a larger enterprise development. Moreover, the development of information technology, information systems development has laid a solid foundation. Enterprise as a social one, the logistics industry as a social service, the logistics industry is bound to create their own information systems.

References

1. Luo, Y.: Study on the current situations and countermeasures for China’s logistic safety. Science Technology and Industry 2006(5), 16–19 (2006)
2. Luo, Z.: Analysis of logistics chain safeguards system. Logistics Sci- Tech 2005(10), 8–10 (2005)
3. Wang, D.: Increasingly important non- tradition security problems. Social Observations 2005(12), 35–37 (2005)
4. Zhang, C., Shan, S.: Brief discussions of logistics safety management. Enterprise Economy 2006(5), 39–40 (2006)
5. Xuequan, L.: Road transport network vulnerability studies. Urban Roads Bridges & Flood Control, 6 (2010)
6. Xifeng, T., Haijun, M., Xu, L.: logistics distribution center location of multi-objective optimization model. Southeast University (Natural Science), 2 (2009)
7. Cheng, C., New Heng, X., Qingmin, Z.: Multi-Agent based on ontology and a reliable supply chain network design model. Computer Integrated Manufacturing Systems, 1 (2011)
8. Song, C.: Urban road network based on reliability optimization model and algorithm. Transport Information and Safety, 3 (2010)

9. Yifeng, Z., Peng, L., Ruhe, X.: Based Safety Reliability food logistics system optimization. Wuhan University of Technology (Transportation Science & Engineering Edition), 6 (2010)
10. Cheng, C., New Heng, X.: Supply Chain Reliability Management Research and Prospect. Chinese Manufacturing Informatizationm, 19 (2011)
11. Cheng, C., New Heng, X.: MAS-based supply chain reliability assessment model. Chinese Manufacturing Informatization, 23 (2011)
12. Wei, B.W., Zhao degrees: Emergency logistics network resiliency Study. Value Engineering, 4 (2008)
13. Qin, X.W., Lixin, T.: Disruption Event oriented reliable logistics system design research progress. Control and Decision, 2 (2010)
14. Lei, W., Xu, L., Mei, L.L., Zheng, Y.-H.: Logistics and distribution center Handling Systems Reliability Modeling. Hoisting and Conveying Machinery, 7 (2009)
15. Tran.: Logistics network reliability and application of the key issues. Central South University (2010)
16. Meiling, H.: Reliability analysis based logistics service supply chain design and coordination of Southwest Jiaotong University (2010)
17. Cheng, C.: Based on Multi-Agent System Reliability Supply Chain Management Model. Nanjing University (2011)
18. Zhang, D.-H.: Logistics Service Supply Chain Coordination Mechanism. Electronic Science and Technology University (2007)
19. Wei, F.: Urban hubs external transport system integrated key issues theory. Southwest Jiaotong University (2010)
20. Yan, Z.: Reliability-based fresh produce logistics network optimization. Southwest Jiaotong University (2009)
21. Po, C.C.: Risk perspective supply chain design optimization model and related issues evaluation. China University of Mining (2010)
22. Zhou.: Based on variable weight multi-objective logistics network traffic allocation. Southwest Jiaotong University (2012)

Pharmaceutical Drugs Logistics, Storage, Security with RFID Sensor Networks

Yin Qun and Zhang Jianbo

Kunming University of Science and Technology, 650093, Kunming Yunnan, China

Abstract. This paper analyzes the management of pharmaceutical drugs problem and the characteristics of RFID sensor networks. Then it describes how to use RFID sensor network technology to solve pharmaceutical drugs logistics, storage security, which improves pharmaceutical logistics service, and enhances pharmaceutical drug safety management to reduce accidents.

Keywords: sensors, logistics, RFID, nRF2401, Ad Hoc, ZigBee.

1 Introduction

According to the "To err is human" published by the Institute of Medicine (IOM) of U.S. National Academy of Sciences in 2000, because of Adverse Events every year in the United States,the number of deaths is between 44,000 people to 98,000 [1-5]. Each year about 7,000 deaths for Adverse Events which are caused by drugs, is more than car accidents, breast cancer, or even AIDS deaths. It is a serious warning, which prompt the U.S. medical to plan to build a more secure medical environment in 10 years, in order to achieve the six goals of "Security", "Effective", "Patient-centered", "Timely Manner", "efficiency" and "Medical equality" for the medical environment [6].

In this paper, we propose the RFID sensor network which is based on the RFID R/W Devices integrated with sensor networks according the fully research of RFID and WSN technologies.This network can not only expand the coverage of RFID system but also can extend the kind of information of RFID information systems. Preliminarily, we apply the RFID sensor networks in the existing logistics, storage systems of pharmaceutical drugs, in order to reduce the incidence of pharmaceutical drugs accidents [7-11].

2 RFID Sensor Networks

2.1 RFID System

A typical RFID system is shown in Figure 1, which is composed of the electronic tag, Read / Write Device, antenna, and computer systems which are responsible for data exchange and management [12-15].

In this paper,the design of the electronic tag based on RFID sensor network is divided into two, namely tag-sensor integrated design (as shown in Figure 1) and tags

Y. Yang, M. Ma, and B. Liu (Eds.): ICICA 2013, Part II, CCIS 392, pp. 262–271, 2013.

of multi-sensor design (as shown in Figure 2) [16]. Tag-sensor integrated design refers to that in the specific environment required for detection or small range of conditions, the electronic tags bundled with sensor and the electronic tags store relevant information (ID, a single sensor, data) [17-21]. It usually placed on the items which need to be identified,and the information stored by the electronic tags usually can be got by Read / Write Device via non-contact mode; The design of multi-sensor tags refer to that in the required environment where exists more factors or broader range, the electronic tags communicate with the sensor network through the antenna. For example, electronic tags can get the data of temperature, humidity, light intensity which provided by the wireless sensor networks through the antenna, then Read / Write Device can get the data from the tags by the wireless transceiver.

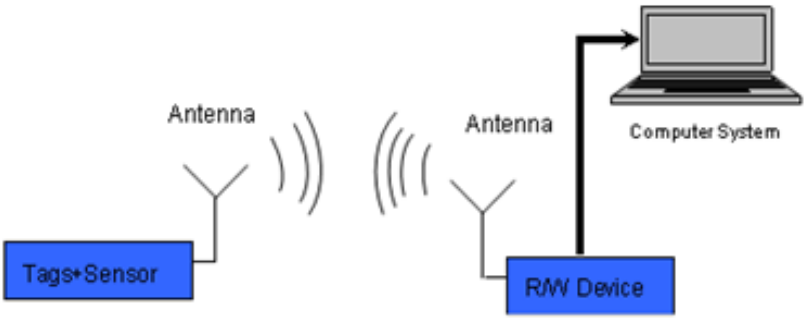


Fig. 1. Design of tag-sensor integratedtags

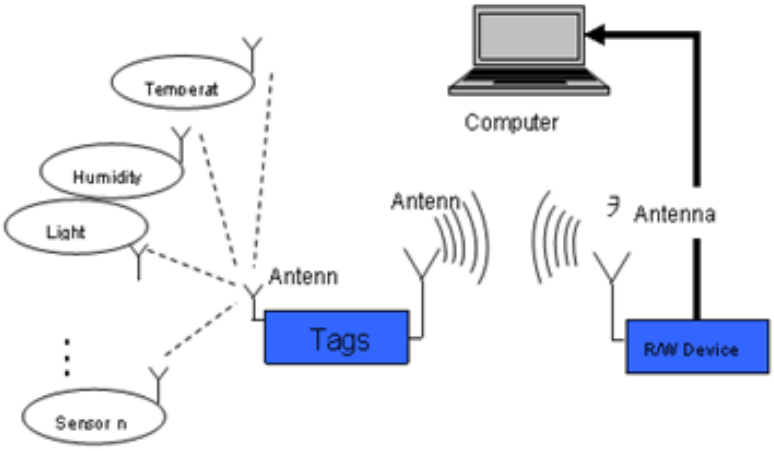


Fig. 2. Design of tags of multi-sensor

2.2 The Design of Hardware Abstraction Model

WSN systems are composed by the large quantities of low-cost sensor nodes which deployed in the detected region, and it is a multi-jump, self-organizing network formed

by wireless communications, aiming at cooperating to perceive, collect and process the object information from the coverage area of network, and send it to observers [22-25]. In the aspects of medical logistics and warehousing, the change of temperature, humidity and light need to be reflected to the medical inspectors.The following shown in Figure 3 is a sensor network system which is suitable for pharmaceutical drugs in the sectors of medical logistics and warehousing.

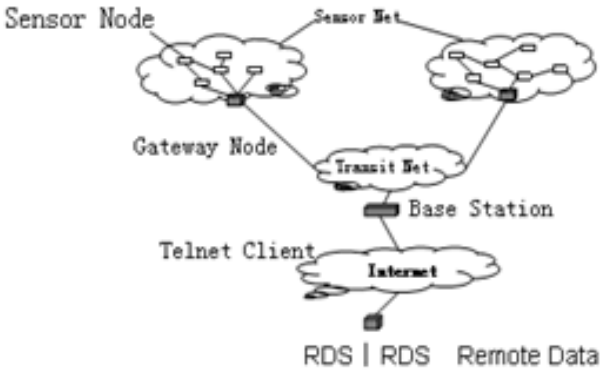


Fig. 3. The WSN system architecture areas

3 The RFID Sensor Networks Which Is Adapted to the Medical Logistics and Warehousing

Tag is tagged items (medical supplies, pharmaceuticals) and traditional logistics for transport [26-30]. Each to a transit station, the tags communicate with R/W Devices automatically, R/W devices will collect the data from tags which stored the transit data (temperature, light, humidity, etc.) and connect to our "business logic services" through the Internet on storage.While the end-user at the destination can view the logistics linked data collected by R/W Devices from time to time. Once an exception is found related to the transportation of drugs (such as temperature, humidity and other parameters beyond the specified range),it will alert and inform the relevant principal for handling. It overcomes the disadvantages of the traditional logistics, because transit environmental parameters can not be detected, the drugs only can be tested when it reaches the destination. But at the destination, generally there isn't a test conditions, if a patient injected with a metamorphic drugs, it is probably that endanger the safety of patients [31-35].While the end-user at the destination can view the logistics linked data collected by R/W devices from time to time. Once an exception is found related to the transportation of goods (such as temperature, humidity and other parameters beyond the specified range) is to alert, inform the relevant principal for handling. It overcomes the drawbacks of traditional logistics, because transit environmental parameters can not be detected, the drugs only can be tested when it reachs the destination. But at the destination, there isn't a test conditions generally. If a patient is injected with metamorphic drugs, it will endanger the safety of patients.

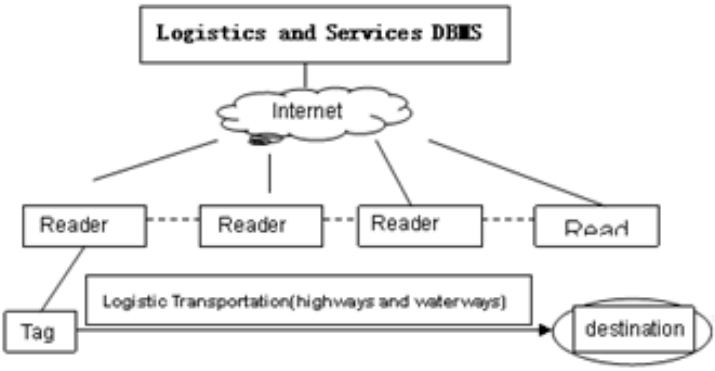


Fig. 4. RFID sensor network structure

We can research into the circulation of the pharmaceuticals by the RFID sensor network. The RFID sensor network has a large number of sensor nodes and R/W Device, sensor nodes can be used to accurately monitor the surrounding information (temperature, humidity, light intensity)of the moving targets, and RFID R/W Device can be used to determine the specific environment value of pharmaceutical drugs, according to which we can take appropriate measures.

4 Sensors and Tags Communication and Principle

4.1 Wireless Ad Hoc sensor network design

Wireless sensor networks as a new Ad hoc networks, compared to traditional wireless network has wide coverage, the formation of convenience, and strong survivability advantage, It is widely used In health care, natural disaster monitoring, military and other field.In this paper, we used ZigBee technology based on a standard of IEEE 802.15.4 as sensors and RFID tags communication technology [36]. ZigBee has a

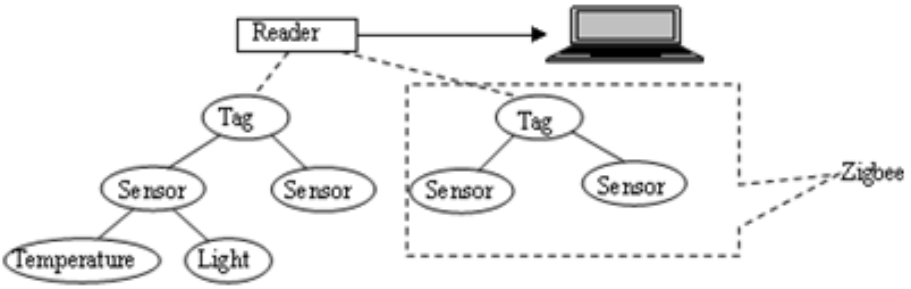


Fig. 5. Tags and sensor network structure based on Zigbee

close, low complexity, low power, low data rate, low-cost wireless communication technology. A RFID tag can be connected to multiple sensors, tags and sensors is linked to be network with ZigBee technology.

4.2 Tag Sensor Hardware Diagram

The hardware design of Tags and sensors is the same basically. It can be that the Tag is the simple sensor.Tags and sensors are used ATMEL microcontroller as the core processor.

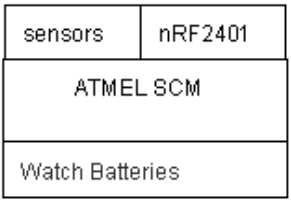


Fig. 6. Sensor hardware diagram

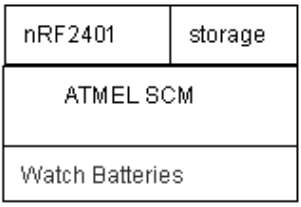


Fig. 7. Tag hardware diagram

Under normal circumstances, the tag is in a standby state. The standby current of nRF2401 is about 32uA.The current of ATMEL SCM is between 4 to 20 mA, A button battery which voltage is 3V, capacity is 210mAh, it can standby 100 hours (5 days) in theory. When combined with software processing, the system can work about 200 hours into hibernation. The following table describes the power consumption of whole system at launch and standby in sleep:

4.3 WSN Communication Protocol Introduced

Micro-node of wireless sensor network is a one-time which requires low-cost nodes and working time as long as possible. Wireless sensor network should not exist special router nodes. Each node is a terminal node and is a router node. Nodes link together by mobile ad hoc networks and communicate by multi-hop routing mechanism.

IC	Power consumption in Emission	Power consumption in standby, receiving	Power consumption when Sleeping
ATMEGA48	3V,14.25mA,42.75mW	3V,11.85mA,35.55mW	1.2V,1mA,1.2mW
nRF2401	3V,8.8mA,26.4mW	3V,11.95mA,35.85mW	3V, 1uA, 0.3mW

Fig. 8. The power consumption of whole system at launch and standby in sleep

Therefore, tag is the node and is the routing node. The tag keeps the data which sensor sends, then it sends data to the RFID reader until the transfer station. Transceiver sensors and RFID chips communicate by nRF2401.

The hardware of the control unit of an active tag uses "ATMEL SCM" and software uses the C programming language. Currently, there are 2.45 GHz communications protocol standard for active RFID systems, active electronic tag and reader communication using a specific communication protocol. The control of the whole communication protocol is based on the MAC (Media Access Control) layer, each tag has a specific ADDR address and a length of 2 byte. Data transmission with a simple message communication between Tag and Reader. Each "message" is made of "command frame" and "data frame" composition. Communication protocol in data message is by the command frame and the data frame composition.

```
typedef struct _cmd_hdr    //define the head of frame
  {
     unsigned char SOF;          //4-bit Identify +16-bit protocol version number
     unsigned char ADDR;         //16-bit MAC address
     unsigned short total_len;   //Command frame length (byte)
     unsigned short Parameter;   //Parameter
     unsigned short checksum;    //8-bit the checksum 0f command frame
     unsigned char EOF;          //8-bit the end logo of command frame
}CMD_HEADER;
```

Fig. 9. Different message's command frame and data frame composition

SOF is frame head mark position, length is 1 byte, its scopefor 00H ~ FFH.

EOF is frame ending sign bit, length is 1 byte, value range for 00H ~ FFH.

ADDR is the label address bits, and the length is 2 byte. Its value ranges for 0000H ~ FFFFH. the total can accommodate 65535 labels address which keeps four address: (0x00,0x00)(0x00,0xff)(0xFF,0x00) (0xFF,0xFF) the actual usable address is 65531.

Data frames (Response) is by seven mark a composition, respectively is SOF, LEN, UID, EXID, CHK and EOF. The UID is the label ID, EXID is to expand the ID, the system has a total of 65531 address, actual can accommodate tag number for 65531 * 255, about 16.71 million addresses. For a small application this tag number is enough. If a large and medium-sized application, the system design obligate 1 byte expansion ID (EXID), enable extended ID a, always tag number about 4.2 billion.

5 Summary

In this paper, medical supplies and pharmaceuticals is managed and tracked by using RFID technology. In this way, counterfeit medicines can be marked. At the same time, medicines can be tracked on drug management. Then, the security problems of the source of drugs are solved.

Since we added the sensor and large-capacity memory to the tag, the tag not only has the function of its own identity, but also the function of a multi-point inspection and full monitoring. The tag can collect and alter sensor data automatically and regularly, which solves problems of detecting and managing drugs in transit. Once the drugs have problems in the way, the tag will record relevant information. When the tag arrive at the destination, managers can quickly know that this group of drugs, goods in transit is affected to be useless.

China is currently not yet standardized coding medicines, pharmaceutical production enterprises, commercial wholesale production and sale of drugs but no one Taiwan and France - the identification of various fields were drawn up their own logistics coding, the result is not intelligence between different fields pass, hampering the effective implementation of logistics management system, resulting in inefficient information processing and circulation. There is no uniform identification code, not been able to query and track the flow of goods, can not determine whether a drug as soon as possible, and in some pharmacies, hospitals frequently encountered buy real return false, as fake, dealt a tremendous difficult. Will not be able to meet in order processing, pharmaceuticals valid management, cargo tracking by lot number and other modern quality management requirements, but also for drug quality control brought great difficulties. Demand from the pharmaceutical product itself and national logistics management requirements in terms of drugs, we must first choose an advanced and scientific coding system to encode for pharmaceutical products, we believe that with the EPC system more appropriate. It can not only be the product code, the key is the combination of energy and RFID. Electronic Product Code EPC system is a system, which is considered to be uniquely identified an effective way of all physical objects. These objects contain trade products, packaging and logistics units and other systems. Although the EPC code itself contains limited identifying information, but it has a corresponding back-end database as a support, the EPC code corresponding to the product information stored in the database can quickly query the information they need.

Use of RFID technology can be used in pharmaceutical production and distribution process, the specific methods of operation as follows: First, manufacturers, wholesalers, retailers can use the unique product code to identify the identity of

pharmaceutical products. Production process in every kind of pharmaceutical products to affix electronic tags, electronic tags recorded the only product code number, product code before using the product identified in the production, in the production of electronic tags written after the completion of the batch to batch number, complete complete electronic medical products code number, as in the next circulation, distribution and recycling unique code. Logistics providers, wholesalers, retailers with manufacturers to provide the reader can test the product in strict legality. Establish through RFID technology such drugs from manufacturers to pharmacies throughout the tracking capabilities to enhance consumer access to drug safety, can effectively eliminate the harm caused by counterfeit medicines, but also to prevent expired drugs into the market. Also in the pharmaceutical supply chain management, the use of RFID technology, in every kind of product loaded on electronic tags, recorded the only product code number that will solve many production processes and marketing aspects. Pharmaceutical producers can accurately grasp the status of products, improve production efficiency, reduce labor costs, shorten product quality, test time, real-time monitoring product manufacturing process all the circumstances, to respond quickly to market, reducing the number of expired products losses. Product coding standard is very basic work, especially for pharmaceutical production and logistics has very important significance, but the specific implementation requires authority and economic strength. The developed world over the years a lot of manpower and material resources, efforts to standardize the work of medical information, has made remarkable achievements. There are many standards have been widely used, it is worth learning from. If NDC National Drug Code (Natona DrugCodes, NDC) which is one of the outstanding representatives. NDC is the U.S. Federal Drug Administration requires the use of medicines coding standard, it includes many drugs meticulous detail, including packaging requirements.

References

1. Hinkka, V., Holmström, J., Främling, K.: RFID tracking in the book supply chain: the transition from postponed to speculative tagging. International Journal of Logistics Research and Applications 15(3), 199–214 (2012)
2. Zhu, Z., Tan, J., Ren, H., Ni, W., Guan, Q.: RFID application in manufacturing: A case study on a pilot RFID project in household appliance production. International Journal of Computer Integrated Manufacturing 25(1), 3–10 (2012)
3. Shi, X., Tao, D., Voß, S.: RFID Technology and its Application to Port-Based Container Logistics. Journal of Organizational Computing and Electronic Commerce 21(4), 332–347 (2011)
4. Bose, I., Lui, A.H., Ngai, E.T.: The Impact of RFID Adoption on the Market Value of Firms: An Empirical Analysis. Journal of Organizational Computing and Electronic Commerce 21(4), 268–294 (2011)
5. Wu, X., Subramaniam, C.: Understanding and Predicting Radio Frequency Identification (RFID) Adoption in Supply Chains. Journal of Organizational Computing and Electronic Commerce 21(4), 348–367 (2011)

6. Liu, C.M., Chen, L.S.: Applications of RFID technology for improving production efficiency in an integrated-circuit packaging house. International Journal of Production Research 47(8), 2203–2216 (2009)
7. Bottani, E., Ferretti, G., Montanari, R., Rizzi, A.: The impact of RFID technology on logistics processes of the fashion industry supply chain. International Journal of RF Technologies: Research and Applications 1(4), 225–252 (2009)
8. Hardgrave, B.C., Aloysius, J., Goyal, S.: Does RFID improve inventory accuracy? A preliminary analysis. International Journal of RF Technologies: Research and Applications 1(1), 44–56 (2009)
9. Bottani, E., Hardgrave, B., Volpi, A.: A methodological approach to the development of RFID supply chain projects. International Journal of RF Technologies: Research and Applications 1(2), 131–150 (2009)
10. Hellstrom, D.: The cost and process of implementing RFID technology to manage and control returnable transport items. International Journal of Logistics Research and Applications 12(1), 1–21 (2009)
11. Gaukler, G.M., Hausman, W.H.: RFID in mixed-model automotive assembly operations: Process and quality cost savings. IIE Transactions 40(11), 1083–1096 (2008)
12. Loebbecke, C., Huyskens, C.: A Competitive Perspective on Standard-Making: Kaufhof's RFID Project in Fashion Retailing. Electronic Markets 18(1), 30–38 (2008)
13. Sigala, M.: RFID Applications for Integrating and Informationalizing the Supply Chain of Foodservice Operators. Journal of Foodservice Business Research 10(1), 7–29 (2007)
14. Loebbecke, C.: Piloting RFID along the Supply Chain: A Case Analysis. Electronic Markets 17(1), 29–38 (2007)
15. Wang, W., Yuan, Y., Wang, X., Archer, N.: RFID Implementation Issues in China: Shanghai Port Case Study. Journal of Internet Commerce 5(4), 89–103 (2006)
16. Hu, L.-C.: Transit non-contact IC card reader application design. Anhui Electronic Information Vocational and Technical College, 6 (2003)
17. Lu, W., Xiang, Q.: RFID-based tracking and locating system for underground personnel. Safe, 1 (2004)
18. Qiu, M., Qun, S.S., Jingsong, W., Hq.: Radio frequency identification system of digital processing core module. Beijing University of Posts and Telecommunications, 6 (2005)
19. Suli, Z., Min, L., Willowy, Z.: Based on USB interface board non-contact IC card reader. Changchun Institute of Technology (Natural Science), 2 (2004)
20. Xian, C., Luo, S.-D., Yang, H.: Radio frequency technology in mine management system application. Changsha Aeronautical Vocational and Technical College, 4 (2005)
21. Zhijie, Z., Jinming, Y., Liang, W.: Marathon time based on radio frequency identification technology system. Journal of Donghua University (Natural Science), 5 (2003)
22. Zhao- Gifted: Transit non-contact IC card reader application design. Microcontrollers & Embedded Systems, 10 (2001)
23. He A Uu: RFID technology application prospects. Information Technology & Standardization, Z1 (2006)
24. Xiaoli, C., Fang, S.Y.: RFID-based Conference Registration System. Electronics Engineers, 7 (2005)
25. Shi, J.H., Cheng, Z., Jiayu, L.: Radio frequency identification technology and its application and development trend. Electronic Technology Applications, 5 (2005)
26. Jing, H., Shimin, Y.: China's pharmaceutical distribution and circulation development agency Discussion. China Pharmacy, 6 (2003)

27. Hu, J.-H., Tianxiang, L., Wei, Y.Z., Lei, G., Qin, H., Wei, Y., Chang, Z., Yu every row., Feng, Q.: pharmaceutical supply chain management system development and application. China Pharmacy, 2 (2005)
28. Jie, X., Dong, J.: Logistics network content analysis. Beijing Jiaotong University (Social Science Edition), 2 (2005)
29. Zhou, Z., Lei, Y., Min, S.: Based on GPS / GIS / RS Integrated Technologies monitoring and management. Systems Engineering, 1 (2003)
30. Jianming, C., Feng, Z.: GO Methodology Based Supply Chain Reliability Analysis. Highway and Transportation Research, 3 (2007)
31. Jing, L.: RFID radio frequency identification technology in the logistics industry Status and Prospects. Gansu Science and Technology, 11 (2006)
32. Wen, L.: Western Regulation Theory and Its Evolution. Teaching and Research, 7 (2002)
33. Weijian, L., Qin, P.: Analysis of RFID applications in the field of logistics. Logistics Technology, 3 (2006)
34. Hao, P., Quanlin, L.: RFID privacy and security of key technologies. In: 2006 International RFID Technology Forum Proceedings (2006)
35. Zhiguo, W., LEARNING IN, Jie, L., Teng, F., Yujing, Y.: Based on RFID & Barcode data fusion technology in the automotive assembly process Application Research. In: Twelfth International Conference on Industrial Engineering and Engineering Management (Two) (2005)
36. Shijie, W., Meijia, Y., Shiliang, W.: Design of Storage Management Information System and Research of the Information Flow (ID: A-014). In: The Proceedings of the 14th International Conference on Industrial Engineering and Engineering Management, Volume A (2007)

Research on Integration of Information Flow and Logistics Based on TPL

Tiangong Zhang

China Academy of Transportation Sciences, Beijing 100029 China

Abstract. In VMI mode, logistics outsourcing to third-party logistics, although vendors can solve logistics capabilities and information issues such as lack of capacity, but also increases the complexity of the operation of the VMI. VMI-based third party logistics operation mode (i.e. VMI & TPL mode), the supply chain information flow and logistics bound to change. For example VMI mode transport, distribution and inventory management changes; information sources, information transmission and information sharing changes. Therefore, for the introduction of third-party logistics VMI works, how to effectively integrate information flow and logistics, third party logistics and VMI achieve the organic integration has become necessary to solve the problem. This article from the system's perspective on VMI & TPL mode information flow and logistics integration have deeply study and to achieve the desired goals.

Keywords: Integration and Research, Vendor-managed inventory (VMI), Inventory and Transportation, Third-part logistics (TPL).

1 VMI Introduction

Supply chain management (SCM)[1-3] is the process of globalization and the future direction of development, which has become the theory and application of research focus. With economic globalization and integration of development, competition among enterprises has jumped out of a single enterprise to compete for markets and customers one-dimensional single-competition model, based on the entire supply chain tend to multidimensional, three-dimensional integrity of the competition. The world's leading supply chain expert Professor Christopher[4-7] pointed out that the 21st century, competition is not competition between enterprises, but the supply chain and supply chain competition on the market and not only the supply chain enterprises.

Supply chain [3] refers to the production and distribution process involved in the raw material suppliers, manufacturers, wholesalers, retailers and final consumers that is supply and demand networks. Vendor managed inventory (VMI) [8-11] refers to the retailer or the user will own inventory business entrusted to a vendor to manage the supply chain management. Implementation of VMI mode, dominated by the retailer's supply chain mechanism into the whole supply chain-driven mechanism, part of the transfer of power from the retailer to the supplier hands of the hands of retailers and suppliers to achieve the sharing of power, this power sharing behind is the long-term interests of both sides depend on each other, interact to form pressure. Currently, VMI

Y. Yang, M. Ma, and B. Liu (Eds.): ICICA 2013, Part II, CCIS 392, pp. 272–281, 2013.

model in medicine, books and other widely used in everyday products. Research on VMI model has been a lot of research into domestic and foreign. Lee et al and Chen et al [4] demonstrated the system at VMI mode between enterprises within the supply chain nodes VMI information systems through information sharing, can reduce the process of information transmission links caused by too much information distortion and bullwhip effect. Waller and other demonstrated to improve the utilization efficiency of VMI mode helps to reduce production costs and inventory costs [5]. WENG solve a joint order policy, optimal pricing and quantity discounts problems [6]. Yan Dong and Kefeng Xu VMI model for the implementation of the economic effects of modeling, analysis and evaluation [7]. Wu Wenxiang in the supply chain to create a simple model based on the discussion of optimal order quantity and short-term, long-term costs, profit changes [8].

This paper analyzes the transportation services transferred between suppliers and retailers supply system obtained when the total profit (costs) and the mode of implementation of VMI supply chain profit (cost) of each node changes and corporate profits (cost) changes, demonstrates the implementation of VMI supply chain model brings improvements in the overall economic effect.

2 Slotting Optimization Manager

Slotting Optimization Manager is used to determine the proper regulation of each product storage methods, storage methods in the appropriate storage space under the bit allocation [12-15]. Slotting Optimization Management pursue different types of equipment and shelves characteristics, goods group, cargo space planning, labor costs and other factors in order to achieve the best built cargo space layout, can effectively control product changes that will maximize cost savings. Slotting Optimization warehouse management as being operational and cost efficiency of mining and the construction of a distribution center or warehouse provides key management operations before preparing. Since few of Warehouse Management System (WMS) and computer systems to support optimal management of storage spaces, so currently about 80% of the distribution center or warehouse can not be properly Slotting Optimization [16-20]. The reason mainly lies in the basis of insufficient data, MIS resources is not yet supported, without the right Slotting Optimization software and methods. Actual demand for modern warehouse, certain of the strength of the professional software company has successfully developed Slotting Optimization warehouse management software. Software via batch of goods, volume, quality control, unmarketable degrees, day picking volume, the amount of data such as date and out of the analysis and calculation, analysis, combined with a variety of strategies, such as correlation rules (Correlation), complementary law (Complementary), compatibility rules (Compatibility) Slotting Optimization, etc., thus greatly reducing the cost of goods layout [21-25].

Whether it is commercial software or self-developed software applications, Slotting Optimization must have the following basic functions: the built-in or custom attribute goods attribute grouping of the goods ; evaluation recommendations for each item of

the best available cargo space and cargo space ; in Category property changes and adjustments in case the state of facilities ; inefficient state from the warehouse into a functional state optimization of movement required to perform configuration ; dimensional graphical analysis or digital output. Slotting Optimization need for a lot of raw data and information you need to know for each commodity chemicals regulation number, description of goods regulations, material type, storage conditions, shelf life, size, weight, carton number, number of boxes per pallet, and even including customer order information. Once collected, the full original data, the choice of what kind of optimization strategy is particularly important. Survey showed that applying some intuition and methods for granted would be misleading, and even lead to the opposite result. Efficient cargo space optimization strategies can increase throughput, improve the use of labor and reduce work-related injuries, better use of space and reduce product damage [26-31].

Based on the logistics management and third party logistics enterprises operating mode of research, especially following a few suggestions put forward as the conclusion. These conclusions are: (1) In this paper, empirical methods were investigated and developed logistics development in our country 's history and current situation, by comparing the results of the relatively slow development of the logistics industry, logistics costs are significantly too high. Therefore, to meet the challenges of e-commerce, is seriously lagging behind the development of the logistics industry has restricted the development of electronic commerce is one important factor. From the above point of view, TPL in China has great potential for development. (2) as a new logistics agency model TPL different from a simple concepts such as warehousing and transportation, and its essence is to use the information flow to integrate logistics. However, the flow of information to be done really have to rely on at least two integrated logistics premise that powerful information systems and logistics system design science. Only with these two basic, TPL order for small quantity and variety of needs for timely and efficient delivery service, and reduce overall logistics costs. (3) scientific logistics system must be in a networked environment using a variety of modern logistics technology and strategy. TPL used in logistics technology, the automatic sorting technology, automated picking system, warehouse operations and computer intelligence technology is the most important technology ; in support of the above techniques, TPL rapid response in order to effectively implement policies and effective customer management. (4) In the network environment, information systems platform design is directly related to the business process design and real-time monitoring of logistics and other important functions are implemented. TPL logistics information system must meet the " customer-centric " thinking, in order to improve their customer service capabilities. Integration of logistics information flow is reflected in the final results : the possibility of using information systems to effectively manage all logistics operations, logistics and operations provide information and data to ensure that the storage, sorting, picking, transportation and other operational aspects of coordination between consistent, and improve overall economic efficiency of logistics. (5) According to the developed logistics development experience and the specific national conditions, to design the whole package on TPL operation. The program is basically used in practice.

The research work embodied in the following aspects : (1) review and summarize the basic theory about the logistics management, which focuses on analysis of the functions of modern logistics, cost, and logistics and distribution center determines the basic form, and particularly pointed out in the TPL our country has great potential for development. (2) established the integration of logistics information flow analysis of the effect of the theoretical model, using rigorous mathematical deduction is derived information flow integration of logistics has tremendous benefits. (3) a detailed analysis of the network environment, the operation of TPL necessary basic techniques and strategies. Which focus on the major logistics operation based on analysis of the automatic sorting technology areas, warehouse operation technology and other modern logistics technology ; addition, the article also detailed analysis of the rapid response and effective customer management strategies for the use of information flow integration of logistics TPL important contribution. (4) specifically studied in the network environment, logistics and information systems design principles to be observed and the basic functions required to achieve and effects. (5) operate on China's general TPL program carried out a detailed investigation, and on this basis, combined with the company's operations in Iraq were quick case study.

Because of the time, in many ways this need for further study. Specifically, these are mainly the following several aspects : (a) from the entire logistics system view of modern logistics technology and the emergence of new logistics enterprises will undoubtedly reduce the overall cost of logistics, but these were exhumed, "the third profit" in suppliers, logistics companies and retailers is how customers ultimately carve up. (2) TPL use of information flow integration of logistics is the embodiment of its most remarkable features of modern features, but this depends on the ability to integrate the factors which, how the relationship between, and so are worthy of study. (3) TPL is the emergence of new things, its effectiveness has yet to be proven. In particular, how the modern logistics operation techniques and strategies combined with our specific, in order to achieve the maximum economic effect is TPL direction of further research.

3 Implementation Conditions of the Third Party Logistics Based on VMI Mode

3.1 The Parties must Establish an Advanced Information Systems

Distributors via electronic data interchange (EDI) of the point-of-sale information to the suppliers, and suppliers via EDI delivery information sent to distributors, while using bar code technology and scanning technologies to ensure the accuracy of the data. In addition, such as inventory management, planning system, and so must use effective information processing technology. If any party suppliers or distributors without a perfect information system, VMI can not be implemented.

3.2 The Parties Must Establish Strategic Partnership

Both sides must develop a certain degree of trust. The implementation of VMI supplier master distributor sales information, inventory information, and many other

confidential information, distributors must confirm that the information is not understood by competitors, and suppliers must also be recognized distributor inventory reduction as a result of such cooperation and increase effective space will not benefit competitors.

3.3 Required to Obtain the Support of the Top Management of the Enterprise

Due to the confidential information that might otherwise only the top management in order to understand, now have to be shared with suppliers and customers; and this strategic partnership may also have an impact on the internal distribution of power, often in contact with distributors such as the implementation of VMI is no longer a marketing staff but the logistics personnel, this transfer of power may affect the cooperation and coordination between the departments, requiring the participation and support of the top leadership.

3.4 The Parties Must Sign a Contract

The terms of the contract, the parties shall agree. These terms include the time of transfer of ownership of the goods, credit conditions, ordering responsibilities, performance indicators.

4 VMI Mode and Integrated Transport Information

In a typical two-stage supply chain is composed by suppliers and retailers, retailer inventory management retailers responsible for inventory management, and order according to their own strategy to determine the timing and amount of orders to suppliers, the supplier is only responsible for within a certain period of time after receiving the order delivered the goods to the retailer. In contrast, in the VMI mode, suppliers and retailers signed a cooperation agreement by the supplier is responsible for managing the retailer's inventory, retailers share sales information and inventory information to suppliers, vendors according to their sales forecast instead of the retailer and the retailer's inventory levels to place orders and is responsible for the distribution of goods to retailers. VMI and traditional inventory management model differs from Table 1.

After implementation VMI, suppliers can understand the customer's inventory levels and sales information in real time and in a timely manner to develop their own production and transportation programs to meet the needs of customers; this can in large part to prevent the occurrence of out of stock, the customers thus greatly reduce safety stock, even to zero inventories. Therefore, in the implementation of the VMI, the supply and demand double} mode of operation of the party will be changed accordingly. A basic VMI system structure by the demand forecasting and replenishment and distribution modules, as shown in Figure 1.

Table 1. The difference method between VMI and traditional inventory management

Factor	Traditional inventory management	Supplier inventory management
Replenishment decisions	The buyer separate decision-making, suppliers perform	Supplier to a buyer's decision-making
information	Do not share	Sharing
Cooperation	Does not close	Close
Supply chain inventory costs	High	Low
Mutually beneficial	Difficult to achieve a win-win	Win-win
Interactive	Low	High
Agreement	Weak	Strength

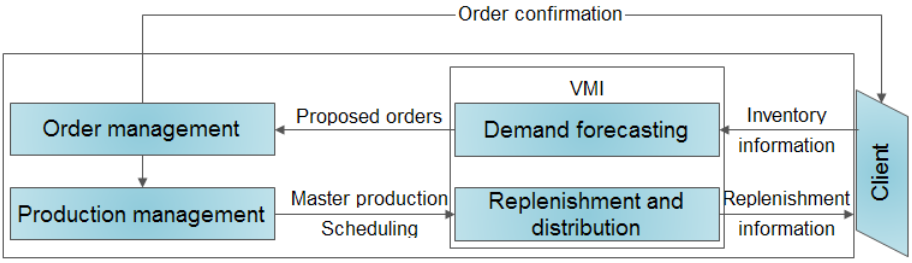


Fig. 1. VMI overall operating structure diagram

5 Integration and Researches on VMI and TPL

In supply chain management, emphasizing the various members of the supply chain should give full play to their own core competitiveness. Suppliers often as a production enterprise, its core competence is the development and production of products, suppliers VMI services are often faced with the problem of lack of logistics, information processing capacity and the low level of customer service, in order to solve their own lack of problems, suppliers are beginning to third-party logistics introduced to the VMI operation. Depending on the logistics operation and management of third-party logistics as the core competitiveness just to meet the logistics needs of the members of the supply chain in the implementation of VMI.

The driving force for the introduction of third-party logistics article VMI implementation strategy grouped into five factors: the environmental driving, speed-driven, efficiency-driven, cost-driven and service-driven. Figure 2 depicts the relationship between the five drivers in the integrated operation of the third-party logistics and VMI. Due to the combined effect of these drivers contributed to the

suppliers and third party logistics cooperation, the two sides working together to enhance the effect of VMI operation.

(l) Environmental driving. As market competition intensifies, the traditional single enterprise competition model has been difficult to enable enterprises to maintain a competitive advantage in the market competition, and information resources increasingly open, breaking the boundaries of the enterprise and the establishment of new cooperation beyond corporate boundaries the emergence of the relationship, supply chain management to meet this trend. In this new situation, the competition among enterprises has gradually evolved into the competition between supply chains. Members of the supply chain at the same time to enhance their core competitiveness, but also ongoing management innovation, VMI mode is in this context that can be widely implemented. At the same time, customers' needs change and improve logistics, third party logistics in order to adapt to market competition must improve logistics operational capability, to expand the scope of services to provide customers with personalized, integrated services, VMI suppliers outsourcing services has become an inevitable.

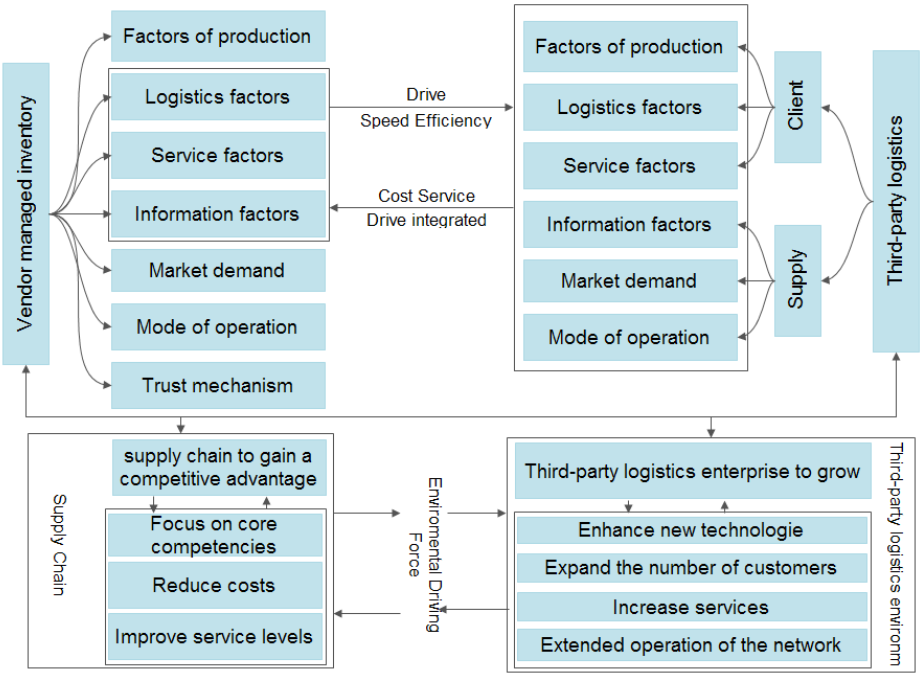


Fig. 2. Integration of third-party logistics and VMI overall operating principle

(2) Speed drive. Time obtain a competitive advantage by improving the reaction speed and operational efficiency of the entire logistics system. As the supplier's logistics capabilities may not be able to meet the requirements of the supply chain to respond quickly. Therefore, VMI operations outsourcing to a professional third-party

logistics service providers, third-party logistics facilities and efficient logistics and operational capacity to achieve the supply chain to respond quickly.

(3) Efficiency drive. In the process of implementation of VMI, the supplier need to invest a lot of money and human resources to enhance their logistics operation, but the effect is not ideal, low operational efficiency but also a waste of resources. Therefore, the VMI implementation process through the introduction of third-party logistics to improve the operational efficiency of the entire supply chain, the rational allocation of resources, so that to maximize the interests of all enterprises in the supply chain.

(4) Cost drivers. Michael Porter, one of three strategies to gain competitive advantage, cost advantage in the fierce market competition, reduce costs and improve profit margins tend to be the preferred target for the pursuit of business. The introduction of third-party logistics VMI operation, on the one hand, providers can reduce operating costs, to focus more on their core business; On the other hand by the third-party logistics enterprises of scale to reduce transport costs in the supply chain operations and inventory costs.

(5) Service driver. Professional social division of labor market demand and personalized service has become a key factor for enterprises to gain competitive advantage; companies must to minimize the cost to meet the expected customer service. Due to the lack of logistics capabilities, difficult for suppliers to provide quality and efficient VMI services to downstream customers, third-party logistics state-of-the-art logistics management experience and professional logistics operation level can help suppliers to get rid of the shackles of the logistics capabilities, enhance customer service levels.

6 The Basic Characteristics of VMI & TPL Integrated Operation

(l) Protocol. VMI & TPL integrated mode of implementation, require suppliers through a specific framework agreement between the third party logistics and retailers to determine the respective responsibilities and obligations, in order to achieve consistency of all enterprises in the concept to improve the operability of the integrated mode.

(2) Cooperation. Mutual trust between suppliers, third-party logistics, as well as retailers is the VMI & TPL integration of the operation of the foundation. Therefore, the three parties need to have a good spirit of cooperation, to establish a close working relationship on this basis, to ensure the successful operation of the supply chain.

(3) Mutually beneficial. The VMI & TPL mode of operation is to consider lowering the cost of the whole supply chain logistics operation on the basis of, the pursuit of the maximization of the interests of the members, in order to achieve a win-win situation by the parties.

(4) customer-oriented. VMI & TPL integrated operation of the main goals is to respond quickly to customer needs, this is also one of the driving factors VMI operation of the introduction of third-party logistics. Customer-oriented and fast response through co-operation among members in the supply chain, that improves customer service levels.

(5) Continuous improvement. In the VMI & TPL functioning mode implementation process, from the perspective of the whole supply chain coordination, and actively take

measures to various problems in the operation of continuous improvement, eliminating waste, the parties to fully share the benefits.

7 Conclusions

The article analyzes the VMI operation of the introduction of third-party logistics driving force on the basis of the study VMI mode, and factors influencing, collaborative forecasting environment VMI & TPL mode of operation, and analyze its operational characteristics. Second, we analyze the information content and features of the VMI & TPL mode on the basis of classification of shared logistics information, and the use of value engineering theory quantitatively investigate the value of logistics information.

References

1. Liu, L.: Composition and structure of the supply chain. Economic Management Press, Beijing (2002)
2. Savelsbergh, M., Jin-Hwa Song: An optimization algorithm for the inventory routing problem with continuous moves. Computers & Operations Research 35(7), 2266–2283 (2008)
3. Mete, H.O., Zabinsky, Z.B.: Stochastic optimization of medical supply location and distribution in disaster management. International Journal of Production Economies (in press, Corrected Proof) (available online November 1, 2009)
4. Javid, A.A., Azad, N.: Incorporating location, routing and inventory decisions in supply chain network design. Transportation Research Part E (in press, Corrected Proof) (available online November 11, 2009)
5. Lee, H., Padmanabhan, P., Whang, S.: Information distortion in a supply chain: the bullwhip effect. Management Science 43(4), 546–558 (1997)
6. Chen, F., Drezner, Z., Ryan, J.K., et al.: Quantifying the bullwhip effect in a simple supply chain:the impact of forecasting, lead times and information. Management Science 46(3), 436–443 (2000)
7. Waller, M., Johnson, M.E., Davis, T.: Vendor-managed inventory in retail supply chain. Journal of Business Logistics 20(1), 183–203 (1999)
8. Weng, Z.K.: Channel coordination and quantity discount. Management Science 41(9), 1509–1522 (1997)
9. Dong, Y., Xu, K.: A supply chain model of vendormanaged inventory. Transportation Research (Part E) 38, 75–95 (2002)
10. Wenxiang, W.: VMI supply chain systems evaluation of the economic effects. Management Review 15(6), 55–58 (2003)
11. Wen-An, L., Zhao degrees: Bullwhip effect of System Dynamics. Southeast University (Philosophy and Social Sciences), S2 (2002)
12. Hu, C., Han, Y.-Q., Bin, W.: Inventory Management Based on System Dynamics Study. Engineering and Engineering Management, 3 (2005)
13. Xin, Z., Shihua, M.: Information Sharing and Collaboration revenue for the two supply chain impact. Journal of Management, 1 (2007)
14. Yan, Z.: High ocean; Introduce a third party logistics VMI model optimization study. Journal of Management (January 2007)

15. Shuqin, C.: Static: Supply chain collaboration and information sharing association studies. Journal of Management, 2 (2007)
16. Weiping, W., Light Yang, Z., Fu, D.: 3-layer supply chain and for VMI short-term effects Game Analysis. Journal of Management, 2 (2007)
17. Yingjun, W., Yajun, G.: Supply chain information flow. Industrial Engineering and Management, 3 (2000)
18. Bin, Z.X., De, Y.: Supply chain system reconfiguration Model. Industrial Engineering and Management, 1 (2002)
19. Ning, W.X., Xinyu, S., Sun, L.-Y.: Supply Chain Information Sharing and Collaboration. Industrial Engineering and Management, 1 (2003)
20. Renhua, Z., Zhicheng, L., Jinlong, Z.: Uncertain environment with a flexible production system. Huazhong University of Technology (Social Sciences), 3 (1995)
21. Chang, Y., Jiang, P., Much, D.: Green Building Supply Chain Operations Reference Model. Sichuan Building Science, 6 (2006)
22. Xiaoming, L.: Business online Management Support System. Anhui Electrical Engineering Professional Technique College, 2 (2004)
23. Peiguang, W., Branch, L.X., Xiaolin: ERP system vendor management best practices. Anhui Electrical Engineering Professional Technique College, 1 (2009)
24. Same, Z.: Enterprises to implement the strategic importance of supply chain management. Anhui Electronic Information Vocational and Technical College, 6 (2007)
25. Yang, L.: Domestic and Land Use Change Overview. Anhui Agricultural Science Bulletin (on a fortnightly), 9 (2010)
26. Meiqin: Computer simulation technology in water resources and environmental issues on the application. Anhui Agricultural Sciences, 4 (2004)
27. Shengjun, Y., Lap, W., Jinchuan, Z., Yifeng, Z.: Radio frequency identification (RFID) technology in the vegetable supply chain Research. Anhui Agricultural Sciences, 6 (2005)
28. Hai, Y., Yamin, T.: Agricultural Industry Chain Management and increase agricultural productivity. Anhui Agricultural Sciences, 12 (2005)
29. Lin, H., Zhongying, L., Jinchuan, Z.: Pig breeding enterprise quality security capabilities Evaluation. Anhui Agricultural Sciences, 7 (2006)
30. Wei, C.W., Tao, B.: Agricultural supply chain partners dynamic evaluation method. Anhui Agricultural Sciences, 26 (2007)

Study on Data Mining in Digital Libraries

Chen Bin

The library of Jianghan University, Wuhan, Hubei, 430056

Abstract. Data mining technology is playing an increasingly important role in information utilizing and extracting. Based on the description of the data mining technology, this paper discusses the application of data mining in the digital libraries from three aspects, describes the necessity of data mining technology in digital libraries, and puts forward a mining system model based on data mining technology in digital libraries.

Keywords: Data Mining, Digital Library, Structure Mining, Content Mining, Usage Log Mining.

1 Introduction

The digital library is a digital information resource which has a variety of media and is rich in content, is a digital information system supported by the introduction of modern high-tech, is the next generation of the Internet information resources management mode, is a kind of service mechanism convenient to provide information for readers [1-5]. It will fundamentally change the present situation that the Internet information is scattered, disorderly and inconvenient to use [6].

At the same time, the significance of the digital library is not only to change the traditional library services, media and methods, it also uses its rich digital resources and network advantages to develop new ideas and tools, increase the functions of digital libraries, and provide the service ways for education, life, and academic research [7-10]. It is predicted that the electronic commerce will become the core services and benefits growth point of digital libraries, thus breaking the manacle of "waiting, relying on, asking for" library funds, which will enable it to obtain the considerable development.

Data mining is a new technology and method of calculation. It has been widely used in scientific discovery, commercial retail and credit management, medical science and other fields, and shows great power [11, 12]. This paper tries to clarify that data mining has broad application of space in the digital library information construction, optimization of automatic information processing, information service quality and the business development, and will show the theurgy in the digital library automation, networking, and intelligent direction process.

2 Overview of Data Mining

Data mining (referred to as DM) is a hotspot in computer science research in recent 10 years [13-16]. It refers to the extraction or mining data or knowledge from a large

Y. Yang, M. Ma, and B. Liu (Eds.): ICICA 2013, Part II, CCIS 392, pp. 282–291, 2013.

number of information. Data mining can be performed in any type of information carrier or storage, such as data warehouse, relational database, transactional databases, object-oriented database, object relation database, spatial database, temporal database, text database, multimedia database, Web database, etc. This new technique is proposed to predict the information hidden from the large-scale database or data warehouse in data, mine the potential modes, automatically predict knowledge and behavior, and automatically discover previously unknown patterns. The knowledge data mining extracts can be expressed as Concepts, Rule, and Pattern, Constraints, and Visualization etc. The data mining process can be divided into 3 stages: data preparation, mining operation, expression and interpretation of the results. The mining process is repeated concise, which cannot do without the participation of the users. Data mining makes the mining of the large data in large databases become easier, and the mining personnel do not need to go through years of training of the statistical or data analysis. There exist certain differences and relations between data mining and knowledge discovery. Generally, data mining is a particular step in the process of knowledge discovery. It uses special algorithms to extract models from the database, and then through the system of interpretation and evaluation module, convert the mode into the knowledge users can understand. However, the generalized data mining is often considered the knowledge discovery process consisting of data preparation, model extraction, knowledge representation and a series of steps. Data mining and information retrieval have certain differences, which main show: information retrieval is driven by goals, and the users need to explicitly put forward the query request, but data mining is random, the result of which is unpredictable. The objective of information retrieval is to help users to find requested documents satisfying the query from a large number of documents, and data mining is to discover the implicit knowledge in the documents [17-21]. But the two are complementary to each other. We can use the research achievements of data mining to improve information retrieval's precision and efficiency and improve the expression of search results, so that the information retrieval system is developed to a new level. OLAP is an analysis tool, but the two have obvious difference. On-line analytical processing technology is not an automatic process. When the user puts forward questions, online analytical processing is responsible for extracting detailed information about this issue. The process of data mining is usually automatic, and the user does not need to present the exact problem, and this is more conducive to discover the hidden knowledge.

Data mining is a completed process which mines unknown, effective and practical information from large databases. Data mining process includes four steps, identification area objects, data preparation, mining process and results expression and analysis, which is shown in Fig. 1. The above four steps is not a linear, and in the actual operation only continuously repeat we can obtain good results. The explanation of the four steps is as follows.

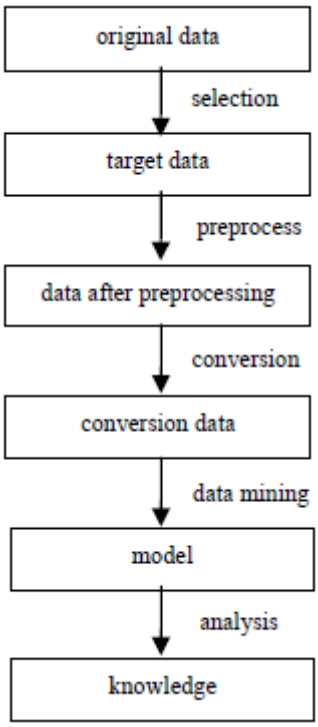


Fig. 1. Data mining process

Currently, more research at home and abroad, and has been used in large-scale investment is MyLibrary system, which is suitable for interactive library service model of personalized service system. It is based on personal interests and learning needs of individual users through the Web pages the way to collect, organize and link individual areas of concern to digital resources, when users visit again MyLibrary, they can get the latest content related to this. This service works, first by professional librarians and library online digital resources or resource type by subject theme for the user to create a resource list, followed by the system to provide users with a login account, the user account logged on from the library website can provide a list of resources needed to select their own resources and other Web resources, adding MyLibrary; addition, MyLibrary also provides the latest communications, personal folders and bookmark functions and other services, users can also customize the services they need items [22-25]. The system is based on the user's registration information and customized content for each user to create a policy file, including the user's personal information and digital resources on behalf of the user to select a list of parameters. This file is stored in the form in Cookie user's computer hard drive or server database, when a user visits a page MyLibrary is extracted, and according to the latest returns customize page content.

Compare home and abroad have been put into use MyLibrary, found abroad in this regard started earlier, the basic realization of system resources customization, page

customization, the latest resources Bulletin, and the Library OPAC interface, personal library management, personal link collection, library catalog and Patron query functions [26-30]; domestic personalized service system is relatively backward lot, some small and medium systems are mostly based on LAS Ⅱ development, can only provide bibliographic queries, Journal Content Search, online booking, online renewal, briefing books, ILL and other simple services, but is limited to school readers, external users can not enter.

Classification model is mainly used to extract important data describe the model, and the model will be based on data from a database into different classes. We can use it for user classification research, including the user's classification, user attributes and characteristics of classification, customer satisfaction analysis. For example, according to the user's personal information and they are customized and personalized service, we can use decision trees, rule induction, neural network classification methods to find out what features users have customized a certain kind of services will, the proportion of such users how much. Its conclusions not only as a basis for proactive push service, expanding the potential user base, but also according to this part of the person's qualifications, professional background determines the appropriate service strategy, for example, if this part of the user base has a certain income, and for online database enduring needs or interest, we can consider them as fixed service object, when buying online database they take policy tilt, database updates regularly send them a message, but also on their questionnaires as improve service basis.

Association model is to describe the data in the database entry the relationship that exists between the rule that one thing appears certain items and some items can be exported in which they appear in the same things that is hidden in the things of the association. For example, we found through association rules, if a certain kind of customized service, then he is likely to be another customized services; Or, link to a network resource links on the majority of people will go to another network resource, then we can put together two services or resources recommended to the user, you can also find out the reasons, in order to develop appropriate service strategy.

Model and serial correlation model is very similar, except that the object is a time series model of the domain distribution, discovery rules and relevant chronological order. For example, in a B / S mode realization of personalized service system, we can use Cookies to record the user has visited resources, which may be found if the user visited links A, then B within 30 min access link may be 48% this is the series model discovery rules. If the conclusion is this: if the user visited A link, then the link to access B may be 48%, which is the association rules.

Clustering is a group into several categories, as similar as possible to make similar individuals without similar between individual differences as large as possible. That is to identify a set of clustering rules, the data is divided into several categories, to achieve " Like attracts like." The difference is that with the classification, clustering from an unknown start, did not know the specific classification criteria, do not know what kind, according to a given cluster parameters (such as similarity, etc.) to decompose, merge. The results obtained by an expert screening, if not met the target, clustering parameters need to be changed, re- clustering. Once target classification rules will get through the clustering parameter. Here, we can cluster a particular user group, find the largest

possible group of similarity, they take a different approach to service and service strategy, and based on user feedback, validate clustering results and service satisfaction.

Regression using historical data attributes to predict future trends. The simplest case, using standard statistical methods, such as linear regression. But in fact the majority of nonlinear problems, such as web access volume changes, the database traffic increased and decreased, due to many factors, the problem becomes complicated. The task is to find the regression model for the accurate description of these changes. Some of these technologies can be used for classification can be used for regression, such as CART (Classifiedon And Regression).

Time series model with existing data sequence to predict the future. From this point of view, with the regression model is very similar, but the regression model does not emphasize the order between data and time series models to consider the temporal characteristics, in particular, the level of the time period to be considered, such as day, week, month, year, etc. sometimes you have to consider the impact of the calendar, such as holidays and so on. For example, for online access to the database will have a certain amount of periodic variation, we can find out the law (or even find out the reasons, such as changes in semester or holidays, etc.), so in a timely manner to adjust focus. Here we must note that the time series models and time series data mining are two different concepts. For time series data mining can be used a variety of models and methods.

We do not really grasp the needs of users and put it accurately represented. User information needs and expression is a complex issue, the British scholar Taylor 's information needs theory classic. He believes that some of the potential user's needs, while others are dominant. From the degree of generalization can be divided into three levels: the common needs of various users, the same part of the user 's needs, the user's individual needs. Here we mainly discuss how to use data mining techniques to extract a certain commonality of demand, thus pushing it as a service to carry out group, choose to buy an electronic database, adjusting service strategy etc. reference standards. Data mining refers to the large amounts of data to identify potential and valuable relationship between model and data (knowledge) of the non-trivial process, which uses machine learning, neural networks, statistical and other methods to extract the rules easy to be understood, to guide business conduct or assist scientific research. Here we were based on the kind of data mining in a digital library for their application to be analyzed and compared.

Digital Library longer period of time in the future is bound to develop rapidly. For existing technologies: portals, from the development of distributed artificial intelligence and distributed systems to multi-agent systems, grid, Which one better promote the development of libraries, we should behold a give users a truly personalized service environment, truly " what the user needs, I will provide what" works. Data mining technology because of its vast information resources in smart performance, without prior knowledge of the objective expert conclusion, a large choice of mining tools and algorithms, personalized digital library services become an integral part of building technical support.

3 The Function of Data Mining in Digital Library's Information Service

3.1 Improve the Speed of Obtaining Information

In order to guarantee the user access to the required information in the response time as short as possible, we can collect the users' special collections (browse mode) as a transaction, record all users' browsing process each time to form the transaction database, and then carry out the following operations on the database: First, use the association rule mining algorithm to find the project (project) set whose access frequency exceeds a given threshold, and then do similarity matching of customer's browsing patterns and frequent items using the classification algorithm, organize the customers who have similar browsing patterns to a server, thereby reducing the number of server cache and transmission pages. Second, find the feature sets in transactional database which is accessed frequently, and store them in the server's knowledge base using the association rules by the association analysis. When users browse a page, the network agent previously connect the associated page according to the rules, so as to improve the response speed. Third, we also can use Web mining to get the sequence pattern of the user's access, and previously transmit the pages users may read according to the prediction.

3.2 Upgrade the Quality of Traditional Information Services of Libraries

As the network develops to the intellectualized direction, information services should also be "intelligentialized", that is, from passive service to active service, from simple information presentation to information generation. The first is the information retrieval service. Traditional retrieval tools lack friendly, understandable and interactivity of the results, and often show a lot of query results in a linear, which makes the users at a loss. Intelligent information retrieval not only support the concept retrieval, fuzzy search, associative retrieval and multilingual retrieval, but also can quickly use the clustering algorithm to cluster analyze the query results, to make them principled and convenient for users' selection, and at the same time determine the further search on this basis. For example, the system prompts "Limited your research to result within one of the following categories", in order to achieve interactive retrieval. The second is the SDI and a novelty searching service. The traditional way of these two kinds of information services commonly used for scientific research is the document retrieval or CD-ROM database. However, in the network era, we cannot ignore the external network, the instant convenient platform for the dissemination of information search, to ensure the credibility of the novelty search results. At the same time, the digital library can use the interest pattern algorithm to judge the potential users. In the service process, we can also use the visualization technique to help users to carry out online real-time information analysis. The third is the information analysis service. Some experts pointed out that "exploratory data analysis of the text (EDA) is the real text mining." It is to directly analyze the text data and the relationships between them,

in order to identify the unknown and useful knowledge of the process. For example, use the text EDA to form some scientific hypothesis. Based on this, the digital library can provide analysis for business users, help them determine the development and competitive strategies, and realize their own income.

3.3 Support the Decisions of the Library

The low level of management is one of the important factors affecting the development of libraries in China. Enhancing the management level depends largely on the decision which is scientific or not. The traditional way of Library Decision mostly relies on the experience to make decisions, which is subjective, one-sided, blind and other problems, and cannot meet the requirements of the times. The data mining technology can provide strong guarantee for the scientific decision of the leadership. Firstly, data mining can collect the internal data and the external information together which is related to the information system of the library, and after processing and conversion, form the centralized and unified and readily available information, to prevent wrong decision due to insufficient information. Secondly, use the data warehouse system OLAP tool to compare to the integration of multidimensional data, review and verify the assumption of the policy decision, to improve the feasibility and reliability of the decision, and make reasonable use of the limited funds, and optimize the allocation of resources to the library. Thirdly, the tools of data mining can find out a latent pattern from the historical data, and automatic forecast on the basis of the model, which is of great significance for the inspiration of the innovative thinking of the library decision makers and the challenge of the information society.

3.4 Support the Library Work

The first is the interview department. Whether the interview department can play the functions affects the utilization ratio of capital and resource levels. How to use the limited funds to buy high quality books and guarantee the scientificalness and rationality of the information resource system of the library is the priority among priorities in the library work. Therefore, accurately positioning the readers demand becomes an important factor for improving the utilization rate of resources. For example, a normal university library should not only meet the special class of lower level readers, but also should meet the demand of doctoral students and instructors, and should also highlight the characteristics of normal universities, in order to solve the needs of different levels of objects. The use of data mining clustering and deviation detection technology, and the method of the readers' questionnaire can provide the fundamental basis for the clustering analysis, analyze the survey results, and understand the needs of readers of the books of all kinds of degrees, and timely feedback to the statistical results of interview department, thereby enhancing the pertinence of the procurement. The second is the circulation. The damaged books due to the side frame as well as fast and frequent circulation is also worthy of our mining. Use the methods of data mining association analysis to analyze the past borrowing data. The books of the large increase should be reserved the position in accordance with the

forecast trend on time. Those who are borrowed frequently and those which are damaged as well as those many readers renew should feed back to the interview department to increase purchasing power in the quantitative way.

3.5 Personalized Information Service

The traditional information service of the library is that the librarians answer users of the information request one by one, and the library provides regular information which is engaged in the research of the users. The response of this kind of passive service mode is slow, which cannot adapt to the speed of current development of knowledge update. The application of data mining technology makes the information service of the library active steering by passive, greatly enhance the efficiency of information services, and the library has become an intelligent, initiative place to provide information. Use the data mining technology on the user's browsing database mining to find the users' usage patterns, and provide personalized service initiatively according to the user's interest model, including:

(1) When we find the new information or bibliographic data, we should timely inform the users;

(2) When a user visits, according to the interest of users, promote related information;

(3) Track the mode of user's interest change, and find the new needs of users;

(4) According to the interests of users, provide the corresponding forecast report, dynamic analysis and so on.

These functions are reflected in the target user's logging home page in the digital library, or integrated into the digital campus' portal, as long as the user customizes the corresponding services, according to the result of data mining, the system will remind the user to view the latest information on the personal home page in E-mail or mobile phone text messages and other forms, and can help the user personalize information space.

3.6 Text Mining Is to Realize the Automation of Information Processing

From the development trend of the digital library, its resources will cover paper documents, electronic books, electronic journals, CD databases, multimedia data and Web data, and each kind of data is in a rapid growth. So the automation requirements of various types of information processing are increasingly urgent. According to the characteristics that the data of Digital Library occupy the vast majority, this paper mainly discusses the application of data mining technology in the automatic processing of text data. The image, audio, animation and other information in the digital library system, due to the limited ability of the author, will not be involved.

4 Conclusion

Today, the digital library begins to take shape and increasingly highlights its existing value. Data mining provides technical support for the organization and management of

the digital resources, the promotion of the expansion of the service quality and service mode, and at the same time, makes the data mining technology and the method research develop both in scope and in depth. However, in addition to the massive text data, the digital library should have a large number of multimedia data. How to develop the technology of data mining still remains issues needing in-depth study to enable it to process the text feature of super-high dimension, accurately describe the characteristics of multimedia information, correctly map the implementation of the concept space, and improve processing speed of all kinds of information and the precision.Digital Library longer period of time in the future is bound to develop rapidly. For existing technologies: portals, from the development of distributed artificial intelligence and distributed systems to multi-agent systems, grid, Which one better promote the development of libraries, we should behold a give users a truly personalized service environment, truly " what the user needs, I will provide what" works. Data mining technology because of its vast information resources in smart performance, without prior knowledge of the objective expert conclusion, a large choice of mining tools and algorithms, personalized digital library services become an integral part of building technical support.

References

1. Xiangyang, Z.: Data mining technology and information services in urban construction archives Applications. Archival Science Communication, 3 (2005)
2. Mo, L.: Web-based data mining technology in digital library application. University Library and Information Science, 4 (2007)
3. Qian, H.: College Library search tools use. Commerce, 8 (2010)
4. Liu, L.: Data mining technology in the “Police Technology Information” database application. Chinese People’s Public Security University (Natural Science), 1 (2006)
5. Wan, J.: Library under the Network Environment User knowledge mining. College Library and Information Forum, 3 (2005)
6. Xiaolan, F.: Data mining application in digital libraries. Heihe Sciences, 2 (2011)
7. Wide, L., first Lee, W.: Data mining applications in Digital Libraries Research, p. 2. New Century Library (2012)
8. Zhihong, S.: Data mining - the Internet age of information services essential skills. Information Development & Economy, 10 (2007)
9. Jinmei: Based on data mining techniques Library. Information Development & Economy, 36 (2008)
10. Hongxia, Y.: Based on data mining technology library management. Information Development & Economy, 12 (2011)
11. In the light, Jing, W.: University Library under the Network Environment Information Service. National Library Science, 4 (2003)
12. Zhao, C.-Y.: Data mining technology and its application in University Library. Beijing College of Political Science, 3 (2007)
13. Quan, Y.G.: Data mining personalized services in the digital library application. Dongguan University of Technology, 3 (2004)
14. Na, C.: Data Mining Research Status and Development Direction. Computer and Information Technology, 1 (2006)

15. Rui, W., Tao, M., Chen, C.: Data Mining and Its Application Status Analysis. Computer Application Technology, 2 (2007)
16. Bo, X.: MPEG-21 framework standards and Intelligent Information Push Technology, television technology, 8 (2002)
17. Xiaodong, L., Liu, D.: Patent overview of data mining. Electronic Technology, S1 (2003)
18. Hui, L., Daiji, H.: Data mining technology development and its application. Gansu Science and Technology, 4 (2006)
19. Shuwan, J.: Library under the Network Environment User knowledge mining. College Library and Information Forum, 3 (2005)
20. Zhao, Z., Yan, F.: Commerce technology research and application of data mining study, p.4. Hunan Economic Management Institute (2001)
21. Branch XI, J., Shun, Y.T.: Web data integration data mining research. Computer Engineering and Design, 8 (2006)
22. Liang, S.J.: Intelligent search engine. Technology Consulting Herald, 15 (2007)
23. The Index System Yang: Data mining technology in the Web Application Discussion. Chinese new technologies and products, 8 (2009)
24. Ju, M., Siu, L., Jie: WEB-based data warehouse and business marketing decision support system design. Market Modernization, 5 (2007)
25. Haili, Y., Emily: Using a data mining methods to solve network congestion problems. Qingdao Technological University, 5 (2005)
26. Zhao, C.-Y.: Data mining technology and its application in University Library. Beijing College of Political Science, 3 (2007)
27. Yu, D., Fuqiang, L.: Web-based data mining system research and design. Electronic Technology, 11 (2008)
28. Shan, Z., Tingting, X.: Web-based data mining in e-commerce application. Xiaogan University, p. 3 (2010)
29. Qin, L.Z.: Web data mining in college construction site in the use of Computer Knowledge and Technology, p. 25 (2011)
30. Yun, S., Han, L., Dong, J., Chen, D.: KDW Summary : Web-based data mining. Computer Engineering (January 2003)

Poultry Disease Data Mining Platform Based on Strategy Pattern

Jia Chen

Chongqing Business Vocational College, Chongqing, 400036 China

Abstract. With the development of the technology of artificial intelligence, statistical analysis and database technology of data mining gradually, many data mining methods are applied to the field of poultry disease research. This paper study the strategy pattern of object oriented software design patterns in data mining application scientific research software platform design and development, and puts forward the design summary of the platform. The paper proposes a data mining study ways of thinking: in order to achieve the disease data encapsulation algorithm package, to achieve a unified interface. Realize different data mining methods in a certain type of disease problems, A certain kind of data mining methods used in different diseases.

Keywords: data mining, poultry disease, statistical analysis.

1 Introduction

Data mining is done by careful analysis of large amounts of data to uncover meaningful new relationships, trends and patterns in the process of its emergence in the late 20th century, 80, is a useful database research in new areas of application value is a cross-discipline, the integration of artificial intelligence, database technology, pattern recognition, machine learning, statistics and data visualization, and other fields of theory and techniques of data mining as a technique, it is in the chasm life cycle stages, requires time and effort to research, development and matures, and eventually accepted by the people [1-6].

Data mining, also known as knowledge discovery in the data, its purpose is to automatically or easily extract the hide or recorded knowledge is represented in the data mode [7-11]. Data mining have applications in the IT, banking, insurance, pharmaceutical and retail industries, as well as biology, astronomy and other scientific fields, and the emergence of a number of commercial data mining software products.

A wide variety of the disease in poultry, aquaculture personnel are mostly non-professional farmers of the junior middle school education. How to suit the medicine to the illness will become research important topic. In addition, case diagnosis; treatment records will have a lot of data processing, research [12-18]. We now need to be designed to achieve flexible and many-to-many relationship between the problems and methods of disease data mining platform: data encapsulation of the disease, the algorithm package to achieve [19-22]: (1) a problem with different

Y. Yang, M. Ma, and B. Liu (Eds.): ICICA 2013, Part II, CCIS 392, pp. 292–301, 2013.

solutions, comparing different method to solve effect; (2) a method is applied to different problems, and to maximize the efficacy of a particular method.

2 Data Mining Concepts and Technical Approach

2.1 Design Patterns and Strategy Pattern

Object-oriented software design pattern is a description of the object class that used to solve a general design problem in a particular scene and communicate with each other, that divided into creational patterns, structural model, and behavior patterns. The strategy pattern is an object behavioral pattern, the core intent is to define a series of algorithms, encapsulate every one of them, and so that they can be used interchangeably [23, 24]. The strategy pattern so that the algorithm can be independent of the changes in the clients that use it.

For use strategy pattern motivation, literature [2] use an example to describe: a document editor, the need to achieve wrap There are many algorithms for a text stream branches would these algorithms hardcoded into the use of their class is desirable for the above problems, according to the idea of policy mode, can define different classes to package wrapping algorithm to avoid an algorithm in this way is called a policy package (Strategy), shown in Figure 1. Which Composition class is responsible for maintaining and updating a text viewer to display text wrap. Wrap strategy than by the Composition class implementation, but by abstract Compositor subclass independently implemented. Compositor various subclasses implement different newline strategy.

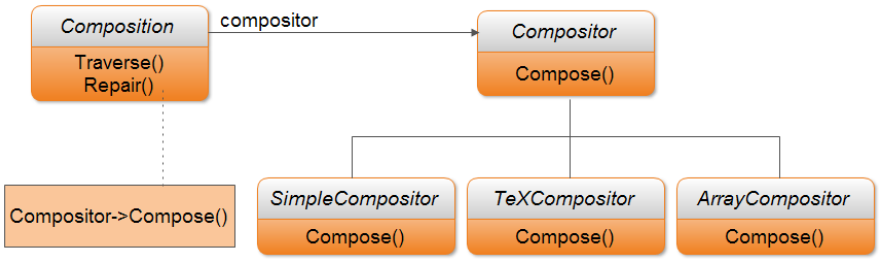


Fig. 1. An instance of strategy

What circumstances is suitable for use strategy pattern; the literature [2] proposed the following:

1 Many related classes just act different.

2 We need to use different algorithm variants.

3 algorithms use data that clients should not know.

4 A class defines a number of actions, and these actions in the operation of this class in the form of a plurality of conditional statements.

2.2 Method of Data Mining

The main methods of data mining include the following six kinds: characterization and distinction, profile association rules, classification and prediction, clustering and outlier analysis, evolution analysis. Each method contains a number of algorithms.

2.3 MVC Framework Based on JSP

As a Web development framework, this server-side architecture is divided into three logical units: the model (M), the view (V), controller (C). Server application is usually divided into the business logic, presentation and request processing. Model corresponding business logic and data, view corresponding controller corresponds to request processing. View JSP pages, Java developers, the controller uses the Servlet class model and controller. The server uses TOMCAT.

3 Design and Analysis of Platform Module

3.1 Data Mining Platform Needs Analysis

This platform is designed and developed disease data mining based on the actual needs, the main requirements are as follows:

(1)Data entry and data preprocess. A wide range of disease data, include structured and semi-structured data, different issues we take different methods. Different methods were used to solve different problems; this platform is required to achieve the entry and logical data structure package.

(2) Flexible data mining method call. The background components need data mining package of each method, unified database interface problem (data) to many-to-many call the relationship between the methods.

(3) The effect of display interface. This platform is required to achieve the computing visualization effects to show users a data mining method to solve the problem of a poultry disease. Differences due to disease problems, data mining methods organizational clarity, so the main methods of data mining organization take GUI interface. Interface includes content, text, icons.

3.2 Data Mining Platform Design and Implementation

In order to achieve data display, data processing, the separation of the database to read and write, give full play to the MVC architecture, the function module (Figure 1) are as follows:

(1) The data entry module.. Workflows: A. users sign up for a new database table, and its registration properties, then standardized coding input attributes to an XML file as a database buffer file; B. semi-formatted data into the database by the user according

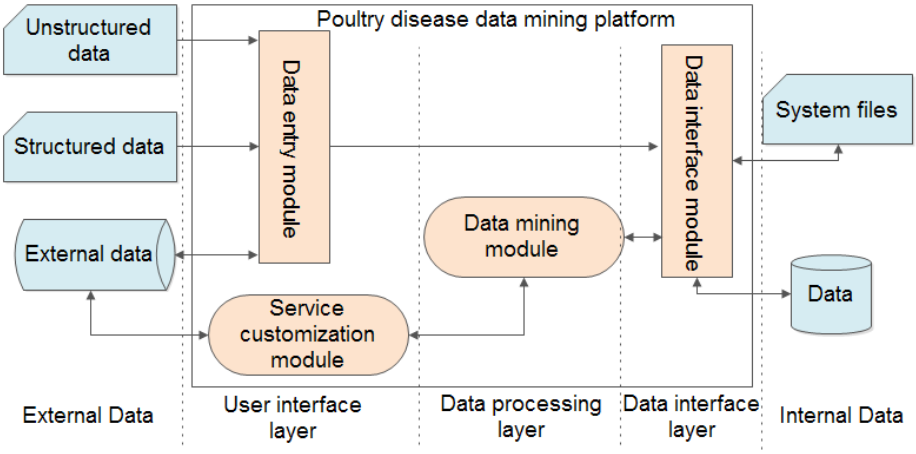


Fig. 2. Platform structure

to The standardized coding indexing for structured data entry, stored in the database. Achieve a unified logical model of the different data package. Operations are involved in any of the database to read and write data interface module is called. The relationship between the data classes is shown in Figure 2: the first layer is an abstract class, the lowest level for the interface, achieved with the Figure 3 DataOpInterface compatible. Each data class inherits from the upper abstract parent class, and achieves Data Interface. This module works in the user interface layer; need to interact with the user and the JSP page as the front display interface, Servlet class background call.

(2) Data mining module. The background can be registered by a series of add components. These components provide a compatible interface for module 1, the encapsulated data, and organizational form shown in Figure 3. The first layer, the second layer is an abstract class. The bottom interface, achieve compatibility with Figure 2 Data Interface. Workflow: the user selected a data service customization module, a data mining method to start the data mining module. The module call data interface module reads the data processing, results delivery service custom module. The various data mining methods class inherits from the upper abstract parent class to achieve DataOpInterface interface. This module works in data processing layer, these components are developed using the Java language does not directly interact with the user, the Servlet class background call.

(3) Services customized modules. To achieve two functions: A. providing a second layer 2, layer 2 (in Figure 3, as well as the following algorithm layer) corresponding to the user to select a function, i.e. an already registered data types and data mining methods have already been registered The select list; B. receiving data mining module results obtained and displayed to the user. Since each of the data mining method processing object from different angles, therefore the results shows different forms, and each method of Figure 3 in the second layer corresponding to this module in a display assembly. This module works in the user interface layer; need to interact with the user and the JSP page as the front display interface, Servlet class background call.

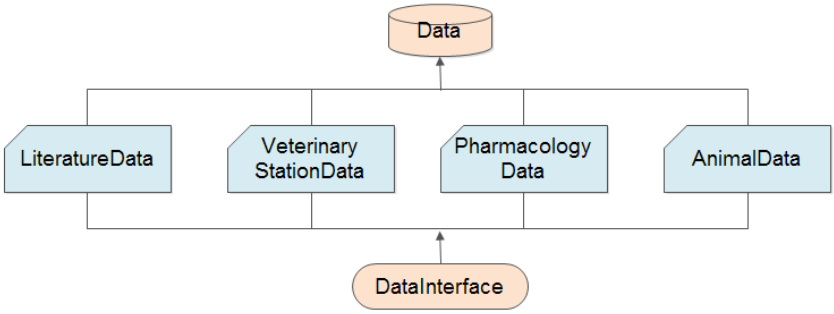


Fig. 3. Data diagram

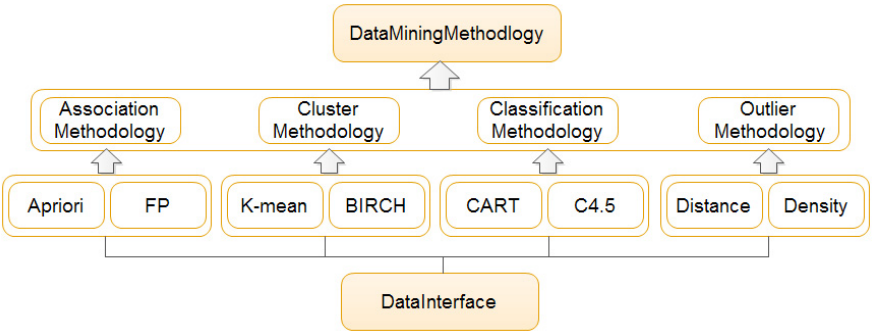


Fig. 4. Data mining method class diagram

(4) Data interface module. Responsible for all to read and write database-related operations, data interface layer. JDBC interface, JSP pages, XML files and other databases written to the back-end database to read data from the back-end database.

(5) Additional Information: (1) data entry module that can be directly applied to structured data interface directly into the database through a data interface module. (2) In order to improve the efficiency of data mining, storage mining results, the introduction of the existing knowledge of disease areas, building the knowledge base in the internal database, to facilitate interaction.

4 Platform Capabilities and Evaluation

4.1 Features and Achieve

(1) Data mining is a new method of poultry disease research, specific issues need to try a variety of data mining methods. Question 1: Through literature study a poultry disease pathogenesis and symptoms. We use the association rules or classification solution. We can input attributes including pathogenesis, symptoms (Fig. 3) data, import documents and indexing of items, and then customization of association rules and classification of relevant methods for data mining.

Fig. 5. Classification of symptoms

Disease	Diagnosis	Confidence level
Increase water	Long-term water shortages	82%
	Coccidia infection early, Bursa, Newcastle	90%
Decrease water	temperature too low	92%
	Near-death or water odor	85%
Stool changes	Red indicates coccidia	75%
	White indicates scours	75%
	White thin feces indicates the ventilation	95%
	Uric acid salts metabolic disorder diseases	96%
	Sulfur signal the cecum hepatitis	92%
	Histomoniasis	90%
	yellow-green->Newcastle disease	82%
	Cholera, typhoid, Spirochetes diseases	74%
	white-cell Protozoan diseases	85%
	Drinking too much	70%
	Too much magnesium ion	83%
	Rotavirus infection	96%
	Chronic coccidiosis	88%

Fig. 6. Results display for association rule

(2) poultry diseases is a new field of data mining application, the same kind of data mining method may be applied to different poultry disease. For example, association rules, Question 1, can also be applied to the problem: through the poultry disease treatment medication data research some kind of disease with the frequency of prescription level.

(3) Results visualization of data mining typically includes text, and icons in two parts. Different methods have deferent display form. So it is necessary to apply the policy mode, using a variety of display components in a unified display interface.

4.2 Performance Evaluation

The platform was developed to provide computer-aided tools for poultry disease research. Data mining, poultry disease research focused on the attempt to get the results

of the various mining methods in different poultry disease problems, assessment of their contributions, rather than a specific performance of the algorithm.

This platform is configured as a dual-core CPU clocked at 2.5GHz over, 4GB of memory, Windows XP, Tomcat 5.0, Oracle 9i server can be successfully completed the 103 magnitude the data scale data mining calculate.

Currently, most of the poultry disease data mining research is experimental in nature, the upcoming individual data mining method used in specific poultry disease research. There have been some poultry disease data warehouse, data mining systems, but is still a commercial data mining tools software areas can not be reached between the problem and the method to achieve the effect of many-to-many flexible switching. In this regard, the platform compared to having advantages.

5 Conclusions

This paper discusses the applications for data mining research and proposed design brief. A poultry disease data mining platform based on its strong flexibility advantages: Problem side and end can achieve openness add, delete, update, and test of any method to solve any problems through service customization avoid the interaction shackles between the issues and approaches to meet the actual requirements.

With China's large-scale , intensive poultry industry, poultry industry has become integrated control of communicable diseases to protect their healthy development is one important factor . In implementing the "prevention, prevention is more important than treatment " on the basis of a comprehensive prevention measures , to further improve and perfect the poultry epidemic prevention system, strengthen poultry disease diagnosis , prevention and treatment technology research and promotion of scientific and technological achievements have become the objective needs . In this paper, a computer network database technology for poultry expert system design, combined with epidemiological and clinical pathology diagnostic techniques to diagnose disease and to provide comprehensive control measures laying poultry and related content query for the poultry industry to provide certain poultry disease diagnosis, prevention support and policy guidance to ensure the healthy development of China's poultry industry .

Click the " combination of multi- symptom diagnosis " , the expert system will pop up window multi- symptom combinations diagnosis expert system China has discovered 77 breeder prone to symptoms of the disease is divided into 16 detailed symptom groups, namely common symptom group prone , head and neck symptomatic group , crown beard symptomatic group , eye symptomatic group , skin feathers symptomatic group , group of respiratory symptoms , digestive symptoms group , neuromotor symptoms group , subcutaneous, muscle disease symptom group , liver, spleen , kidney disease becomes symptomatic group , stomach, intestine lesions symptomatic group , heart, brain , neuropathy symptom group , bone , bone marrow lesions symptomatic group , bursa , tubal disease symptom group , lung, trachea dirty symptomatic group and anatomical lesions and other lesions symptomatic group . Symptoms The symptoms of each group under the entry with sick chickens as symptoms pictures you sick at the time of diagnosis based on the symptoms of chicken first click into the corresponding disease group , select the expert system gives you

observe chickens matches a series of symptoms , and then press the " diagnostics" button , the expert system will give the diagnosis, diagnose chickens may be suffering from the disease onset nonaggressive and probability , and the expert system will prompt you , in order to achieve diagnosed you should carefully observe whether there other symptoms , if you are diagnosed with symptoms consistent chickens can be selected again , perform diagnostics again , so again , you'll get expert system gives you accurate results . Meanwhile expert system also provides you the details of the disease , prevention measures and the color spectrum symptoms , they work with you to understand and learn . Detailed operating methods and sample window see below under various combinations of symptoms diagnosis of chicken diseases diagnosed instance operation.

Click the " main symptom diagnosis according to Law" , the expert system will pop up the window according to the main symptom diagnosis , combined with multi-symptom diagnostic method is similar to the expert system has been found in 77 breeders of disease -prone symptoms detailed breakdown for the 16 symptoms groups , each group of symptoms symptoms of entry under the sick chicken with possible symptoms of pictures you sick chickens at the time of diagnosis according to the main symptom of disease into the corresponding group , select the expert system gives consistent with your observations to chickens the main symptoms , the expert system will be given the right window with the main symptoms associated with other symptoms , you need to select at least one related symptoms , the expert system to diagnose , press the " diagnostics" button , the expert system will give the diagnosis, diagnosis chickens may be suffering from the disease onset nonaggressive and probability , and the expert system will prompt you , in order to achieve diagnosed you should carefully observe whether some other symptoms, if you are diagnosed with symptoms consistent chickens can be selected again , perform diagnostics again , so again, you'll get expert system gives you accurate results . Meanwhile expert system also provides you the details of the disease , prevention measures and the color spectrum symptoms , they work with you to understand and learn . Detailed operating methods and sample window see below under the main symptom diagnosis diagnostic chicken diseases instance operation.

You click on " suspicion nonaggressive diagnostic method based on " expert system will pop up the window under suspicion nonaggressive diagnosis expert system has been found in 77 breeders of disease based on risk factors are divided into five groups, namely chickens viral diseases , chicken bacterial infections, parasitic disease of chickens , chicken poisoning disease, nutritional deficiencies and metabolic chicken disease . If you have some experience on the chicken disease diagnosis , chickens might have been initially diagnosed as suffering from a disease , you can use " under suspicion nonaggressive diagnostic " method validation and re- diagnosis. Click on the left you suspect nonaggressive disease group corresponding button, then click on suspicion nonaggressive , expert systems will be given the right window the symptoms associated with the disease , select the symptoms consistent with your observations , and then press the " Diagnostics" button , the expert system will given diagnosis, diagnose chickens may be suffering from the disease onset nonaggressive and probability , and the expert system will prompt you , in order to achieve diagnosed you should carefully observe whether some other symptoms, if you are diagnosed with symptoms consistent with chickens can be selected again , perform diagnostics again ,

so again , you'll get expert system gives you accurate results . Meanwhile expert system also provides you the details of the disease , prevention measures and the color spectrum symptoms , they work with you to understand and learn . Detailed operating methods and sample window see below under suspicion nonaggressive instance chicken diseases diagnosis diagnostics operation .

In the diagnosis results, the user presses take " preventive measures " button to be displayed and printed out the diagnosis of the disease 's details: overview, epidemiology, clinical symptoms, pathological changes , differential diagnosis , prevention and control measures and other information, and the corresponding color symptoms of disease patterns , ease of control observation Huanxu symptoms .

Another of the medical consultation system is configured with animal medical records, prescriptions and a variety of physiological and biochemical indexes registration , print, search functions , so that you understand and diagnose disease and drug sales inquiries and other relevant circumstances, is more applicable to the veterinary clinic diagnostics division .

References

1. Jing, L.S.: Nanping City Disease Control Problems and Countermeasures. In: Fujian Provincial Animal Husbandry and Veterinary Institute 2009 Annual Conference Proceedings (2009)
2. Qiang, G., Yu, L., An, X.: Guangdong Province in the current main problems Disease Control and Countermeasures. In: 21st Century Technological Progress and Social and Economic Development (first volume) (1999)
3. Pan, Y., Zhao, Y., Chen, H.-J., Jiang, Z., Xi, A., Chang, L., Junong, Z.: Mountain pastures sheep rearing management utility supporting technology study. Fourth China Sheep Industry Development Conference Proceedings (2007)
4. Yu, H.: Right scale pig farms feeding and management procedures to build a scientific discussion. In: China Animal Husbandry and Veterinary Institute Pig Disease Prevention and Control of Infectious Diseases Credits will be the Third Symposium (2008)
5. Yuan, Z.: Commercial laying feeding management and feed processing on egg safety implications. In: 2009 China Industry Development Conference Proceedings Laying Hens (2009)
6. Division, J., Yang, D., Chen, G., Lu, Y.: Goose off -season production capacity of feeding and management. In: Third China waterfowl Development Conference Proceedings (2009)
7. Wang, H., Guo, X., Ling, H., Sun, L.: analysis of the security situation in raw milk. In: Chinese Dairy Association 2007 Proceedings (Volume) (2007)
8. Wu, Z., Xie, J., Fan, D.: Lactating dairy management point. In: 2009 Xinhua Cup Greet the 60th Anniversary of Animal Husbandry and Veterinary Science and Technology Proceedings (2009)
9. Tian, L.-M., Guo, X.: On Breastfeeding Fawn feeding and management. In: 2003 National Information Exchange Deer Industry Compilation of Materials (2003)
10. Pan, Y., Yuan, Z., Mao, Chen, Z., Yu, F.: Pigs introduced shortly after the introduction and breeding management. In: 2005 Liaoning Provincial Animal Husbandry and Veterinary Research Institute Pig Annual Conference Proceedings (2005)
11. Du, Y., Zoo, W.: Giant panda breeding management and disease prevention research. Huazhong Agricultural University (2008)

12. Members of the military: Dogs feeding and management of disease prevention and investigation and research. Nanjing Agricultural University (2005)
13. Hu, X.: Piglets after weaning multisystemic wasting syndrome prevention technology research. Nanjing Agricultural University (2005)
14. Min, S.: Dairy farming district business model and rearing management technology research. China Agricultural University (2004)
15. Meng, F.: Hazard Analysis and Critical Control Point (HACCP) system in the pollution-free chicken in the promotion and application of. Sichuan Agricultural University (2004)
16. Liu, C.: Southern DHI system applied research. Hunan Agricultural University (2006)
17. Chao: Cow cell pattern research and management standard formulation. Henan Agricultural University (2009)
18. Wenju, C.: Supplementation on production performance and perinatal Yak Yak Growth. Qinghai University (2009)
19. Tang, H., stable: Yongzhou scale pig farms infected with gastrointestinal parasites Investigation and Prevention Measures. Hunan Agricultural University (2008)
20. Zhao, Y.: Perinatal pre Dietary NFC / RDP levels on dairy nutrient digestibility and blood biochemical indices of. Inner Mongolia Agricultural University (2009)
21. Han, J.-W., Kamber, M.L.: Data mining concepts and techniques second edition. lsevier Pte Ltd. (2007)
22. Gamma, E.: Design patterns elements of reusable object oriented software. Pearson Education (2002)
23. Agrawal, R.: Data mining: Crossing the chasm [R/OL], `http://www.almaden.ibm.com/cs/people/ragrawal/papers/kdd99_chasm.ppt` (November 20, 2002)
24. Inmon, W.H.: Building the data warehouse, 3rd edn. John Wiley &Sons Inc., New York, 2002.31-145

Green Elderly Community and Elderly Apartment Based on Environmental Resource Superiority

Fang Cheng

Urban Planning Department of College of Landscape Architecture of Nanjing Forestry University, Nanjing 210000, China

Abstract. At present, the suburb retirement products with environmental resource superiority are gradually popular because of the heavy urban pollution. The paper analyzes the types of environmental resources and distinctive design methods. Combined with white sand source program in Kunming, the paper practices design methods of green elderly community and elderly apartment based on environmental resource.

Keywords: environmental resources, elderly community, elderly apartment.

1 Introduction

At present, the urban pollution is severe, the traffic is overcrowded, and the environment is noisy [1-5]. Public activity places and training places are deficient, and the living conditions of the elderly is worse and worse. The suburb retirement products are popular with the elderly who have self-care ability, long for free moving and enjoy free life in nature [6-10]. The retirement products take various environmental resources as the advantage including ecologic resources such as forest and water resources. They provide green environment and green infrastructures which can't be achieved in the city for the elderly, and the elderly don't need to commute between the place of residence and workplace in order to work, so the elderly can enjoy the resources. But the complicated natural conditions make it difficult to plan the parcels with environmental resource superiority [11-14]. How to convert the defect into the advantage, how to combine the site conditions and make it become green elderly community and green elderly apartment fitting for the elderly is very difficult.

2 Types of Environmental Resources for Elderly Community and Planning Challenge

General lack of existing residential center for senior citizens, the elderly lack of cultural activities. Both houses, old people lack many residential activities in public space, so many older people feel that life is monotonous, lonely, lonely, there is the feeling of isolation. According to the Shanghai survey shows that 17 % of households in the elderly amateur life monotonous and loneliness have a sense of loss, bad state of mind.

Y. Yang, M. Ma, and B. Liu (Eds.): ICICA 2013, Part II, CCIS 392, pp. 302–312, 2013.

The city is currently conducting a large-scale residential construction, but also in the planning process does not consider supporting facilities for the elderly, which is rapidly increasing aging population are incompatible. Overseas experience shows that other large cities, residents live in a certain scale of the group, should be equipped with a certain degree of scale facilities for the elderly. Currently residential design specification, how much has been achieved to build a primary school population needs to reach the population in need to build a nursery norms, but there is not much to be built within the scope of the provisions of an annual activity center. Just out of high style Lei Feng developers voluntarily, far can not meet the needs of our aging. For example, a district in the development and construction of the Department, the configuration of the old center, but consumers stay a few years later, developers will be elderly activity center to be transferred to temporary use room, the room had to be fixed rental profits. The old man to unite the night also hard floors, sleeping in elderly activity center, day and night shifts guarding " their homes." But because there is no mandatory requirement, and the purchase contract, the developer will not be committed to providing a permanent center for senior citizens, when this stalemate. Lack of medical care for the elderly living and corresponding institutions, elderly care and the lack of appropriate and timely medical services. Some new residential area far away from the urban downtown area, and its own facilities without sound, so the elderly doctor, shopping, very inconvenient, elderly people 's basic right to life can not be guaranteed.

The lack of easy elderly residential outdoor gathering space, the lack of mutual exchanges to promote the conditions for the elderly, but also the lack of promotion for the elderly and non- elderly people interact with each other conditions. Most elderly people do not just want to get economic support, life care, is more important to get spiritual comfort, fear of being left out of society. Elderly living environment requirements is not merely confined to the " beautiful environment ", the favorite gathering place for the elderly is more lively residential lots on both sides of the main road, the old posts from years of work after retiring, and do not want to reduce or even lost and social interaction. Therefore, in elderly residential environmental design aspects, not only to simply focus on the beauty of the environment, but also to create ease of blending the elderly and social life, to facilitate activities for the elderly outdoor space environment.

Residential area convenient for the elderly in their daily lives is not enough thoughtful considerations. For example many residential district on the downstairs level with many, but not on the downhill, making the elderly wheelchair inaccessible suitable residential activity venue.

Should be carried out within a large residential paving and planting a beautiful lawn, shade trees can be missing, great radiant heat in summer, is not conducive to the elderly in the park activities.

Residential district lack of public toilets, the elderly easy urinary urgency, out activities, rather to worry about, so had relatively passive outdoor activities to reduce the time and frequency of their physical and mental health of the elderly is also unfavorable.

In the continuous development of society, the impact of fast-paced modern life, the followed the way of family pension for thousands of years, has been with the children 's work and life pressure increases severely challenge to parents to take care of themselves Xinyou Liuzaishenbian is weak, today's old apartment has been regarded as part of the people preferred to rely on old-age, in people older apartments will this "social pension house" when generating strong demand, the paper design for apartments for the elderly elderly interaction behavior, interaction needs of research, so that the design of apartments for the elderly, especially in the top to meet the basic living conditions, improve the utilization of the environment, the elderly in a more " easy to apply " an environment better suited apartments for the elderly living in the moment feelings of other elderly people exchanges with the surrounding environment and the interaction effects, the ability to find a sense of belonging and family fun of later life. The innovation of this paper is the " interaction design 'concept used in environmental design apartments for the elderly, in order to expand and extend to other aspects of environmental design, not only for environmental design of apartments for the elderly to provide new ideas for design and research methods, more so in the theory and practice of environmental design has become more enriched and improved.

2.1 Forest Resources

There are widespread forest lands in the suburb. The forest lands form natural air anion bar which is suitable for the elderly to make long-term or short-term life and cultivation. But how to plan and construct and design architecture without destroying forest resources, how to perfect the forest fire prevention system and make the corresponding planning layout and architectural design is a planning challenge.

2.2 Water Resources

Water resources include streams, reservoirs, mountain streams and ponds. Water can form good natural landscape, which not only plays a good role of landscape convalescence therapy for some diseases, but also is an essential important resource to construct healing gardens. While making use of water resources, the influence of natural disasters including flood control needs to be considered, security defense measures need to be taken. And water resources need to be applied flexibly combined with architectural design.

2.3 Landscape Resources

Landscape resources mean mountain resources and valley resources. Mountain resources provide unique dwelling environment. But the greatest challenge of design is how to solve site elevation difference, how to make reasonable non-barrier design and provide convenient transportation services and medical care for the elderly. Valley resources integrates with the mountain resources closely. The river valley with the

function of flood discharge is suitable for constructing valley landscape and is not suitable for constructing buildings.

3 Planning Principles of Elderly Community Based on Environmental Resources

The greatest advantage of elderly community based on environmental resources is green and ecological environment which can't be provided by the increasingly crowded city. So how to construct the community which is suitable for behavior characteristics and living features of the elderly on the basis of making use of natural environment is the emphasis of formulating planning principles and planning strategies.

More generally, the planning principles of the elderly community based on environmental resources are as follows.

(1) Natural ecology. Planning layout and landscape construction based on respecting the nature conforms to the taste request of the elderly, natural beauty and plain beauty. Natural forest and natural food are planned and arranged in combination with natural forests. Air farms, air footpaths and Stream Rivers are designed in the central valley, which forms appropriate natural scenery of dynamic and static combination. And the landscape facilities can be used to construct elderly farm to provide natural foods.

Building construction and landscape construction should obey the ecological planning design requirements, maximize the selection of local materials and energy, and create the elderly community in harmony with nature. Perfect ecological infrastructure should be established, and natural energy and the local materials should be used, which realizes low-carbon first.

Safe and convenient. Establishing the safety guarantee system giving priority to slow traffic, establishing traffic control and community management system with the group as the unit, and ensuring the safety of the elderly. Safe footpaths and safe groups are combined with natural landscape of the community to plan and design safe slow roads, which may be the footpaths along the hillside or air corridors connecting with the groups and common service facility. Community management with the group as the unit is good for association and is safer.

Establishing safe and rapid transportation network and social services network, and providing the most convenient living pattern.

Health and beauty. Distributing medical service aid station with the group as the unit, establishing comprehensive hospitals, setting quick calling system and transport system, and guaranteeing the health of the elderly. Health services and sports fitness facilities are allocated in natural air anion bar to ensure the body-building request of the elderly.

4 Building Design Strategies of the Elderly Community Based on Environmental Resources

Internationally, the establishment of the centralized elderly community has been for a long time. In America, the people over 55 years old can buy or rent the aging community. The

aging community which can take the way of purchase is mostly the apartment or residential villas, and the aging community which can take the way of lease is mostly the agedness flat with service facilities.According to the circulation of the elderly, the elderly community can be divided into self-care habitation, assistant habitation and nursing habitation. The elderly community changes evidently in the past 20 years. More and more people realize that the cognition of the requirement of the elderly, social solicitude, individualized nurse, the concern of the friends and the family, and graceful natural rehabilitation environment can guarantee the healthy living needs of the elderly. The core of the aging house based on environmental resource superiority is to solve the residential requirement, rehabilitation demands and medical needs of the elderly with the help of environmental resources.

Firstly, we can combine with the site condition to make land suitability analysis and land value analysis. As the site condition of environmental resource parcel is more complicated, we can use GIS to analyze elevation difference, slope and runoff. Generally speaking, the self-care aging house has low requirement on the site, and can pursue privacy and site characteristics. But assistant and nursing dwelling form has high requirement on the site, so it should be close to the access of the site and selects the section with flat relief for layout.

In order to reduce the influence of building infrastructure on natural resources, we can use the group layout pattern and make flexible complement in housing cluster, residential quarter and residential district according to the level of public service facility. The groups should be cascaded by using safe and flexible balking traffic systems, and diversified and energetic exercise places should be provided for the elderly, which is the important public association space of the elderly.

5 Design Cases of Green Elderly Community and Elderly Apartment Based on Environmental Resource Superiority

5.1 Project Context

The project is located in Qingyun community of subdistrict office of Panlong District in Kunming, Yunnan province, borders on east belt expressway and is in the northern ecological barrier of Kunming. The project is 9.1 kilometers away from the center of Kunming, and is 5 kilometers away from Kunming Changshui Airport. The project includes river valley land and valley land, and is typical suburb elderly community land with environmental resource superiority.

The designed site covers 86.6 hectares. A river valley crosses over the valley land scattered across the north and south of river valley from east to west. The traffic condition in the site is worse, but the vegetation condition is good, as shown in Figure 1.

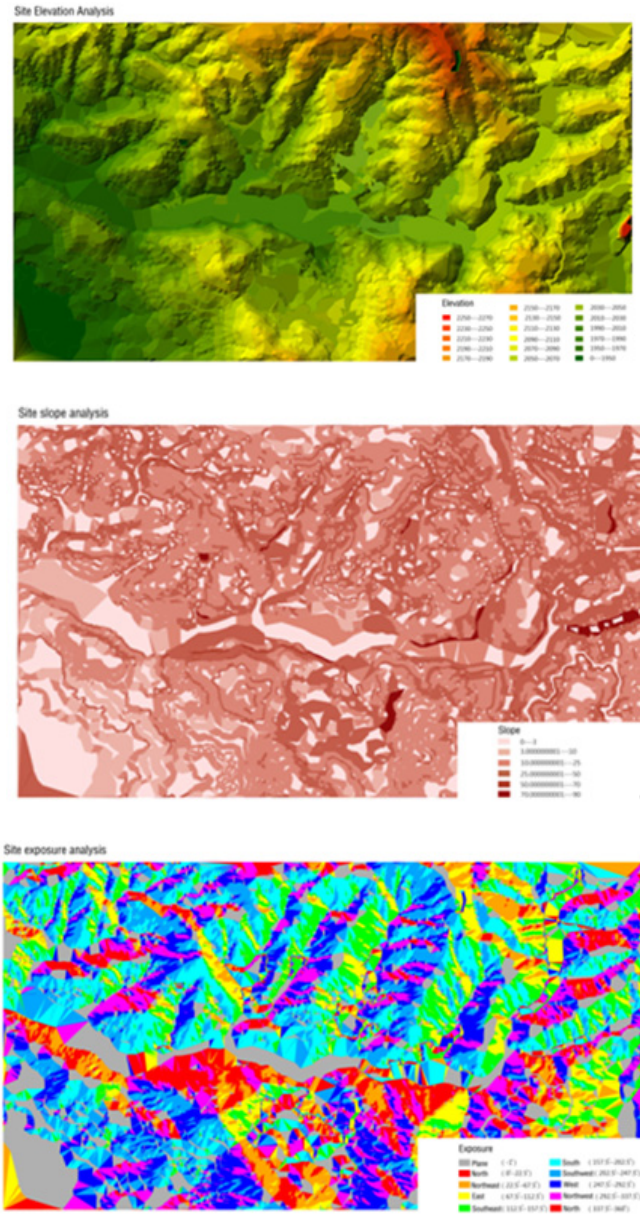


Fig. 1. River valley and land value analysis

(d)

Fig. 1. (*Continued.*)

5.2 Site Analysis

The influence of natural environmental factor and human factors makes the sections with severe terrain changes form in the site including steep and cliff of displacement. GIS software is used to analyze the site. There is about 200 meters of height difference from the west entrance of the site to the east valley. The relief of the river valley is flat, and there is great topography in the joint part of the river valley and mountain, which makes the planning design difficult. The falling gradient and exposure of the site is analyzed, and the levels and scope of the construction land are divided.

The light, transportation, landscape and privacy of the blocks of the site are graded. The river valley with great commonality in the middle of the site is planned for ecological corridor, and the south hillside block with good lighting and landscape is designed for CCRC elderly community. The north hillside block with bad lighting and good resources of forest vegetation is designed for holiday retirement community which meets different old-age support demands.

5.3 Planning Concept

The designed site is located in the valley. A long trench connects the valleys scattered among the mountains. The lands are in the mountains,which makes it quite,beautiful and natural.Long and narrow valleys have phoenix, and the lands scattered in the mountains are like color twill of phoenix. From ancient times to the present, the phoenix means everything goes well, harp and ring. The program combines the site topography and planning design analysis for planning layout and landscape intention design, which tries to construct the distinct landscape elderly community with ecological landscape as the core resource.

5.4 Planning Structure and Function District

The planning structure include ecological corridor, public service and six districts. The construction of ecological corridor reflects the environmental resource superiority in the area. The public service includes arranging general hospital, elderly recreational facilities, supermarket and elderly shopping district. Three groups are three groups of the comprehensive aging community, and include self-care, assistant and nursing aging house.

The site includes six functional areas, scenery cohesive zone, demonstration zone, recuperation health area, holiday endowment area, comprehensive endowment area and public service area.

5.5 Slow Traffic System

The project combines natural scenery to design diversified pedestrian road system, slow landscape road. Long and narrow footpath is from the main entrance, which creates a quiet atmosphere. Air footpath is after reaching the first landscape node, and the line of sight is suddenly enlightened. The footpath solves the linkage of each group and site elevation difference,and can smoothly enter public service area. Waterfront boardwalk has beautiful landscape and fresh air. Perfect footpath connects with open space, which not only provides the best exercise wat for the elderly, but also is the place of association.

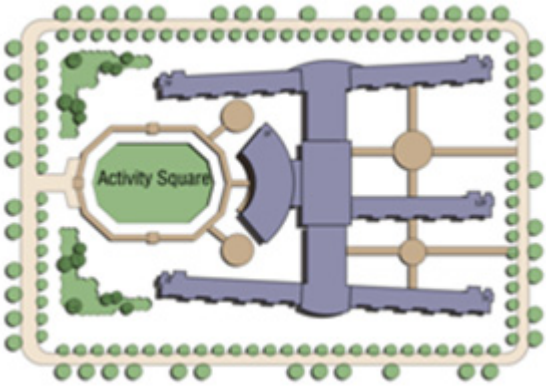
5.6 Architectural Design Guideline

Combined with site conditions, the aging house is distributed with the group as the unit. The buildings in the group are connected with long corridors and surrounds the core activity space. The northern and southern architectures are residential buildings connecting with the northern middle corridors. Public activity space is arranged in the long corridors connecting the architectures, which includes medical services, small supermarket and property management office in the bottom layer, the nursing station, public living room, public restaurant and activity rooms, which satisfies the exchanges demand of the elderly, as shown in Figure 2.

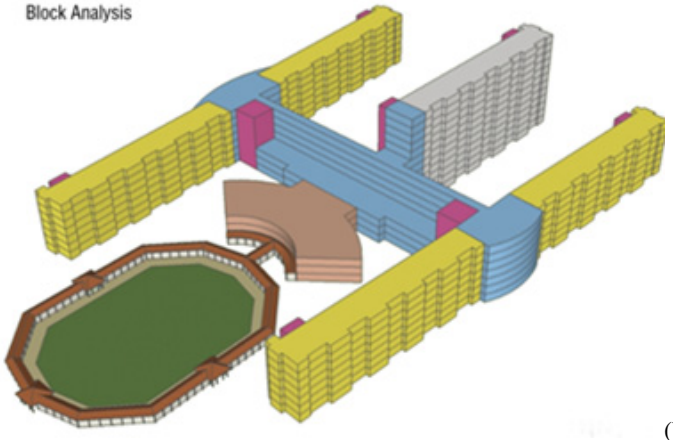
The architecture with the group as the unit can be flexibly combined according to site topography. The products take the transportation as the core, and consider the orientation and landscape for flexible reverse.

While designing the architectures, non-barrier design and the facilities for the physical condition of the elderly need to be noticed such as the anti-slip process at the bottom, the setting of continuous roof, rest seat and message board.

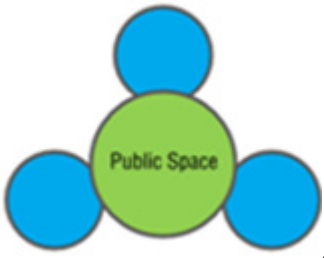
Three-level public service equipment system is setted in the architecture. The first-level service system giving priority to the core area of the entire region includes general hospital, supermarket and elderly commercial district. The second-level public service facility is setted in the entrance and enters of the group and includes community



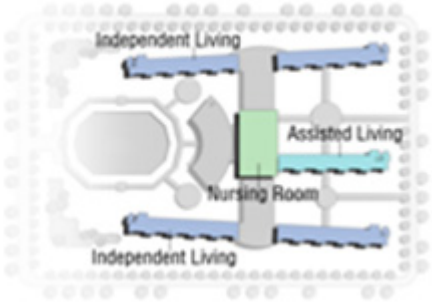
(a)



(b)



(c)



(d)

Fig. 2. Architectural design guideline

medical service and small convenience store. The third-level public service facility is setted on the first floor of each unit and includes nursing station, small restaurant and audiovisual studio.

The above cases practice the design methods of elderly community and elderly building based on environmental resources, which develops a road considering natural health and comfortable convenience for the endowment living mode of the elderly.

6 Conclusion

In the continuous development of society, the impact of fast-paced modern life, the followed the way of family pension for thousands of years, has been with the children 's work and life pressure increases severely challenge to parents to take care of themselves Xinyou Liuzaishenbian is weak, today's old apartment has been regarded as part of the people preferred to rely on old-age, in people older apartments will this " social pension house" when generating strong demand, the paper design for apartments for the elderly elderly interaction behavior, interaction needs of research, so that the design of apartments for the elderly, especially in the top to meet the basic living conditions, improve the utilization of the environment, the elderly in a more " easy to apply " an environment better suited apartments for the elderly living in the moment feelings of other elderly people exchanges with the surrounding environment and the interaction effects, the ability to find a sense of belonging and family fun of later life. The innovation of this paper is the " interaction design 'concept used in environmental design apartments for the elderly, in order to expand and extend to other aspects of environmental design, not only for environmental design of apartments for the elderly to provide new ideas for design and research methods, more so in the theory and practice of environmental design has become more enriched and improved.

References

1. Bingwen, W., Na, Z.: Urban residential landscapes explore the outdoor environment. In: Henan Construction Industry Outstanding Proceedings (2007)
2. Sun, H.: Ocean: Tianjin apartments for the elderly Planning and Construction - About baodi Planning of apartments for the elderly. In: Planning for Innovation: 2010 China Urban Planning Annual Conference Proceedings (2010)
3. Zhang, W.Q., Li, J.: New Era Senior Apartment Design. Civil Engineering Academic Library (Volume 11) (2009)
4. Bing, Baoqing, M.: Nanjing certain apartments for the elderly elderly diet, living habits survey. In: Progress and Elderly Nutrition Elderly Nutrition Feeding Norms Symposium Advances in Medical Nutrition Therapy of Diabetic Nephropathy Classes Compilation (2011)
5. Chen: Zhangjiagang City apartments for the elderly elderly food intake Survey. In: Progress and Elderly Nutrition Elderly Nutrition Feeding Norms Symposium Progress of Diabetic Nephropathy Medical Nutrition Therapy Classes Compilation (2011)
6. Maimex good, Yaqin, Z., Wei, L.: Beilin apartments for the elderly Situation, Problems and Countermeasures. Shaanxi Gerontology Newsletter. 2011 Section 1 (total of 85) (2011)

7. Chao, L., Yixiang, C.: “Old apartment” urban social pension is an important way. In: Gerontology Proceedings (two) (1998)
8. Hongyu, W.: Private apartments for the elderly confused. In: Jiangsu Ageing Research Selected Papers (2000-2004) (2005)
9. Xuan, R., Hong, L., Council: Modern planning theory under the influence of urban settlements Space Environment Analysis. In: Ecological Civilization under the Town and Country Planning - 2008 China Urban Planning Annual Conference Proceedings (2008)
10. Jisheng, Z.: Snoop Macao Environmental Theatre Development. New Drama Study Series (2009)
11. Xia, S., Yao, H., Left, Y.: Ecological community outdoor environment design. Sichuan Environment (June 2004)
12. Liu, L., Xiao, B.: Elderly community outdoor environment Survey - to Beijing Oriental Sun City as an example. Anhui Agricultural Sciences, 27 (2008)
13. Zhou, W.-X., Cassie: Pleasant · · Ye pleasant person - How to Create High -quality residential area outside environment. Huazhong Architecture (May 2003)
14. Lin, L., Sun, Y.: Residential area planning on aging some thinking. South Architecture (March 1998)

Research on Rule of Project Risk Transfer

Li Xiaosong[1,2], Lv Bin[1], Zhan Ming[1], and Han Li[1]

[1] The Third Research Office, China Defense Science and Technology Information Center, Beijing, China
[2] Departments of Equipment and Command, Equipment Academy, Beijing, China

Abstract. Various project risk factors are not isolated from each other. There is a mutual influence among the factors as well as an interaction between the factors and their consequences. Research on the transfer laws between them can get more comprehensive insight in project risks. In this paper, analyzed the interactions between various project risk factors, elaborated 4 kinds of basic project risk transfer structure, discussed horizontal and vertical transfer modes of project risks, built Markov model of project risks transfer, and provided a case to verifying the model.

Keywords: Project Risk, Risk Transfer, Markov Model.

1 Introduction

"Domino effect" suggests that a small initial energy can produce a chain reaction in an interconnected system, which can be extended that a risk factor could aggregates huge risk after a series of risk transfer [1-5].

According to the "butterfly effect" concept, some tiny differences seem to be irrelevant, but they could produce large variations [6-10]. In the project risks, random movement of single risk can not cause great loss to itself, however, "butterfly effect" could appear by risk transfer and therefore cause enormous impact on the goal of project.

So, isolatedly analyze single project risk factor's own changes is unreasonable, we should fully thought about the interaction between various project risk factors and the risk transfer effect, thus project risks can be comprehensively and deeply understood [11-14]. There is apparent interrelationship among project risk factors and interaction between the risk factors and their consequences, the uncertainties and influences caused by some risk factors can amplify those from other risk factors, and then change the value of the risk consequences [15-18]. In the paper, that some project risk factors changed make other risk factors and their risk consequences changed are called project risk transfer [19-23].

2 Project Risk Factors

Project is temporary work to create a unique product, service or achievement. Project has the following basic characteristics:

Y. Yang, M. Ma, and B. Liu (Eds.): ICICA 2013, Part II, CCIS 392, pp. 313–322, 2013.

1. Project development is to achieve a specific target.
2. Project restrict by budget, time and resource.
3. Project is one-time activity.
4. Project is customer-centric.

Project risk, refers to the possibility of occurrence and adverse consequences In the process of project. Because the project involves a wide scale, high technology, long life cycle, spend money and more complex internal structure and external linkages widely, leading to project uncertainty is growing, project risk increased greatly.

Project risk size is determined by the following formula:

$$l = p \times q$$

l express the size of project risk, p express the size of risk occurrence possibility, p express the size of risk adverse consequences.

There are 3 project risk factors, namely "Schedule Risk", "Performance Risk" and "Cost Risk". The schedule risks, are caused by unreasonable schedule plan or other reasons, when those risks occured the project need to postpone the schedule. The performance risks, are caused by factors as insufficiency demand, difficult technique, and unclear feasibility analysis etc., when those risks occured the project could't achieve the technique and operation requirement. The cost risks, are caused by unreasonable cost target or some design technique problems etc., when those risks occured the project should increase cost budget.Figure 1 shows the relationship among various risk factors and the relationship between risk factors and risk consequences. The coefficient of each path is the value of interaction between factors.

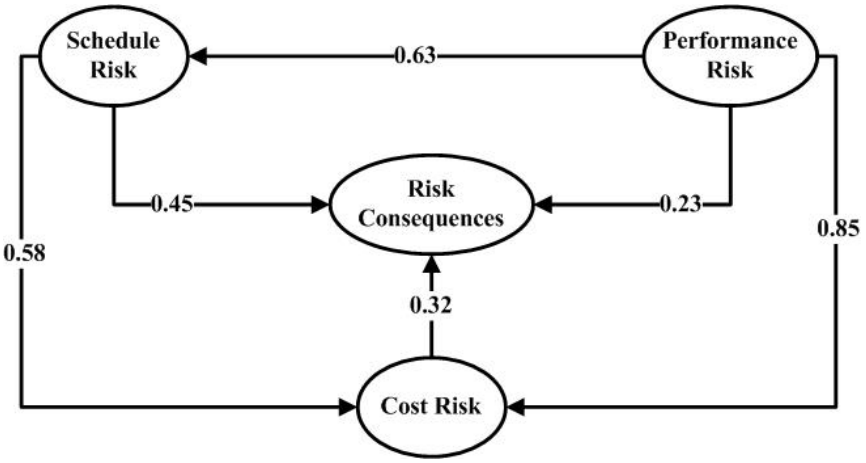


Fig. 1. The Relationship among Various Risk Factors and the Relationship between Risk Factors and Risk Consequences

3 Qualitative Analysis of Project Risk Transfer

3.1 Basic Structure of Project Risk Transfer

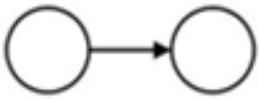
There are 4 main kinds of basic structures of project risk transfer, as shown in Figure 2[1-3].

1. “Serial” Transfer Structure: the structure of risk transfer look like a chain, various risk factors transfer in turn.

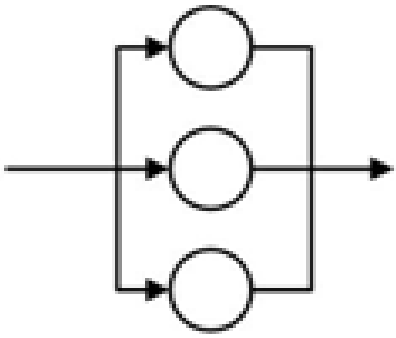
2. “Parallel” Transfer Structure: two or more risk factors transfer in parallel without affecting each other;

3. "And" Transfer Structure: the risk factors transfer like sector, in this structure, when all the risk factors occured at the lower level, the lower risk factors can be transferred to risk factors at the upper level;

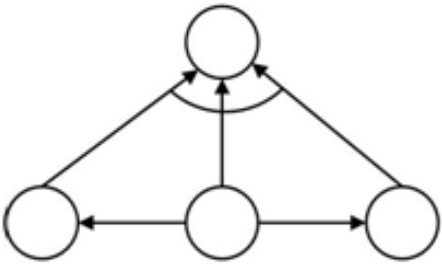
4. "Or" Transfer Structure: the risk factors transfer like sector, in this structure, when one risk occured at the lower level, the lower risk factors can be transferred to risk factors at the upper level.



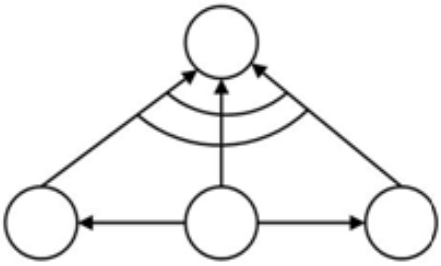
(a) Serical Transfer Structure



(b)Parallel Transfer Structure



(c) “And” Transfer Structure



(d) “Or” Transfer Structure

Fig. 2. 4 Kinds of Project Risk Transfer Structures

3.2 Project Risk Transfer Types

There are 2 main kinds of project risk transfer Types.

3.2.1 The Vertical Project Risk Transfer Type

The vertical project risk transfer type is that the risk factors and risk consequences change with the time by their own laws without being influenced by other risk factors. In short, if no risk control measure is taken, the risk factors and risk consequences will be passed down vertically, and the value of risk won't change. The vertical project risk transfer obeys the rules of serial structure, the risk factors and risk consequences will pass along with time.

3.2.2 The Horizontal Project Risk Transfer Type

The horizontal project risk transfer type means that the interaction among the various risk factors and influence between risk factors and their consequences at a time, the pathway of the horizontal project risk transfer is shown in Figure 1.

1. Horizontal Schedule Risk Transfer

The architecture of horizontal schedule risk transfer is composed of "serial" structure and "or" structure, as shown in Figure 3.

There are 2 main horizontal transfer paths of schedule risk, one is from schedule risk to risk consequences, the other is from schedule risk to cost risk to risk consequences.

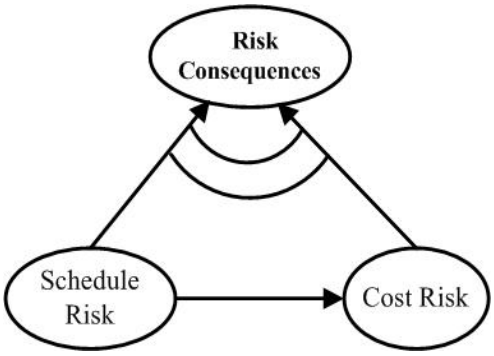


Fig. 3. Horizontal Schedule Risk Transfer

2. Horizontally Performance Risk Transfer

The architecture of horizontally performance risk transfer consists of "serial" structure and "or" structure, as shown in Figure 4.

There are 3 main paths of horizontally performance risk Transfer. The first path is from performance risk to risk consequences, The second path is from performance risk to cost risk to risk consequences, The third path is from performance risk to schedule risk to risk consequences.

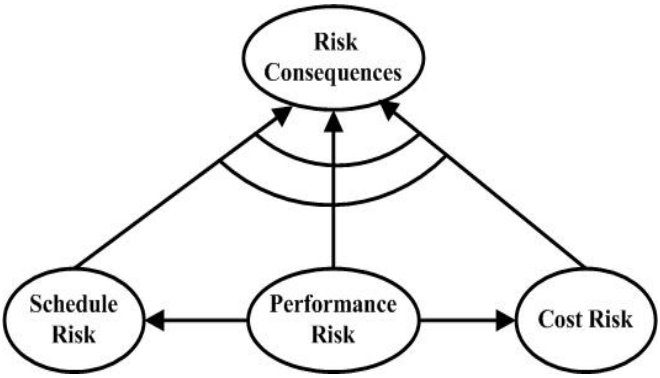


Fig. 4. Horizontally Performance Risk Transfer

3. Horizontal Cost Risk Transfer

The architecture of horizontal cost risk transfer is a kind of serial structure, as shown in Figure 5. The path of horizontal cost risk transfer is from cost risk to risk consequences.

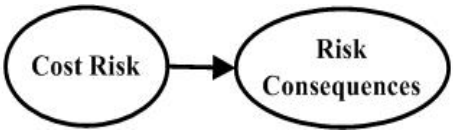


Fig. 5. Horizontally Cost Risk Transfer

4 Quantitative Analysis of Project Risk Transfer

Project risks transfer can be seen as a random variable, from a time series perspective, project risks transfer is a random process [4-10]. Therefore, it can use the theory of stochastic processes to study project risk transfer. The Markov random process model is used in the paper to quantitatively research project risk transfer, the model is described below.

1. Define the state variables as $X(t) = \{1, 2, 3, 4\}$, which are respectively the value of "Schedule Risk", "Performance Risk","Cost Risk" and "Risk Consequences" at t moment.

2. Assume that the state of risk is independent of each other but only related to former state, and managers don't take any effective measures to control risk and let things slide.

3. Let $P_{ij} = P\{X(t) = j \mid X(t_0) = i\}$ $i, j = 1,2,3,4$ indicate the transfer probability which the i risk at t_0 moment transfer to the j risk at t moment, which reflects the dynamic risk change with time.

4. One-step transfer matrix $P^{(1)}$ can be gotten through P_{ij}, so the n-step transfer matrix can be described as follows: $P^{(n)} = (p^{(1)})^n$

5. The research goal of project risk transfer is to know about how will the risk value changes when it starts from a "known state" and follows several risk transfers. So the prediction model of project risk transfer can be described as: $s(n) = s(0) \times P^{(n)}$, $s(0)$ represent as the value of various risk factors and risk consequences at the initial state; $s(n)$: the value of various risk factors and risk consequences after n-steps risk transfer; $P^{(n)}$: n-step transfer matrix.

6. When t tends to be infinity, there is a balance point $s = s(t) = s(t-1)$ where the value of project risks are tends to stabilize.

Supposed that initial value of “schedule risk” is 0.5, “performance risk” is 0.75, “cost risk” is 0.6 and risk consequence is 0, $s(0) = (0.5, 0.75, 0.6, 0)$ The relationship of risk factors and risk consequences can be get based on Figure 1, as shown in Table 1.

Table 1. The transfer relationship of risk factors and risk consequence

	Schedule Risk	Performance Risk	Cost Risk	Risk Consequences
Schedule Risk	1	0	0.58	0.45
Performance Risk	0.63	1	0.85	0.23
Cost Risk	0	0	1	0.32
Risk Consequences	0	0	0	1

The one-step transfer matrix of $P^{(1)}$ can be obtained.

$$P^{(1)} = \begin{matrix} 0.6135 & 0 & 0.2387 & 0.225 \\ 0.3865 & 1 & 0.3498 & 0.115 \\ 0 & 0 & 0.4115 & 0.16 \\ 0 & 0 & 0 & 0.5 \end{matrix}$$

The n-step transfer matrix of $P^{(n)}$ in the equipment acquisition can also be gotten, So the value of risk factors and risk consequences at $t = n$ moment can be calculated by the formula $s(n) = s(0) \times P^{(n)}$. The weekly changed value of various risk factors and risk consequences can be obtained in the 22 weeks, as shown in Figure6 and 7.

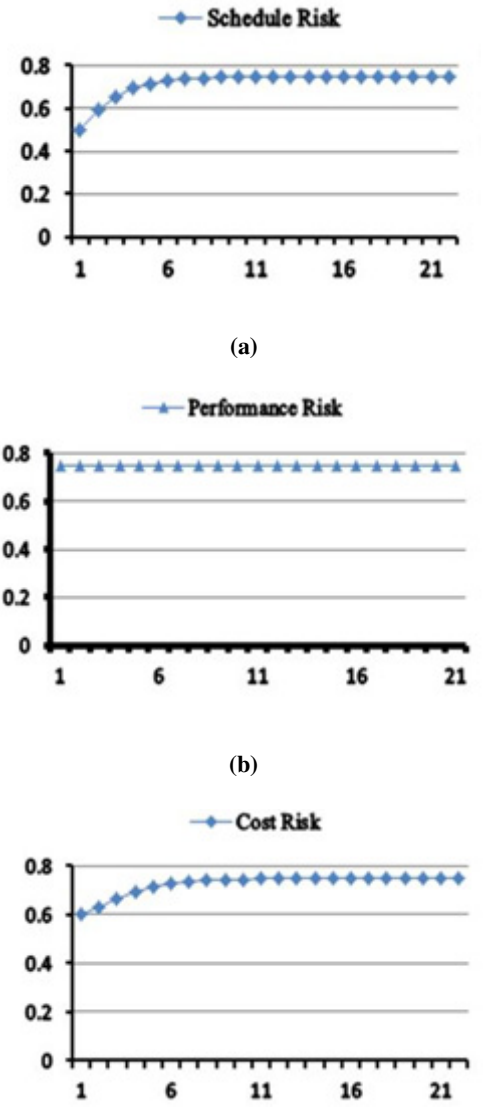


Fig. 6. The changed value of various risk factors

Figure 6 and Figure 7 illustrate that the value of various risk factors and risk consequences change with risk transfers, among them, the value of schedule risk, cost risk and risk consequences tend to increase because of continuing risk transfer from other risk factors, and then tend to stabilize at 21 weeks or so. The value of performance risk unchanged because there aren't impact by other risk factors. Figure 6 and Figure 7

can also indicate that the risk factors and risk consequences will tend to be stabilized when risks have being passed for a period.

The change of risk factors and risk consequences are shown in following :

Before 6 weeks,schedule risk is impact by performance risk, so it rapid growth from 0.5 to 0.75. From 6-11 weeks, schedule risk showing a steady growth from 0.75 to 0.78. After 11 weeks, schedule risk is stabilized.

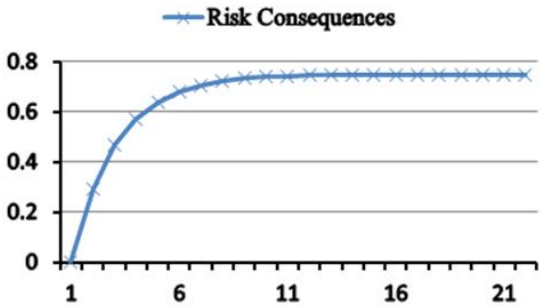


Fig. 7. The changed value of risk consequences

Performance risk was not impact by other risk factors, so it unchanged which remain 0.75.

Before 7 weeks,cost risk is impact by performance risk and schedule risk, so it rapid growth from 0.6 to 0.76. From 7-11 weeks, cost risk showing a steady growth from 0.76 to 0.78. After 11 weeks, cost risk is stabilized.

Before 7 weeks, risk consequences are impact by schedule risk,performance risk and cost risk, so it rapid growth from 0 to 0.62. From 7-11 weeks, Risk consequences showing a steady growth from 0.62 to 0.78. After 11 weeks, Risk consequences are stabilized.

5 Conclusion

This paper analyzed the structures and types of project risk transfer, constructed the Markov model of project risk transfer, simulated dynamic change of project risk transfer. As we know, The project risks are much random, while Markov model is a effective method to study the random process, which only need present dynamic information without a lot of historical data, so it good means to study project risk transfer.

These conclusions are based on the premise that managers don't take any risk control measures when risks occur, in fact, managers always adopt appropriate strategies previously and take some necessary measures to control and mitigate risks, when both the influences caused by risk transfer and the consequences of risk control are needed to consider in the research on the dynamic risk changes, which will be researched in the next stage.

References

1. Cunbin, L.: Project Risk Element Transmission Theory and Application, pp. 37–40. China Water Power Press, Beijing (2009)
2. Li, M., Guangming, Z., Ping, L.: Research on the Transmission Mechanism and Transmission Path of Supply Chain Risk. Sci-Tech Information Development & Economy 31, 96–98 (2007)
3. Li, X., Cai, W.: The Application of Risk Matrix to Software project Risk management. 2009 International Forum on Information Technology and Applications 5, 480–483 (2009)
4. Su, X.-K., et al.: Stochastic Simulation and Risk Analysis of the Xiang-jiang River Water Quality. Journal of Hunan University (Natural Sciences) 33(2), 106–109 (2006)
5. Li, C., Wang, J.: Model of Generic Project Risk Element Transmission Theory Based on Date Mining. Journal of Central South University of Technology 15(1), 539–546 (2008)
6. Zhuokui, Z., Huichan, C.: Random process, vol. 140. Xidian University Press, Xian (2003)
7. Yang, C.: Analysis Model of Cooperatively Technical Innovation Risk. In: Shen, G., Huang, X. (eds.) ECWAC 2011, Part II. CCIS, vol. 144, pp. 37–43. Springer, Heidelberg (2011)
8. Yang, P.C., Weeb, H.M., Liu, B.S., Fong, O.K.: Mitigating Hi-tech Products Risks due to Rapid Technological Innovation. Omega 39 (2011)
9. Narasimhan, R., Talluri, S.: Perspectives on risk management in supply chains. Journal of Operations Management 27(2), 114–118 (2009)
10. Chen, Y.J., Seshadri, S.: Supply chain structure and demand risk. Automatica (Oxford) 42(8), 1291–1299 (2006)
11. Chang, C.: A critical analysis of recent advances in the techniques for the evaluation of renewable energy projects. International Journal of Project Management 31(7), 1057–1067 (2013)
12. Dave, S., Sooriyabandara, M., Yearworth, M.: System behaviour modelling for demand response provision in a smart grid. Energy Policy 61, 172–181 (2013)
13. Espinoza, D., Morris, J.: Decoupled NPV: a simple, improved method to value infrastructure investments. Construction Management and Economics 31(5), 471–496 (2013)
14. Yu, C., Lam, K.: A Decision Support System for the determination of concession period length in transportation project under BOT contract. Automation in Construction 31, 114–127 (2013)
15. Kuo, Y., Lu, S.: Using fuzzy multiple criteria decision making approach to enhance risk assessment for metropolitan construction projects. International Journal of Project Management 31(4), 602–614 (2013)
16. Park, A., Chang, C.: Impacts of construction events on the project equity value of the Channel Tunnel project. Construction Management and Economics 31(3), 223–237 (2013)
17. Zhang, X., Chen, S.: A systematic framework for infrastructure development through public private partnerships. IATSS Research 36(2), 88–97 (2013)
18. Abbasianjahromi, H., Rajaie, H.: Developing a project portfolio selection model for contractor firms considering the risk factor. Journal of Civil Engineering and Management 18(6), 879–889 (2012)
19. Di Lorenzo, G., Pilidis, P., Witton, J., Probert, D.: Monte-Carlo simulation of investment integrity and value for power-plants with carbon-capture. Applied Energy 98, 467–478 (2012)

20. Ke, Y., Wang, S., Chan, A.: Risk management practice in China's Public-Private Partnership projects. Journal of Civil Engineering and Management 18(5), 675–684 (2012)
21. Khanzadi, M., Nasirzadeh, F., Alipour, M.: Integrating system dynamics and fuzzy logic modeling to determine concession period in BOT projects. Automation in Construction 22, 368–376 (2012)
22. Lee, H.: An integrated model for planning development projects using ACO and construction simulation. Civil Engineering and Environmental Systems 28(4), 285–300 (2011)

Client Credit Evaluation Index System of Logistics Enterprises Based on Fuzzy Comprehensive and Analytic Hierarchy Process

Yong Luo[1,2] and Zhiya Chen[1]

[1] School of Traffic and Transportation Engineering, Central South University, Changsha, Hunan, 410075, China
[2] Department of Management Engineering, Hunan Engineering Polytechnic, Changsha, Hunan, 410151, China
luoyong@guigu.org

Abstract. Much accounts receivable is occurred in the transactions of logistics enterprises, so the judgment of risk is necessary for operation of the logistics enterprises. Client Credit Evaluation Index System of Logistics Enterprises construction is studied in this paper. The evaluation index system of client credit for logistics enterprises is constructed by using the method of fuzzy comprehensive, and the index weight is determined by Analytic Hierarchy Process (AHP). An application example is studied, and the results are compatible with the facts. It illustrates the effectiveness of the proposed evaluation index system of client credit.

Keywords: Logistics Enterprises, Fuzzy comprehensive, Analytic Hierarchy Process, Client Credit Evaluation.

1 Introduction

In current market situation, cash transaction mode is gradually decreasing in rate on sums. Most of transactions are that supplying service is the first step and then collecting cash. These are the credit transactions. Credit management is a solution of market expansion. It is also a very common way for expanding the logistics market.

Much accounts receivable is occurred in the transactions of logistics enterprises. One reason is the production enterprises and consumer goods wholesale enterprises occupy large amounts of money of the logistics enterprises, even seriously arrear of logistics service payment to develop market and expand blindly. The logistics enterprises will lead to financial deterioration, operation difficulties, and even bankruptcy. The other reason is that the production enterprises and consumer goods wholesale business also will meet the logistics providers of poor reputation, and even encounter the phenomenon of cheat goods. So the judgment of risk is necessary for operation of the logistics enterprises. The credit management of logistics enterprises is requirement. To new contact clients, logistics enterprises should grasp its business scale according to its condition of faith. So client credit evaluation of logistics enterprises is the first step for credit management.

Y. Yang, M. Ma, and B. Liu (Eds.): ICICA 2013, Part II, CCIS 392, pp. 323–332, 2013.

2 Construction of Index System

2.1 Client Credit Evaluation Index System Construction Based on Fuzzy Comprehensive Method

2.1.1 Evaluation Index Selection

According to the characteristics of logistics enterprise, the design requirements of credit evaluation index system of logistics enterprises management mainly include: ① The content of index system should be comprehensive. Not only the basic of default risk capacity evaluation of financial, management and service level should be included, historical records of credit and regional area credit conditions should also be included; ②Index structure should be distinctive nuance, avoiding repetition and disordered; ③Index should be convenient for computer data processing.

Character, Capital, Capacity of Management, Collateral and Condition of Business are included in client credit evaluation of logistics enterprises. Based on the above three requirements, a three layer progressive structure index system is constructed. The first layer is target layer, and the target is client credit evaluation index of logistics enterprise; the second layer is criterion layer, and character, capital structure, capacity of management, collateral and condition of business are considered as five elements of criterion; the third layer is index layer, according to the 5C element analysis and the expert investigation method, the following ten index factors are determined: corporate reputation, breach of contract; the quick ratio, the ratio of liabilities to assets; human resource level, profit ability; clarity of property rights, guarantee condition; regional credit status, regional legal environment.

2.1.2 Evaluation Standard

Evaluation standard determination is the difficulty step in constructing evaluation index system. The method of expert consultation is introduced in Evaluation standard determination. Through several times of consultation to the logistics enterprise managers and university management expert in Hunan Province, specific indicators are selected according to the difference among different evaluation object characteristics. Then, the client credit evaluation standard is established as shown in Table 1.

Table 1. Client credit evaluation index of logistics enterprise

Target layer A	Standard layer B	Index layer C	Standard value	Unit	Source of information
Client credit evaluation elements of logistics enterprise	B_1 character	C_1 corporate reputation*	≥6		peer evaluation
		C_2 default records	≤2	times / year	company records
	B_2 capital	C_3 quick ratio	≥200	%	financial statements
		C_4 ratio of liabilities to assets	≤60	%	financial statements

Table 1. (*Continued.*)

	B_3 capacity of management	C_5 human resource level	⩾60		peer evaluation
		C_6 profit ability	⩾6		financial statements
	B_4 collateral	C_7 clarity of property rights*	⩾6		accepted data
		C_8 guarantee condition	⩾10	%	company to negotiate
	B_5 condition of business	C_9 regional credit status*	⩾6		accepted data
		C_{10} regional legal environment *	⩾6		

The above indexes with '*' indicators such as corporate reputation are qualitative indicators which using 10-point score scale.

2.1.3 Fuzzy Comprehensive Evaluation Method

The above evaluation index is represented by Vector B. The fuzzy relation matrix $R_i(i=1,2,3,4,5)$ is constructed through the comment terms set $V=\{v_1,v_2,v_3,v_4,v_5\}$, where r_{ij} is the degree of membership about evaluation indicator i belongs to reviews of grade j. Combined with the Analytic Hierarchy Process (AHP)[2] to determine the weight vector W, and index vector B is calculated by $B=W\bullet R$. Fuzzy synthetic value A is calculated by $A=B\bullet V(V=10,7.5,5,2.5,0)$, where A is a comprehensive fuzzy evaluation results less than 10, and it is used to evaluate client credit status of logistics enterprises.

2.2 Index Weight Determination by AHP

2.2.1 Analytic Hierarchy Process

The AHP method is put forward by Professor T. L. Saaty[2]. AHP is a simple, flexible and practical multiple criteria decision making method for the quantitative analysis of qualitative problems. The characteristic of the method is that the complex question of various factors is well-organized by dividing into interconnected orderly level. The opinions of expert and the objective judgments of analysts are combined directly and effectively, according to certain objective reality subjective judgment. The importance of two elements in a same layer comparing is described quantitatively. Then, the sequence of weights importance of each level element relative is calculated by mathematical method. All elements of the relative weights is computed and sorted through the total sequencing of all layers.

2.2.2 Index Selection

(1) Constructing the hierarchical structure model

There are both quantitative and qualitative indicators in the client credit evaluation index system of logistics enterprise. Therefore, a comparable index is provided in order

to give out the final scientific conclusion. The index weight is decided democratically by expert group, and the index weight coefficient is determined by AHP. The hierarchical structure model of client credit evaluation index system[3] for logistics enterprise is constructed based on the AHP, as shown in Table1.

The process of application of AHP to calculate the weight of index:

Step 1: Hierarchical structure model construction

A recursive and orderly hierarchy model[3, 4] is constructed according to the analysis about each client credit influence factor of logistics enterprise, and the influence and subordination relationship among each factors. The model is shown in Fig. 1.

Step 2: Judgment matrix construction

The importance of each element is quantified by judgment matrix. In this process, a judgment matrix *R* is formed using DELPH method. The questionnaire are filled by the logistics management experts from logistics enterprises and institutions. Each expert gives a quantitative judgment for each layer of constituent elements by comparing the importance between two elements.

$$R = \left\{ r_{ij} \right\}_{nn}, i, j = 1, 2, \cdots, n \tag{1}$$

Experts give the judgment value of relative importance r_{ij} for various elements, where r_{ij} values generally range of *1*, *3*, *5*, *7* and *9* which represents the degree of importance between r_i and r_j. Larger of the number represents that higher degree of importance.

step 3: Hierarchical ranking

The relative importance of influence is obtained through the judgment matrix from high to low layer of various elements, and then the importance of each element is sorted from high to low layer for each layer. That determines the weight of all the elements in the right layer. It can be called hierarchical ranking. The geometry average method is applied for hierarchical ranking. The equation is shown as follows:

$$W_i = \overline{W}_i \times \frac{1}{\sum_{i=1}^{n} \overline{W}_i}, \quad i = 1, 2, \cdots, n \tag{2}$$

$$\overline{W}_i = \sqrt[j]{\prod_{j=1}^{n} r_{ij}}, \quad i, j = 1, 2, \cdots, n \tag{3}$$

step 4: Total sorting of layers

The integrated weight of the upper element, namely total sorting of layers, is computed based on the weight hierarchical ranking of each level. The equation is shown as follows:

$$W_i = \sum W_C \bullet W_{\mathbf{B}}, \quad i = 1, 2, \cdots, n \tag{4}$$

step 5: Consistency test

Consistency test to the evaluation results of judgment matrix is necessary for the validation of total sequencing results. Consistency is defined as an evaluation index of scoring rationality. If the score is obviously unreasonable, then sort result is defined as invalid. Complete consistency is not possible to achieve because the result of judgment matrix is composed of experts experience fuzzy quantification. So if the consistency variation is in the allowable range, the consistency is effective. T.L.Saaty proposed a random consistency ratio, referred as $C{\bullet}R$. When $C{\bullet}R$ <0.1, it shows that the judgment matrix is consistency. Otherwise, the judgment matrix should be adjusted and recalculated until $C{\bullet}R < 0.1$. $C{\bullet}R$ can be calculated by the following equation:

$$C \cdot R = \frac{C \cdot I}{R \cdot I} \tag{5}$$

Where $R{\bullet}I$ in Eq. 5 is a given ratio, it's relative to order number of the judgment matrix, it commonly values as shown in Table 2.

Table 2. $R{\bullet}I$ Value table

N	1	2	3	4	5	6	7	8	9	10
R·I	—	—	0.58	0.90	1.12	1.24	1.32	1.41	1.45	1.49

Where $R{\bullet}I$ is the consistency index, it can be calculated by the following equation:

$$C \cdot I = \frac{\lambda_{\max} - n}{n-1} \tag{6}$$

Where $\lambda_{\max}$ is the maximum characteristic root of judgment matrix. The approximation calculation equation of the maximum characteristic root is shown as follows:

$$\lambda_{\max} = \frac{1}{n}\sum_{i=1}^{n}\frac{(B \bullet W_B)_i}{(W_B)_i} \tag{7}$$

Where B is given judgment matrix

n is judgment matrix order

W_{Bi} is column vector of relative weight

(2) Judgment matrix of evaluation index weight construction by AHP

An expert team is consisted of fifteen logistics management experts from colleges and universities, logistics industry association and logistics enterprises. Expert opinions are repeatedly consulted and summarized, which tends to consistency. Expert opinions are collected, and then the calculation results are as follows.

① The judgment matrix $A-B$ shows the relative importance comparison among factors of criterion layer relative to the total target of client credit evaluation of logistics enterprises.

$$R_{AB}=\begin{array}{c} \\ B_1 \\ B_2 \\ B_3 \\ B_4 \\ B_5 \end{array}\begin{array}{c} \begin{array}{ccccc} B_1 & B_2 & B_3 & B_4 & B_5 \end{array} \\ \begin{bmatrix} r_{11} & r_{12} & r_{13} & r_{14} & r_{15} \\ r_{21} & r_{22} & r_{23} & r_{24} & r_{25} \\ r_{31} & r_{32} & r_{33} & r_{34} & r_{35} \\ r_{41} & r_{42} & r_{43} & r_{44} & r_{45} \\ r_{51} & r_{52} & r_{53} & r_{54} & r_{55} \end{bmatrix} \end{array}=\begin{bmatrix} 1 & 1/3 & 1/5 & 3 & 2 \\ 3 & 1 & 1/2 & 5 & 4 \\ 5 & 2 & 1 & 7 & 6 \\ 1/3 & 1/5 & 1/7 & 1 & 1/2 \\ 1/2 & 1/4 & 1/6 & 2 & 1 \end{bmatrix}$$

② The judgment matrix B_1-C, B_2-C, B_3-C, B_4-C and B_5-C show that the relative importance comparison among interior influencing factors according to character, composition of capital, capacity of management, collateral and condition of business, respectively.

$$R_{B_1C}=\begin{array}{c} \\ C_1 \\ C_2 \end{array}\begin{array}{c} \begin{array}{cc} C_1 & C_2 \end{array} \\ \begin{bmatrix} r_{11} & r_{12} \\ r_{21} & r_{22} \end{bmatrix} \end{array}=\begin{bmatrix} 1 & 1/2 \\ 2 & 1 \end{bmatrix}$$

$$R_{B_2C}=\begin{array}{c} \\ C_3 \\ C_4 \end{array}\begin{array}{c} \begin{array}{cc} C_3 & C_4 \end{array} \\ \begin{bmatrix} r_{33} & r_{34} \\ r_{43} & r_{44} \end{bmatrix} \end{array}=\begin{bmatrix} 1 & 2 \\ 1/2 & 1 \end{bmatrix} \qquad R_{B_3C}=\begin{array}{c} \\ C_5 \\ C_6 \end{array}\begin{array}{c} \begin{array}{cc} C_5 & C_6 \end{array} \\ \begin{bmatrix} r_{55} & r_{56} \\ r_{65} & r_{66} \end{bmatrix} \end{array}=\begin{bmatrix} 1 & 3 \\ 1/3 & 1 \end{bmatrix}$$

$$R_{B_4C}=\begin{array}{c} \\ C_7 \\ C_8 \end{array}\begin{array}{c} \begin{array}{cc} C_7 & C_8 \end{array} \\ \begin{bmatrix} r_{77} & r_{78} \\ r_{87} & r_{88} \end{bmatrix} \end{array}=\begin{bmatrix} 1 & 1/5 \\ 5 & 1 \end{bmatrix} \qquad R_{B_5C}=\begin{array}{c} \\ C_9 \\ C_{10} \end{array}\begin{array}{c} \begin{array}{cc} C_9 & C_{10} \end{array} \\ \begin{bmatrix} r_{99} & r_{9,10} \\ r_{10,10} & r_{10,10} \end{bmatrix} \end{array}=\begin{bmatrix} 1 & 2 \\ 1/2 & 1 \end{bmatrix}$$

(3) Evaluation index system sorting

Geometric average method is applied to calculate the importance coefficients W_{Bi} of $B_i(i=1,2,3,4,5)$ relative to A according to judgment matrix $A-B$. The equations to calculate the relative weight coefficients are Eq. 2 and Eq. 3, and the process of calculation is shown as Table 3.

Table 3. The detailed process of calculation W_{Bi}

A	B_1	B_2	B_3	B_4	B_5	M_{Bi}	$\overline{W_{Bi}}$	W_{Bi}
B_1	1	1/3	1/5	3	2	2/5	0.8326	0.1185
B_2	3	1	1/2	5	4	30	1.9744	0.2810
B_3	5	2	1	7	6	420	3.3470	0.4763
B_4	1/3	1/5	1/7	1	1/2	1/210	0.3432	0.0488
B_5	1/2	1/4	1/6	2	1	1/24	0.5296	0.0754
$\sum$							7.0267	1.0000

The calculation results in Table 3 shows that capacity of management (B_3) is the first Influence factor to client credit evaluation (A) of logistics enterprise, the second is the composition of capital (B_2), and the last is the character (B_1). It is the important basis for evaluating for client credit layer of logistics enterprises. The calculation of the judgment matrix $B-C_1, B-C_2, B-C_3, B-C_4, B-C_5$ and The results of client credit evaluation index weight of logistics enterprise are shown in Table 4.

Table 4. Client credit evaluation index weight of logistics enterprise

	B_1	B_2	B_3	B_4	B_5	Index weigh t
	W_{B1} =0.1185	W_{B2} =0.2810	W_{B3} =0.4763	W_{B4} =0.0488	W_{B5} =0.0754	
C_1	0.4311					0.0511
C_2	0.5689					0.0674
C_3		0.5689				0.1599
C_4		0.4311				0.1211
C_5			0.6081			0.2896
C_6			0.3919			0.1867
C_7				0.3444		0.0168
C_8				0.6556		0.0320
C_9					0.5689	0.0429
C_{10}					0.4311	0.0325
$\sum$	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

The weights in the above table show that the importance sequence of the credit layers are as following: human resource level C_5, profit ability C_6, quick ratio C_3, ratio of liabilities to assets C_4, default records C_2, corporate reputation C_1, regional credit status C_9, regional legal environment C_{10}, guarantee condition C_8 and clarity of property rights C_7.

(4) Consistency test

Consistency test for weight evaluation results of judgment matrix is necessary after the index weight is calculated. Eq. 5, Eq. 6 and Eq. 7 are used for the consistency test. Consistency is tested by calculating $C \bullet R$, $C \bullet I$ and $\lambda_{\max}$, respectively. According to the definition of consistency index, maximum characteristic root of judgment matrix is calculated. Then the consistency of each judgment matrix is tested separately.

Firstly, the consistency of judgment matrix A is tested. The calculation process of maximum characteristic root is shown as follows:

$$\lambda_{\max}=\frac{1}{n}\sum_{i=1}^{n}\frac{(B\bullet W_B)_i}{(W_B)_i}=\frac{1}{5}\sum_{i=1}^{5}\frac{\left(\begin{bmatrix}1 & 1/3 & 1/5 & 3 & 2\\ 3 & 1 & 1/2 & 5 & 4\\ 5 & 2 & 1 & 7 & 6\\ 1/3 & 1/5 & 1/7 & 1 & 1/2\\ 1/2 & 1/4 & 1/6 & 2 & 1\end{bmatrix}\bullet\begin{bmatrix}0.1185\\ 0.2810\\ 0.4763\\ 0.0488\\ 0.0754\end{bmatrix}\right)_i}{\left(\begin{bmatrix}0.1185\\ 0.2810\\ 0.4763\\ 0.0488\\ 0.0754\end{bmatrix}\right)_i}=\frac{1}{5}\times 25.4402=5.0880$$

The consistency test is shown as follows:

$$C\cdot I=\frac{5.0880-5}{5-1}=0.022\quad,\quad n=5\quad,\quad \text{so}\quad RI=1.12\quad,$$

$$C\cdot R=\frac{C\cdot I}{R\cdot I}=\frac{0.022}{1.12}=0.0196<0.1$$

The consistency of judgment matrix A is well, so the calculation is reasonable.

Because the order of judgment matrix $B_i(i=1,2,3,4,5)$ is 2, the consistency test is unnecessary.

2.2.3 Evaluation Set Description of Evaluation Index

The evaluation model is applied to evaluate the client credit layers of logistics enterprise. The evaluation set of evaluation index system for each criterion can be referred as $V=\{v_1,v_2,v_3,v_4,v_5\}$, in which $v_1=\{10,\text{represents "best"}\}$, $v_2=\{7.5,\text{represents "better"}\}$, $v_3=\{5,\text{represents "good"}\}$, $v_4=\{2.5,\text{represents "bad"}\}$, $v_5=\{0,\text{represents "worse"}\}$. The following comment set is summed up through referring to relevant information and investigating several logistics enterprises managers and university logistics experts in Changsha city, which shown in Table 5. The fuzzy relation matrix[5, 6] $R_i(i=1,2,3,4,5)$ is constructed according to the above comment set and combining with the evaluation index system. The elements r_{ij} of R represents the degree of evaluation index i attached to comment j. The index vector B is calculated by $B=W\bullet R$. Fuzzy synthetic value A is calculated by $A=B\bullet V(V=10,7.5,5,2.5,0)$, where A (comprehensive fuzzy evaluation results) is less than 10.

$$A=B\bullet V=W\bullet R\bullet V=\begin{pmatrix}w_1 & w_2 & \cdots & w_{10}\end{pmatrix}\bullet\begin{bmatrix}r_{1,1} & \cdots & r_{1,5}\\ \vdots & \ddots & \vdots\\ r_{10,5} & \cdots & r_{10,5}\end{bmatrix}\bullet\begin{pmatrix}10 & 7.5 & 5 & 2.5 & 0\end{pmatrix}\quad(8)$$

The highest score of the model is 10 marks, while the lowest score is 0 marks. It is Class-A clients when scored 8 ~ 10 points, Class-B clients are scored 5 ~ 7, and Class-C clients are scored 1 ~ 4. The line of credit is determined by the clients' credit rating. The logistics enterprises can adjust credit factors according to the actual situation, and determine the weight of Specific Indicators. For example, it is difficult for non-listed companies to obtain financial information, so other factors should be replaced or reduced characteristic factors.

3 Application Examples and Evaluation

A total of ten employees in a company department are investigated in order to test the performance of the evaluation index system. They are asked to evaluate credit condition of monthly billing client F, and detailed results are shown in Table 5.

Table 5. The evaluation set and credit evaluation rating score of client F (%)

Specific Indicators		Standards of grading					Weight
		10	7.5	5	2.5	0	
C_1	corporate reputation	best	better	good	bad	worse	0.0511
		70	20	10	0	0	
C_2	default records	default record number within the last 12 months					0.0674
		0	1	2	3	4	
		0	100	0	0	0	
C_3	quick ratio	quick ratio of client- quick ratio of industry average					0.1599
		20%	10%	0	-10%	-20%	
		0	100	0	0	0	
C_4	ratio of liabilities to assets	ratio of liabilities to assets of client- ratio of liabilities to assets of industry average					0.1211
		20%	10%	0	-10%	-20%	
		0	100	0	0	0	
C_5	human resource level	best	better	good	bad	worse	0.2896
		80	10	10	0	0	
C_6	profit ability	best	better	good	bad	worse	0.1867
		10	70	20	0	0	
C_7	clarity of property rights	most clear	more clear	clear	basic clear	not clear	0.0168
		70	20	10	0	0	
C_8	guarantee condition	guarantee ratio					0.0320
		20%	15%	10%	5%	0	
		0	100	0	0	0	
C_9	regional credit status	best	better	good	bad	worse	0.0429
		50	40	10	0	0	
C_{10}	regional legal environment	best	better	good	bad	worse	0.0325
		60	20	10	10	0	
$B = W \bullet R$		33.883	57.729	8.063	0.325	0	1
$A = B \bullet V$		812.925					8.1293

The credit rating score of client F in the above table is 8.1293. The result of evaluation method coincides with the facts by comparing the actual situation. So client F is one of the company's good-quality customers.

4 Conclusions

In this paper, the client credit evaluation index system of logistics enterprise is constructed based on the fuzzy comprehensive analysis method. The index weight factor is determined by using the AHP method. An application example is detailed illustrated, and the checking results are compatible with the facts. Therefore, the index system can be used in practice. But the AHP method to determine weight is kind of subjectivity, further study and discuss should be followed so that each index weight coefficient is more objective and scientific.

References

1. Jin, Z.: Based on the fuzzy comprehensive analysis on Hangzhou logistics system of green degree evaluation. Logistics Technology (3), 100–102 (2011)
2. Liu, J.: Forestry enterprise credit management, pp. 40–57. School of economics and management, Northeast Forestry University, Harbin (2009)
3. Zhiwei, Y.: The warehousing performance evaluation of logistics enterprises based on the entropy method. Logistics technology (3), 89–90 (2011)
4. Luo, Y., Zhao, C., Xue, Z.: Analysis of client credit management of logistics enterprises in China. Enterprise Technology Development (4), 86–88 (2007)
5. Ning, Y.: The research on credit risk guard of logistics enterprises. School of Traffic and Transportation Management, pp. 22–23. Dalian Maritime University (2010)
6. Ji, Y.-D., Yang, X.-J.: Research on the construction of comprehensive evaluation index system for credit risk of enterprises. Engineering and Management (11), 102–105 (2011)

Research on Real Time Data Warehouse Architecture

Rui Jia[1], Shicheng Xu[1], and Chengbao Peng[1,2]

[1] Neusoft Corporation, Xinxiu Road, Shenyang, P.R. China
[2] Northeastern Universities, Wenhua Road, Shenyang, P.R. China
{jia.r,xushicheng,pengcb}@neusoft.com

Abstract. Real time data warehouse is the research hotspots of data warehouse. It expands the application scope of data warehouse and provides real-time decision-making system for business users. This paper describes the concepts of real time data warehouse and proposes a real time data warehouse architecture which is based on real-time cache storage. The architecture consists of three main components: real-time data capture and integration, business event management component and view materialization decision. There are two key technologies: real-time data extraction and materialized view decision-making. This paper describes existing solutions and their shortcomings, then proposes feasible technical solutions: real-time data extraction based on transaction log analysis and materialized view estimation model with time factor.

Keywords: real time data warehouse, data cube, real time storage, materialized view.

1 Introduction

With the development of information technology, massive data have been generated, enterprise need to analysis data efficiently and accurately. Data warehouse, online analysis, business intelligence and data mining are developing, these technologies help enterprise to analyze data and make businesses decision. Traditional data warehouse system use historical data aggregation and analysis to provide strategic decision making, long-term planning, and product management for corporate decision makers. However, enterprise hope data warehouse to provide real-time strategic decision making, such as real-time marketing, personalized service, but the traditional data warehouse technologies can't meet these needs.

Real-time data warehouse is a new data warehouse architecture which is based on the traditional data warehouse development. Real-time means detect and capture the changed data from business systems in time, and load data into the data warehouse. Users can access and query real-time data warehouse to make tactical decision analysis. Currently there have some academic research and product development about real-time data warehouse. Literature [1] proposes active data warehouse base on ODS and data warehouse concept, which can provide both strategic and tactical decision-making for enterprise. Other representative research results [2-4] on active data warehouse are also introduced. Literature [5-8] describes the challenge of data

Y. Yang, M. Ma, and B. Liu (Eds.): ICICA 2013, Part II, CCIS 392, pp. 333–342, 2013.

capture-efficient ETL and proposes their solutions. Data Integration is one solution, and Change Data Capture is the key technology for data integration.

This paper compares traditional data warehouse architecture and real-time data warehouse architecture, and propose a new real-time data warehouse architecture which is based on real-time data cache. The paper focus on two key technologies: real-time data extraction and real-time view of decision choice, propose specific and feasible technical solutions: real-time data extraction based on log analysis and real-time view estimation model which has fine timeliness and effectiveness.

2 Traditional Data Warehouse Architecture and Problem

Traditional data warehouses usually consists of business data system, ETL tool, data warehouse, business intelligence and online analysis tools, etc.

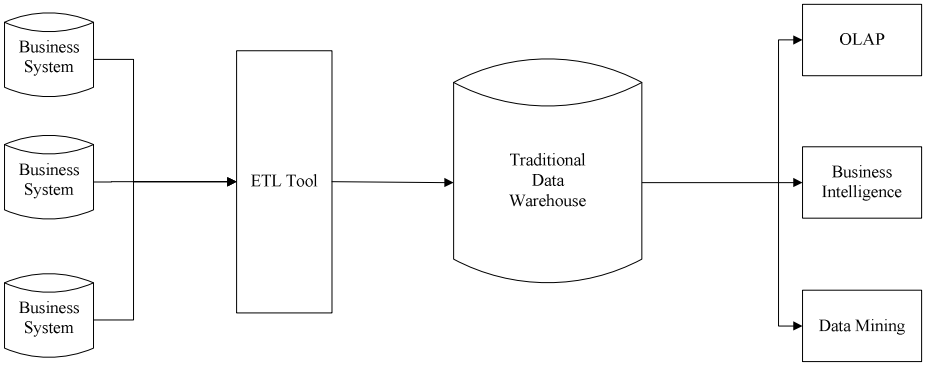


Fig. 1. Traditional data warehouse architecture

ETL tool extract data from business systems by using batch mode at scheduled time, then clean and transform data, finally load the processed data into data warehouse by using batch mode. Business Intelligence tool provides decision analysis, business reports and data mining base on the data warehouse.

The real-time demand requires synchronizing changed data from business system to data warehouse for analyzing. This brings many technical problems which traditional data warehouse has to solve [9-12].

Traditional ETL tool extract data in batch form, and need to monopolize data warehouse for certain time. But real-time update can't tolerate data warehouse being monopolized for long time. Data warehouse update and user query are usually occur at the same time, this should be concern especially.

In order to improve analysis performance, traditional data warehouse usually calculate lots of aggregation data, but data will be updated frequently in real-time environment which causes pre-calculated data become invalid. Also, user query usually needs to execute many operations to complete analysis task, so real-time data update will cause analysis result inconsistent problem.

In traditional data warehouse design, in order to improve resource utilization, response performance, also consider data warehouse stability, it will try to reduce query I/O number. But this conflicts with real-time data warehouse frequent update.

The goal of real-time data warehouse design is to provide decision making and abnormal event handling functionality for managers. In the case of balancing data warehouse burden, the important functionality of real-time data warehouse is response sensitive information and makes corresponding treatment in time.

3 Real-Time Data Warehouse Architecture

This paper proposes real-time data warehouse architecture, which use real-time data storage method and capture business system changed data continuously, use real-time data extraction/integration method to load data into real-time data storage [3]. Use consistent dimension and consistent fact table of dimension modelling theory to make data structure model in real-time data stored and data structure model in historical data warehouse are the same. So, real-time data warehouse can integrate real-time data and historical data seamlessly, provide unified logical view, guarantee the consistency of query results.

This real-time data warehouse architecture includes real-time data capture and load, data integration, business event management and view materialization decision. Figure 1 shows real-time data warehouse architecture.

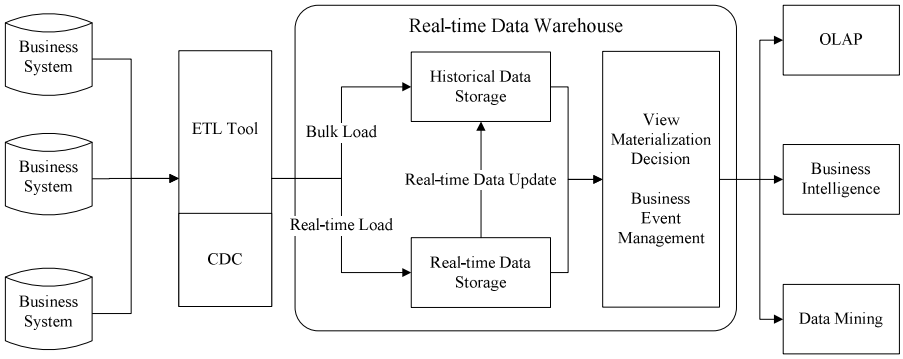


Fig. 2. Real-time data warehouse architecture

Real-time data extraction component uses incremental data capture method to get the changed data from business system, and load changed data into real-time data storage via push way.

Real-time data integration component synchronizes the real-time changed data into historical data storage regularly according to settings.

Business Event Management component is responsible for recording and managing data change events of the business system, updating dimension and fact tables in real-time data storage, providing service for view materialization decision.

View Materialization Decision component selects and materializes views according to query frequency and view effective estimation model, which promotes query efficiency.

4 Real-Time Data Capture and Integration

In traditional data warehouse system, design ETL task based on data warehouse dimension model. ETL task contains business data extraction, data transformation/cleaning, loading rules and task scheduling. ETL engine runs ETL task according to the pre-setting scheduling time and cycle, extracts data from data business system, process data based on the conversion rules, loads data into data warehouse eventually.

In traditional data warehouse, ETL task needs to handle large amount of data and process data with batch mode, which impacts performance of operational system and data warehouse a lot. So, ETL task scheduling cycle is usually 24 hours, also ETL task is triggered at the time when business system load is relatively low. But this solution brings problem: scheduling cycle is too long which make difference between traditional data warehouse and business system, this cause online analysis system cannot provide accurate and consistent analysis results for uses.

Real-time data warehouse must ensure data consistent and analytical results are accurate, this requires real-time data warehouse to load changed data in time when data in business system are changed by transaction processing, in order to meet user's real-time analysis and decision-making requirements. Therefore, real-time data warehouse should be able to notice and extract changed data accurately, do data integration in time [13-16]. The key technical points of the real-time data capture and integration are as follows:

Design good real-time data extraction method

In traditional data warehouse, incremental data extraction technology is divided into trigger, whole table contrast and log analysis. Trigger method needs to invade business system to set triggers to capture real-time changed data, but continuous real-time extraction will bring great load to business system. Whole table contrast method has to copy the needed table from business systems, do additional complex calculation to capture changed data. But this method consumes lots of storage when business systems table data is large, also interact with business system frequently will bring great load. Log analysis method analyze system log, locate and extract changed data based on transaction records. This method doesn't need to invade business system or additional storage space, which makes it suitable for real-time data warehouse field.

Real-time data capture component uses data extraction method which is based on log analysis. This method monitor system operation log. When transaction operations change business data, locate the changed data by analyzing transaction operation logs, build and deliver message, synchronize and load the changed data into the real-time data storage. Some business systems provide transaction log structure, and then real-time capture component can analyze it. But some other business systems don't

provide transaction log structure or content, and then real-time data capture component cannot analyze logs.

Ensure sequence and consistency of real-time data
After extracting data, real-time data capture component integrates real-time data continuously. During data message handling and transmission, message sequence maybe out of order because of message processing or network transmission problems. The disorder problem causes real-time data in cube are inconsistent, and leads to inaccurate query analysis results.

Real-time data capture component adds sequence attribute in data message and builds sequence model to control message order. The transmission message structure includes real-time data location metadata, real-time data and sequence, Message{Meta, Data, Seq}. Location metadata attribute indicates the table which has been changed and the specific changed location. Data sequence attribute indicates the data changed time and the place in transaction.

When delivering data messages, data capture component decides message priority according to the location and sequence metadata. The data which are used by analysis application frequently or has higher sequence will be handled with higher priority.

When receiving data messages, data integration component allocates resource reasonable, organizes cleaning and conversion steps effectively, loads higher priority data firstly, finds out the messages which are needed to be updated firstly in real-time storage according to application access probability and location metadata. This method improves message processing efficiency, promotes load performance of real-time data integration component, and guarantees the consistency of data integration order.

Design good real-time data continuous load method
In real-time data warehouse application scene, application query and data warehouse update are processing simultaneously, this will cause conflict between update and query. When business data changes and data warehouse update frequency are very high, and the upper applications are analyzing at the same time, then this conflict will be more obvious.

In order to solve query-update conflict problem [16] and provide efficient real-time data continuous loading, this paper proposes the real-time data warehouse architecture which puts user real-time processing requirement into system design consideration. Real-time data warehouse provides filter functionality. User can remove the data which are not needed to be analyzed by setting filter rules, only transmit the real-time integration needed data, reduce data volume that need to be processed, improve system load performance. Real-time data integration component is responsible for receiving real-time data messages, putting messages into different queues according to location metadata and sequence, using batch processing and parallel loading to provide real-time data continuous integration loading.

5 View Materialization Decision Component

View materialization is an effective way to promote the execution efficiency because it saves the query results for re-using to improve query efficiency. The principle of the

view materialization is to divide queries into different groups according to their attributes, calculate and save results in advance, which uses extra storage space for time saving. System will response next query and fetch results from the pre-saved results directly, this will shorten the response time greatly. But view materialization needs to use extra storage space, and the materialized views need to be updated when business system tables have been changed. View materialization selection needs to balance between query response costs and update maintenance cost. Real-time data warehouse will gain optimal query response time if all views have been materialized, but this requires a large number of extra storage space and has highest update maintenance cost. Real-time data warehouse will response each query by accessing fact tables and dimension tables to get data if no view has been materialized, this doesn't need any extra storage space and minimizes update maintenance cost, but average query response time is too long.

View materialization selection balances between query response costs and update maintenance cost [4]. View Materialization Decision Component puts user query sets and extra view storage space into consideration. It chooses to materialize certain views, which will balance between query performance and update maintenance cost, maximize real-time data warehouse performance.

5.1 Data Cube Model

Data cube [9] is an important way to implement data warehouse analysis. Data cube attributes are divided into dimension attributes and measure attributes. Dimension attributes are the properties for observing data objects, and measure attributes are the properties for reflecting data object characteristics. Multi-dimensional data analysis is the process to acquire data along with different dimensions. In data cube, different dimension combinations construct sub-cubes with different granularity, and different dimension and measure value combinations construct different data units. Different sub-cubes or data units will be used according to query and analysis.

Data cube model is constituted by dimensions and fact measures [10]. The representation is Cube={D_1, D_2, .., D_n, F}. F represents fact table. Dimension D_i is represented as D_i= {H_{i1}, H_{i2}, .., H_{im}, All}, H_{ij} ($1 \le j \le m$) represents the hierarchy of D_i, All represents the top aggregation hierarchy. Dimension hierarchy means that measure data in higher dimension hierarchy H_{ij} can be generated by aggregating measure data in lower dimension hierarchies H_{ik} ($k \le j$).

Data cube contains large amount of data. If data cube has N dimensions and each dimension has no hierarchy, then the number of sub-cube will be 2^n. If each dimension has complex hierarchies or dimension number is big, then the storage space that data cube need will be very large. In theory, all queries can be serviced by accessing and calculating the most fine grit sub-cube. But the calculation and aggregation cost for large amount of data is quite high, so we need to involve certain strategies to improve query response time and update efficiency of data cube. Pre-calculate aggregation sub-cube is a common method. It calculates different sub-cube view with different granularities in advance, and user query can fetch result data directly, but it brings storage space increment and maintenance complexity.

Therefore, cube view calculation strategy has to consider several major factors: cube storage space, query response time and update maintenance cost.

Data cube query is to select corresponding dimension hierarchies and data range according to the user's input query, slice or dice data cube to obtain the measure data sets which match the query. Because cube views contain various combinations of dimension and hierarchy, so each query can find the corresponding sub-cube views. Cube view is based on dimensions and measures of different hierarchies. It is represented as $V=\{(D_1, H_1, R_1), (D_2, H_2, R_2), .., (D_k, H_k, R_k)\}$, D represents dimension, H represents dimension hierarchy, R represents data value range of dimension. Minimum granularity view is generated by calculating lowest dimension and measure hierarchies, and they are the basis view of other cube views. The high-level cube views can be generated by aggregating number of minimum granularity views.

Cube query is a process to select certain corresponding cube views essentially, it is represented as $Q=\{(D_i, H_i, R_i), (D_j, H_j, R_j), .., (D_k, H_k, R_K,)\}$, $D_{i,j,k}$, $H_{i,j,k}$ and $R_{i,j,k}$ represents the corresponding dimension, hierarchy and data range [12]. Cube queries are inter-dependencies between with each other, high-level query result can be generated by combining low-level query results. For query Q and Q`, if result of Q` can be generated by result of Q, this indicates that $Q` \leq Q$.

User query set is represented as $Qset=\{Q_1, Q_2, .., Q_n\}$. Cube view which will be accessed by Qset is $\{(V_{11}, V_{12}, .., V_{1\alpha}), (V_{21}, V_{22}, .., V_{2\beta}), .., (V_{n1}, V_{n2}, .., V_{n\gamma})\}$. We calculate the number that each view has been accessed, and represent as $\{(V_1, C_1), (V_2, C_2), .., (V_m, C_m)\}$, C_i represents the count that Vi has been accessed in the statistical period. The query frequency of V_i is defined as:

$$Query(V_i) = \frac{C_i}{\sum_{i=1}^{n} C_i} \tag{1}$$

Using the same statistical method, the update frequency of V_i is defined as:

$$Update(V_i) = \frac{U_i}{\sum_{i=1}^{n} U_i} \tag{2}$$

5.2 View Materialization Decision Making Algorithm

View materialization decision-making evaluates view revenue with constraints of time and space, selects certain views to do materialization operation in the principle of view revenue maximization in order to improve query efficiency [13]. Another thing needs to be considered is, business data has changed will cause cube data to be changed, this makes materialized view and cube are inconsistent, and materialized views have to be updated. View materialization decision-making needs to consider both query revenue of materialization view and update maintenance cost, makes balance between these two factors, selects the optimal materialization view sets which meet user needs best.

View materialization decision-making usually uses greedy algorithm to select materialized view set. Greedy algorithm starts from an initial solution, approaches to finial solution gradually in order to find better solution as soon as possible. Greedy algorithm will stop when approaching process can't continue to move forward. Greedy algorithm divides feasible solution construction into different stages. It selects partial optimum solution at each stage, and expects these partial solutions will construct global optimum solution. Greedy algorithm may get the best solution, but it gets only approximate optimum solution at most time.

Paper [14] proposes a greedy algorithm B(v, S) which selects the view with maximum revenue every time. The revenue means the materialization effect that view v can offer for cube materialized view sets S. This algorithm proposes that for any sub-view w of view v, if the cost of using v to calculate w is lower than any other view in S, then use v to calculate w can bring some value. Paper [15] proposes view space revenue model, which involves the revenue that per-unit space of cube view can offer. Define $B_s(V)=B(v, S)/S(v)$, S(v) is the real-time storage space that v has occupied.

Paper [11] proposes an evaluation model which involves query and update frequency of materialized view to improve the original cost model. The model is defined as:

$$CB_s(v) = \frac{Query(v) * B(v,S)}{Update(v) * S(v)} \tag{5}$$

Real-time data in data warehouse are time-efficient, and query results will be different since time and operator have been changed. This causes the original cube view sets are no longer fit for current query, which causes the decline of query response performance. So view materialization sets should make necessary adjustments in time, remove unavailable materialized views and rebuild materialized views according the changes.

The main drawback of above-mentioned view evaluation model is without putting time factor in to consideration, which brings attenuation to cube view when selecting cub views to do materialization treatment. The cube views which are accessed more frequently should have higher efficiency, but the cube views which are visited less frequently should have relatively lower efficiency.

In this paper, we propose a new cube view benefit evaluation model which adds time factor into benefit estimation. This model adjusts materialized view sets dynamically according to the long-term and short-term query frequency of cube view, and short-term query frequency has higher weight. This model uses the materialized cube views which are accessed frequently in recent time to replace the materialized cube views whose view benefit have reduced too much. Query`(v) is the query frequency of materialized cube view v in recent statistical period T. Update`(v) is the update maintenance frequency of materialized cube view v in recent statistical period T. The new cube view benefit evaluation model with time-efficient is:

$$TB_s(v) = \frac{(\alpha * Query(v) + \beta * \overline{Query}(v)) * B(v,S)}{(\lambda * Update(v) + \delta * \overline{Update}(v)) * S(v)} \quad (4)$$

Wherein, $\alpha + \beta = 1$, $\gamma + \delta = 1$. Operator indicates the importance degree of time factor for this cube view benefit evaluation model by setting different weight values.

6 Conclusion

This paper researches real-time data warehouse field and proposes a new real-time data warehouse architecture which is based on real-time data storage. The real-time data warehouse architecture contains real-time data capture and integration, business event management component and view materialization decision component. This paper focuses on analyzing two key technologies: real-time data extraction/integration and view materialization decision-making. It proposes real-time data extraction technology which is based on log analysis to ensure integrated performance and query consistency. It also proposes view benefit evaluation model solution which is based on data cube and time-efficient to achieve balance between real-time query performance and data warehouse update maintenance.

References

1. Brobst, S., Rarey, J.: The five stages of an active data warehouse evolution. Teradata Magazine 3(1), 38–44 (2001)
2. Schrefl, M., Thalhammer, T.: On making data warehouses active. In: Kambayashi, Y., Mohania, M., Tjoa, A.M. (eds.) DaWaK 2000. LNCS, vol. 1874, p. 34. Springer, Heidelberg (2000)
3. Majeed, F., Mahmood, M.S., Iqbal, M.: Efficient data streams processing in the real time data warehouse. In: 3rd IEEE International Conference on Computer Science and Information Technology, vol. 5, pp. 57–61 (2010)
4. Lin, Z., Zhang, D., Lin, C., Lai, Y., Zou, Q.: Performance Optimization of Analysis Rules in Real-Time Active Data Warehouses. In: Sheng, Q.Z., Wang, G., Jensen, C.S., Xu, G. (eds.) APWeb 2012. LNCS, vol. 7235, pp. 669–676. Springer, Heidelberg (2012)
5. Jie, S., Yubin, B., Jingang, S.: A Triggering and Scheduling Approach for ETL in a Real-time Data Warehouse. In: IEEE 10th International Conference on Computer and Information Technology, pp. 91–98 (2010)
6. Song, G., Yang, D., Lin, Z., Tang, S., Wang, T., Xie, K.: Active real time data warehouse concepts, problems and applications. Journal of Computer Research and Development 44(suppl.), 441–446 (2007)
7. Qi, W.: Research of Real-time Data Warehouse Architectur. Eastern Liaoning University Journal 15(1) (2008)
8. Jörg, T., Dessloch, S.: Near Real-Time Data Warehousing Using State-of-the-Art ETL Tools. In: Castellanos, M., Dayal, U., Miller, R.J. (eds.) BIRTE 2009. LNBIP, vol. 41, pp. 100–117. Springer, Heidelberg (2010)
9. Hou, D., Lu, C., Liu, Q., Zhang, W.: Data cube computation methods overview. Computer Science (2008)

10. Qi, W., Xu, B., Tang, H.: Materialized view selection in Data Cube. Henan University Journal 31(1), 20–24 (2001)
11. Ren, J., Li, Z., Zong, J.: Data warehouse materialized view selection method research. Computer Research and Development 43(suppl.), 621–625 (2006)
12. Tang, H., Zou, L.: Dynamic selection of Multidimensional materialized views. Software Journal 13(6), 1090–1096 (2002)
13. Huang, Z., Xue, Y., Wen, J., Cai, J., Wen, W.: DSSMV - Dynamic Selection Strategy of Materialized Views of Multi-Dimensional Data. Computer Science 32(7), 363–368 (2005)
14. Harinarayan, V., Rajaraman, A., Ullman, J.D.: Implementing data cubes efficiently. In: Proc. of the 1996 ACM SIGMOD Int'l Conf. on Management of Data, pp. 205–227. ACM Press, New York (1996)
15. Amit, S., Deshpande, P.M.: Materialized view selection for multidimensional datasets. In: Proc. of the 24th Int'l VLDB Conference, San Francisco, pp. 488–499 (1998)
16. Thiele, M., Lehner, W.: Evaluation of Load Scheduling Strategies for Real-Time Data Warehouse Environments. In: Castellanos, M., Dayal, U., Miller, R.J. (eds.) BIRTE 2009. LNBIP, vol. 41, pp. 84–99. Springer, Heidelberg (2010)

Regional Road Network Shortest Path Algorithm in Vehicle Navigation Systems

Ying Li[1], Aimin Yang[1], and Donghui Li[2]

[1] College of Sciences, Hebei Union University. 063009 Tang'shan, China
liying7909@163.com
[2] Tianjin supply section, Beijing Railway Bureau, 064000 Tang'shan, China

Abstract. According to the characteristics of the regional road network, analyze the connectivity among the road sections, and draw to seek the regional road network the shortest path algorithm between two nodes. First, the paper will introduce the research background and significance of the vehicle navigation system, and analyze the research status and development prospects of the issue; secondly, build the regional road network model according to the transportation network map; then further describes Floyd algorithm of the shortest path, and gives the C language code; finally, take actual traffic map in Tangshan city of Hebei province as network model, the paper achieves purpose by the use of Floyd algorithm and provides a scientific and rational calculation basis for system realization.

Keywords: the shortest path, regional road network, vehicle navigation, Floyd algorithm.

1 Introduction

The vehicle navigation system which uses a certain positioning and navigation technology can make the vehicle that is in the process display the physical location of real-time, and plan the best driving route from the beginning to the end, and guide the vehicle traveling the best path. The path planning based on the regional road network map is a process which plans the best driving path before the vehicle move or during the driving. In order to meet the actual requirements, the path planning should have a property of fast and optimality [1-5]. The best path selection requires certain evaluation criteria, and an intuitive and desirable criterion for the shortest path is that the distance from the starting point to the target point is the shortest. The shortest path problem is an important issue in network research. The Dijkstra algorithm, the Floyd algorithm and Kruskal's algorithm are widely recognized as the optimum solution undoubtedly [6-12]. Here we will introduce the Floyd algorithm which can directly find the shortest path between any two points in the network, it is also called distance matrix method for power multiplication or interpolation point method， and the algorithm was proposed in 1962 by Floyd-Warshall [13-18]. Although it is more complicated than the Dijkstra algorithm, it is widely used as it applies to the solving of the shortest path between any two vertices in the network.

Y. Yang, M. Ma, and B. Liu (Eds.): ICICA 2013, Part II, CCIS 392, pp. 343–353, 2013.

2 Build the Regional Road Network Model

Regional transport hub is made up of the national highway and the river, but the focus is the collection of the national highway. The traffic map is composed of a number of intersecting and conjoint national highways, and it composes intertwined and complex regional transportation network map. The location relations among national highway in the transportation network map are quite complex, a national highway may intersect or connect with a number of national highways connected, and the mode is very complicated. In order to avoid considering topological relationship between national highways, this paper will take districts and counties within the transport network map as one of the object of analysis, and analyze the other object - national highway; taking all districts and counties as the nodes we split the national highway and make it become sections. In this way, the entire network diagram will be made up of the district and sections [19-20]. We define the various districts and counties as the network node and define the sections as the sides of the network [21-30]. In a regional transportation network map we can draw characteristics from the natural angle as follows:

In a city map in the scale of 1:10000(or the scale is more larger), the sections (network side) are approximate straight lines or the radians is small (Particularly in the planned modern metropolis)

The network topology relationship is complex

The side is usually two-way through.

According to the above city transport network characteristics, we can do the following analysis and assumes as follows:

1. All the sections (sides) are straight line. For the sections with a larger radian, in order to solve included angle between the arcs conveniently and make the included angle correctly reflect the relative position between the arcs, we can reduce the radian value by adding a node at the inflection point in the sections, and it is shown in figure 1. The radian between node 1 and 2 is larger, we add node 3 on the section, and then the original sections split two sections with relatively smaller radian. It is shown in Fig 1. That is to say add a node in the network and divide a side in two and each side can be seen as straight line and we can calculate the angle, the other properties(e.g. length) of the side are still characterized by the arc attributes (such as curve length);

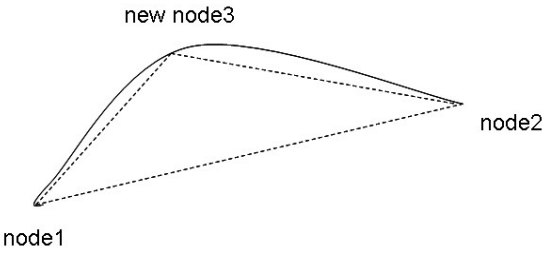


Fig. 1. Add a node to reduce the radian

2.The sides are all two-way through.

3. Nodes are the real nodes

Regional transportation network map is a vectorizable regional road traffic map. Vectorize the regional road traffic map through Google map, and extract the nodes and edges from the regional transportation network. That is the section cross-points and sections in the map.

3 The Main Idea of the Floyd Algorithm

The Floyd algorithm is a kind of algorithm which is used to solve the shortest path length by matrix calculation when having defined a network weight matrix. The network weight matrix is sometimes called the distance matrix of the network.

Suppose there should be the graph G (V, E) which was stored by the adjacency matrix cost, set a two-dimensional array A used to store the length of the shortest path between the current vertex in addition, the component $A[i][j]$ indicates the shortest path between the current vertex v_i and the vertex v_j. The basic idea of Freud algorithm is that a recursive can generate a matrix sequence A_0, A_1, …, A_k ,…, A_n, the $A[i][j]$ indicates the shortest path length where from the vertex v_i and the vertex v_j the path to go through the vertex number is not greater than K.

Initially $A_{-1}[i][j] = cost[i][j]$, when seeking the shortest path length from the vertex v_i to the vertex v_j on the path through the vertex number is not greater than (k +1), two cases to be considered, one case is that the path go without the number of vertices (k +1);another case is that the path from the vertex v_i to the vertex v_j which goes via vertices (k +1).Then, the path can be divided into two, one is the shortest path goes from vertex v_i to vertex v_{k+1}, another is the shortest goes from vertex v_{k+1} to vertex v_j, the shortest path length is equal to the sum of these two paths. The smaller value in both cases through the required path where the vertex number is not greater than (k +1) from vertex v_i to vertex v_j is the shortest path.

The Freud thought can be described by using the following expression:

$$\begin{aligned} A_{-1}[i][j] &= cost[i][j] \\ A_{k+1}[i][j] &= \min\{A_k[i][j], A_k[i][k+1], A_k[k+1][j]\} \\ &\qquad (-1 \le k \le n-2) \end{aligned} \tag{1}$$

The formula is an iterative expression, A_k indicates the shortest path including (k+1) vertex that we have considered from vertices 0,1, ..., k (k +1) ,as so on; $A_k[i][j]$ indicates the shortest path going from vertex v_i to vertex v_j which has including vertices 0,1, ..., k (k +1) . On this basis, then we consider the vertex (k +1), the shortest path including each vertex can be calculated considering the vertex (k +1), so we can get the value A_{k+1} .The shortest path from vertex v_i to vertex v_j will

include a vertex in every iteration. The valve $A_{n-1}[i][j]$ we got after n times of iterations is the shortest path from vertex v_i to vertex v_j considering all the vertices.

If $A_k[i][j]$ has been calculated, the path length which goes from vertex i to vertex(k+1) is $A_k[i][k+1]$, the path length which goes from vertex i to vertex j is $A_k[i][j]$, the path length which goes from vertex (k+1) to vertex j is $A_k[k+1][j]$. Now we take vertex (k+1) into account, if $A_k[i][k+1]+A_k[k+1][j]<A_k[i][j]$, then the path which goes from vertex i to vertex j will change into the path which goes from vertex i to vertex(k+1), then from vertex (k+1) to vertex j, $A_{k+1}[i][j]=A_k[i][k+1]+A_k[k+1][j]$, Otherwise we does not need to modify the path which goes from vertex i to vertex j.

Besides, we save the shortest path with two-dimensional array "path", which has relation with the number of iterations. When calculating $A_k[i][j]$, the array "$path[i][j]$" save the previous number of the shortest path from vertex v_i to vertex v_j in which the number of the intermediate nodes is not greater than K .When the calculating has finished, we will get the shortest path which goes from vertex v_i to vertex v_j by tracing back to two-dimensional array "path". If $path[i][j]=-1$, there is no intermediate nodes.

For example, given a directed graph in Fig. 2, we will calculate the shortest distance including each vertex.

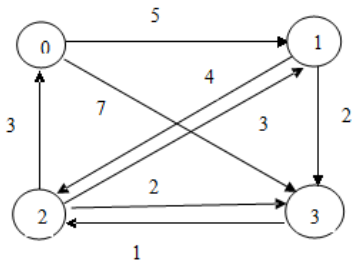


Fig. 2. Directed Graph

Initially:

$$A_{-1}=\begin{bmatrix}0 & 5 & \infty & 7\\ \infty & 0 & 4 & 2\\ 3 & 3 & 0 & 2\\ \infty & \infty & 1 & 0\end{bmatrix},\quad path_{-1}=\begin{bmatrix}-1 & -1 & -1 & -1\\ -1 & -1 & -1 & -1\\ -1 & -1 & -1 & -1\\ -1 & -1 & -1 & -1\end{bmatrix} \tag{2}$$

Take v_0 into account, $A_0[i][j]$ indicates the shortest path length which goes from the vertex v_i to the vertex v_j via the vertex v_0. By comparison, it does not affect the original shortest path, thus we can get the following array:

$$A_0 = \begin{bmatrix} 0 & 5 & \infty & 7 \\ \infty & 0 & 4 & 2 \\ 3 & 3 & 0 & 2 \\ \infty & \infty & 1 & 0 \end{bmatrix}, \quad path_0 = \begin{bmatrix} -1 & -1 & -1 & -1 \\ -1 & -1 & -1 & -1 \\ -1 & -1 & -1 & -1 \\ -1 & -1 & -1 & -1 \end{bmatrix} \tag{3}$$

Take v_1 into account, $A_1[i][j]$ indicates the shortest path length which goes from the vertex v_i to the vertex v_j via the vertex v_1. By comparison , we modify the path to $v_0 - v_1 - v_2$, the path length is 9 which is shorter than before, next we will change $A_0[0][2]$ into 9 and change $path_0[0][2]$ into 1,therefor we will get the following arrays:

$$A_1 = \begin{bmatrix} 0 & 5 & 9 & 7 \\ \infty & 0 & 4 & 2 \\ 3 & 3 & 0 & 2 \\ \infty & \infty & 1 & 0 \end{bmatrix}, \quad path_1 = \begin{bmatrix} -1 & -1 & 1 & -1 \\ -1 & -1 & -1 & -1 \\ -1 & -1 & -1 & -1 \\ -1 & -1 & -1 & -1 \end{bmatrix} \tag{4}$$

Similarly,we will get the following arrays:

$$A_2 = \begin{bmatrix} 0 & 5 & 9 & 7 \\ 7 & 0 & 4 & 2 \\ 3 & 3 & 0 & 2 \\ 4 & 4 & 1 & 0 \end{bmatrix}, \quad path_2 = \begin{bmatrix} -1 & -1 & 1 & -1 \\ 2 & -1 & -1 & -1 \\ -1 & -1 & -1 & -1 \\ 2 & 2 & -1 & -1 \end{bmatrix} \tag{5}$$

$$A_3 = \begin{bmatrix} 0 & 5 & 8 & 7 \\ 6 & 0 & 3 & 2 \\ 3 & 3 & 0 & 2 \\ 4 & 4 & 1 & 0 \end{bmatrix}, \quad path_3 = \begin{bmatrix} -1 & -1 & 3 & -1 \\ 3 & -1 & 3 & -1 \\ -1 & -1 & -1 & -1 \\ 2 & 2 & -1 & -1 \end{bmatrix} \tag{6}$$

The array A_3 is the shortest distance between each vertex.

Here we list the main C language code:

```
typedef struct
int edges[MAXV][MAXV];
int n,e;
}MGraph;
void Floyd(MGraph g)
int A[15][15],path[15][15];
int i,j,k,n=g.n;
for(i=0;i<n;i++)
for(j=0;j<n;j++)
A[i][j]=g.edges[i][j];
path[i][j]=-1;
for(k=0;k<n;k++)
for(i=0;i<n;i++)
for(j=0;j<n;j++)
if(A[i][j]>(A[i][k]+A[k][j]))
A[i][j]=A[i][k]+A[k][j];
path[i][j]=k;
```

4 Simulation and Analysis

Make simulation experiment by Floyd algorithm on the Visual Studio2008 simulation platform. We will make use of the real map of 14 counties of Tangshan city to demonstrate the problem. It is shown in Fig. 3.

Considering that the path of the district is non-isotropic, so we can abstract it as a non-isotropic graph to solve the problem, the shortest path algorithm model between any pair of vertices is what we want to solve right now. In the non-isotropic graph, 1 represents the Yutian County, 2 represents Fengrun District, 3 represents Zunhua City, 4 represents Qianxi County, 5 represents Qianan City, 6 represents Fengnan District, 7 represents Lunan District, 8 represents Lubei District, 9 represents Kaiping District, 10 represents Guye District, 11 represents Luan County, 12 represents Tanghai County, 13 represents Luannan County, 14 represents Laoting County. The

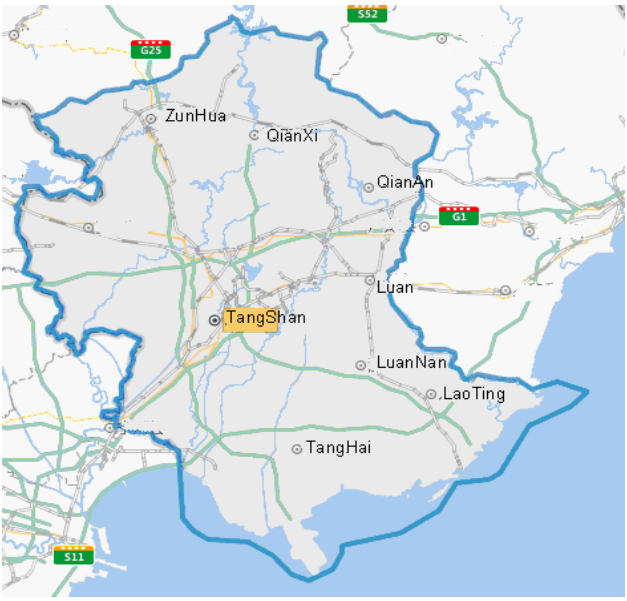


Fig. 3. The Road Net in TangShan

connection line between two nodes shows that there is a direct access between them, and the value above the line represents the length of the path between them, the unit of the values is Km. It is shown in Fig. 4:

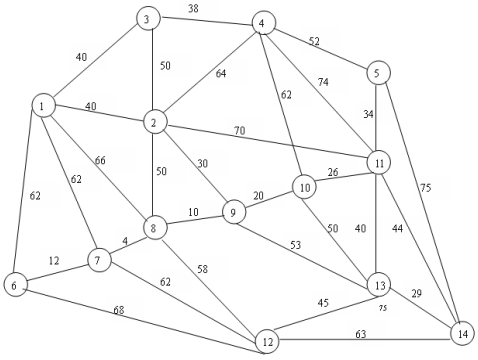


Fig. 4. The non-isotropic graph of Tangshan district

We can list in the initial adjacency matrix according to the fig. 4:

$$\begin{pmatrix}
0 & 40 & 40 & \infty & \infty & 62 & 62 & 66 & \infty & \infty & \infty & \infty & \infty & \infty \\
40 & 0 & 50 & 64 & \infty & \infty & \infty & 50 & 30 & \infty & 70 & \infty & \infty & \infty \\
40 & 50 & 0 & 38 & \infty & \infty & \infty & \infty & \infty & \infty & \infty & \infty & \infty & \infty \\
\infty & 64 & 38 & 0 & 52 & \infty & \infty & \infty & \infty & 62 & 72 & \infty & \infty & \infty \\
\infty & \infty & \infty & 52 & 0 & \infty & \infty & \infty & \infty & \infty & 34 & \infty & \infty & 75 \\
62 & \infty & \infty & \infty & \infty & 0 & 12 & \infty & \infty & \infty & \infty & 68 & \infty & \infty \\
62 & \infty & \infty & \infty & \infty & 12 & 0 & 4 & \infty & \infty & \infty & 62 & \infty & \infty \\
66 & 50 & \infty & \infty & \infty & \infty & 4 & 0 & 10 & \infty & \infty & 58 & \infty & \infty \\
\infty & 30 & \infty & \infty & \infty & \infty & \infty & 10 & 0 & 20 & \infty & \infty & 53 & \infty \\
\infty & \infty & \infty & 62 & \infty & \infty & \infty & \infty & 20 & 0 & 26 & \infty & 50 & \infty \\
\infty & 70 & \infty & 74 & 34 & \infty & \infty & \infty & \infty & 26 & 0 & \infty & 40 & 44 \\
\infty & \infty & \infty & \infty & \infty & 68 & 62 & 58 & \infty & \infty & \infty & 0 & 45 & 63 \\
\infty & \infty & \infty & \infty & \infty & \infty & \infty & \infty & 53 & 50 & 40 & 45 & 0 & 29 \\
\infty & \infty & \infty & \infty & 75 & \infty & \infty & \infty & \infty & \infty & 44 & 63 & 29 & 0
\end{pmatrix}$$

The shortest path length between each node will be showed by the following matrix A:

$$A = \begin{pmatrix}
0 & 40 & 40 & 78 & 130 & 62 & 62 & 66 & 70 & 90 & 110 & 124 & 123 & 152 \\
40 & 0 & 50 & 64 & 104 & 56 & 44 & 40 & 30 & 50 & 70 & 98 & 83 & 112 \\
40 & 50 & 0 & 38 & 90 & 102 & 94 & 90 & 80 & 100 & 110 & 148 & 133 & 154 \\
78 & 64 & 38 & 0 & 52 & 108 & 96 & 92 & 82 & 62 & 72 & 150 & 112 & 116 \\
130 & 104 & 90 & 52 & 0 & 106 & 94 & 90 & 80 & 60 & 34 & 119 & 74 & 75 \\
62 & 56 & 102 & 108 & 106 & 0 & 12 & 16 & 26 & 46 & 72 & 68 & 79 & 108 \\
62 & 44 & 94 & 96 & 94 & 12 & 0 & 4 & 14 & 34 & 60 & 62 & 67 & 96 \\
66 & 40 & 90 & 92 & 90 & 16 & 4 & 0 & 10 & 30 & 56 & 58 & 63 & 92 \\
70 & 30 & 80 & 82 & 80 & 26 & 14 & 10 & 0 & 20 & 46 & 68 & 53 & 82 \\
90 & 50 & 100 & 62 & 60 & 46 & 34 & 30 & 20 & 0 & 26 & 88 & 50 & 70 \\
110 & 70 & 112 & 74 & 34 & 72 & 60 & 56 & 46 & 26 & 0 & 85 & 40 & 44 \\
124 & 98 & 148 & 150 & 119 & 68 & 62 & 58 & 68 & 88 & 85 & 0 & 45 & 63 \\
123 & 83 & 133 & 112 & 74 & 79 & 67 & 63 & 53 & 50 & 40 & 45 & 0 & 29 \\
152 & 112 & 156 & 110 & 75 & 108 & 96 & 92 & 82 & 70 & 44 & 63 & 29 & 0
\end{pmatrix}$$

The shortest path between each vertex will be showed by the following matrix path:

$$path = \begin{pmatrix} -1 & -1 & -1 & 2 & 3 & -1 & -1 & -1 & 1 & 8 & 1 & 6 & 8 & 12 \\ -1 & -1 & -1 & -1 & 10 & 8 & 8 & 8 & -1 & 8 & -1 & 8 & 8 & 12 \\ -1 & -1 & -1 & -1 & 3 & 0 & 8 & 8 & 1 & 3 & 3 & 8 & 8 & 10 \\ 2 & -1 & -1 & -1 & -1 & 9 & 9 & 9 & 9 & -1 & -1 & 9 & 9 & 10 \\ 3 & 10 & 3 & -1 & -1 & 10 & 10 & 10 & 10 & 10 & -1 & 12 & 10 & -1 \\ -1 & 8 & 0 & 9 & 10 & -1 & -1 & 6 & 7 & 8 & 9 & -1 & 8 & 12 \\ -1 & 8 & 8 & 9 & 10 & -1 & -1 & -1 & 7 & 8 & 9 & -1 & 8 & 12 \\ -1 & 8 & 8 & 9 & 10 & 6 & -1 & -1 & -1 & 8 & 9 & -1 & 8 & 12 \\ 1 & -1 & 1 & 9 & 10 & 7 & 7 & -1 & -1 & -1 & 9 & 7 & -1 & 12 \\ 8 & 8 & 3 & -1 & 10 & 8 & 8 & 8 & -1 & -1 & -1 & 8 & -1 & 10 \\ 1 & -1 & 3 & -1 & -1 & 9 & 9 & 9 & 9 & -1 & -1 & 12 & -1 & -1 \\ 6 & 8 & 8 & 9 & 12 & -1 & -1 & -1 & 7 & 8 & 12 & -1 & -1 & -1 \\ 8 & 8 & 8 & 9 & 10 & 8 & 8 & 8 & -1 & -1 & -1 & -1 & -1 & -1 \\ 12 & 12 & 10 & 10 & -1 & 12 & 12 & 12 & 12 & 10 & -1 & -1 & -1 & -1 \end{pmatrix}$$

We can get the shortest path between any two counties (or districts) of Tangshan District from the above results. If we want to go to Laoting County from Fengrun District, we have to find the shortest path between V_2-V_{14}.We find that $A[1][13]=112$,$path[1][13]=12$,$path[1][12]=8$, $path[1][8]=-1$, so we learn that the shortest path length from Fengrun District to Laoting County is 112 Km, the path is 2-9-13-14.That is to say, the shortest path when we set out from Fengrun District to Laoting County is via Kaiping District and Luannan County.

5 Conclusion

As a classical algorithm to solve the shortest path problem, the Floyd algorithm is also widely used in the field of logistics industry, production control, transportation, tourism and so on. This article only study the application of Floyd algorithm on the vehicle navigation system for traffic and road network in the plain areas, and the algorithm is usually able to get an optimal solution, so this algorithm has a strong practical. We can do some modifications to the algorithm for road conditions of different areas, or add other idea of the algorithm into this algorithm, and we will also receive a satisfactory navigation.

References

1. Ahmed, Y.H.: A genetic algorithm for finding the k shortest paths in a network. Egyptian Informatics Journal 11(2), 75–79 (2010)
2. Campos, R., Ricardo, M.: A Fast Algorithm for Computing Minimum Routing Cost Spanning Trees. Computer Networks 52(17), 3229–3247 (2008)
3. Desaulniers, G., Soumis, F.: An efficient algorithm to find a shortest path for a car-like robot. IEEE Transactions on Robotics and Automation 11(6), 819–828 (1995)
4. Gao, L.-Q., et al.: The Application of Harmony Search Algorithm for Solving Shortest Path Problems. Journal of Northeastern University (Natural Science) 32(6), 769–772 (2011)

5. Hribar, M.R., Taylor, V.E., Boyce, D.E.: Implementing parallel shortest path for parallel transportation applications. Parallel Computing 27(12), 1537–1568 (2001)
6. Li, S.-J., et al.: Study of Shortest Path Based on Weighted Identification S-graph. Computer Engineering 37(15), 46–48 (2011)
7. Parichart, P., Dongjoo, P., Laurence, R.R., et al.: Dynamic and stochastic shortest path in transportation networks with twocomponents of travel time uncertainty. TransportationResearch Part C: Emerging Technologies 11(5), 331–354 (2003)
8. Shioda, S., Ohtsuka, K., Sato, T.: An Efficient Network-wide Broad-casting Based on Hop-limited Shortest-path Trees. Computer Networks 52(17), 3284–3295 (2008)
9. Yang, J.-Q., et al.: Problem solving of shortest path between the institutions Based on Freud algorithm. Scienfic & Technical Information of Gansu 39(5), 28–29 (2010)
10. Yue, Q.-J., et al.: Design and Realization of Floyd Algorithm System Based on the Optimation of the Shortest Path Problem. Journal of Gansu Normal Colleges 15(5), 41–43 (2010)
11. Chaorui, W.: Graph Theory, 2nd edn. Beijing University of Technology Press, Beijing (1997)
12. Yingchun, L.: The research for application of geography information system. Zhejiang Institute of Silk Texture, HangZhou (1999)
13. Xu, Z., Zhang, N., et al.: Data Structure. Higher Education Press, Beijing (1981)
14. MapInfo Corporation, MapBasic Reference Guide. MapInfo Corporation (1997)
15. Lu, J., Wang, W.: Index system of urban road network planning. Journal of Traffic and Transportation Engineering 14 (2004)
16. Zhang, G., Lu, J., Yan, K.: Urban Road Network Capacity mathematical model and algorithm. Computer and Communications 1 (2007)
17. Dai, J., Fu, Y.: Comprehensive Evaluation of Urban Road Network. Journal of Beijing Institute of Civil Engineering and Architecture 4 (1997)
18. Zhao, S., Wang, S.-M., Xiong, T.-W.: Urban road network capacity and the supply and demand balance analysis model. Communications Standardization 7 (2004)
19. Ye, P., Chen, X., Cui, X.: From differentiate into fusion - urban road network structure evolution of planning concepts. Urban Planning Forum 5 (2010)
20. Zhang, L., Qiao, Z., Chen, P.: On the main city of Chongqing idea of Intelligent Transportation System. Technology of Highway and Transport 3 (2013)
21. Mao, Y., Min, R., Ping, T., Fulin, Z.: Urban road traffic-induced structural vibration measurements and numerical simulations. Chinese Journal of Applied Mechanics 3 (2013)
22. Qin, X., Zhang, B.: Road alignment technology in the Beijing road data management system for construction planning application. Remote Sensing for Land & Resources 2 (2013)
23. Lan, L., Mo, Q.: Mountain City Static Traffic Analysis and Countermeasures. Journal of Chongqing Jiaotong University (Social Sciences Edition) 3 (2013)
24. Shi, C.K., Yan, Z.: Urban context with a combination of fast and slow traffic design strategy - to Metropolis Glory Guihu overall urban design as an example. Special Zone Economy 5 (2013)
25. Sun, C.-M.: Improve the quality of city sidewalks conservation Discussion. Shanxi Architecture 15 (2013)
26. Kong, X., Yang, H.-L., Huang, Y.-C.: Mainland China's export-oriented manufacturing activities Location Selection and Layout Optimization. Scientia Geographica Sinica 5 (2013)
27. Lingqi, K.: Urban planning and urban road intersections to improve countermeasures. Urban Roads Bridges & Flood Control 5 (2013)

28. Pengfei, C.: Urban Transportation Planning Problems Analysis and Countermeasures. Urban Roads Bridges & Flood Control 5 (2013)
29. Xie, H., Yu, X.-H., Yan, K.-F.: Urban road traffic network system capacity assessment model. China Journal of Highway and Transport 3 (2012)
30. Dong, A.-H., Yang, L.: Geographic Information System technology in a city road traffic injuries in applied research. Journal of Environmental & Occupational Medicine 2 (2012)

DM Data Mining Based on Improved Apriori Algorithm

Yongping Wang, Yanfeng Jin, Ying Li, and Keming Geng

Shi Jiazhuang Post & Telecommunication Technical College, 050021,
Shijiazhuang, China
{mingyuexin_1981,yanfengjin}@163.com,
{liyingyz,gengkeming}@green.sjzpc.edu.cn

Abstract. Association rules are the key technology in data mining; it has a very broad applying foreground in many industries. An improved association rules algorithms based on Apriori was proposed in this paper. And it will be used in direct mail data mining. By analyzing the normative database of users' sets, we can get item set which satisfy the minimal support degree, and form the rule set. We can get more accurate DM data mining results than other methods by testing the post DM database. Experiments indicate the validity of the method.

Keywords: data mining, association rules, apriori algorithms, support degree.

1 Introduction

This template, modified in MS at present the DM service in China is in a stage of coexistence of carpet bombing and database marketing. Although the direct mail enterprises started to build database, Look for target customers and develop direct mail service by using database, Since most of the companies are unable to grasp the right customer data, or the screening and application of the existing customer data is so weak, many enterprises still mail lot of mails by collecting as many of the addresses, and carry out the "carpet" of the direct mail service. As the database marketing idea gradually been accepted by DM enterprises, they began to think about how to establish a long-term and stable and promoting feedback mechanism with customers, and improved the interaction between customers directly. As industry leader of direct mail in china, data accumulation of China post is more and more huge, the traditional data query, statistics and analysis, has not satisfied for the needs of the development of DM. The emerging data mining technology can well solve the problem, so the study of the application of data mining in the industry has very important practical significance.

Association rules are the key technology in data mining, and it was proposed firstly by R.Agrawal [1-3]. Association rules are to give a set of items (Item) and a record collection, and then we can derive the correlation between the Items by anglicizing the record collection. Association rules are widely used in business, health insurance, financial services, the judiciary, etc. In this paper, the association rules algorithm –Apriori algorithm in data mining has been improved and applied to the bill classifying, and get more accurate results.

Y. Yang, M. Ma, and B. Liu (Eds.): ICICA 2013, Part II, CCIS 392, pp. 354–363, 2013.

2 Association Rules

2.1 The Basis of Data Mining

Data mining is the extraction of knowledge from large database. The knowledge usually is implicit, unknown and potential information. The task of Data mining is to discover knowledge, including the following knowledge categories: general type of knowledge, reflects the common knowledge of similar transactions; feature-based knowledge, reflecting the knowledge characteristics of the of all aspects of things; differences in knowledge, reflecting attribute differences between different things in knowledge; relational knowledge, reflecting the dependence or association between objects of knowledge; predictive knowledge, based on historical and current data; deviation from the type of knowledge, reveals things deviate from the normal phenomenon[4-6].

2.2 The Definition of Association Rules

Definition 1: if I = { i1, i2, …, im} is a set composed by m different attributes, given a database D, which each record T is a set of properties in I, $T \subseteq I$, T has a unique identifier TID. If $X \subseteq I$ and $X \subseteq T$, then the record T includes X. A association rules are shaped like $X \Rightarrow Y$, which $X \subseteq I$, $Y \subseteq I$, $X \cap Y = \Phi$. The conditions of Association rule $X \Rightarrow Y$ are:

(1) It has the support degree S. That is at least S % record in database D contains $X \cup Y$.

(2) It has confidence C. That is at least C % X records in database D also contain Y.

Usually A association rules are expressed as $X \Rightarrow Y(S\%, C\%)$.

$$\text{Support degree: } S\% = \frac{\text{The Number of Transactions}(X \cup Y)}{\text{The Number of Transactions}(D)}$$

$$\text{Confidence: } C\% = \frac{\text{The Number of Transactions}(X \cup Y)}{\text{The Number of Transactions}(X)}$$

2.3 The General Steps of Association Rules

There are many algorithms of association rules, no matter what data structure it use, how its complexity and efficient, they all can be divided into the following steps:

(1) Preprocess the data relevant to the task. According to the requirements of specific issues, Manipulating the database and form the standardized database D.

(2) Calculate all item sets which meet the minimum support degree, the large item sets.

(3) Generate the rule to meet the minimum confidence, form the rule set R, explain and output R [7-9].

2.4 Data Mining and Association Rules

Data Mining is a technique for discovering useful information from large databases. Analyzing the data and extracting useful information can be potentially very profitable to a business. For example, if a seller can find the association between two products, critical decisions in pricing or product placements can be made in order to promote the business. By this way, the seller can concentrate the marketing efforts on every subset of customers who are very likely to buy the associated products.

Association Rules are used for discovering regularities between products in big transactional databases. A transaction is an event involving one or more of the products (items) in the business or domain; for example buying of goods by a consumer in a super market is a transaction. A set of items is usually referred as "itemset", and an itemset with "k" number of items is called "k-itemset".

The general form of an association rule is X => Y, where X and Y are two disjoint itemsets. The "support" of an itemset is the number of transactions that contain all the items of that itemset; whereas the support of an association rule is the number of transactions that contain all items of both X and Y. The "confidence" of an association rule is the ratio between its support and the support of X.

A given association rule X => Y is considered significant and useful, if it has high support and confidence values. The user will specify a threshold value for support and confidence, so that different degrees of significance can be observed based on these threshold values.

2.5 Data Structures

We create the following data-structure classes. An Itemset is just a list of strings; it can be used to represent a transaction or any set of items. The ItemsetCollection is list of itemsets; it can be used to represent a transactional database or any group of itemsets. The AssociationRule class represents an instance of a generated association-rule.

```
public class Itemset : List<string>
{ public double Support { get; set; }}
public class ItemsetCollection : List<Itemset>
public class AssociationRule
public Itemset X { get; set; }
public Itemset Y { get; set; }
public double Support { get; set; }
public double Confidence { get; set; }
Finding Large Itemsets using Apriori Algorithm.
```

The first step in the generation of association rules is the identification of large itemsets. An itemset is "large" if its support is greater than a threshold, specified by the user. A commonly used algorithm for this purpose is the Apriori algorithm.

The Apriori algorithm relies on the principle "Every non-empty subset of a larget itemset must itself be a large itemset". The algorithm applies this principle in a bottom-up manner. Let Li denote the collection of large itemsets with "i" number of

items. The algorithm begins by identifying all the sets in L1. Each item that has the necessary support forms a large 1-itemset and included in L1, other itemsets are dropped from consideration. This process of retaining necessary itemsets only is called "pruning". The set of itemsets used to find Li is called candidate itemsets (Ci).

The collection L2 can be constructed by considering each pair of sets in L1 and retaining only those pairs that has enough support. In general, having constructed Li, the collection Li+1 is constructed by considering pairs of sets, one from Li and another from L1 and eliminating those for which the support is smaller. This procedure is continued until all large itemsets up to the desired maximum size have been obtained or no further pruning is possible.

The following code implements the Apriori algorithm.

```
public static ItemsetCollection DoApriori(ItemsetCollection db, double supportThreshold)
Itemset I = db.GetUniqueItems();
ItemsetCollection L = new ItemsetCollection(); //resultant large itemsets
ItemsetCollection Li = new ItemsetCollection(); //large itemset in each iteration
ItemsetCollection Ci = new ItemsetCollection(); //candidate itemset in each iteration
//first iteration (1-item itemsets)
foreach (string item in I)
{Ci.Add (new Itemset () {item}) ;}
//next iterations
int k = 2;
while (Ci.Count != 0)
//set Li from Ci (pruning)
Li.Clear ();
foreach (Itemset itemset in Ci)
{itemset.Support = db.FindSupport(itemset);
If (itemset.Support >= supportThreshold)
Li.Add (itemset);
L.Add (itemset);
//set Ci for next iteration (find supersets of Li)
Ci.Clear ();
Ci.AddRange (Bit.FindSubsets (Li.GetUniqueItems (), k)); //get k-item subsets
k += 1;    }
return (L);}
```

The FindSubsets ( ) function defined in Bit class is used to find all the subsets of a given set of items. This is explained in more detail in this article.

2.6 Finding Association Rules

Having found the set of all large itemsets from the input database, the next task is to find the required set of strong association rules. An association rule is "strong" if its confidence value is greater than a user-defined threshold. The association rules are created by combining each large itemset with each of its subsets. The strong rules are published as result and others are dropped.

```
public static List<AssociationRule> Mine(ItemsetCollection db, ItemsetCollection
L, double confidenceThreshold)
List<AssociationRule> allRules = new List<AssociationRule> ();
foreach (Itemset itemset in L)
ItemsetCollection subsets = Bit.FindSubsets (itemset, 0); //get all subsets
foreach (Itemset subset in subsets)
double confidence = (db.FindSupport(itemset) / db.FindSupport(subset)) * 100.0;
if (confidence >= confidenceThreshold)
AssociationRule rule = new AssociationRule ();
rule.X.AddRange(subset);
rule.Y.AddRange(itemset.Remove(subset));
rule.Support = db.FindSupport(itemset);
rule.Confidence = confidence;
if (rule.X.Count > 0 && rule.Y.Count > 0)
allRules.Add(rule);
return (allRules);
```

3 Apriori Algorithm Research

3.1 Traditional Apriori Algorithm

There are many algorithms of association rules, but most of them are the interpretation and improvements of the classic Apriori algorithm. Apriori is a breadth-first algorithm, by scanning the database D more times to find all the frequent item sets, consider only one pass with the same length K in each scanning named K - Project set. In the first scanning, calculate the database D by using apriori algorithm, generate all the length of a frequent item set. In each follow-up scanning, generate all new candidate sets based on the past found, this is the potential frequent item sets, then scans the database D, calculating the support degree of Candidate set. Finally, the project focused on determining which candidate truly become a frequent item set. Repeat this process until no new frequent item sets. The key efficient algorithm is to generate smaller candidate sets. It uses a basic nature: that is, any subset of frequent item sets must also be frequent item sets. This property is now inherited by most of association rules [10-12]. The specific Apriori algorithm is described as follows:

```
Procedure Apriori ()
Begin
L1= {large1-itemsets; For {k=2; L [k-1] <>nil;k++}do
Begin
C[k] =Apriori-gen (L [k-1]) ;
For all transactions t ∈ D do
Begin
C[t] = subset (C[k], t) ; For all candidates C ∈ C[t] do ccount + + ;
End;L[k] = {C ∈ C[ k ] 0 ccount > = minsup} ;
End;Gen - rules (L[k]) ;
```

End;
Function apriori – gen (L [k-1]) ;
Insert into C[k],Select Cp[1],Cp[2],⋯,Cp[k-2],Cp [k-1],Cq[k-1] From L[k-1] CpL[k-1]
Cq
Where Cp[1]=Cq[1],Cp[2]=Cq[2],⋯,Cp[k-2]= Cq[k-2],Cp[k-1]<Cq[k-1];
If (k-1)-subset c of C[k], c L [k-1] then delete c from C[k];
Return C[k];

3.2 Incomplete Set Algorithms

Incomplete set algorithms are a variation of condensed representations whereby a reduced result set is produced that can provide useful (although incomplete) information about dataset inferences, reducing analysis by discovering incomplete information about the complete set of valid itemsets within D. There are two significant types of incomplete set algorithms: sampling and maximal valid sets. Sampling algorithms analyze only a portion of the dataset and Maximal Frequent Set algorithms (MFS) identify only those valid itemsets for which no valid supersets exist. Sampling works by reducing |D|, and MFS works through the introduction of new pruning strategies.

Maximal Frequent Set (MFS) algorithms identify all itemsets within D for which no valid supersets exist. Therefore, although the result set implies all valid itemsets through the identification of the search space boundary, the actual support and hence inference strength of internal itemsets (those within the search space boundary) remain unknown. The implementation of MFS algorithms result in reduced analysis time because MFS properties enable the inclusion of additional pruning strategies that facilitate search space reduction.

—Upward Closure Principle (UCP). Given a valid itemset A with a set of possible extensions B, if the itemset C|C = A∪B is valid, then C becomes the terminal MFS of A and no further exploration of A's supersets is required.

—Superset Checking (An extension of UCP that avoids direct counting of A). If A is subsumed by or is a subset of an existing MFS, then no further exploration of A's supersets is required.

—Parent Equivalence Pruning. Given an extension-based analysis algorithm incorporating dataset projection, then for each child itemset generated in the enumeration-set tree, the projected object set is compared with that of its parent. If they match, the child can replace the parent node.

Maxminer extends Apriori through the inclusion of UCP and subsumption testing to derive the set of maximal frequent sets. Furthermore, MaxMiner is underpinned by an an enumeration-set tree instead of a hash-tree within which items are dynamically sorted in order of increasing support.

Pincer Search also extends Apriori through the introduction of a bidirectional search that has proven efficient in the discovery of long maximal valid itemsets. The algorithm uses a typical Apriori bottom-up search but extends it by incorporating additional

pruning through the use of a Maximal Frequent Candidate Sets (MFCS) that approaches the valid search space border and hence the discovery of MFS from the top-down.

Given that initially MFCS contains a single itemset of the union of all items, analysis proceeds by iteratively generating the next bottom-up valid set V_K, refining MFCS and then pruning V_K, or V_{K+1} candidates, based on those members of MFCS determined valid. For example, given E = {a, b, c, d, e} and V_1 = {a, b, d, e}, then MFCS is refined, ensuring that no member of MFCS is a superset of an invalid itemset which is provided by $\overline{V}_K \mid \overline{V}_K = C_K - V_K$. Thus MFCS1 = {abde} as item c is determined invalid.The refined members of MFCS are then validated, those found valid are MFS and all subsets within C_{K+1} are removed from further consideration. Thus if {abde} is found valid, then through pruning $V_1 = \emptyset$, and {abde} is identified as the only MFS.

This process of MFCS refinement and C_{K+1} pruning greatly reduces |C| and facilitates MFS convergence, however, in some cases, it can be preemptive and recovery of some candidates may be required. The recovery process is achieved by creating additional C_{K+1}'s for each MFCS member where it is a superset of the $k-1$ prefix of a V_k^i. Given this, a new candidate is created by merging each item in the MFCS that occurs after the last item of the $k-1$ prefix with the $k-1$ prefix. Once recovery has been undertaken, regular Pincer Search analysis continues with the derivation of V_{K+1}.

Max-Eclat and Max-Clique are based on the Eclat and Clique algorithms, however, by using either a DFT or hybrid, DFT-BFT, approach, all MFS are efficiently identified. Using DFT, processing of each decomposed lattice begins with the identified potential maximal itemset whose validity (support) is determined from tidList intersection. If valid, the processing of that lattice is complete, otherwise each lattice subset is checked at the next level of decomposition until all maximal valid itemsets have been identified. The benefit of Max-Clique over Max-Eclat is that the identified potential maximal itemsets are more refined as maximal cliques identify smaller sublattices than the equivalence class method used within Eclat. Hence less traversal is often required to discover all MFS using Max-Clique.

3.3 Improved Apriori Algorithm

DHP (Direct Hash and Pruning) algorithm proposed is trying to use a Hash filter C2 to improve the performance of Apriori in 2 times. They think that C2 is usually the biggest, most of the time consumption of the algorithm to generate frequent 2 - Project set. In fact, C2 is not necessarily the largest, especially the average length of trade is longer and the number of project is less.

In the first scanning, DHP construct a Hash filter. For each 2 – item sets including in a certain transaction, the Hash bucket count corresponding to the item sets plus 1. When scanning the end of this trip, set to get its an upper limit of support count for each 2 –

Project in the transaction database. And then to generate the C2 with L1, Hash the each candidate set. If its support count in hash table is less than the minimum support degree, then delete it. Experiments show that the first scanning time of DHP is much longer than apriori because of construction of Hash filter, but the second scanning time is much faster than apriori obviously. Thus C2 can be expressed as a simple count of two-dimensional array. In this way it will significantly reduce the demand for memory and function call overhead. Any time savings from C2 due to the use of Hash filters will be lost, because of the construction of Hash filters, the using of hush tree and the support degree counting [13-15].

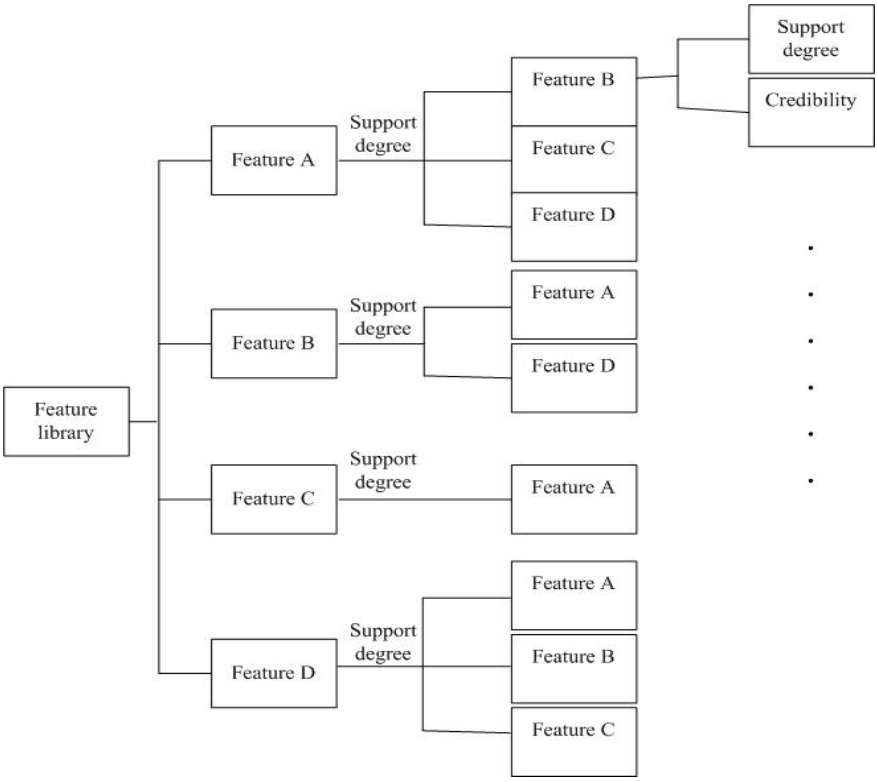


Fig. 1. Schematic diagram of algorithm

4 Experiment

4.1 Experiment Steps

The experiment is supported by financial statements data. The improved apriori algorithm proposed in 3.2 is applied to bill classification. The specific steps are described as follows:

Build the database with the financial statements data of China post. Each data record contains a variety of customer information, including basic personal information, consumer information and records of all kinds of trade details and so on.

Classfy the date in self-building database into different group by using the improved association rules algorithm in 2.2, according to consumer spending and the consumption frequency, the relationship between the two properties are classified.

Composed different databases according to the new data classes.

The accuracy of the financial marketing up to 95% by using the new classification database.

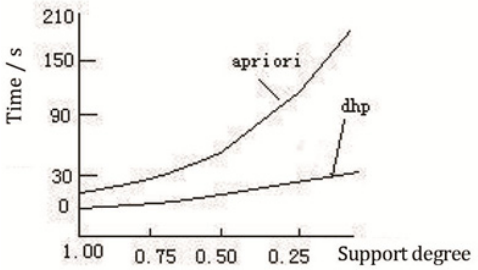


Fig. 2. Time-support degree at different algorithm

4.2 Experimental Results

The self-building database was classified 4 different categories according to different attributes by using the apriori algorithm in 3.2. The first one is the low income category, second category is middle-income class, and the third category is the middle-high-income class, the fourth for the high income group. From a practical point of view, the classification of the results obtained and the actual survey results are very similar. It can be 95%. And contrast with the traditional apriori algorithm, the results of this paper are more accurate and realistic.

5 Conclusion

In this paper, association rules in data mining Apriori algorithm is improved, and the algorithm applied to the postal bills category. By analyzing a collection of users' database, we can get item sets which meeting the minimum support degree, and then form the rule sets. Use of the postal billing data for testing, users get accurate classification results. Compared with the traditional apriori algorithm the results obtained in this paper is more accurate and more realistic.

References

1. Han, J., Pei, J., Yin, Y.: Mining: Frequent patterns without candidate generation. In: Proc. 2000 ACM-SIGMOD Int. Conf. Management of Data (SIGMOD 2000), pp. 1–12 (May 2000)

2. Kamber, M., Han, J., Chang, J.: Metarule-guided mining of multi-dimensional association rules using data cubes. In: Proc. 1997 Int. Conf. Khowledge Discovery and Data Mining (KDD 1997), pp. 207–210 (1997)
3. Kleinberg, J., Papadimitriou, C., Raghavan, P.: Segmentation problems. In: Proceedings of the 30th Annual Symposium on Theory of Computing. ACM (September 1998)
4. Krishnan, B.: Multi-resolution Methods for Data mining and Visualizations 5 (2003)
5. Novobilski, A.J.: Mining Bayesian Networks to Forecast Adverse Out comes Related to Acute Coronary syndrome 1 (2004)
6. Baxt, W.G.: A Neural Network Aid for the Early Diagnosis of Acute Cardiac Ischemia Inpatients Presenting to the Emergency Department with Chest Pain. Ann. Emerg. Med. 40, 595–597 (2002)
7. Luna, J.M., Romero, J.R., Ventura, S.: Design and behavior study of a grammar-guided genetic programming algorithm for mining association rules. Knowledge and Information Systems 32(1), 53–76 (2012)
8. Yu, H., Wen, J., Wang, H., Jun, L.: An Improved Apriori Algorithm Based On the Boolean Matrix and Hadoop. Procedia Engineering 15, 1827–1831 (2011)
9. Cao, X.: An Algorithm of Mining Association Rules Based on Granular Computing. Physics Procedia 33, 1248–1253 (2012)
10. Liao, S.-H., Chen, Y.-J., Deng, M.-Y.: Mining customer knowledge for tourism new product development and customer relationship management. Expert Systems with Applications 37(6), 4212–4223 (2010)
11. Liao, S.-H., Chen, J.-L., Hsu, T.-Y.: Ontology-based data mining approach implemented for sport marketing. Expert Systems with Applications 36(8), 11045–11056 (2009)
12. Shang, E., Ye, L., Fan, X., Tang, Y., Duan, J.: Discovery of Association Rules between TCM Properties in Drug Pairs by Association Mining between Datasets and Probability Tests. World Science and Technology 12(3), 377–382 (2010)
13. Jiang, Y., Wang, J.: An Improved Association Rules Algorithm Based on Frequent Item Sets. Procedia Engineering 15, 3335–3340 (2011)
14. Qodmanan, H.R., Nasiri, M., Minaei-Bidgoli, B.: Multi objective association rule mining with genetic algorithm without specifying minimum support and minimum confidence. Expert Systems with Applications 38(1), 288–298 (2011)
15. Yu, K.-M., Zhou, J.: Parallel TID-based frequent pattern mining algorithm on a PC Cluster and grid computing system. Expert Systems with Applications 37(3), 2486–2494 (2010)

Multi-Point Shortest Path in the Complex Road Network Based on Floyd Algorithm

Ying Li, Donghui Li, and Yaping Yu

College of Sciences, Hebei Union University, Tangshan, 063009, China
liaying7909@163.com
Tianjin Supply Section, Beijing Railway Bureau, Tang'shan, 064000, China
Department of Basic Courses, Tangshan College, Tang'shan, 063000, China

Abstract. The problem of shortest path about multi-point which is based on Complex network model is a classical algorithm in graph theory; the theory is to find the optimal path among each pair of vertices. Firstly, the paper will introduce the research background and significance of the shortest path problem, and analyze the research status and development prospects of the subject; then, the paper will further described Floyd algorithm of the shortest path problem, and give an example based on this algorithm; Finally, the paper will take main campus map of Hebei united university as a network model, Providing a scientific and rational basis for calculating for the realization of the system by the application of the Floyd algorithm.

Keywords: The Shortest Path, Weighted Identification, Net Model, Floyd Algorithm.

1 Introduction

The shortest path is an important issue in graph theory, its purpose is to find the shortest path from vertices of the given weighted graph, and it has many important applications [1, 2]. For example, the shortest path algorithm can be applied to the vehicle scheduling and vehicle navigation in the field of intelligent transport systems [3, 4]; In the field of computer networks [5, 6] and communications, the shortest path algorithm can be used for routing [7, 8]. Although there are a lot of algorithms on the shortest path, the Dijkstra algorithm, the Floyd algorithm and Kruskal's algorithm are undoubtedly the most widely used among all the algorithms [9-12]. Here the paper will introduce the Floyd algorithm which can directly calculate the shortest path between any two networks, the Floyd algorithm is also called distance matrix method for power multiplication or insertion point method, which was proposed in 1962 by Floyd-Warshall. Although it is more complicated than the Dijkstra algorithm, it applies to the calculating of the shortest path between any two in the network, so the application of the algorithm is very extensive [13-16].

2 The Main Idea of the Floyd Algorithm

The Floyd algorithm is a kind of algorithm which is used to solve the shortest path length by matrix calculation when having defined a network weight matrix.

Y. Yang, M. Ma, and B. Liu (Eds.): ICICA 2013, Part II, CCIS 392, pp. 364–372, 2013.

The network weight matrix is sometimes called the distance matrix of the network [17-20].

Suppose there should be the graph G (V, E) which was stored by the adjacency matrix cost, set a two-dimensional array A used to store the length of the shortest path between the current vertex in addition, the component $A[i][j]$ indicates the shortest path between the current vertex v_i and the vertex v_j. The basic idea of Freud algorithm is that a recursive can generate a matrix sequence A_0, A_1, ..., A_k ,..., A_n, the $A[i][j]$ indicates the shortest path length where from the vertex v_i and the vertex v_j the path to go through the vertex number is not greater than K.

Initially $A_{-1}[i][j] = cost[i][j]$, when seeking the shortest path length from the vertex v_i to the vertex v_j on the path through the vertex number is not greater than (k +1), two cases to be considered, one case is that the path go without the number of vertices (k +1);another case is that the path from the vertex v_i to the vertex v_j which goes via vertices (k +1). Then, the path can be divided into two, one is the shortest path goes from vertex v_i to vertex v_{k+1}, another is the shortest goes from vertex v_{k+1} to vertex v_j, the shortest path length is equal to the sum of these two paths. The smaller value in both cases through the required path where the vertex number is not greater than (k +1) from vertex v_i to vertex v_j is the shortest path.

The Freud thought can be described by using the following expression

$$A_{-1}[i][j] = cost[i][j] \tag{1}$$

The formula is an iterative expression, A_k indicates the shortest path including (k+1) vertex that we have considered from vertices 0,1, ..., k (k +1) ,and so on; $A_k[i][j]$ indicates the shortest path going from vertex v_i to vertex v_j which has including vertices 0,1, ..., k (k +1) . On this basis, then we consider the vertex (k +1), the shortest path including each vertex can be calculated considering the vertex (k +1), so we can get the value A_{k+1}.The shortest path from vertex v_i to vertex v_j will include a vertex in every iteration [21, 22]. The valve $A_{n-1}[i][j]$ we got after n times of iterations is the shortest path from vertex v_i to vertex v_j considering all the vertices.

$$A_{k+1}[i][j] = \min\{A_k[i][j], A_k[i][k+1], A_k[k+1][j]\} \quad (-1 \le k \le n-2) \tag{2}$$

If $A_k[i][j]$ has been calculated, the path length which goes from vertex i to vertex(k+1) is $A_k[i][k+1]$, the path length which goes from vertex i to vertex j is $A_k[i][j]$, the path length which goes from vertex (k+1) to vertex j is $A_k[k+1][j]$ .Now we take vertex (k+1) into account, if $A_k[i][k+1] + A_k[k+1][j] < A_k[i][j]$, then the path which goes from vertex i to vertex j will change into the path which goes from vertex i to vertex(k+1),then from vertex (k+1) to vertex j, $A_{k+1}[i][j] = A_k[i][k+1] + A_k[k+1][j]$, Otherwise we does not need to modify the path which goes from vertex i to vertex j.

Besides, we save the shortest path with two-dimensional array “path”, which has relation with the number of iterations [23-25]. When calculating $A_k[i][j]$, the array

" $path[i][j]$ " save the previous number of the shortest path from vertex v_i to vertex v_j in which the number of the intermediate nodes is not greater than K .When the calculating has finished, we will get the shortest path which goes from vertex v_i to vertex v_j by tracing back to two-dimensional array "path". If $path[i][j]=-1$, there is no intermediate nodes [26-29].

For example, given a directed graph in Fig. 1, we will calculate the shortest distance including each vertex.

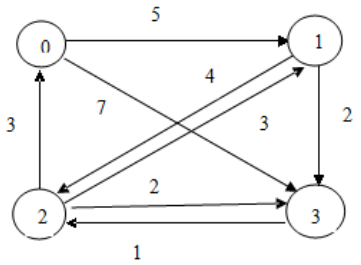


Fig. 1. Directed graph initially

$$A_{-1}=\begin{bmatrix}0&5&\infty&7\\ \infty&0&4&2\\ 3&3&0&2\\ \infty&\infty&1&0\end{bmatrix}\ path_{-1}=\begin{bmatrix}-1&-1&-1&-1\\ -1&-1&-1&-1\\ -1&-1&-1&-1\\ -1&-1&-1&-1\end{bmatrix} \tag{3}$$

Take v_0 into account, $A_0[i][j]$ indicates the shortest path length which goes from the vertex v_i to the vertex v_j via the vertex v_0. By comparison, it does not affect the original shortest path, thus we can get the following array:

$$A_0=\begin{bmatrix}0&5&\infty&7\\ \infty&0&4&2\\ 3&3&0&2\\ \infty&\infty&1&0\end{bmatrix}\ path_0=\begin{bmatrix}-1&-1&-1&-1\\ -1&-1&-1&-1\\ -1&-1&-1&-1\\ -1&-1&-1&-1\end{bmatrix} \tag{4}$$

Take v_1 into account, $A_1[i][j]$ indicates the shortest path length which goes from the vertex v_i to the vertex v_j via the vertex v_1. By comparison, we modify the path to $v_0-v_1-v_2$, the path length is 9 which are shorter than before, next we will change $A_0[0][2]$ into 9 and change $path_0[0][2]$ into 1, and therefore we will get the following arrays [30, 31]:

$$A_1=\begin{bmatrix}0&5&9&7\\ \infty&0&4&2\\ 3&3&0&2\\ \infty&\infty&1&0\end{bmatrix}\ path_1=\begin{bmatrix}-1&-1&1&-1\\ -1&-1&-1&-1\\ -1&-1&-1&-1\\ -1&-1&-1&-1\end{bmatrix} \tag{5}$$

Similarly, we will get the following arrays:

$$A_2 = \begin{bmatrix} 0 & 5 & 9 & 7 \\ 7 & 0 & 4 & 2 \\ 3 & 3 & 0 & 2 \\ 4 & 4 & 1 & 0 \end{bmatrix} path_2 = \begin{bmatrix} -1 & -1 & 1 & -1 \\ 2 & -1 & -1 & -1 \\ -1 & -1 & -1 & -1 \\ 2 & 2 & -1 & -1 \end{bmatrix} \tag{6}$$

The array A_3 is the shortest distance between each vertex.

3 The Application of the Shortest Path Algorithm in the Path Information of the Campus

With the continuous development of the society, now the size of university campus is constantly expanding, the information is constantly enriched too. But it also brings some difficulties to the people who want to understand the school, especially when the new students register or others visit the campus, it is difficult for them to check the campus' information. So, it is necessary to develop a kind of intelligent campus information inquiry system.

However, the information query on campus path is a very important function module. Although many college campuses have placed road signs now, but those road signs only mark out the current location which can only tell us where we can go to ,and which can only mark a rough direction. So the road signs can't satisfy the visitor's needs well, the system we want to design can make the visitors learn the exact location of the place rapidly they want to query, and the system can show the bird's-eye view, it can also provide the shortest path of the place where the visitors want to go according to the visitor's current location. The approach is both intuitive and efficient, and it achieves the functionality of the tour guides.

3.1 Modeling

Now, we use the map of Hebei United University headquarters to demonstrate the problem. Considering that the path of the campus is non-isotropic, so we can abstract it as a non-isotropic graph to solve the problem, the shortest path algorithm model between any pair of vertices of is what we want to solve right now. In the non-isotropic graph, 1 represents the front door, 2 represents the administrative building, 3 represents the auditorium, 4 represents the building of Science and Technology, 5 represents metallurgy building,6 represents the comprehensive building, 7 represents laboratory building, 8 represents the fifth teaching building, 9 represents girls apartment, 10 represents the fourth teaching building , 11 represents library ,12 represents the canteen ,13 represents the ruins of the Tangshan earthquake,14 represents the sports ground , 15 represents boys apartment. The connection line between two nodes shows that there is a direct access between them; the value above the line represents the length of the path between them. Shown in Fig. 2:

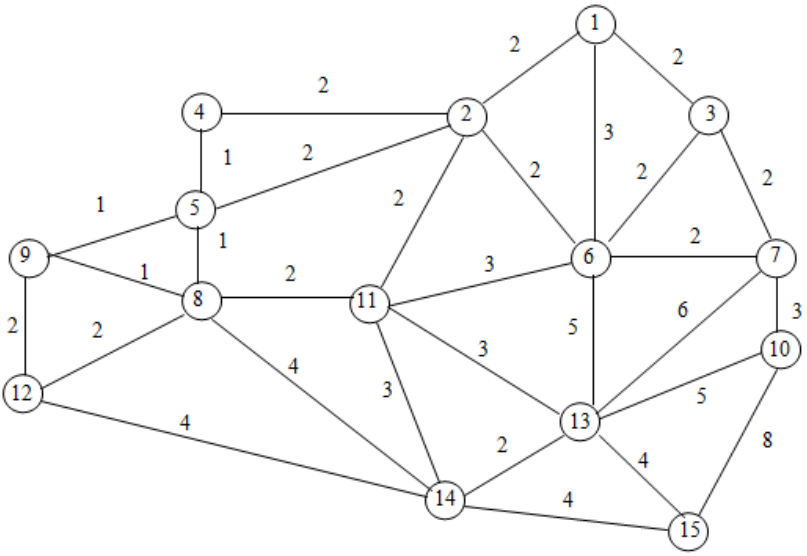


Fig. 2. The abstracted non-isotropic graph

3.2 The Solution of the Model

We can list in the initial Adjacency matrix according to the Fig. 2:

$$\begin{bmatrix}
0 & 2 & 2 & \infty & \infty & 3 & \infty & \infty & \infty & \infty & \infty & \infty & \infty & \infty & \infty \\
2 & 0 & \infty & 2 & 2 & 2 & \infty & \infty & \infty & \infty & 2 & \infty & \infty & \infty & \infty \\
2 & \infty & 0 & \infty & \infty & 2 & 2 & \infty & \infty & \infty & \infty & \infty & \infty & \infty & \infty \\
\infty & 2 & \infty & 0 & 1 & \infty & \infty & \infty & \infty & \infty & \infty & \infty & \infty & \infty & \infty \\
\infty & 2 & \infty & 1 & 0 & \infty & \infty & 1 & 1 & \infty & \infty & \infty & \infty & \infty & \infty \\
3 & 2 & 2 & \infty & \infty & 0 & 2 & \infty & \infty & \infty & 3 & \infty & 5 & \infty & \infty \\
\infty & \infty & 2 & \infty & \infty & 2 & 0 & \infty & \infty & 3 & \infty & \infty & 6 & \infty & \infty \\
\infty & \infty & \infty & \infty & 1 & \infty & \infty & 0 & 1 & \infty & 2 & 2 & \infty & 4 & \infty \\
\infty & \infty & \infty & \infty & 1 & \infty & \infty & 1 & 0 & \infty & \infty & 2 & \infty & \infty & \infty \\
\infty & \infty & \infty & \infty & \infty & \infty & 3 & \infty & \infty & 0 & \infty & \infty & 5 & \infty & 8 \\
\infty & 2 & \infty & \infty & \infty & 3 & \infty & 2 & \infty & \infty & 0 & \infty & 3 & 3 & \infty \\
\infty & \infty & \infty & \infty & \infty & \infty & \infty & 2 & 2 & \infty & \infty & 0 & \infty & 4 & \infty \\
\infty & \infty & \infty & \infty & \infty & 5 & 6 & \infty & \infty & 5 & 3 & \infty & 0 & 2 & 4 \\
\infty & \infty & \infty & \infty & \infty & \infty & \infty & 4 & \infty & \infty & 3 & 4 & 2 & 0 & 4 \\
\infty & \infty & \infty & \infty & \infty & \infty & \infty & \infty & \infty & 8 & \infty & \infty & 4 & 4 & 0
\end{bmatrix}$$

The specific solution procedure will not be listed, because we have made a detailed introduction about the Flody algorithm, here we list the main C language code:

```
typedef struct
int edges[MAXV][MAXV];
int n,e;
}MGraph;
void Floyd(MGraph g)
int A[15][15],path[15][15];
int i,j,k,n=g.n;
for(i=0;i<n;i++)
for(j=0;j<n;j++)
A[i][j]=g.edges[i][j];
path[i][j]=-1;
For (k=0;k<n;k++)
For (i=0; i<n;i++)
For (j=0; j<n;j++)
If (A[i][j]>(A[i][k]+A[k][j]))
A[i][j]=A[i][k]+A[k][j];
Path[i][j]=k;
```

The model will be solved in Visual C + + 6.0 environments:

The shortest path length between each vertex will be showed by the following matrix A:

$$A = \begin{bmatrix} 0 & 2 & 2 & 4 & 4 & 3 & 4 & 5 & 5 & 7 & 4 & 7 & 7 & 7 & 11 \\ 2 & 0 & 4 & 2 & 2 & 2 & 4 & 3 & 3 & 7 & 2 & 5 & 5 & 5 & 9 \\ 2 & 4 & 0 & 6 & 6 & 2 & 2 & 7 & 7 & 5 & 5 & 9 & 7 & 9 & 11 \\ 4 & 2 & 6 & 0 & 1 & 4 & 6 & 2 & 2 & 9 & 4 & 4 & 7 & 9 & 10 \\ 4 & 2 & 6 & 1 & 0 & 4 & 6 & 1 & 1 & 9 & 3 & 3 & 6 & 5 & 9 \\ 3 & 2 & 2 & 4 & 4 & 0 & 2 & 5 & 5 & 5 & 3 & 7 & 5 & 6 & 9 \\ 4 & 4 & 2 & 6 & 6 & 2 & 0 & 7 & 7 & 3 & 5 & 9 & 6 & 8 & 10 \\ 5 & 3 & 7 & 2 & 1 & 5 & 7 & 0 & 1 & 10 & 2 & 2 & 5 & 4 & 8 \\ 5 & 3 & 7 & 2 & 7 & 5 & 7 & 1 & 0 & 10 & 3 & 2 & 6 & 5 & 9 \\ 7 & 9 & 5 & 9 & 9 & 5 & 3 & 10 & 11 & 0 & 8 & 11 & 5 & 7 & 8 \\ 4 & 2 & 5 & 4 & 3 & 3 & 5 & 2 & 3 & 8 & 0 & 4 & 3 & 3 & 7 \\ 7 & 5 & 9 & 4 & 3 & 7 & 9 & 2 & 2 & 11 & 4 & 0 & 6 & 4 & 8 \\ 7 & 5 & 7 & 7 & 6 & 5 & 6 & 5 & 6 & 5 & 3 & 6 & 0 & 2 & 4 \\ 7 & 5 & 9 & 6 & 5 & 6 & 8 & 4 & 5 & 7 & 3 & 4 & 2 & 0 & 4 \\ 11 & 9 & 11 & 10 & 9 & 9 & 10 & 8 & 9 & 8 & 7 & 8 & 4 & 4 & 0 \end{bmatrix}$$

The shortest path between each vertex will be showed by the following matrix path

We can get the shortest path between any two points on campus from the above results. If we want to reach the fourth teaching building from the front door, we have to find the shortest path between $v_1 - v_{10}$, we find out that $A[0][9]=7$, $path[0][9]=6$, $path[0][6]=2$, $path[0][2]=-1$, so we learn that the shortest path length from the front door to the fourth teaching building is 7, the path is $v_1 - v_3 - v_7 - v_{10}$. That is to say, the shortest path when we set out from the front door is through the auditorium and then through the laboratory building.

$$path = \begin{bmatrix} -1 & -1 & -1 & 1 & 1 & -1 & 2 & 4 & 4 & 6 & 1 & 7 & 10 & 10 & 12 \\ -1 & -1 & 0 & -1 & -1 & -1 & 5 & 4 & 4 & 6 & -1 & 7 & 10 & 10 & 12 \\ -1 & 0 & -1 & 1 & 1 & -1 & -1 & 4 & 4 & 6 & 5 & 7 & 5 & 10 & 12 \\ 1 & -1 & -1 & -1 & -1 & 1 & 5 & 4 & 4 & 6 & 7 & 7 & 10 & 7 & 13 \\ 1 & -1 & -1 & -1 & -1 & 1 & 5 & -1 & -1 & 6 & 7 & 7 & 10 & 7 & 13 \\ -1 & -1 & -1 & 1 & 1 & -1 & -1 & 4 & 4 & 6 & -1 & 7 & -1 & 10 & 12 \\ 2 & 5 & -1 & 5 & 5 & -1 & -1 & 5 & 5 & -1 & 5 & 7 & -1 & 10 & 12 \\ 4 & 4 & 4 & 4 & -1 & 4 & 5 & -1 & -1 & 6 & -1 & -1 & 10 & 7 & 13 \\ 4 & 4 & 4 & 4 & -1 & 4 & 5 & -1 & -1 & 6 & 7 & -1 & 10 & 7 & 13 \\ 6 & 6 & 6 & 6 & 6 & 6 & -1 & 6 & 6 & -1 & 6 & 13 & -1 & 12 & -1 \\ 1 & -1 & 5 & 1 & 7 & -1 & 5 & -1 & 7 & 6 & -1 & 7 & -1 & -1 & 12 \\ 7 & 7 & 7 & 7 & 7 & 7 & 7 & -1 & -1 & 13 & 7 & -1 & 13 & -1 & 13 \\ 10 & 10 & 5 & 10 & 10 & -1 & -1 & 10 & 10 & -1 & -1 & 13 & -1 & -1 & -1 \\ 10 & 10 & 10 & 7 & 7 & 10 & 10 & -1 & 7 & 12 & -1 & -1 & -1 & -1 & -1 \\ 12 & 12 & 12 & 13 & 13 & 12 & 12 & 13 & 13 & -1 & 12 & 13 & -1 & -1 & -1 \end{bmatrix}$$

4 The Conclusion

The effect of the Floyd algorithm which was used in the calculation of the dense graph is better [9, 10]. Using this algorithm, we can calculate not only the figure whose edge weight is plus, but also the figure whose edge weight is negative. The algorithm is simple and effective, and it is one of the classical algorithms that can be used to solve the shortest path. The ways of solving the shortest path problem are various, and every one has advantages and disadvantages. We can take the appropriate algorithm flexibly for different problems. Even if we use the same algorithm, we can do some modifications on the algorithm itself, or we can integrate the idea of the other algorithm with the algorithm we are using, etc. Thus, we will get the satisfied effect about the solving.

References

1. Li, S.-J., et al.: Study of Shortest Path Based on Weighted Identification S-graph. Computer Engineering 37(15), 46–48 (2011)
2. Gao, L.-Q., et al.: The Application of Harmony Search Algorithm for Solving Shortest Path Problems. Journal of Northeastern University (Natural Science) 32(6), 769–772 (2011)

3. Hribar, M.R., Taylor, V.E., Boyce, D.E.: Implementing parallel shortest path for parallel transportation applications. Parallel Computing 27(12), 1537–1568 (2001)
4. Parichart, P., Dongjoo, P., Laurence, R.R., et al.: Dynamic and stochastic shortest path in transportation networks with two components of travel time uncertainty. Transportation Research Part C: Emerging Technologies 11(5), 331–354 (2003)
5. Desaulniers, G., Soumis, F.: An efficient algorithm to find a shortest path for a car-like robot. IEEE Transactions on Robotics and Automation 11(6), 819–828 (1995)
6. Ahmed, Y.H.: A genetic algorithm for finding the k shortest paths in a network. Egyptian Informatics Journal 11(2), 75–79 (2010)
7. Campos, R., Ricardo, M.: A Fast Algorithm for Computing Minimum Routing Cost Spanning Trees. Computer Networks 52(17), 3229–3247 (2008)
8. Shioda, S., Ohtsuka, K., Sato, T.: An Efficient Network-wide Broad-casting based on Hop-limited Shortest-path Trees. Computer Networks 52(17), 3284–3295 (2008)
9. Yue, Q.-J., et al.: Design and Realization of Floyd Algorithm System Based on the Optimation of the Shortest Path Problem. Journal of Gansu Normal Colleges 15(5), 41–43 (2010)
10. Yang, J.-Q., et al.: Problem solving of shortest path between the institutions Based on Freud algorithm. Scienfic & Technical Information of Gansu 39(5), 28–29 (2010)
11. Burrough, P.A., McDonnell, R.A.: Principles of Geographical Information Systems. Oxford University Press, New York (1998)
12. Dijkstra, E.W.: A note on two problems in connexion with graphs. Numerische Mathematik 1, 269–271 (1959)
13. Filho, G.G., Samet, H.: A linear iterative approach for hierarchical shortest path finding. Computer Science Department CS-TR-4417, University of Maryland, College Park, MD (November 2002)
14. Floyd, R.W.: Algorithm 97: Shortest path. Communications of the ACM 5(6), 345 (1962)
15. Fredman, M.L., Tarjan, R.E.: Fibonacci heaps and their uses in improved network optimization algorithms. Journal of the ACM (JACM) 34(3), 596–615 (1987)
16. Freeman, H.: Computer Processing of Line-Drawing Images. ACM Computing Surveys (CSUR) 6(1), 57–97 (1974)
17. Gargantini, I.: An effective way to represent quadtrees. Communications of the ACM 25(12), 905–910 (1982)
18. Goldberg, A.V., Harrelson, C.: Computing the shortest path: A search meets graph theory. In: Proceedings of the Sixteenth Annual ACM-SIAM Symposium on Discrete Algorithms, Vancouver, British Columbia, January 23-25 (2005)
19. Goldberg, A.V., Werneck, R.F.: Computing point-to-point shortest paths from external memory. In: ALENEX 2005: Proceedings of the 7th Workshop on Algorithm Engineering and Experiments, Vancouver, BC, Canada (Janyary 2005)
20. Gupta, S., Kopparty, S., Ravishankar, C.: Roads, codes, and spatiotemporal queries. In: Proceedings of the Twenty-Third ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems, Paris, France, June 14-16 (2004)
21. Guttman, A.: R-trees: a dynamic index structure for spatial searching. In: Proceedings of the 1984 ACM SIGMOD International Conference on Management of Data, Boston, Massachusetts, June 18-21 (1984)
22. Henzinger, M.R., Klein, P., Rao, S., Subramanian, S.: Faster shortest-path algorithms for planar graphs. Journal of Computer and System Sciences 55(1), 3–23 (1997)
23. Hjaltason, G.R., Samet, H.: Ranking in Spatial Databases. In: Egenhofer, M., Herring, J.R. (eds.) SSD 1995. LNCS, vol. 951, pp. 83–95. Springer, Heidelberg (1995)

24. Hjaltason, G.R., Samet, H.: Incremental distance join algorithms for spatial databases. In: Proceedings of the 1998 ACM SIGMOD International Conference on Management of Data, Seattle, Washington, United States, June 01-04, pp. 237–248 (1998)
25. Hjaltason, G.R., Samet, H.: Distance browsing in spatial databases. ACM Transactions on Database Systems (TODS) 24(2), 265–318 (1999)
26. Hutchinson, D., Maheshwari, A., Zeh, N.: An external memory data structure for shortest path queries. Discrete Applied Mathematics 126(1), 55–82 (2003)
27. Jing, N., Huang, Y.-W., Rundensteiner, E.A.: Hierarchical Encoded Path Views for Path Query Processing: An Optimal Model and Its Performance Evaluation. IEEE Transactions on Knowledge and Data Engineering 10(3), 409–432 (1998)
28. Kolahdouzan, M.R., Shahabi, C.: Voronoi-based k nearest neighbor search for spatial network databases. In: VLDB 2004: Proceedings of the 30th International Conference on Very Large Data Bases, Toronto, Canada, pp. 840–851 (September 2004)
29. Micikevicius, P.: General parallel computation on commodity graphics hardware: case study with the all-pairs shortest paths problem. In: PDPTA 2004: Proceedings of the International Conference on Parallel and Distributed Processing Techniques and Applications, Las Vegas, NV, USA, pp. 1359–1365 (June 2004)
30. Mitchell, J.S.B., Mount, D.M., Papadimitriou, C.H.: The discrete geodesic problem. SIAM Journal on Computing 16(4), 647–668 (1987)
31. Nelson, R.C., Samet, H.: A consistent hierarchical representation for vector data. ACM SIGGRAPH Computer Graphics 20(4), 197–206 (1986)

Modeling and Simulation of Semi-Active Suspension for Vehicle

Zhao Chang Li, Zhang Zhu Lin, and Li Ai Juan

[1] Shandong Jiaotong University, Jinan, China
[2] Key Laboratory of Vehicles Detection, Dignosis & Maintenance Technology, Jinan, China
1661589041@qq.com

Abstract. In order to shorten the development cycle of semi-active suspension, a batch processing modeling method is studied. 1/4 of the semi-active suspension model is constructed based on AMESim software in this modeling method. Through changing model's parameters, the simulation results are compared and the system's performance is analyzed. The compared simulation results showed that this paper's modeling method is feasible and right, using this method to development semi-active suspension can short development cycle and the semi-active suspension's performance can be improved.

Keywords: Semi-active Suspension, Batch process, Modeling, Simulation.

1 Introduction

Suspension is the power transmission connection device's general term between the vehicle body and the wheel. It delivers the bearing force, the traction force, braking force, lateral force and the torque to the vehicle body and make the vehicle can move forward[1, 2]. Several forces and torque are generated by the road works on the wheels. The suspension system is an important constituent of vehicle and the performance of the suspension influent the vehicle's smooth performance, handling stability and security[3, 4]. Therefore the design of suspension system is an important part of the whole vehicle's design. As computer technology's improving, simulation technology plays an important role in the study of the vehicle's suspension system. A lot of special software applies to the mechanical components' motion simulation, such as Adams and Matlab. Some of the studies have achieved a lot of research results.

AMESim (Advanced Modeling Environment for Performing Simulations of Engineering Systems) is a simulation modeling software developed by France's IMAGINE company. It provides a perfect comprehensive simulation environment and flexible solution method for engineering system. The engineering system includes fluid, liquid, gas, machinery, control, electromagnetic, and so on [5, 6]. The simulation software has rich model library, users can build a custom module according to the actual physical system or simulation model [7-10].

This paper build the vehicle's AMESim simulation model of 1/4 model semi-suspension system. The suspension's parameters is changed by this batch process

Y. Yang, M. Ma, and B. Liu (Eds.): ICICA 2013, Part II, CCIS 392, pp. 373–383, 2013.

method, through the analysis of the simulation results, the influence of several parameters on the suspension performance is obtained. The paper will play a role of guide for the design of the vehicle's suspension system.

2 The Model

In order to study the active suspension system's performance, the 1/4 semi-active suspension model is set to the research objective [11-14]. The simplified model of two freedom degrees suspension can be expressed in the Fig. 1. According to the vehicle theory knowledge, the differential equation of suspension double weight system is

$$\begin{aligned} m_2\ddot{z}_2 + C\left(\dot{z}_2 - \dot{z}_1\right) + K\left(z_2 - z_1\right) &= 0 \\ m_1\ddot{z}_1 + C\left(\dot{z}_1 - \dot{z}_2\right) + K\left(z_1 - z_2\right) + K_t\left(z_1 - q\right) &= 0 \end{aligned} \tag{1}$$

Where m2 defines the 1/4 sprung mass, m1 is the unsprung mass, k is the suspension spring stiffness, kt is the wheel spring stiffness, C is the damper.

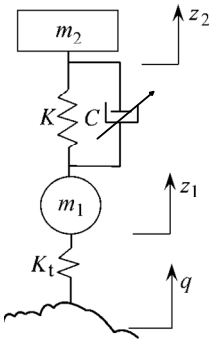


Fig. 1. The 1/4 semi-active suspension model

Our countries' high grade highway pavement is divided into three scopes of A, B and C grades basically [15-18]. Where the B and C grades road have a larger proportion than A grade road. The road input signal is white noise random road input integral signal. When the driving velocity is steady, the road velocity's input is white noise in the time domain. When the velocity is constant (v=20 m/s), the vehicle's vertical power spectrum is constant, so the surface profile can be generated by the white noise. The white noise is generated by an integrator. The random road's white noise input model is set by the AMESim software and the model is presented as Fig. 2.

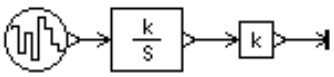


Fig. 2. The road's input signal

3 The Simulation and Analysis

According to the 1/4 semi-suspension model, the components' sub-model is constructed in the AMESim's model library [19-21]. The system's simulation sub-model is shown as Fig. 3. The absorber of the passive suspension is

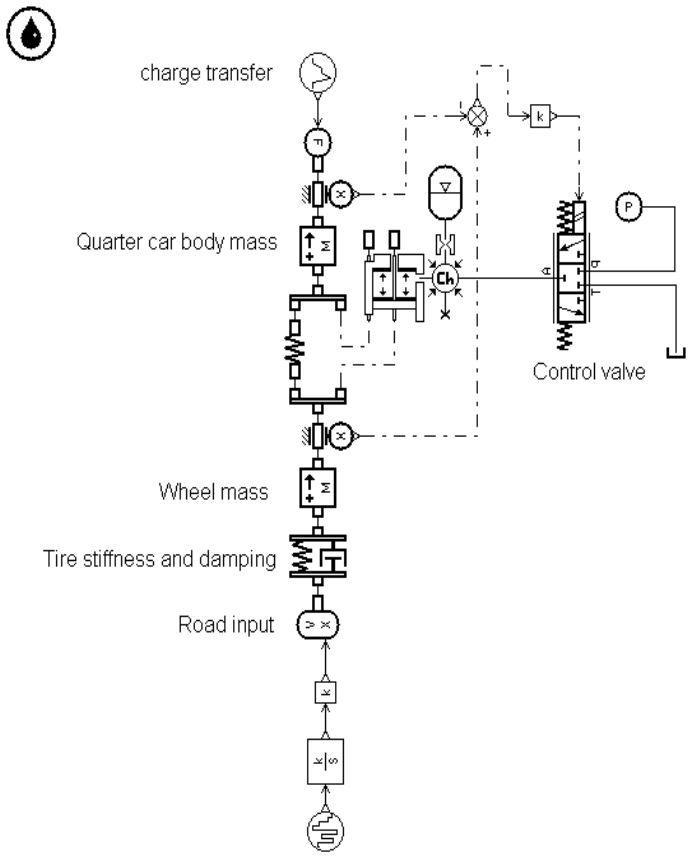


Fig. 3. The AMESim model of the vehicle's semi-active suspension

Replaced by hydraulic power component in the model. The absorber's damp is changed through the control of the hydraulic power component. The suspension system is optimized.

Under the sub-model pattern, the ideal model is selected for the system's components [22-25]. The constant fuel supply pressure p=70 bar; the diameter of the hydraulic cylinder piston is 5 mm.

Under the running mode, the running time is set to 5s and the time interval is set to 0.01s. After the running mode is set to dynamic, the simulation begins. The simulation results of the vehicle body's vertical displacement and acceleration time domain just as Fig. 4.

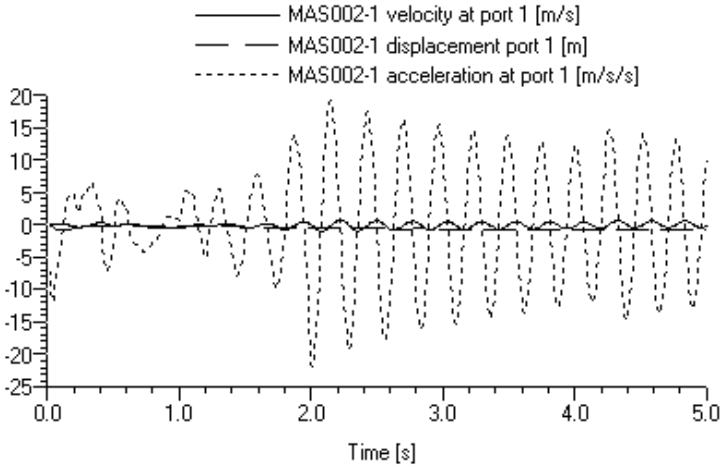


Fig. 4. Simulation results of vehicle's body, velocity and acceleration

4 The Influence of Suspension Parameters for Performance

In order to obtain the influence sensitivity of several parameters on the suspension performance, the hydraulic cylinder and servo valve's parameters are regulated through the batch mode [26, 27]. The response of different parameters on the acceleration can be obtained through the simulation.

4.1 The Influence of Piston Diameter Is Modified and the Simulation is Run

The results are presented as Fig. 5 and Fig. 6.

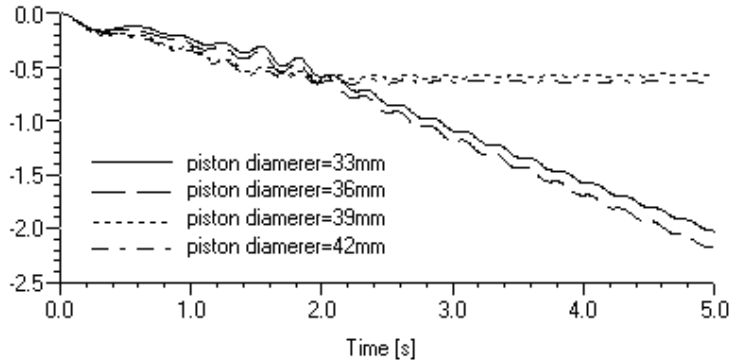


Fig. 5. Simulation results of vehicle body's displacement under different piston diameters

Fig. 5 and Fig. 6 showed that as the piston diameter increases, the system damping is reduced and the steady state error is increased. But in order to ensure the system's reliability, it is difficult to make the piston diameter very small in the practical application.

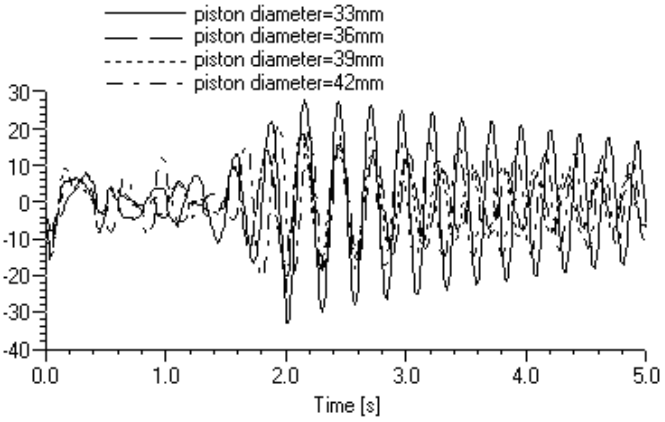


Fig. 6. Simulation results of vehicle body's acceleration under different piston diameters

4.2 The Influence of the Piston Rod Diameter

The piston rod diameter is modified and the simulation is run, the simulation results are shown as Fig. 7 and Fig. 8.

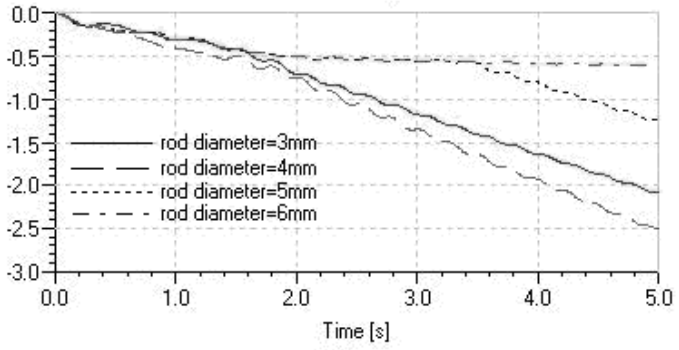


Fig. 7. Simulation results of vehicle body's displacement under different piston rod diameters

Fig. 7 and Fig. 8 showed that as the piston rod diameter increases, the system damping is reduced and the steady state error is increased.

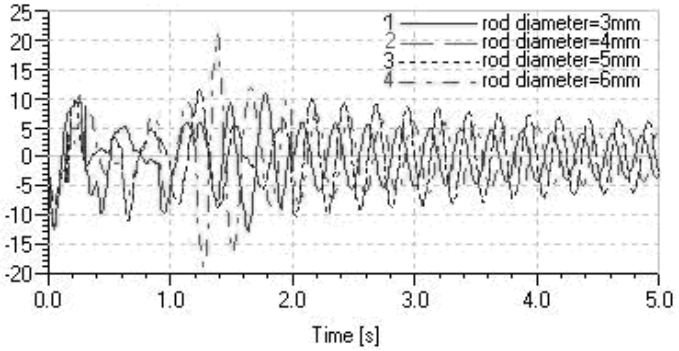


Fig. 8. Simulation results of vehicle body's acceleration under different piston rod diameters

4.3 Changing the Accumulator Gas Volume

Fig. 9 and Fig. 10 is the accumulator volume on the influence of the body's displacement and velocity. The Fig. 9 and Fig. 10 showed that: the larger the accumulator air volume, the bigger the body displacement. Taking into consideration of the steering stability, the accumulator volume should not be too big.

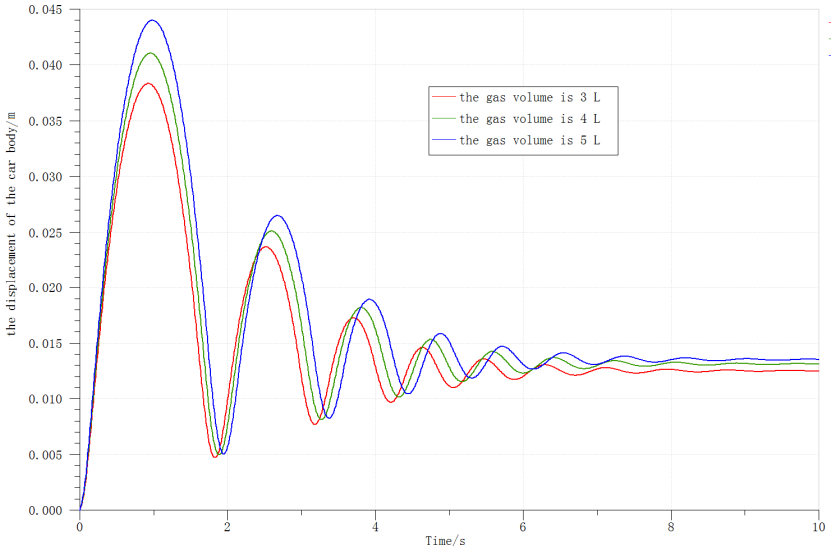


Fig. 9. Accumulator volume on the influence of body displacement

4.4 Changing the Accumulator Gas Initial Pressure

Fig. 11 and Fig. 12 are the accumulator initial pressure on the influence of body displacement and velocity. Fig. 11 and Fig. 12 showed that: the smaller the initial pressure of the accumulator, the bigger displacement and speed of the body. At the same time, the accumulator initial pressure couldn't be too big, because the damping effect and the ride comfort may be affected if the accumulator's initial pressure is too big.

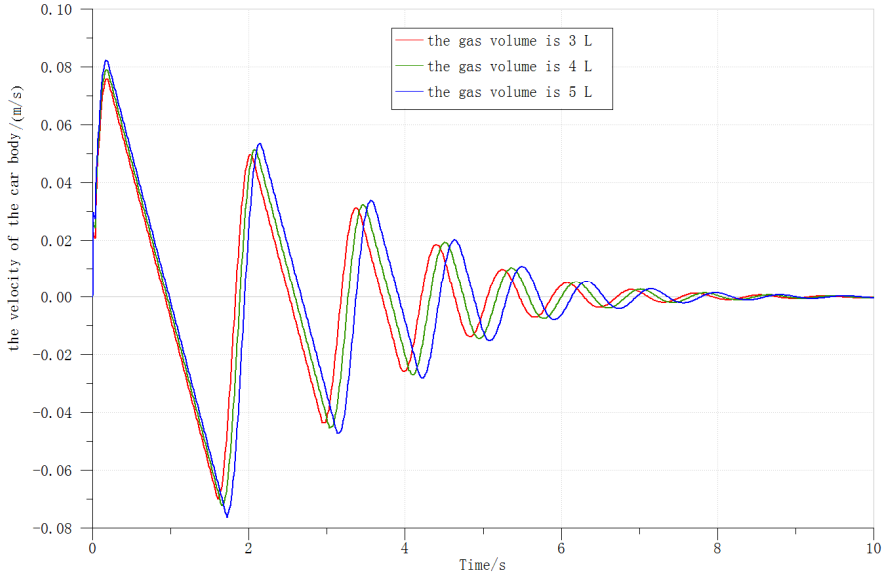


Fig. 10. Accumulator volume on the influence of body velocity

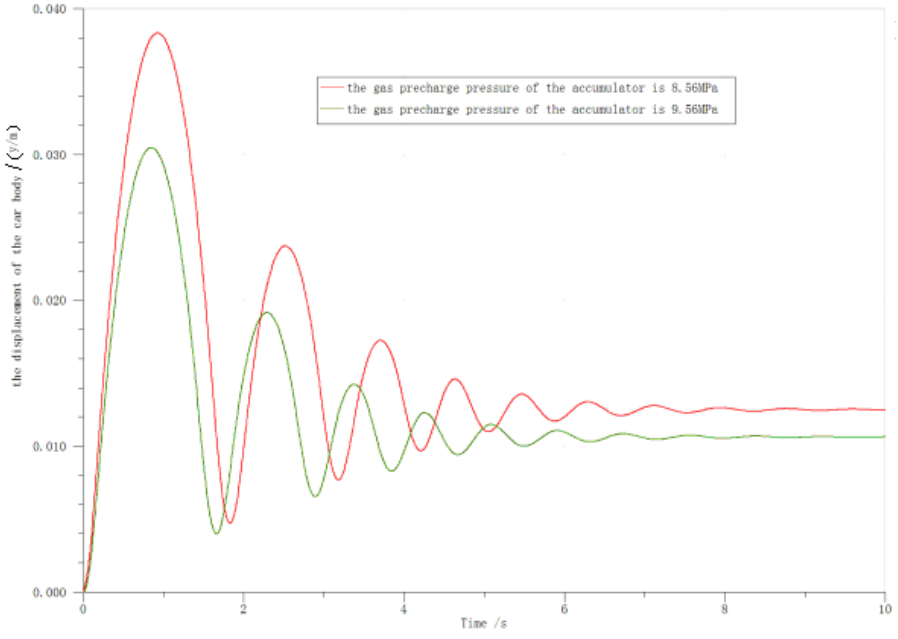


Fig. 11. Accumulator initial pressure on the influence of body displacement

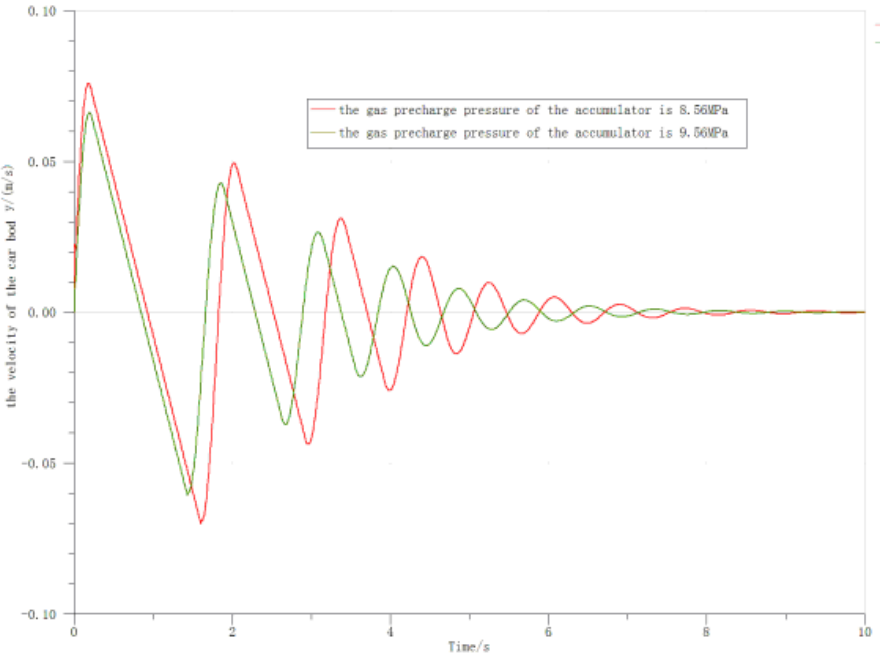


Fig. 12. Accumulator initial pressure on the influence of body velocity

The comparison simulation results showed that the hydraulic suspension damper specification parameters directly determine the hydraulic suspension's performance under white noise random road input signal.

5 Conclusions

In this paper a vehicle's semi-suspension model is built and the simulation analysis is implemented. The conclusions can be summarized as follows:

(1) The AMESim simulation software can obtain favorable effect in the hydraulic suspension modeling and simulation. It has great help for system optimization and parameter selection, the analysis of the design effect and efficiency can be improved.

(2) The batch process method of simulation and analysis is feasible. The batch process mode is used to provide the comparison results of different influence factor.

(3)The simulation results showed that the hydraulic suspension damper specification parameters directly determine the performance. The batch process method is right and feasible, using this method to development semi-active suspension can short development cycle and the semi-active suspension's performance can be improved.

Evaluation of vehicle ride comfort is an important indicator of performance of the car, how to improve vehicle ride comfort has been a car designer hot issues of concern. The assembly of the vehicle suspension system as one plays dampen road shock

attenuation due to body vibration caused by uneven road surface such as the role, and its performance is good or bad ride comfort of the vehicle has a very important impact. Semi-active suspension because not only overcome the performance limitations of passive suspension and active suspension to overcome the high cost, high energy shortcomings, with low energy consumption, low cost, can significantly improve suspension performance advantages, such as research hot spots. This article relies on key research projects, Zhejiang Province, "multi-vehicle chassis platform sharing of digital technology and application development", using ADAMS / View and MATLAB / Simulink software relative to the joint, the semi-active suspension multibody system dynamics model and its control system model combining joint simulation. Firstly, the use of multi-body dynamics software ADAMS View module to establish a vehicle suspension system of multi-body system dynamics model. Based on fuzzy control theory, using MATLAB software to design two-input single-output fuzzy controller. Fuzzy control and PID control both the advantages of a semi-active suspension design fuzzy PID controller. Joint ADAMS / View created in two degrees of freedom suspension system dynamics model and MATLAB / Simulink control system model created by ADAMS / Controls interface suspension multi-body dynamics model into MATLAB, the co-simulation. With body acceleration, suspension dynamic deflection, dynamic displacement of the wheel evaluation, the random road input and sinusoidal road input conditions, respectively, passive suspension, semi-active suspension fuzzy control, fuzzy PID control has been semi-active suspension Joint Based on ADAMS and MATLAB simulation results show that: semi-active suspension system is superior to the passive suspension system; compared with fuzzy control, fuzzy PID controller has better control effect, greatly reducing the road to face body vibration and shock, a significant improvement in the vehicle's ride. Co-simulation approach avoids the cumbersome vehicle dynamics equations to establish and control system transfer function to solve the semi-active suspension mathematical model of the problem, can reduce the physical prototype testing session of the trial, the number of trials in the product development phase, reducing design costs and shorten the design cycle for complex vehicle system control and simulation provides a new idea. This semi-active suspension for co-simulation study on the research and development of semi-active suspension system with some guidance and reference value.

References

1. Maric, D.M., Meier, P.F., Estreicher, S.K.: Mater. Sci. Forum 83-87, 119 (1992)
2. Yanhua, S., Junfeng, Z., Jue, Y., et al.: Research on test and simulation of hydro-pneumatic suspension. IEEE Consumer Electronics, Communications and Networks 80, 678–681 (2011)
3. Longxin, Z., Wenming, Z.: Research on simulation and experiment of hydro-pneumatic suspension with single gas cell. Journal of Mechanical Engineering 45, 290–294 (2009)
4. You, Z., Fan, Y., Tao, S.: Simulation of nonlinearhydro-pneumatic suspension for comfort. Computer Simulation 21, 157–159 (2004)

5. Wamer, B., Raltheja, S.: An analytical and experimental investigation of friction and gas spring characteristics of racing car suspension dampers. SAE962548, pp. 489–502
6. Konghui, G., Yuhang, C., Yehai, Y.: Modeling and simulation of a hydro-pneumatic spring based on internal characteristics. In: International Conference Mechanic on Automation and Control Engineering, pp. 5910–5915. IEEE, USA (2011)
7. Yu, Z.S.: Theory of Automobile. China Machine Press, China (2009)
8. Chen, W., Sun, C.: Automotive Semi-Active Suspension System Using Neural Network Adaptive PID Control. Anhui Institute of Technology (April 1997)
9. Qing, D., Lin, Y.M., Cao, S., Zhao, Y.: Automotive semi-active suspension control hydrocarbon Study. Machine Tool & Hydraulics (February 2003)
10. Chen, Mou, Y., Wen, X., Haobin, J., Chen, Y.: Based on TS fuzzy model semi-active suspension control study. Jiangsu University, Natural Science (May 2004)
11. Li, M.: Based on PID control and optimization of the main suspension design. Machinery Design & Manufacture (August 2007)
12. Wang, H., Shi, L., Liu, W., Sun, J.: Nonlinear vehicle suspension system Active Vibration Control. Mechanical Strength (March 2000)
13. Wang, H., Guo, L.: Fuzzy Immune PID nonlinear vehicle suspension control strategy and simulation study. Mechanical Strength (June 2008)
14. Zhang, D., Xia, Z., Guo, S., Bo, Z.: School; Vehicle Semi-Active Suspension Fuzzy Control and Simulation. Machine Design (September 2008)
15. Chen, W.Q., Li, B.: Based on multi-body model vehicle semi-active suspension control method. Journal of Mechanical Engineering (January 2004)
16. Li, R., Yu, M., Chen, W., Rong, L., Dong, X.M.: Magnetorheological damper on vehicle suspension vibration control. Journal of Mechanical Engineering (June 2005)
17. Chen, X.: LAW FORUM; Anhu Ping; Baiyue Fei; fuzzy control car simulation of semi-active suspension. Mechanical Research & Application (May 2006)
18. Wei, Zhao, B.: Precision air hydrostatic drive system rail assembly design and analysis. In: Fifteenth Session of the Fluid Power and Mechatronic Control Engineering Conference Proceedings (2011)
19. Gao, W.T., Qiao, M., Zhoa, F., Lu, Z.: Whole; automotive steering wheel kingpin inclination angle design calculations. In: 2004 "Anhui Manufacturing Development," Dr. Technology Forum Proceedings (2004)
20. Ren, G.: HFC6700 minibuses Braking Process Simulation. In: Anhui fifth "Hing Wan Light" Youth Annual Meeting Proceedings (Engineering Volume) (2005)
21. Yu, S., Zheng, Q., Chen, L.: Based on ADAMS / Car Ride Vehicle Ride Comfort Simulation Study. In: Theory and Practice of Modern Agriculture - Dr. Anhui Modern Agricultural Science and Technology Forum Proceedings (2007)
22. Wang, L., Du, J.: Based on state feedback pole assignment 4WS system design and simulation. In: Beijing Mechanics will be the 18th Annual Conference Proceedings (2012)
23. Hu, C.: Applicable Scope; Yan Hui; a new electric vehicle smart charging machine development. In: Fourteenth Session of the 19th Annual Meeting of China Association for Science and Venue: Electric Vehicle Charging and Discharging Technology Conference Proceedings (2012)
24. Wei, J., Hao, S.: On the clutch slave cylinder bracket Optimization. In: 2011 Western Auto Industry & Academic Forum" and the Society of Automotive Engineers, Sichuan fourth Ninth Annual Conference Proceedings (2011)

25. Rui, H.: Primary and secondary one integral frame Improvement. In: 2011 Western Auto Industry & Academic Forum and the Society of Automotive Engineers, Sichuan fourth Ninth Annual Conference Proceedings (2011)
26. Rong, G.E., Chen, Q.: XU a village; air suspension buses for handling stability of multi-body dynamics modeling and simulation. In: Eighth Henan Automotive Engineering Technology Symposium, under (2011)
27. Jie, L., Li, C.: Floating space manipulator posture, joint coordinated motion blur Sliding Mode Variable Structure Control. In: Twenty-Sixth Chinese Control Conference (2007)

Research on Tracking Filters Algorithm in Waveform Selection Application

Yu Daobin[1], Wu Yanhong[1], and Zhu Weigang[2]

[1] Academy of Equipment
Department of Information Equipment
Beijing, China
[2] Academy of Equipment Department of Photoelectricity Equipment, Beijing, China
1711784450@qq.com

Abstract. The problem of target tracking in waveform selection application is considered in this paper. Firstly, we compare the current tracking algorithms and put emphasis on the integral-free ones .Secondly, basic model for waveform selection is constructed and we focus on the common nonlinear problem. By comparing the Unscented Kalman filters, Cubature Kalman filters and Square-root Cubature Kalman filters, an modified SCKF is proposed so as to increase the tracking properties. Two simulations and analysis for tracking is made for testing, and the result shows that modified SCKF performs better in improving estimation accuracy.

Keywords: Tracking filter, waveform selection, Kalman filter.

1 Introduction

In 2006, the signal processing expert Simon Haykin, puts forward the concept of the cognitive radar [1](CR) for the very first time, and announce several articles[2-5]later on realization of each part. It's one of the main directions in radar's development which is a waveform selection application and becomes a general agreement that radars systems capable of waveform selection are becoming practical for operational use such as in the CR. The problems involved in fully utilizing such capability are manifold and here we address the tracking filter only, where the problems about the nonlinear tracking and short processing time caused by the selection are under our discussion.

We are faced with a lot of nonlinear tracking problem. The optimal method for solving this problem encounter the calculation of posterior probability density, involving describe the integral of conditioned posterior probability density in multi-dimension. Among many suboptimal tracking filter's algorithms, the extended Kalman filter (EKF)[6] is widely used. Its basic thought is to match second-order of Taylor series, converting the problem to be a linear one. When the nonlinearity becomes severely, tacking errors will amount to a large scale. The integral-free algorithm is proposed then, and .the unscented Kalman filter (UKF)[7] approximating the distribution of the random variable through a certain set of weighted points., capture the statistics characteristic after the nonlinear transformation. The parameter should be set properly so as to get a better effect. Particle filter (PF) [8, 9] is another method

Y. Yang, M. Ma, and B. Liu (Eds.): ICICA 2013, Part II, CCIS 392, pp. 384–396, 2013.

using Monte Carlo simulation to produce large amount of particles for calculation. The number of particles will be the obstacle for tracking timely in our application. Quadrature Kalman filter (QKF) [10] obtain the variable's statistics characteristic through numerical integral by calculating Gauss-Hermitical matrix. Its limitation lies in the improvement of the dimension. Cubature Kalman filter (CKF) capture the posterior mean and covariance numerical integral by the cubature principle. It is precise and has a simple realization. The SCKF essentially propagates square-root factors of the predictive and posterior error covariance and avoid matrix square-rooting operations on the basis of CKF [4].

In this paper, we compare three typical integral-free tracking filter algorithms which is suitable for the short time waveform selection application. Then we propose a Modified SRKF to optimize original part of tracking and simulate in details for approval. LFM waveforms are selected so as to adjust the appropriate parameter with the minimum errors in range and speed. When ignoring the effect of clutter and environmental jamming, the result shows that it's advantageous to use modified SRKF in waveform selection application.

2 Modeling for Target Tracking

2.1 System Introduction

It is necessary to have a clear understanding of the frames that arise in the tracking system, one associated with the sensor, the other with the tracker. In active (as opposed to passive) tracking scenarios the sensor must perform signal processing within the time duration of a received pulse to obtain estimates of a possible target's range, radial velocity, azimuth angle, and elevation angle, etc. Conversely, the tracker is performing information processing based on outcomes of the signal processing and must operate on an interpulse time frame. The basic framework of the waveform optimal system is depicted in figure1 [11].

2.2 Signal Introduction and Relation with Tracking

We arc now in a position to outline the mathematical models used for both the signal (intrapulse) processing and the information (interpulse) processing in the tracking system. The first model of interest is the model for intrapulse processing. Let a narrow-band transmitted pulse be represented by

$$s_T(t) = \sqrt{2}\,\mathrm{Re}\left\{\sqrt{E_T}\,s(t)e^{j\omega_c t}\right\} \tag{1}$$

Where $s(t)$ is the complex envelope, ω_c is the carrier frequency (radians/sec), and E_T is the energy of the transmitted pulse. The received signal due to a single target can be written as

$$s_R(t) = \sqrt{2}\,\mathrm{Re}\left\{[\sqrt{E_R}e^{j\phi}s(t-\tau)e^{jvt} + n(t)]e^{j\omega_c t}\right\} \tag{2}$$

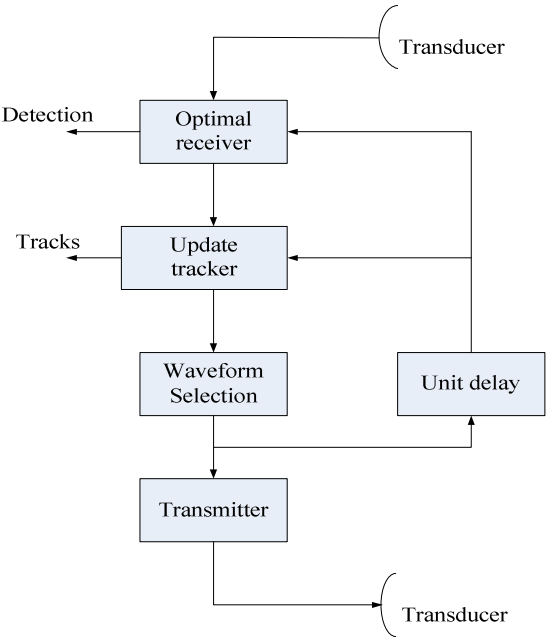


Fig. 1. Frames of waveform selection system

Where ϕ is a random phase shift, E_R is the energy of the reflected pulse, $n(t)$ is zero-mean complex white Gaussian noise. The parameter τ and v are target time delay and Doppler shift, respectively. If the target is in one dimensional motion, we use r and $\dot{r}$ to represent its range and speed, and the relationship between τ and v is

$$r = \tau c/2, \dot{r} = vc/(2\omega_c) \tag{3}$$

For the received signals, we get the baseband signal $s_R(t)$, and after the radar's matching filter, the ambiguous function can be written as

$$A(\tau, v) = \int_{-\infty}^{\infty} s\left(t - \frac{\tau}{2}\right) s^*\left(t + \frac{\tau}{2}\right) e^{-j\omega t} dt \tag{4}$$

It's the two dimension function of the variables τ and v. In our application we need to combine the waveform's parameter and tracking equation. The covariance in measurement equation contains environmental information can be applied for the linkage. $N(\theta_k)$ stands for the covariance where is the transmitted waveform's parameter vector, The relationship is

$$N(\theta_k) = TJ^{-1}T^T \tag{5}$$

Where θ_k stands for the waveform's parameter vector. J is the Fisher Information matrix about τ and v in $A(\tau, v)$, and J^{-1} is the Cramer Rao low bound(CRLB) of the unbiased estimation.. For a ideal outputs of the matching filters, the $N(\theta_k)$ can be replaced by the value of CRLB, getting the corresponding update in the tracking equation.

3 Tracking Algorithms

3.1 Comparision between CKF and UKF

Same as the EKF, UKF is a recursive MMSE(Minimum Mean Square Error) estimator. But unlike the EKF, UKF does not approximate the nonlinear state and measurement equations. It uses the true nonlinear model of state and/or measurements equation but approximates the probability density functions(pdf) of the state vector. This density is still Gaussian, but is specified by a set of deterministically chosen sample (or sigma) points. The sigma points completely capture the true mean and covariance of the Gaussian density and when propagated through the nonlinear system, capture the posterior mean and covariance accurately to the second order for any non-linearity [7].

The UKF and the CKF share a common property. They use a weighted set of symmetric points. Fig.2 shows the spread of the weighted sigma-point set and the proposed cubature points set [4], respectively in the two-dimensional space; the points and their weights are denoted by the location and the height of the stems, respectively. However, as shown in Fig.2, they are fundamentally different-for the sigma-point set the stem at the center is highly significant as it carries more weight, whereas the cubature-point set does not have a stem at the center. The even number of CKF's points is suitable for the square root operation in contrast with the UKF, and it affects the complexity of tracking algorithm and makes sure the non-negative definite matrix of the covariance during the processing.

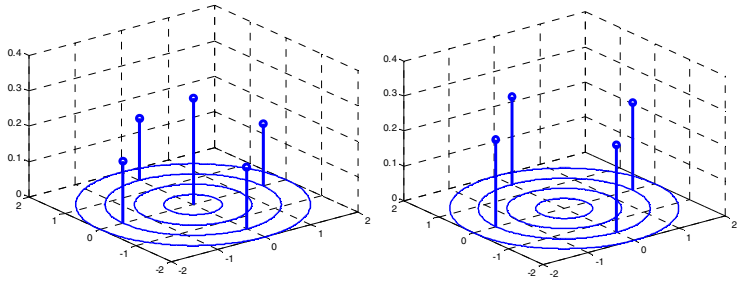


Fig. 2. Distribution of sigma points in UKF(left) and CKF(right)

3.2 Comparision between SCKF and CKF

A common feature shared by CKF and SCKF is that they all apply the Spherical-Radial cubature rules. Nonlinear filtering in the Gaussian domain reduces to whose integrands

a problem of how to compute integrals, are all of the form nonlinear function × Gaussian density. Specifically, we consider an integral of the form

$$I(f) = \int_{R^n} f(x)\exp(-x^T x)dx \tag{6}$$

defined in the Cartesian coordinate system. To compute the above integral numerically we take the following two steps: (a) We transform it into a more familiar spherical-radial integrations form (b) subsequently, we use a three-degree spherical-radial rule method is similar to 1-D case for 2-D parameter estimation, which is indispensable for measured data. if $\mathbf{x} = r\mathbf{y}, \mathbf{y}^T\mathbf{y} = 1$, the integral can be written as:

$$I(f) = \int_0^\infty \int_{U_n} f(r\mathbf{y}) r^{n-1} \exp(-r^2) d\sigma(\mathbf{y}) dr \tag{7}$$

Where U_n denotes the surface of the sphere, $\sigma(\bullet)$ is the spherical surface measure or the area element on U_n

$$I = \int_0^\infty S(r) r^{n-1} \exp(-r^2) dr \tag{8}$$

$$S(r) = \int_{U_n} f(ry) d\sigma(y) \tag{9}$$

We apply m_r points in Gauss-Hermits rules and m_s points in Spherical rules obtaining $m_r \times m_s$ points Spherical-Radial cubature rules

$$\int_0^\infty S(r) r^{n-1} \exp(-r^2) dr = \sum_{i=1}^{m_r} a_i S(r_i) \int_{U_n} f(ry) d\sigma(y) \tag{10}$$

$$\int_{U_n} f(ry) d\sigma(y) = \sum_{i=1}^{m_r} b_i f(ry_j) \tag{11}$$

$$I(f) = \int_{R_n} f(x)\exp(-x^T x)dx \approx \sum_{j=1}^{m_s}\sum_{i=1}^{m_r} a_i b_j f(r_i y_j) \tag{12}$$

Define $m_r = 1, m_s = 3$, and we get three-degree Spherical-Radial cubature rules.

According to the rule above, nonlinear function × Gaussian density can be approximately estimated as

$$I(f) = \int_{R_n} f(x) N(x; 0, I_n) dx = \sum_{i=1}^{2n} w_i f(\xi_i) \tag{13}$$

When $x \sim N(x;\mu,\Sigma)$, obtaining:

$$\int_{R_n} f(x)N(x;\mu,\Sigma)dx \approx \sum_{i=1}^{2n} w_i f(\sqrt{\Sigma}\xi_i + \mu) \tag{14}$$

where $\xi_i = \sqrt{n}[1]_i, w_i = \frac{1}{2n}, i = 1,2,\ldots,2n; [1] \in R_n,$ and

$$[1]_i \in \left\{ \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} ;\cdots, \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \\ \vdots \\ 0 \end{pmatrix}, \cdots, \begin{pmatrix} 0 \\ 0 \\ \vdots \\ -1 \end{pmatrix} \right\}$$

SCKF apply the above rule as well, and it has two features as opposed of CKF in (a) the least-squares method for the Kalman gain and (b) matrix triangular factorizations (e.g., the QR decomposition) for covariance updates. The least-squares method avoids computing a matrix inversion explicitly, whereas the triangularization essentially computes a triangular square-root factor of the covariance without square-rooting a squared-matrix form of the covariance. All of the steps can be deduced directly from the CKF except for the update of the posterior error covariance; hence we derive it in a squared-equivalent form of the covariance.

3.3 Algorithm of Modified SCKF

The computational complexity of the SCKF in terms of flops, grows as the cube of the state dimension, hence it is comparable to that of the CKF or the UKF. We may reduce the complexity significantly by (a) manipulating sparsity of the square-root covariance carefully and (b) coding triangularization algorithms for distributed processor-memory architectures. Base on the above thought, we propose a modified SCKF to improve the property of the algorithm.

Traditional SCKF algorithm

(1)Update in time

The Original formula keeps unchanged. We directly compute the estimated predicting errors covariance and its square-root. The formula is written as:

$$\begin{cases} \chi^*_{k|k-1} = \frac{1}{\sqrt{2n}}[X^*_{1,k|k-1} - \hat{x}_{k|k-1} \quad X^*_{2,k|k-1} - \hat{x}_{k|k-1} \\ \cdots X^*_{2n,k|k-1} - \hat{x}_{k|k-1}] \\ S_{k|k-1} = QR\{[\chi^*_{k|k-1} \quad \sqrt{Q_k}]\} \end{cases} \tag{15}$$

Where $QR\{\bullet\}$ stands for QR decomposition

(2)Update in measurement

$$X_{i,k|k-1} = S_{k|k-1}\xi_i + \hat{x}_{k|k-1} \tag{16}$$

$$Z_{i,k|k-1} = h(X_{i,k|k-1}) \tag{17}$$

$$\hat{z}_{k|k-1} = \frac{1}{2n}\sum_{i=1}^{2n} Z_{i,k-1|k-1} \tag{18}$$

$$\begin{cases} \eta_{k|k-1} = \frac{1}{\sqrt{2n}}[Z_{1,k|k-1} - \hat{z}_{k|k-1} \quad Z_{2,k|k-1} - \hat{x}_{k|k-1} \\ \cdots Z_{2n,k-1|k-1} - \hat{z}_{k|k-1}] \\ S_{zz,k|k-1} = QR\{[\eta_{k|k-1} \quad \sqrt{R_k}]\} \end{cases} \tag{19}$$

$$\begin{cases} \chi_{k|k-1} = \frac{1}{\sqrt{2n}}[X_{1,k-1|k-1} - \hat{x}_{k|k-1} \quad X_{1,k-1|k-1} - \hat{x}_{k|k-1} \\ \cdots X_{2n,k-1|k-1} - \hat{x}_{k|k-1}] \\ P_{xz,k|k-1} = \chi_{k|k-1}\eta_{k|k-1}^T \end{cases} \tag{20}$$

(3) Calculate the Kalman gain, state estimation, state estimated covariance

$$K_k = (P_{xz,k|k-1} / S_{zz,k|k-1}^T) / S_{zz,k|k-1} \tag{21}$$

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + K_k(z_k - \hat{z}_{k|k-1}) \tag{22}$$

$$\begin{cases} U = K_k S_{zz,k|k-1} \\ S_{k|k} = QR\{[\chi_{k|k-1} - K_k Z_{k|k-1} \quad U]\} \end{cases} \tag{23}$$

We can obviously see that the SRKF is easier than the CKF because the error covariance is used as the recursive estimation. We then discuss the validity and instantaneity that the waveform selection needed, so the optimal method proposed on the basis of the tradition algorithm and called modified SRKF.

In formula (15)(19)and (23),we use the QR decomposition as the triangularization. This is a wise choice but isn't a simple one. We can get the unique decomposition that $P = QR,\ where\ P \in C^{m\times n}, Q \in C^{m\times r}, R \in C^{r\times n}$ In the equation $Q^H Q = E$

and $rankR = r$. By exploiting the transposition of R , the process of updating becomes simple and available. Here we modify the process of decomposition.

In solving linear equation set, we tend to use a linear transformation to simplify the factors matrix. The same operation is used in the error covariance P, get $P = P^{(0)} = L_1 P^{(1)} = L_1 L_2 P^{(2)} = \cdots = L_1 L_2 \cdots L_{n-1} P^{(n-1)}$ By Theories we know that when the order principal minor $\Delta_k \neq 0, k = 1, 2, \ldots, n$, the decomposition is unique, and we set $L_1 L_2 \cdots L_{n-1} = L, A^{(n-1)} = U$, and $P = LU$, where L is a unit lower triangular matrix, and U is an upper triangular matrix. This is a linear decomposition, and the covariance ascends linearly as well.

Considering the simple operation of LU decomposition and symmetry operation of QR decomposition, we combine them to compute together. The covariance increases to a bound where indicate a shift of triangular decomposition from LU to QR. Formula (15) is modified to be

$$\begin{cases} \chi^*_{k|k-1} = \dfrac{1}{\sqrt{2n}}[X^*_{1,k|k-1} - \hat{x}_{k|k-1} \quad X^*_{2,k|k-1} - \hat{x}_{k|k-1} \\ \cdots X^*_{2n,k|k-1} - \hat{x}_{k|k-1}] \\ S_{k|k-1} = LU\{[\chi^*_{k|k-1} \ \sqrt{Q_k}]\} \ \ when\, Tr(P_{k|k}) < eps \\ S_{k|k-1} = QR\{[\chi^*_{k|k-1} \ \sqrt{Q_k}]\} \ \ when\, Tr(P_{k|k}) > eps \end{cases} \tag{24}$$

Formula (19) and (23) is modified like this. The eps is set at a low level relatively and the trace of P indicates the total tracking error [11]. In tracking problems, most measurements are dependent of each other, that it satisfies the order principal minor in most instances. For the waveforms selection application, we choose the range and radial speed as the measurements, and they are assumed dependent of each other before tracking. So we can apply the modified SRKF algorithm for the tracking problems.

4 Simulation and Analysis

4.1 Simulation (A)

We elaborate all of the algorithms above in details. Consider the waveform selection application where the state model is defined by the pair of difference equations in discrete-time, and a single station which located at (0, 0) to observe the uniform linear moving target. Firstly we compare these four algorithms and choose the range as the measurement only. The model is can be written as

$$\begin{cases} X(k) = \Phi X(k-1) + \Gamma u(k) \\ Z(k) = \sqrt{(x(k) - x_0)^2 + (y(k) - y_0)^2} + v(k) \end{cases} \tag{25}$$

Where $\Phi=\begin{bmatrix}1 & T & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & T\\ 0 & 0 & 0 & 1\end{bmatrix}$, $\Gamma=\begin{bmatrix}T^2/2 & 0\\ T & 0\\ 0 & T^2/2\\ 0 & T\end{bmatrix}$, the covariance of u(k) is set to be Q= $\omega\times$diag([1,1]), ω is a adjustable parameter and its range is ω<<1, here we set $\omega=10^{-5}$, means the system error can be ignored. The covariance of $v(k)$ is set R=200. In the simulation, the moving target and the relevant parameter of radar is set: the period of scanning is T=1s and the observation keep to N=60 times, the original state vector of the target [30000, 100, 40000, 200]. It moves far away from us and the errors ascend as time goes on. We apply UKF, CKF, SCKF and Modified SCKF in this simple tracking problem and get the result below. The simulation is run by MATLAB and 100 times Monte Carlos operation is used for precision. The time cost and the Root Mean Square Error (RMSE) for these four algorithms is showed in Fig.3.

The result shows that the CKF is better than UKF, but coat more time, and the improvement in RMSE is not so obvious. SCKF is an improved algorithm and it can perform better in time and precision, and it's more stable during the tracking process. The modified SCKF is outstanding in performance, we can see that the RMSE is low in the original phase and it can work fast due to its simple transformation. What we can't see is that the divergence will happen as time goes on for the RMSE's linearity and it will affect the target's tracking accuracy.

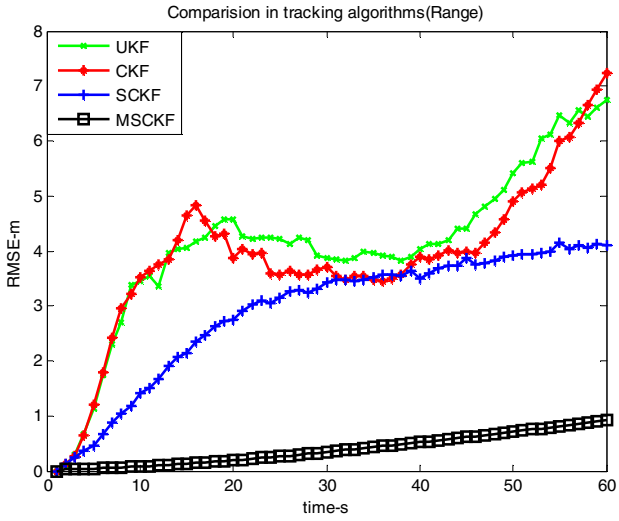


Fig. 3. Comparision in precision of four algorithms in single measurement

Table 1. Comparisons of time cost in four algorithms(run 100 times)

Algorithm type	Time cost
UKF	13.7755
CKF	14.2072
SCKF	12.4073
M-SCKF	11.7031

4.2 Simulation(B)

The simulation on waveform selection is performed. We use the up-sweep and down-sweep linear frequency modulated (LFM) chirp of Gaussian amplitude modulation to conduct the simulation. A two-dimension waveform parameter vector $\theta = [\lambda, b]$, based on which the grid can be expressed as:

$$P = \{\lambda \in [\lambda_{\min} : \Delta\lambda : \lambda_{\max}], b \in [-b_{\max} : \Delta b : -b_{\min}] \bigcup [b_{\min} : \Delta b : b_{\max}]\}$$

Where $\Delta\lambda$ and Δb denote the step-size chirp rate and pulse width parameter, respectively.

Then we make the simulation where most conditions of radar and target stay unchanged. We add the measurement where the range and radial speed is observed at the same time, and Z(k) becomes a two-dimensions vector. We set

$$\lambda_{\min} = 1us, \lambda_{\max} = 10us, \Delta\lambda = 1us$$

$$b_{\min} = 5\times10^8, b_{\max} = 30\times10^8, \Delta b = 5\times10^8$$

So we have 10×12 different LFM waveforms to select.

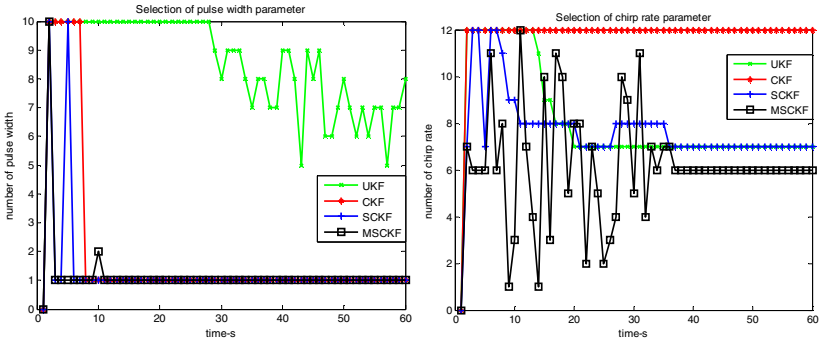


Fig. 4. Changes of parameter(pulse width and chirp rate) in waveform selection application

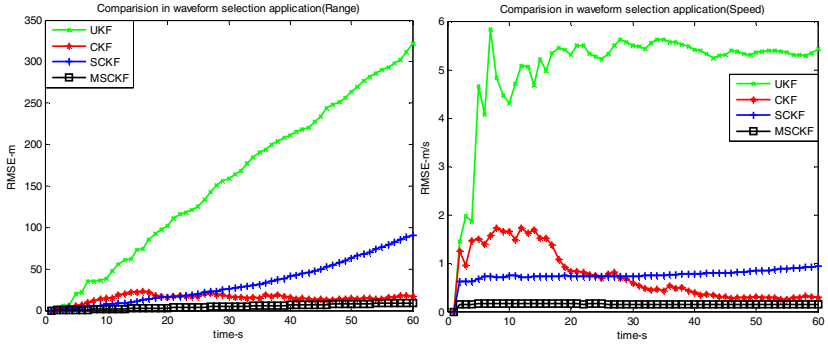


Fig. 5. Precision's comparision of four algorithms in two measurements

The result shows that the modified SCKF has the best performance. Firstly, CKF, SCKF and MSCKF differ in the original part of pulse width selection. They all prefer to the NO.1 pulse width after 10 seconds. Chirp rate is changing frequently for MSCKF, where high-demands for the radar are necessary. What we need to pay attention to is that the end part of SCKF and MSCKF are stable for they take the same QR decomposition. Secondly, the changes in waveform's parameter are frequent so as to make the adjustment which fit to the environment. Cubature rules can deal with this problem, and UKF with a low covariance of sigma points performs poorly. CKF and SCKF are valid algorithms for waveform selection problem. SCKF is stable in predicting the speed, and it costs less time as well, but perform poorly in range prediction. In this example the modified SCKF has the best performance. Though the judgments in decomposition cost a little time compared with SCKF, but the errors in range and speed is the lowest of all.

Table 2. Comparisons of time cost in four algorithms (run 100 times)

Algorithm type	Time cost
UKF	12.1897
CKF	12.5154
SCKF	10.4917
M-SCKF	10.6602

5 Summary

In this paper, waveform selection system is introduced at first, where the tracking algorithms for it are discussed in details. A modified algorithm is proposed so as to highly improve the tracking precision via making optimized choice from two kinds of decompositions. At the same time, two simulations offered to confirm our calculation in the environment of Matlab. The results show that modified SCKF is better than the other algorithms in tracking performance for waveform selection application.

In many powerful noise from the band capable of measuring values in the mathematical random estimation tools, the Kalman filter can be said to be the best known and most commonly used one. Kalman filter is used Rudolph E.Kalman named. In 1960, Kalman published his famous paper describing a recursive solution to the problem of discrete data linear filtering papers. This paper was published, has benefited from the rapid development of digital computers, but also because of their simplicity and robustness, the Kalman filter to get the attention of researchers, and soon to be applied in various fields, in particular, in autonomous or assisted navigation. Kalman filter is a sequence of states of the dynamic system of linear minimum variance estimation algorithm, Kalman prediction process is the previous state of the system is estimated on the basis of the state of the system the next time, that is, each simply save the system the last time the state, therefore, only a very small storage space to store system state, which also makes the Kalman filter computation and real good. Kalman filter consists of a basic set of mathematical formulas described, to achieve a forecast - update mode estimator, which in a number of pre-conditions, enabling the estimation error covariance is minimized, in this sense, Karl Mann filter is an optimal estimator. To achieve precision guided radar function requires precise tracking and measuring the motion parameters of moving targets, and in the tracking measurement process, requires a lot of noise in the measurements required to extract the data that affect the tracking accuracy SNR The most important factor, in the same SNR, in order to further improve tracking accuracy, the need to take effective target motion characteristics of filtering algorithms. This paper discusses the principles and characteristics of Kalman filtering, moving target tracking Measurement parameters and parameter selection, and verified by simulation Kalman filter to improve the measurement accuracy for tracking performance. In this paper, precision guidance systems, in order to improve the precision of the target motion parameters of tracking and measuring performance, improve the measurement accuracy, using the Kalman filter algorithm, respectively, distance and speed, angle and angular velocity joint processing, process noise and measurement noise excitation located is zero mean Gaussian noise, according to the system characteristics to choose the appropriate state noise matrix and measurement noise covariance matrix, effectively improve the measurement accuracy. Better solution to improve accuracy, compatibility measuring real-time, to ensure effective algorithm convergence problems.

References

1. Haykin, S.: Cognitive radar: a way of the future. IEEE Signal Processing Magazine 23, 30–40 (2006)
2. Haykin, S.: Cognitive dynamic systems IEEE IC-ASSP, 1369–1372 (2007)
3. Haykin, S., Xue, Y., Davidson, T.N.: Optimal Waveform Design for Cognitive Radar IEEEAsilomar, 3–7 (2008)
4. Arasaratnam, I., Haykin, S.: Cubature Kalman Filters. IEEE Transaction on Automatic Control 54, 1254–1269 (2009)
5. Haykin, S., Zia, A., Arasaratnam, I., et al.: Cognitive tracking radar. IEEE Xplore, 1467–1470 (2010)

6. Anderson, B.D.O., Moors, S.B.: Optimnl Filtering. Prenfice-Hall, Eaglewood Cliffs (1979)
7. [7] Julier, S.J., Uhlmann, J.K.: Unscented filtering and nonlinear estimation. Proc. of the IEEE 92, 1684–1688 (2004)
8. Gordon, N.S., Salmond, D.S., Smith, A.F.M.: Novel approach to nonlinear/non-Gaussian Bayesian state estimation. Proc. Inst. Elect. EnR. F 140, 107–113 (1993)
9. Cappe, O., Godsill, S.J.: An recent advances in sequential Monte Carlo. Proc. IEEE 95, 899–924 (2007)
10. Ito, K., Xiong, K.: Gaussian filters for nonlinearfiltering problems. IEEE Trans. Automat. Control 45, 910–927 (2000)
11. Kershaw, D.J., Evans, R.J.: Optimal waveform selection for tracking systems. IEEE Transactions on Information Theory 40, 1264–1268 (1994)

Accurate Assessment Model for Multi-system GNSS

Zhongliang Deng, Guanyi Wang, Zhongwei Zhan, and Hui Dong

BUPT, No.10, Xitucheng Road, Haidian District, Beijing, China
{dengszhl,guanysi_wang,azhongawei,donghaui}@bupt.edu.cn

Abstract. Traditional methodology and mathematical model neglected the influence of secondary code which is a commonly component in new signal structure onto the assessment of multi-system GNSS compatibility and did not quantificational give the condition under which the model for the signals with long code and high data rate can be simplified. In this paper, through the exploration of Galileo E5 signals, the influence of secondary code onto the spectral separation coefficient (SSC) was discussed. Based on that conception, we modified the existing formula and quantificational presented a threshold of data rate and length of primary and secondary code so that the traditional methodology can be replaced by a simplified model without gigantic calculating time and quantity. The power spectral density (PSD) and SSC of assumed signals which meet the threshold above was respectively calculated. The simulation results prove the correctness of the proposed assessment model.

Keywords: Compatibility, Secondary Codes, Spectral Separation Coefficient (SSC), PSD.

1 Introduction

The inter-system and intra-system interference between four different navigation systems (GPS, Galileo, GLONASS, Compass) are getting worse because of sharing of the same frequency bands. Under these circumstances, the multi-system GNSS compatibility is urgent to study thereby coordinating the allocation of the precious navigation frequency L-bands [1] between corresponding countries.

Aiming at acquisition, carrier tracking and PVT calculation process, original methodologies use effective $(C/N_0)_{eff}$ ratio and its degradation directly related to an integral called spectral separation coefficient (SSC) [2] at the receiver input to evaluate multi-system GNSS compatibility. Doctor J. Betz firstly proposed that conception and provided detailed derivation [3]. Doctor F.Souall developed that conception and derived code tracking spectral sensitivity coefficient (CT_SSC) [4] to analyze the compatibility reflecting the code tracking process. However, traditional methods substitute power spectral density (PSD) of carriers for the PSD of practical signals as literature [5] did. Considering the inaccuracy using that substitution, Doctor F.Souall presented the impact of spreading code and data rate onto the practical PSD and derived an accurate mathematical model [6] on the ION GNSS Meeting in 2003 [7].

This paper starts by a brief description of the effective (C/N0)eff ratio based on the SSC, complemented by the explanation of three types of navigation signals and their

Y. Yang, M. Ma, and B. Liu (Eds.): ICICA 2013, Part II, CCIS 392, pp. 397–409, 2013.

PSD expression that can be encountered. After testing and verifying the influence of data rate onto the actual PSD, we will propose the idea that the secondary code [7] in the new signal structure could affect the accuracy of calculating actual PSD and then we will improve the formula provided by F.Souall. In the final part, a threshold is being presented so that an ideal PSD model could be used to replace the complicated actual PSD model to rapidly calculate the SSC.

2 Compatibility Assessment Model

According to the definition provided by International Committee on Global Navigation Satellite Systems (ICG) ,the compatibility of GNSS refers to the case that the position, navigation or timing service of a single system will not cause unbearable interference to the service of other systems, when the different navigation systems are joint with each other or working alone. Because of the different types of signals sent by one or more navigation systems on the same frequency band, there exists inter-system and intra-system interference in the definition of compatibility. Usually adopt the degradation of effective $(C/N_0)_{eff}$ ratio to express the interference to desired signals caused by jamming signals.

$$I_{\text{int}ra/\text{int}er} = \sum_{j=1}^{N_{\text{int}re}/N_{\text{int}er}} C_J K_{is} / \int_{-\beta_R/2}^{\beta_R/2} G_s(f)df \tag{1}$$

$$K_{js} = \int_{-\beta_R/2}^{\beta_R/2} G_j(f)G_s(f)df \tag{2}$$

Table 1. SSC on the 1575.42MHz frequency band

SSC[dB/Hz]		Desired signals			
		C/A	L1C	E1OS	B1C
Jamming signals	C/A	[-61.80]	[-61.80]	[-68.15]	[-61.80]
	L1C	[-61.80]	[-65.35]	[-65.40]	[-65.35]
	E1OS	[-68.15]	[-65.40]	[-65.46]	[-65.41]
	B1C	[-61.80]	[-65.35]	[-65.41]	[-65.36]

K Stands for the spectral separation coefficient (SSC) which reflects the overlap and interference degree when calculate the effective $(C/N0)_{eff}$ ratio. β_R Means the receiver processing bandwidth. G_j And G_s respectively stands for the PSD of jamming signals and desired signals which is usually substituted by the PSD of carrier for simplification. $N_{\text{int}er}$ and $N_{\text{int}ra}$ respectively represents the number of visible satellites in a single constellation of a navigation system. G_j is the carrier power related to the time t . C represents the power of desired signals. N_0 is the noise power. Table 1 shows the SSC of signals between three different navigation systems on the L1 frequency band (1575.42MHz) using expression (2). The receiver processing bandwidth β_R equals to 24MHz.

3 Power Spectral Densities of Navigation Signals

3.1 Navigation Signals Model in Time Domain

According to expression (1), effective $(C/N_0)_{eff}$ ratio depends on the accurate SSC value. Likewise, SSC value depends on the accurate PSD of received signals.

The baseband model (including one desired signal, M jamming signals) in time domain for the received navigation signals after down-conversionare is expressed below.

$$s(t)=\sqrt{P_0}d_0(t)g_0(t)+\sum_{i=1}^{N_{\text{int}}}\sqrt{2P_i}d_i(t-\tau_{i,0})g_i(t-\tau_{i,0})\cos(\Delta\omega_{i,0}t+\varphi_{i,0})+n(t) \tag{3}$$

Where P_i is the received signal power. $d_i(t)$ is the data stream with the symbol duration time T_d and data rate $f_d=1/T_d$. $g_i(t)$ is the pulse-modulated code whose expression is:

$$g_i(t)=w(t)\times\sum_{k=-\infty}^{\infty}c_{i,k}\phi(t-kT_c)=w(t)\times(\phi(t))*\sum_{k=-\infty}^{\infty}c_{i,k}\delta(t-kT_c) \tag{4}$$

In the expression, $w(t)$ stands for a time-window of length T, $\phi(t)$ is the pulse shape associated to the spreading codes $c_{i,k}$ with chip duration T_c and code length N.

3.2 Power Spectral Densities Corresponding to the Navigation Signals Model in Time Domain

There exist three main categories of signal PSDs [6] in total. Navigation signal PSDs of the first type are continuous and only depend on the wave carrier PSDs. In this case, the spreading codes equal to the non-periodical ideal pseudo random codes without data modulation. So PSDs of the first type can be regarded as the ideal PSDs as expression (5) and (6) shows:

$$G_{BPSK(N)}(f)=\frac{1}{Nf_{nom}}\sin c^2\left(\frac{f}{Nf_{nom}}\right) \tag{5}$$

$$G_{BOC(M,N)}(f)=\frac{1}{Nf_{nom}}\left(\tan\left(\frac{\pi f}{aMf_{nom}}\right)\bullet\sin c\left(\frac{f}{Nf_{nom}}\right)\right)^2 \tag{6}$$

Navigation signal PSDs of the second type are composed of spreading codes and carrier. The spectral lines are separated by the inverse of the spreading code periods with the same parameters described in 3.1 as expression (7) shows:

$$G_x(f)=\left(\frac{1}{LT_c}\right)^2\left(\sum_{k=-\infty}^{\infty}\delta\left(f-\frac{k}{LT_c}\left|DFT(c_k)\right|^2\Phi^2\left(\frac{k}{LT_c}\right)\right)\right) \tag{7}$$

In the expression, the interval of the spectral lines is $1/LT_c$ These navigation signals without data modulation is called pilot signals such as Galileo E1C signals,E5a_Q signals,Compass B2a_Q signals.

Navigation signal PSDs of the third type have data codes modulated. The spectral lines is substituted by $(1/f_d)\bullet \sin c^2(f/f_d)$ function with the data rate f_d for the impulse function. The corresponding PSD derivation is given by:

$$G_x(f)=\left(\frac{1}{f_d}\right)\left(\frac{1}{LT_c}\right)^2\left(\sum_{k=-\infty}^{\infty}\left|DFT(c_k)\right|^2\Phi^2\left(\frac{k}{LT_C}\right)\sin c^2\left(\left(f-\frac{k}{LT_C}\right)/f_d\right)\right) \tag{8}$$

3.3 Examples

The examples of the PSDs of the first and third type discussed in 3.2 are presented in Fig. 1 below. The examples illustrate the PSD of the GPS C/A codes on frequency band 1575.42MHz with a data rate of 50Hz. A BPSK carrier is used with a chip rate of 1.023MCps and a code period of 1ms. Known from Fig. 1, the PSD has the same shape with the PSD of BPSK carrier and its main lobe width is twice of the code rate. The maximum value difference between the PSD of first type (ideal PSD) and third type (practical PSD) is as much as 20dB, which is caused by C/A code with low data rate and short spreading, codes length. We can tell from the enlarged PSD that the practical signal is composed of overlapped comb spectrums of the data codes with the shape of $\sin c$ function. The interval of the spectral lines is 1000Hz decided by $1/LT_c$ (see expression (8)). The reason why there exists a difference between the two PSDs is the amplitude variation of the main and side lobe on each spectral line.

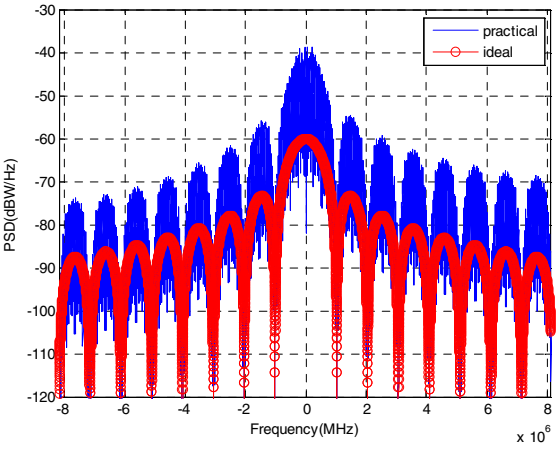


Fig. 1. PSD of the ideal and practical GPS C/A signals

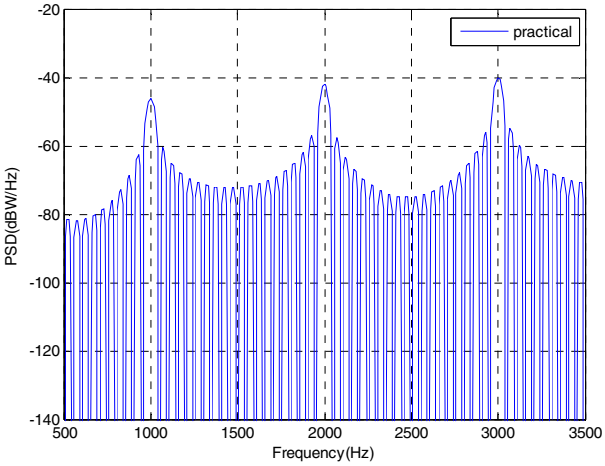


Fig. 2. Enlarged PSD of the practical GPS C/A signals from Fig. 1

4 Improved Power Spectral Densities of Navigation Signals

4.1 Influence of Data Rate on the SSC

In the previous section was shown that the data rate has the function of smoothing the practical PSD which determines the value of SSC. In that case, the accuracy of the SSC depends on the specific data rate.

Fig. 3 illustrates the variations of SSC value calculated by the PSD of two C/A codes. The rate has been respectively set to 50bps, 250bps and 500bps. The abscissa means the Doppler values representative of the satellite dynamics with regard to a user

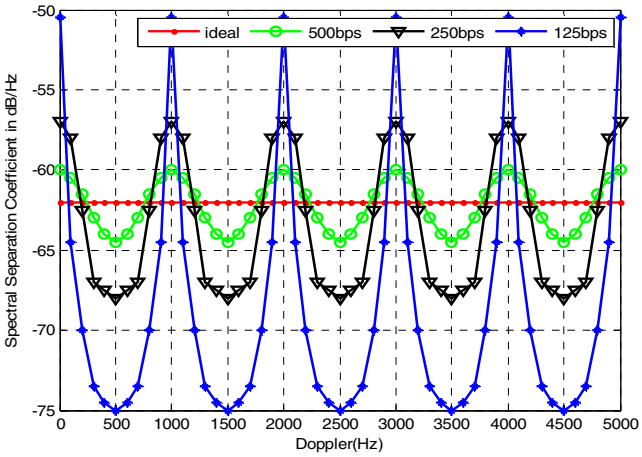


Fig. 3. Variation of SSC value of GPS C/A signals with different data rate

on ground. Expression (8) was used to calculate the PSDs with the same code rate and length. The receiver bandwidth is 22MHz. Fig. 3 also presents the ideal SSC value when the data codes are regarded as the pseudo-random codes.

For the low data rate, the variation of SSC value shows an obviously periodical characteristic (the period equals to the spreading code period) with the shift of Doppler and has a relatively large variance. With the increase of the data rate, SSC value approaches to a smooth line when the symbol duration time T_d is much smaller than the code period LT_c. So the PSD of the carrier should not be substituted for the PSD of the practical signals to calculate the SSC value for the signal with low data rate such as C/A code, otherwise it will cause an unbearable error as much as 15dB.

4.2 Influence of Secondary Codes Length on the SSC

The existing analysis has proved the influence of the data rate to the SSC value. However, the corresponding mathematical model does not reflect the secondary code which is a commonly component in the new signal structure.

Known from the open ICD,every spreading code of GPS and Galileo signals on L1 and L5 frequency bands adopts a hierarchical structure with a combination of primary codes and secondary codes. Fig. 4 illustrates the typical structure of Galileo E5a_i navigation signals considering the secondary codes.

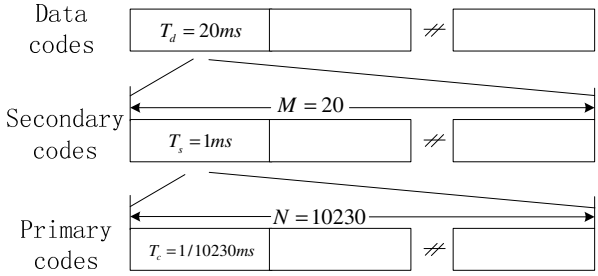


Fig. 4. Structure of Galileo E5a_i navigation signals

4.2.1 Consider Secondary Codes as Data Codes with a Higher Data Rata

The relation between the primary codes, secondary codes and data codes is xor. After adding the secondary codes, each data chip is separated into 20 pieces because of the xor operation with the secondary codes which means the data rate has been improved 20 times. In section 4.1, we have known that there existed a 10dB difference when the data rate was respectively set to 125bps and 500bps. In a similar way, the existence of the secondary codes which is regarded as a faster data stream will inevitably influence the accuracy of SSC value. Here, we use Fig. 5 to illustrate the distinction between the PSD of two different signals with and without the secondary codes. The formula of the third type (expression (8)) is used to calculate the PSD. A data rate of 50Hz is applied. The primary code with code period 1ms is combined with the secondary code with a code length 10. The shape of the carrier is as same as the shape of BOC (1,1).

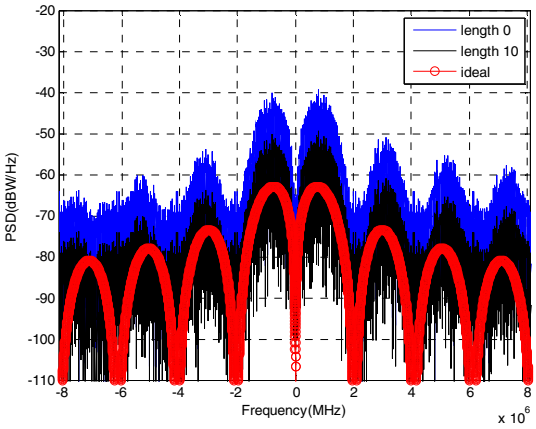
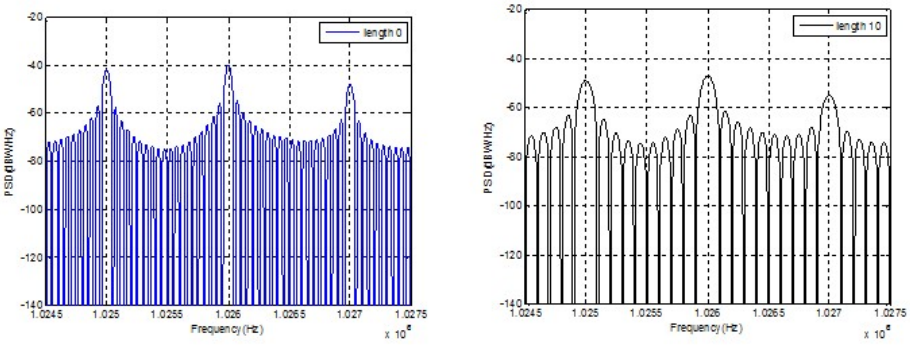


Fig. 5. PSD of BOC (1,1) with secondary code length 0 and 10



(a) secondary code length 0 **(b)** secondary code length 10

Fig. 6. Enlarged PSD of the BOC(1,1) from Fig. 5

Known from the enlarged Fig. 6, PSD of signal with a secondary code has a wider main lobe bandwidth which occupies more spaces than the side lobe does. Therefore the side lobe has a smaller impact on the SSC value than the main lobe. It will only be the variation of the main lobe amplitude which affects the calculation of SSC value when the width of main lobe reaches to certain degree.

4.2.2 Consider Secondary Codes as Primary Codes with a Longer Code Period

The hierarchical spreading code which is discussed in the section above can be considered in a different way. The addition of the secondary codes makes the code period of the primary code expand to 20 times from 1ms to 20ms.

We assume the length of the secondary code is M and the length of the primary code is L. The chip duration of the secondary code is T_c. Therefore the code period of the

hierarchical spreading codes is MLT_c . In this case, the expression (8) could be modified as below:

$$G_x(f)=\left(\frac{1}{f_d}\right)\left(\frac{1}{MLT_c}\right)^2\left(\sum_{k=-\infty}^{\infty}\left|DFT(c_k)\right|^2\Phi^2\left(\frac{k}{MLT_C}\right)\sin c^2\left(\left(f-\frac{k}{MLT_C}\right)/f_d\right)\right) \quad (9)$$

Fig. 8 is the illustration of PSDs calculated by the expression (9) with the same parameters as Fig. 6 except the code length of the secondary codes. The length of the secondary codes is respectively set to 5 and 10.

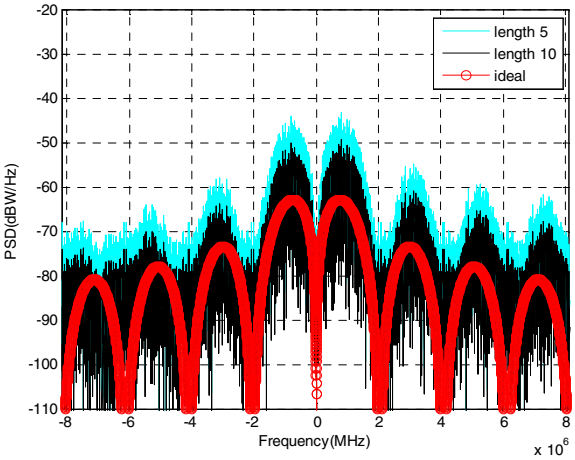
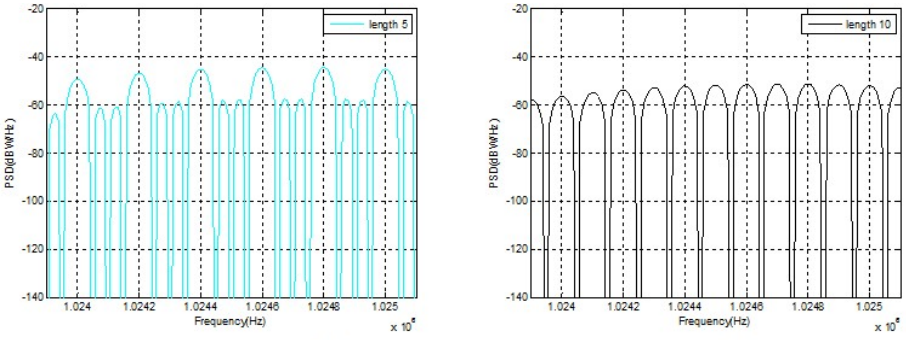


Fig. 7. PSD of BOC (1.1) with secondary code length 5 and 10



(a) secondary code length 5 **(b)** secondary code length 10

Fig. 8. Enlarged PSD of the BOC(1,1) from Fig. 7

We can tell from Fig. 7 that PSD of the signal with a secondary code length 10 is closer to the ideal PSD than the PSD of the signal with a secondary code length 5. The enlarged Fig. 8 shows that the interval of the spectral lines is 0.1 KHz and 0.2 KHz. With a stability of data rate (the wide of main lobe), smaller the interval of the spectral

lines is, smoother the amplitude variation of the PSD. Therefore the PSD with different accuracy will certainly affect the SSC value.

5 Threshold Condition for the Substitution of Ideal PSD for Practical PSD

5.1 Algorithm Analysis

If we adopt formula (9) to express PSD and then calculate SSC value of all kinds of signals with different Doppler shift before we get the degradation of effective $(C/N_0)_{eff}$ ratio under different time and different location,the calculating time and quantity will be tremendous.

The PSD of practical navigation signals are composed of overlapped spectrums of the data codes. From the analysis above, we know that it is the amplitude variation of PSD of data codes which could be as much as 25dB that causes the inaccuracy of SSC value. Hence, the algorithm we adopt here is smoothing this amplitude variation and making it approximate to a straight line by increasing the data rate (increase the main lobe bandwidth) or decreasing the interval of spectral lines $1/MLT_c$ to certain threshold so that the practical PSD could be replaced by the ideal PSD without considering the influence of the data codes and spreading codes.

The influence of side lobe can be neglected when the amplitude of main lobe is much bigger than that of the side lobe. Here, we use bipolar non-return-to-zero pulse sequence to express the data code stream of different rates. Fig. 9,Fig. 10 and Fig. 11 shows the PSD of data codes with three different rates (500bps, 800bps, 1000bps). The interval of the spectral lines is set to a constant value 1 KHz. The black line above in each of the figure is the sum of the overlapped amplitudes on every spectral line.

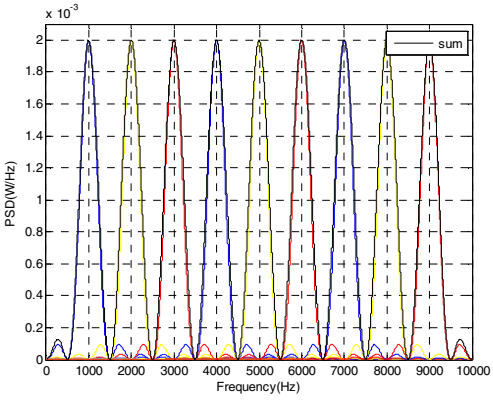


Fig. 9. PSD of the data codes of 500bps

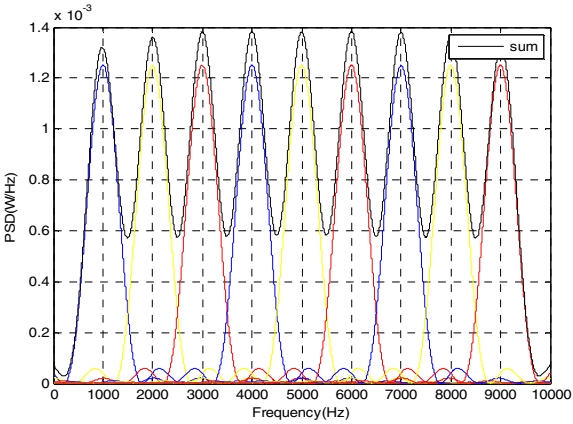


Fig. 10. PSD of the data codes of 800bps

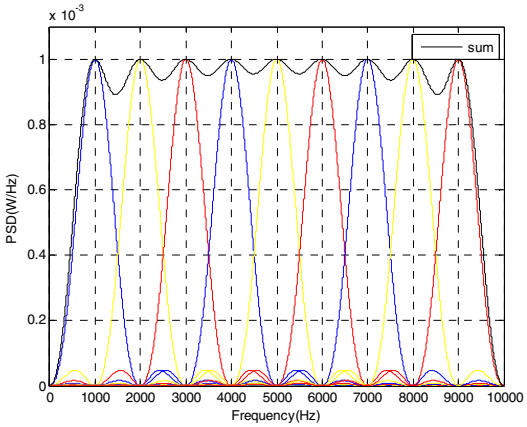


Fig. 11. PSD of the data codes of 1000bps

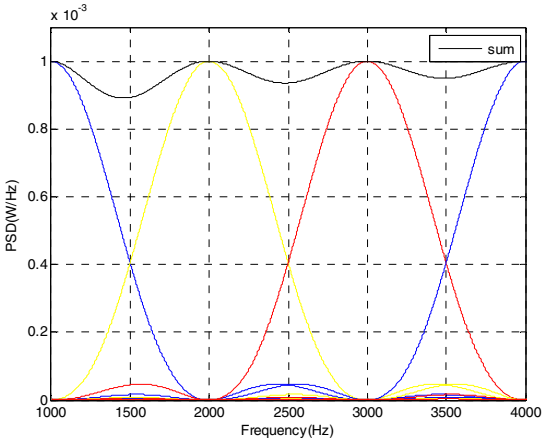


Fig. 12. Enlarged PSD of the data codes from Fig. 11

As seen in Fig. 11 and the enlarged Fig. 12, one half of the main lobe bandwidth exactly equals to the interval of adjacent spectral lines and the sum of amplitudes nearly approaches to a straight line.

In the improved formula (9), the interval of spectral lines is $1/MLT_c$, the data rate is f_d, the bandwidth of the main lobe is $B_d = 2f_d$. We can deduce a relational expression from the previous analysis.

$$\frac{1}{MLT_c} \le \frac{B_d}{2}, or, T_d \le MLT_c \tag{10}$$

When the parameters in formula (9) meet this threshold condition, the amplitude variation of PSD of the data codes will no longer fluctuate periodically with the Doppler shift. Therefore the ideal PSD calculated by expression (5) and (6) could be substituted for the practical PSD calculated by expression (9) so that the calculating time and quantity could be vastly reduced.

Note that this relationship also demonstrates the relation that the amplitudes will approach to a straight line with a smaller T_d (bigger f_d) or bigger M which further proves the conclusions we have discussed in Section 4.

5.2 Simulation

We assume GPS C/A signal with an added secondary code exactly satisfies the relation (17) with parameters $B_d = 2f_d = 50$, $M = 20$, $N = 1023$, $T_c = 1/1023$. Firstly draw the PSD of ideal GPS C/A signal by using the ideal model without data and spreading code modulation and the PSD of assumed GPS C/A signals as Fig. 13 shows. Then calculate the variations of SSC value with the PSD of assumed GPS C/A signals by using the improved formula (9) as Fig. 14 shows.

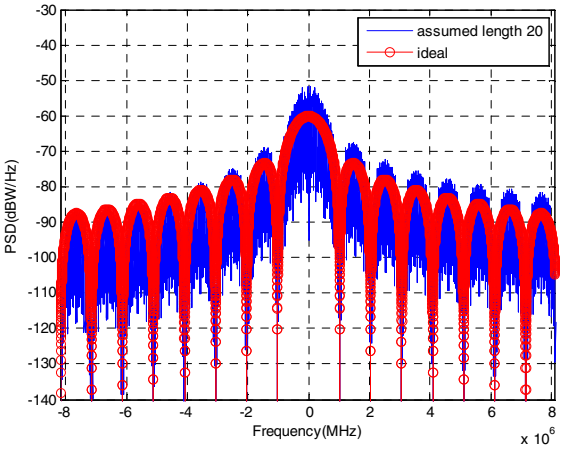


Fig. 13. PSD of the ideal and assumed GPS C/A signals

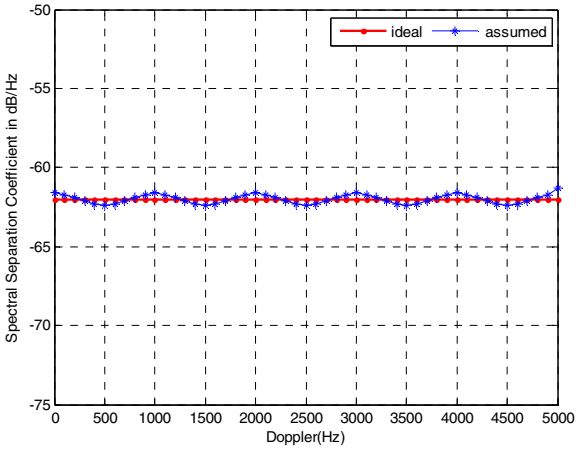


Fig. 14. Variation of SSC value of the ideal and assumed

The illustration shows that when the data rate and secondary code length meet the threshold in relation (17), the shape and amplitude of the practical PSD is approximate to that of the ideal PSD. Moreover, the variations of SSC value approach to a smooth line instead of a periodical fluctuation with the shift of Doppler. Therefore, it proves the correctness of the algorithm for the threshold we proposed. Not only can we ensure the accuracy of assessing the multi-system GNSS compatibility, but also can we greatly decrease the calculating time and quantity.

6 Conclusion

This paper firstly verified the influence of data rate on practical PSD and the SSC value before demonstrating the secondary codes in the new signal structure using Galileo E5a_i signal as an example. Then we modified the existing formula by adding a factor after analyzing the impact of the secondary codes length on the practical PSD. A threshold of data rate and length of primary and secondary codes had been deduced so that the PSD of data codes could approximately approach to a straight line without a characteristic of periodical fluctuation. Therefore when satisfied with the threshold condition, it is possible to use the ideal PSD instead of the practical PSD to calculate the SSC value and the corresponding inter-system and intra-system degradation of effective $(C/N_0)_{eff}$ ratio. At the meantime, tremendous calculating time and quantity could also be reduced with a high accuracy of assessing the multi-system GNSS compatibility in this way.

References

1. Hein, G.W.: Towards a GNSS system of system. In: Proceedings of the 20th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS 2007), Fort Worth, TX, vol. 1, pp. 564–571 (2007)
2. Betz, J.W.: Binary Offset Carrier Modulation for Radionavigation. Navigation: Journal of The Institute of Navigation Winter 1, 673–677 (2001)
3. Betz, J.W.: Effect of Partial-Band Interference on Receiver Estimation of C/N0. In: Proceedings of the 14th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS 2001), vol. 1, pp. 786–789 (2001)
4. Soualle, F., Burger, T.B.: Radio Frequency Compatibility Criterion for Code Tracking Performance. In: Proceedings of the 20th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS 2007), Fort Worth, TX, vol. 1, pp. 231–243 (2007)
5. Titus, B.M., Betz, J.W., Hegarty, J., et al.: Intersystem and Intrasystem Interference Analysis Methodology. In: Proceedings of the 16th International Technical Meeting of the Satellite Division of the Institute of Navigation (ION GNSS 2003), Oregon, Porland, vol. 1, pp. 888–893 (2003)
6. Soualle, F., Burger, T.B.: Impact of Galileo Spreading Code Selection and Data Rate onto Navigation Signal Interference. In: Proceedings of the 16th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS 2003), Oregon, Porland, vol. 1, pp. 325–343 (2003)
7. Hu, X., Tang, Z., Zhou, H., Huang, X.: Analysis on Design Principles of GPS and Galileo Signal Structure. Systems Engineering and Electronics (10), 2285–2293 (2009)

Quality and Safety Warning Model of Entropy Edible Forest Products Based on Extension Theory

Peng Zhang, Baoyong Li, Zhongping Luan, and Zhongsu Ma

College of Biological and Agricultural Engineering, Jilin University, Changchun, Jilin, China 130025
peng-999zhang@163.com
Forestry College of Beihua University, Jilin, China 132013
Jilin Forestry Survey and Planning Institute, Jilin, China 132013
Forestry Survey and Design Institute of Jilin Province, Changchun, Jilin, China 130025

Abstract. In order to ensure the safety of our food, so that preventive measures, we must constantly improve our country's food security early warning system. Food is security early warning multi-level, multi-dimensional and dynamic requirements, using the entropy weight and Extension theory of edible forest products Extension Theory Based on Entropy quality and safety warning model. Food safety data, combined with a forestry enterprise and quantitative evaluation of the status of food security early warning validation the reasonable and applicability of the model.

Keywords: entropy weight, Extension theory, edible forest products, quality and safety, early warning.

1 Introduction

Consumption of forest products market there is a range of security issues, without exception, sounded the alarm for the consumption of forest product safety. "Forewarned is forearmed, without prejudging the waste, food quality and safety management lie in prevention, food security early warning research: Dong-Ling Zhang et al [1] for the language involved in the safety of agricultural products in the assessment and early warning information and panel data, a warning method based on a variety of discrete choice models. Chapter Durbin et al [1, 2] to the actual food safety monitoring data for samples combined with BP neural network is constructed based on monitoring data and the BP neural network model of food security early warning. From the point of view of existing research, qualitative analysis of food security than more quantitative evaluation of small; quantitative evaluation studies on early warning indicators built more of a corporate warning analysis less.

Extension theory is to the multivariate data created by our mathematician Mr. CAI Wen quantifies decision-making theory [3-5] as a new dynamic analysis model can effectively overcome index warning, experts warning, subjectivity analog warning and one-sidedness, and can be quantitatively warning status of food security which is given

Y. Yang, M. Ma, and B. Liu (Eds.): ICICA 2013, Part II, CCIS 392, pp. 410–419, 2013.

to improve the science and the reality of the early warning. Extension theory has been applied to the ecological environment evaluation irregular flight management warning and environmental quality comprehensive evaluation earthquake prediction comprehensive evaluation [7], there is no application to the consumption of forest product safety warning study. In this paper, the safety factors affecting the consumption of forest products, the establishment of fresh forest product safety warning indicator system, the use of the entropy the theory and ethnics thinking, build Extension Theory Based on Entropy food security early warning model with forest products, combined with Jilin Province 20 12 Years of a forestry enterprise data in an empirical study.

2 Food Security Early Warning of Forest Products Model Construction

2.1 Rating and Evaluation Matter Element Model of the Object

The Extension Element evaluation method [8] the matter-element model of the j-th level is expressed as

$$R_j = \left(N_j, C, V_j\right) = \begin{vmatrix} N_j, c_1, V_{j1} \\ c_2, V_{j2} \\ \vdots \quad \vdots \\ c_n, V_{jn} \end{vmatrix} = \begin{vmatrix} N_j, c_1, \langle v_{j\min 1}, v_{j\max 1} \rangle \\ c_2, \langle v_{j\min 2}, v_{j\max 2} \rangle \\ \vdots \qquad \vdots \\ c_n, \langle v_{j\min n}, v_{j\max n} \rangle \end{vmatrix} \tag{1}$$

In the formula (5-1): N_j Represents the j-th grade physical; c Matter Element feature set; c_i Said grade N_j The i-th feature, i = 1,2,..., n; interval $V_{ji} = \langle v_{j\min i}, v_{j\max i} \rangle$ Be c_i About characterized Required by the magnitude of the range,[6, 7] i.e., the data range of each level on the corresponding feature taken - classical domain. R said matter-element model of all grades, $R \supset R_j$, then we say $V_i = \langle v_{\min i}, v_{\max i} \rangle$ Matter Element N About c_i Taken the magnitude of the section of domains,

$$R = \left(N_j, C, V_j\right) = \begin{vmatrix} N_j, c_1, V_1 \\ c_2, V_2 \\ \vdots \quad \vdots \\ c_n, V_n \end{vmatrix} = \begin{vmatrix} N_j, c_1, \langle v_{j\min 1}, v_{j\max 1} \rangle \\ c_2, \langle v_{j\min 2}, v_{j\max 2} \rangle \\ \vdots \qquad \vdots \\ c_n, \langle v_{j\min n}, v_{j\max n} \rangle \end{vmatrix} \tag{2}$$

Use N_e Said to be evaluated object element, said N_e About characterized c_i Eigenvalue detection data or values to be evaluated on various characteristics of the matter element model R_e Said:

$$R_e = (N_e, C, V_e) = \begin{vmatrix} N_e, & c_1, & v_{e1} \\ & c_2, & v_{e2} \\ \vdots & \vdots & \\ & c_n, & v_{en} \end{vmatrix} \tag{3}$$

2.2 Entropy Weight

Determine security early warning system cannot be separated from the index weight, the weight to determine the objective of the method is directly related to the warning result of the good and bad, the evaluation of the entropy applied to different decision-making process in the evaluation of the effect of the case is an ideal scale [9-11] this paper present entropy weight method to determine the evaluation index weight. The steps are as follows:

Judgment matrix R assumes the evaluation object is divided into m Level both n evaluation indicators, Construction of m rows and n columns.

(1) R matrix normalized

To Normalized r_{ij}, there is the need to distinguish between indicators nature. If the indicator is positive indicators, i.e. the better the value, using equation (4); When indicators for reverse index, the smaller the value, the better, with a formula (5)

$$x_{ij} = \frac{r_{ij} - \min\limits_{i}(r_{ij})}{\max\limits_{i}(r_{ij}) - \min\limits_{i}(r_{ij})} \tag{4}$$

$$x_{ij} = \frac{\max\limits_{i}(r_{ij}) - r_{ij}}{\max\limits_{i}(r_{ij}) - \min\limits_{i}(r_{ij})} \tag{5}$$

Here $\max\limits_{i}(r_{ij})$ Represents the *i-th* indicators within the maximum value of the degree of association of each grade, $\min\limits_{i}(r_{ij})$ Compared to the minimum.

(2) Calculate p_{ij}Proportion of indicators i Lower Dir. j levels index value

$$p_{ij} = x_{ij} / \sum_{i=1}^{m} x_{ij} \tag{6}$$

(3) Calculated indicators entropy rights H_i

$$H_i = -\frac{1}{ln(m)}\sum_{i=1}^{m}\left(p_{ij}ln(p_{ij})\right) \tag{7}$$

Where m is the number of levels. If $p_{ij} = 0$, So $p_{ij}ln(p_{ij})$ = 0.

(4) Calculate the weights of indicators i w_i.

$$w_i = \frac{1 - H_i}{m - \sum_{i=1}^{m} H_i} \tag{8}$$

2.3 Levels of Assessment

Extension with grading the degree of approximation of the distance function and associated function is on the evaluation object level. Calculated as follows:

Step 1 Level Matter Element N Characterized in that must meet the necessary conditions c_i [12, 13] Take the object to be evaluated R_e about c_i the magnitude of v_i Make an assessment.

(1) If $v_i \notin V_{ji}$, then that N Does not satisfy the necessary conditions, c_i Unqualified object.

(2) If $v_i \in V_{ji}$ and proceeds to the next step.

Step 2 Determine pending assessment objects on each level of associate degrees, so

$$\rho(v_{ei}, V_{ji}) = \left| v_{ei} - \frac{v_{j_{\min} i} + v_{j_{\max} i}}{2} \right| - \frac{v_{j_{\max}\ i} - v_{j_{\min} i}}{2} \tag{9}$$

$$\rho(v_{ei}, V_i) = \left| v_{ei} - \frac{v_{\min i} + v_{\max i}}{2} \right| - \frac{v_{\max i} - v_{\min i}}{2}$$

Object first to be evaluated i Characteristic for levels j the correlation function:

$$K_j(v_{ei}) = \begin{cases} \dfrac{\rho(v_{ei}, v_{ji})}{\rho(v_{ei}, V_i) - \rho(v_{ei}, V_{ji})}, \rho(v_{ei}, V_{ji}) \neq \rho(v_{ei}, V_i) \\ \dfrac{-\rho(v_{ei}, v_{ji})}{\left|V_{ji}\right|}, \rho(v_{ei}, V_{ji}) = \rho(v_{ei}, V_i) \end{cases} \tag{10}$$

$K_j(v_{ei})$ Is Higher, indicating v_{ei}, Have V_{ji} More properties, and vice versa, said V_{ji} And L v_{ei}, less similar.

Step 3. Using the entropy weight to be evaluated object N About Level j The degree of association.

$$K_j(N) = \sum_{i=1}^{n} w_i K_j(v_{ei}) \tag{11}$$

Step 4 rating, if

$$K_{\bar{j}}(N) = \max_{j \in \{1,2 \cdots m\}} K_j(N), \tag{12}$$

3 Food the Forest Product Safety-Warning Model and Analysis

3.1 Divided Edible Forest Products Safety Class

According to the national food safety assessment system [11], food security early warning P = {the giant police → heavy police → police → light → no alarm}, I = {huge police}, II = {heavy police no police in}, III = {police} IV = {light alarm}, V = { }, and I, II, III, IV, V ∈ P judgment for any p ∈ P belongs I, II, III, IV V, and calculate the degree of association.

3.2 Construction of Edible Forest Products Safety Warning Indicator System

Food security early warning indicators of forest products quality selection principle [12] [13], based on extensive research and seek expert advice, proposed edible forest products quality and safety of early warning indicators system (see Figure 1).

Consumption of forest products quality and safety warning indicator system framework.

3.3 Determination of the Threshold of Warning Indicators

Based on the consumption of forest products, the level of security and early warning indicator system, countries, industries and local regulations, quality standards of forest products, refer to historical data, to determine the security level warning indicators corresponding threshold (see Table 1).

Table 1. Consumption of forest product table threshold security early warning indicators

Early warning indicators	Alarm level	Giant police	Heavy police	Police and	Light alert	No police
Processing raw material security level	1, microbiological testing pass rate	85 or less	(85,90]	(90, 95]	(95, 98]	98 and above
	2, the pass rate of detection of pesticide residues	85 or less	(85,90]	(90, 95]	(95, 98]	98 and above
	3, the detection of heavy metals pass rate	85 or less	(85,90]	(90, 95]	(95, 98]	98 and above
Processing materials safety standards	4, sampling pass rate of food additives	Below 80	(80, 85]	(85,90]	(90, 95]	More than 95
	5 products, including food additives sampling pass rate	Below 80	(80, 85]	(85,90]	(90, 95]	More than 95
	6, health and safety of the process water pass rate	Below 80	(80, 85]	(85,90]	(90, 95]	More than 95
	7, packaging and other accessories, security pass rate	Below 80	(80, 85]	(85,90]	(90, 95]	More than 95
To processing environmental health safety management level	Health Supervision pass rate of processing enterprises	Below 80	(80, 85]	(85,90]	(90, 95]	More than 95
	Processing enterprises production equipment disinfection pass rate	Below 80	(80, 85]	(85,90]	(90, 95]	More than 95
	10, employing physical examination pass rate	85 or less	(85,90]	(90, 95]	(95, 98]	98 and above
Processing technology security management level	11 engineers and technicians accounted for the proportion of employees	Below 6.0	(6.0, 7.5]	(7.5, 9.0]	(9.0,10.5]	10.5 or more
	12, the proportion of investment funding for technology research and development	Less than 0.3	(0.3,0.4 5]	(0.45,0.65]	(0.65,1]	1 or more

3.4 Calculation and Result Analysis

Jilin Province in 2012, a forestry enterprise data (see Table 2) as to calculate the sample application model.

Table 2. The main indicators of the objects to be evaluated

Index	Numerical	Index	Numerical
The microbial detection pass rate (%)	98	Packaging and other accessories the security pass rate (%)	92
Detection of pesticide residues pass rate (%)	96	Processing enterprises Health Supervision pass rate (%)	91
Detection of heavy metals pass rate (%)	99	Processing enterprises production equipment disinfection pass rate (%)	90
Food additives product the qualified sampling rate (%)	95	The employees physical examination pass rate (%)	98
Products containing food additives sampling pass rate(%)	96	Engineering and technical personnel accounting for the proportion of employees (%)	8
Process water and health security pass rate (%)	96	Proportion of the funds for technology research and development (%)	0.4

The first step, matter-element classical domain and section domain confirmation.

According to Table 1, classic domain is fed with the Security Alarm Assessment of forest products grading standard's range.

$$R_{01}=\begin{vmatrix} N_1, C_1, \langle 98,100\rangle \\ C_2, \langle 98,100\rangle \\ C_3, \langle 98,100\rangle \\ C_4, \langle 95,100\rangle \\ C_5, \langle 95,100\rangle \\ C_6, \langle 95,100\rangle \\ C_7, \langle 95,100\rangle \\ C_8, \langle 95,100\rangle \\ C_9, \langle 95,100\rangle \\ C_{10}, \langle 98,100\rangle \\ C_{11}, \langle 10.5,100\rangle \\ C_{12}, \langle 1,100\rangle \end{vmatrix} \quad R_{02}=\begin{vmatrix} N_2, C_1, \langle 95,98\rangle \\ C_2, \langle 95,98\rangle \\ C_3, \langle 95,98\rangle \\ C_4, \langle 90,95\rangle \\ C_5, \langle 90,95\rangle \\ C_6, \langle 90,95\rangle \\ C_7, \langle 90,95\rangle \\ C_8, \langle 90,95\rangle \\ C_9, \langle 90,95\rangle \\ C_{10}, \langle 95,98\rangle \\ C_{11}, \langle 9,10.5\rangle \\ C_{12}, \langle 0.65,1\rangle \end{vmatrix} \quad R_{03}=\begin{vmatrix} N_3, C_1, \langle 90,95\rangle \\ C_2, \langle 90,95\rangle \\ C_3, \langle 90,95\rangle \\ C_4, \langle 85,90\rangle \\ C_5, \langle 85,90\rangle \\ C_6, \langle 85,90\rangle \\ C_7, \langle 85,90\rangle \\ C_8, \langle 85,90\rangle \\ C_9, \langle 85,90\rangle \\ C_{10}, \langle 90,95\rangle \\ C_{11}, \langle 7.5,9\rangle \\ C_{12}, \langle 0.45,0.65\rangle \end{vmatrix}$$

Similarly, R_{04}, R_{05}, R_{06} is Other classic domain.

Section domain range based on the evaluation of the indicators, in most cases, the consumption of forest products safety warning indicators grading standards of all section domain Rp:

$$R_P = \begin{vmatrix} N_1 - N_5, C_1, \langle 0,100 \rangle \\ C_2, \langle 0,100 \rangle \\ C_3, \langle 0,100 \rangle \\ C_4, \langle 0,100 \rangle \\ C_5, \langle 0,100 \rangle \\ C_6, \langle 0,100 \rangle \\ C_7, \langle 0,100 \rangle \\ C_8, \langle 0,100 \rangle \\ C_9, \langle 0,100 \rangle \\ C_{10}, \langle 0,100 \rangle \\ C_{11}, \langle 0,100 \rangle \\ C_{12}, \langle 0,100 \rangle \end{vmatrix}$$

The second step, using the formulas (9), (10) the calculated correlation degree of each of the features and the corresponding level of safety pre alarm (Table 3).

Table 3. Edible forest product safety warning features and the corresponding level association

Index						
Pass rate of microbial detection	98	-0.8667	-0.8	-0.6	0	0
Pass rate of detection of pesticide residues	96	-0.7333	-0.6	-0.2	0.3333	-0.3333
The qualified rate of heavy metals	99	-0.9333	-0.9	-0.8	-0.5	0
Product sampling pass rate of food additives	95	-0.75	-0.6667	-0.5	0	0
Products containing food additives sampling pass rate	96	-0.6	-0.7333	-0.6	-0.2	-0.3333
Process water hygiene and safety pass rate	96	-0.6	-0.7333	-0.6	-0.2	-0.3333
The security pass rate of packaging and other accessories	92	-0.6	-0.4667	-0.2	0.3333	-0.2727
Processing enterprises Health Supervision pass rate	91	-0.55	-0.4	-0.1	0.125	-0.3076

Table 3. (*Continued.*)

Index						
Processing enterprises production equipment disinfection pass rate	90	-0.5	-0.3333	0	0	-0.3333
The employees physical examination pass rate	98	-0.8667	-0.8	-0.6	0	0
Engineering and technical personnel accounting for the proportion of employees	8	-0.2	-0.0588	0.6667	-0.1111	-0.2381
R & D funding for technology research and development investment proportion	0.4	-0.2	0.1429	-0.1111	-0.3846	-0.6

The third step, the equation (6), (7), (8) and Table 3 to determine the relevant indicators weight (see Table 4).

Table 4. List of index weights

Index	Weights	Index	Weights
Pass rate of microbial detection	0.0852	The security pass rate of packaging and other accessories	0.07 60
Pass rate of detection of pesticide residues	0.0723	Processing enterprises Health Supervision pass rate	0.0655
The qualified rate of heavy metals	0.1327	Processing enterprises production equipment disinfection pass rate	0.059 7
The qualified sampling rate of the food additive	0.0785	The employees physical examination pass rate	0.0852
Products containing food additives sampling pass rate	0.0930	Engineering and technical personnel accounting for the proportion of employees	0.099 4
Process water hygiene and safety pass rate	0.0930	The proportion of investment funding for technology research and development	0.059 4

Finally, using equation (10), and Step 3 calculates the event on the level of correlation degree:

$K_1(N) =$ -0.6383, $K_2(N) =$ -0.5668, $K_3(N) =$ -0.4684, $K_4(N) =$ -0.0798, $K_5(N) =$ -0.2061

Step 4 shows that by the K (N), the maximum value corresponding to the level for events corresponding warning level, in this case the maximum correlation degree-0.0798 belonging low warning level, i.e. the risk of the forestry enterprise in the minor alarm status.

4 Conclusion

(1) The model combines entropy theory and data information based on the early warning indicators to determine the weight, enhanced weight objectivity; using extension theory, consumption of forest product safety warning analysis, to provide a basis for the consumption of forest products safe and effective warning.

(2) For the empirical study of a forestry enterprise security early warning indicators data, early warning results with the actual security situation is basically the same, and thus show that the model has a good practical value in the consumption of forest product safety warning.

References

1. Zhang, D.-L., Sheng, Q., Yang, Z.: Risk assessment and early warning model of quality and safety of agricultural products: Shandong vegetable exports demonstration base. Systems Engineering Theory and Practice 30(6), 1125–1131 (2010)
2. Durbin, Areas of Communications; Xu, J-J.: Food security early warning model based on monitoring data and BP neural network Journal of Agricultural Engineering 26 (1), 221–226 (2010)
3. Tsai: Wen-matter element model and its application, pp. 1–63. Science Press, Beijing (1994)
4. Pan, J.-H., Feng, Z.D.: Based on Entropy Matter Element Extension Model of the middle reaches of the Heihe River ecological environment vulnerability assessment. Ecology and Rural Environment 24(1), 1–4 (2008)
5. Yao, Y., Zhu, J.: Irregular flight management early warning model based Extension correlation function. Southwest Jiaotong University 43(1), 101–106 (2008)
6. Wang, J.: Investigation; Yuan, Y.: Extension evaluation method in the comprehensive evaluation of the quality of the environment. Hohai University 30(1),15–18 (2002)
7. Jiang, C., Shan, T.: Natural aging. In strong earthquake prediction evaluation matter element model and its application. System Engineering Theory and Practice 18(1), 135–138 (1998)
8. Wen, T., Yang, C., Wei, L.: early. Extension engineering methods Beijing, vol. 12(5), pp. 207–209. Science Press (1997)
9. Qiu, W.: Management decision entropy learning and its application Beijing, vol. 15(8), pp. 168–169. China Electric Power Press (2010)
10. Yoshiaki, Y.: Wave front-flatness evaluation by wave front correlation information entropy method and its application for adaptive confocal microscope. Optics Communications 232(6), 91–97 (2004)
11. Xiao, A.: Food security early warning system under the multi-view. China Soft Science 11(6), 150–160 (2008)
12. Xiao, T.: Pure Evaluation of food security early warning system design. Food Science and Technology 6(11), 152–155 (2005)
13. Xiao, J., Ma, Z.S.: The research on food security risk early warning under supply chain environment. In: International Conference On Management Science and Engineering, vol. 9(20), pp. 855–861 (2009)

A Prediction Model Based on Time Series Data in Intelligent Transportation System

Jun Wu[1,2], Luo Zhong[2], Lingli Li[2], and Aiyan Lu[2]

[1] School of Computer Science, Hubei University of Technology,
1 Lijiadun Road, Wuhan, Hubei, P.R. China
[2] School of Computer Science and Technology, Wuhan University of Technology,
122 Luoshi Road, Wuhan, Hubei, P.R. China

Abstract. Intelligent Transportation System has a new kind of complicated time series data which would be the traffic flow, average speed or some other traffic condition information at the same time period. All above data is useful and important for our traffic system which includes the traffic flow prediction, tendency analysis or cluster. With the development in time series analysis model and their applications, it is important to focus on how to find the useful and real-time traffic information from the Intelligent Transportation System. Using this method of building models for the Intelligent Transportation System is the way to solve the traffic prediction problem and make control of the massive traffic network.

Keywords: Time Series Data, Prediction Model, Data Mining, Intelligent Transportation System.

1 Introduction

Intelligent Transportation System (ITS) [1] is becoming a very significant enabling technology in many sectors. Recent advances in hardware development have enabled the creation of widespread traffic networks. Currently, there are many ongoing projects that use Intelligent Transportation System (ITS) for traffic environmental monitoring and data acquisition applications, such as wildlife tracking, habitat monitoring, and building monitoring. Intelligent Transportation System establishes a large-scale, in real-time, accurate and efficient transport management system to integrate people, vehicles and traffic road unified closely. But there are still some difficult problems during the development, such as limited storage, low network bandwidth, poor inter-node communication, limited computational ability, and low power capacity still persisting. Several techniques [2-5] have been proposed to alleviate the problem of limited power at the network level and at the data management level. Another methods [6,7,8] at the data management level is in-network query processing or aggregation.

With in-network aggregation[9,10], a part of the computational work of the aggregation is performed within the sensor node before it sends the results out to the network. It can be easily illustrated by the following simple example of a sensor network used to monitor the average or the maximum temperature in a building [11,12]. The default way to implement this is to have each sensor send its temperature

Y. Yang, M. Ma, and B. Liu (Eds.): ICICA 2013, Part II, CCIS 392, pp. 420–429, 2013.

reading up the network to the base station, with intermediate nodes responsible for just routing packets. In network aggregation, communication among sensor nodes is structured as a (routing) tree with the base station as its root. In this scheme, each node would incorporate its own reading together with the average computed so far by its children. As such, only one packet needs to be sent per node and each intermediate node computes the new average temperature before sending information further up the network. As a result, being able to transmit less data (because of aggregation instead of having to forward all the packets) will reduce energy consumption at the sensor nodes. However the urban traffic develop rapidly, we could not only rely on the old Intelligent Transportation System, but also need to build some useful models to adapt to our urban traffic development to improve the efficient use of traffic data.

This paper would focus on the topic of the prediction of daily traffic information which has characteristics of time series. The majority of this paper will be concerned with discussions of this point which focus on the time series analysis in prediction model. Firstly we discuss about the time series character and the data in Intelligent Transportation System (ITS), and it is easy to find their similarity. Furthermore it is necessary to find the corresponding model for our Intelligent Transportation System (ITS) which could adapt to the traffic development and especially to solve how to use this time series traffic data to make the prediction with our complicated traffic data. In this paper, we have tried two models, one is the Autoregressive integrated moving average (ARIMA) model, and the other is Generalized Regression Neural Network (GRNN) model. At the end, from comparison the prediction result between these two models we found the better model for our traffic time series data.

2 Time Series Data

Business, economic, engineering and environmental data are often collected in roughly equally spaced time intervals, for example, hour, week, month, or quarter. In many problems, such time series data may be available on several related variables of interest.

A time series system is a collection of a space of input series, a space of output series, and an operation carrying an input series into an output series.

Suppose $X(t) \quad (t = 0, \pm 1, ...)$ denotes an input series and $Y(t) \quad (t = 0, \pm 1, ...)$ the corresponding output series.

Then a common time series system has the form as following:

$$Y(t) = \mu + \sum_{u=-\infty}^{\infty} a(t-u)X(u) + \varepsilon(t) \qquad (t = 0, \pm 1, ...)$$

For some sequence of filter coefficients $a(u) \quad (u = 0, \pm 1, ...)$, for some constant, and for some zero mean noise series $\varepsilon(t) \quad (t = 0, \pm 1, ...)$.

This problem of time series system identification is that of determining characteristics of the system from corresponding stretches of input and output series, as the following definition:

$$\{X(t), Y(t)\} \qquad (t = 0, ..., T-1)$$

Firstly suppose that the series $X(.)$ and $\varepsilon(.)$ are stationary and independent.

Secondly Let $c_{XY}(u) = \mathrm{cov}\{X(t+u), Y(t)\}$

Denote the cross covariance function of the two series and let the auto covariance functions $c_{XX}(u)$ and $c_{YY}(u)$ be defined similarly.

Thirdly the time series system leads to the relationship $c_{XY}(u) = \sum_v a(v)c_{XX}(v+u)$

For suitable $a(.)$

Let $f_{XY}(\lambda) = (2\pi)^{-1}\sum_u c_{XY}(u)e^{-iu\lambda}$

Denote the cross-spectrum of the series $X(.)$ with the series $Y(.)$ and make corresponding definitions of the power spectra $f_{XX}(\lambda)$, $f_{YY}(\lambda)$

Let $A(\lambda) = \sum_u a(u)e^{-iu\lambda}$, denote the transfer function of the filter.

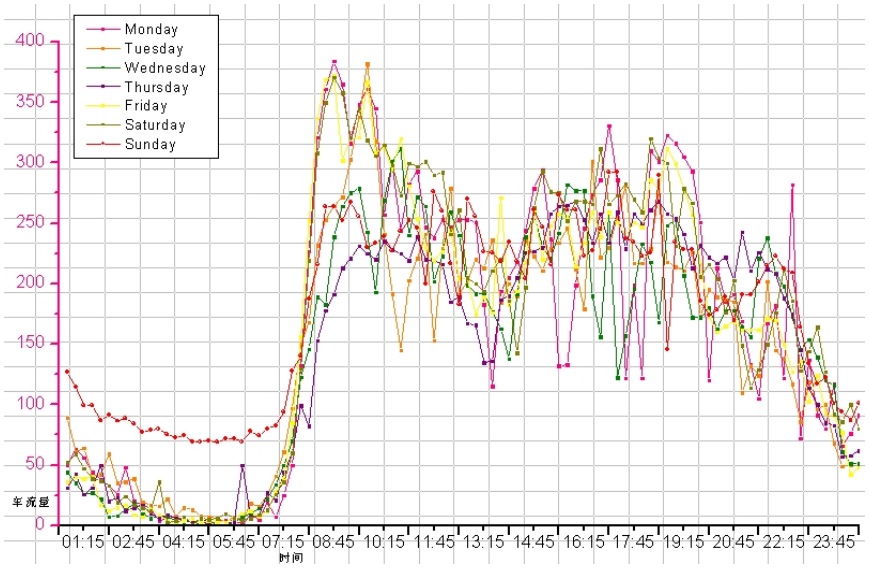


Fig. 1. Daily traffic time series sample from Monday to Sunday

Then the relationship leads to $f_{XY}(\lambda) = A(-\lambda)f_{XX}(\lambda)$ or, if $f_{XX}(\lambda) \neq 0$, to $A(\lambda) = f_{YX}(\lambda)\{f_{XX}(\lambda)\}^{-1}$.

The parameter $f_{YX}(\lambda)\{f_{XX}(\lambda)\}^{-1}$ is called the complex regression coefficient of the series $Y(.)$ on the series $X(.)$ at frequency λ. It provides the transfer function of the best linear filter for predicting the series $Y(.)$ from the series $X(.)$ [13].

The time series [14] is a stochastic process that varies over time, usually observed at fixed intervals. There are a lot of information which belong to time series, for example daily temperature, rainfall, monthly unemployment levels, the annual income and the traffic data. All above are the typical examples of the time series. Let's have a look of

our traffic time series sample which be shown in Fig 1, and Fig 2. This is the weekly example which includes 288*7 points for every 5 min to get.

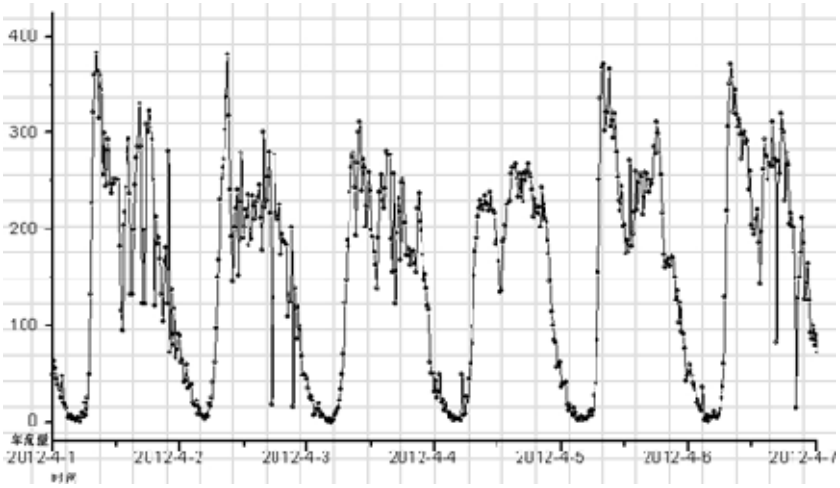


Fig. 2. Weekly traffic time series sample from the Intelligent Traffic System

As far as time series data are concerned, distinctions can be made as to whether the data are discrete-valued or real-valued, uniformly or non-uniformly sampled, uni-variate or multivariate, and whether data series are of equal of unequal length. Non-uniformly sampled data must be converted into uniformed data before clustering operations can be performed. This can be achieved by a wide range of methods, from simple down sampling based on the toughest sampling interval to a sophisticated model and estimation approach.

From the view of data mining, it is possible to get these assortments which include Cluster, Classification and Prediction with the time series data. Clustering is necessary when no labeled data are available regardless of whether the data are binary, categorical, numerical, interval, ordinal, relational, textual, spatial, temporal, image, multimedia, or mixtures of the above data types. The goal of clustering is to identify structure in an unlabeled data set by objectively organizing data into homogeneous groups where the within-group-object similarity is minimized and the between-group-object dissimilarity is maximized. A hierarchical clustering method works by grouping data objects into a tree of clusters. There are generally two types of hierarchical clustering methods: agglomerative and divisive. Agglomerative methods start by placing each object in its own cluster and then merge clusters into larger and larger clusters, until all objects are in a single cluster or until certain termination conditions such as the desired number of clusters are satisfied. Divisive methods do just the opposite. A pure hierarchical clustering method suffers from its inability to perform adjustment once a merge or split decision has been executed.

Model-based methods assume a model for each of the clusters and attempt to best fit the data to the assumed model. There are two major approaches of model-based methods: statistical approach and neural network approach. An example of statistical approach is Auto Class, which uses Bayesian statistical analysis to estimate the number

of clusters. Two prominent methods of the neural network approach to clustering are competitive learning, including ART and self-organizing feature maps. Unlike static data, the time series of a feature comprise values changed with time. Time series data are of interest because of its pervasiveness in various areas ranging from science, engineering, business, finance, economic, health care, to government. Given a set of unlabeled time series, it is often desirable to determine groups of similar time series. These unlabeled time series could be monitoring data collected during different periods from a particular process or from more than one process. The process could be natural, biological, business, or engineered. Works devoting to the cluster analysis of time series are relatively scant compared with those focusing on static data. However, there seems to be a trend of increased activity. From all above analysis, it is easy to find two of the reasons for analyzing and modeling such special series jointly are:

(1) To understand the dynamic relationships among them. They may be contemporaneously related, one series may lead the others or there may be feedback relationships.

(2) To improve accuracy of predictor. When there is information on one series contained in the historical data of another, better predictor can result when the series are modeled jointly.

3 Prediction Model

3.1 ARIMA Model

The Autoregressive integrated moving average (ARIMA) [15] model is a precise forecasting model for short time periods. However, in our society today, due to factors of uncertainty from the integral environment and rapid development of new technology, we usually have to forecast future situations using little data in a short span of time.

Step1. The ARIMA (p,d,q) model is described by parameters in Eq. (1) to (4):

$$\tilde{\Phi}_{\mathrm{p}}(B)W_t = \tilde{\theta}_q(B)a_t \quad (1);$$

$$W_t = (1-B)^d(Z_t - \mu) \quad (2);$$

$$\tilde{W}_t = \tilde{\varphi}_1 W_{t-1} + \tilde{\varphi}_2 W_{t-2} + \ldots + \tilde{\varphi}_p W_{t-p} + a_t - \tilde{\theta}_1 a_{t-1} - \tilde{\theta}_2 a_{t-2} - \ldots - \tilde{\theta}_1 a_{t-q} \quad (3);$$

Where Z_t are observations, $\tilde{\varphi}_1,\ldots,\tilde{\varphi}_p$ and $\tilde{\theta}_1,\ldots,\tilde{\theta}_q$ are fuzzy numbers in the model, so that we could modify this Equation:

$$\tilde{W}_t = \tilde{\beta}_1 W_{t-1} + \tilde{\beta}_2 W_{t-2} + \ldots + \tilde{\beta}_p W_{t-p} + a_t - \tilde{\beta}_{p+1} a_{t-1} - \tilde{\beta}_{p+2} a_{t-2} - \ldots - \tilde{\beta}_{p+\mathrm{q}} a_{t-q} \quad (4);$$

Step2. A general ARIMA formulation model is selected to the traffic data in Intelligent Transportation System (ITS). This selection is of the main characteristics of the 5min traffic time series.

Step3. A model is identified for the chosen data

Step4. The parameters of this model are estimated. The optimum solution of the parameter $\alpha^* = (\alpha_1^*, \alpha_2^*, ..., \alpha_{p+q}^*)$ and the residual s a_t.

Then it is easy to get the fuzzy ARIMA model and its new parameters which shown as Eq.(5):

$$\tilde{W}_t = \langle \alpha_1, c_1 \rangle W_{t-1} + ... + \langle \alpha_p, c_p \rangle W_{t-p} + a_t - \langle \alpha_{p+1}, c_{p+1} \rangle a_{t-1} - ... - \langle \alpha_{p+q}, c_{p+q} \rangle a_{t-q} \quad (5) ;$$

Step5. If the hypotheses of the model are validated, go to the next step, otherwise go to Step3 to refine the model.

Step6. The model would be ready for prediction.

The Autoregressive integrated moving average (ARIMA) has been widely and successfully applied to various systems such as social, economic, financial, scientific and technological, agricultural, industrial, transportation, mechanical, meteorological, ecological, hydrological, geological, medical, military, etc., systems.

The historical data must be less than what the ARIMA model requires which limits its application. The fuzzy regression model [16-18] is an interval forecasting model suitable for the condition of little attainable historical data. In order to make the model include all possible conditions, the spread is wide when data includes a significant difference or bias. The purpose of this paper is to combine the advantages of the fuzzy regression and ARIMA models to formulate the fuzzy ARIMA model and to full the limitations of fuzzy regression and the ARIMA model.

3.2 GRNN Model

The ANN modeling consists of two steps: the first step is to train the network; the second step is to test the network with data, which were not used for training. The processing of adaptation of the weights is called learning. During the training stage the network uses the inductive-learning principle to learn from a set of examples called the training set [19]. Learning methods can be classified as supervised and unsupervised learning. In supervised learning, for each input neuron there is always an output neuron.

As the request of intelligence traffic system is on the way, and as the training speed and predict precision is being higher and higher, the application of Neural Network in traffic field is an unshakable wise choice. Next the algorithm based on the ANN modeling [20] will be given.

In the previous traffic volume forecast by neural network, the sigmoid function usually acts as the approximate base function of forward propagation neural network. However for the lack of suitable theory frame, the topological definition of forward propagation neural network already revealed the flaw in many places. From research works and applications of Artificial Neural Network, it is easy to find that Neural Network has special advantage in traffic field, in which the most important is that it can deal with the non-linear problem perfectly in this field. The Generalized Regression Neural Network is special good at solving the traffic flow forecasting. The architecture for the Generalized Regression Neural Network (GRNN) is shown in Fig 3.

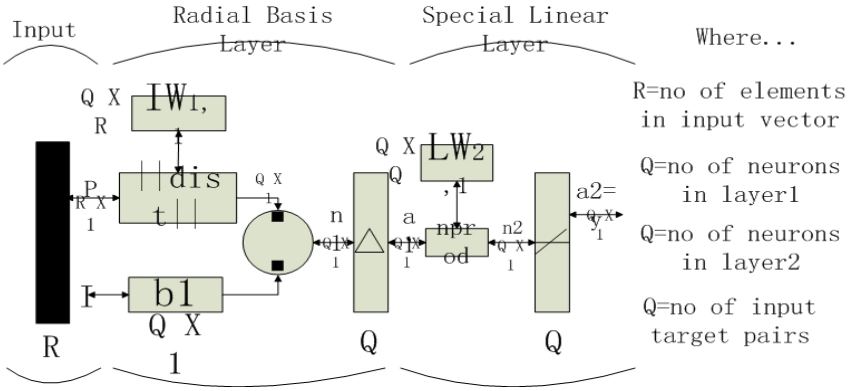


Fig. 3. The architecture of Generalized Regression Neural Network (GRNN)

The first layer is just like that for input of the neural networks. It has as many neurons as there are input/ target vectors. Again, the first layer operates just like the new be radial basis layer described previously. Each neuron's weighted input is the distance between the input vector and its weight vector, calculated with dist. Each neuron's net input is the product of its weighted input with its bias, calculated with net prod. Each neuron's output is its net input passed through radial bas. If a neuron's weight vector is equal to the input vector (transposed), its weighted input will be 0, its net input will be 0, and its output will be 1. If a neuron's weight vector is a distance of spread from the input vector, its weighted input will be spread. The second layer also has as many neurons as input/target vectors, but here LW{2,1} is set to T.

A larger spread leads to a large area around the input vector where layer 1 neurons will respond with significant outputs. As spread becomes larger the radial basis function's slope becomes smoother and several neurons can respond to an input vector. The network then acts as if it is taking a weighted average between target vectors whose design input vectors are closest to the new input vector. As spread becomes larger more and more neurons contribute to the average, with the result that the network function becomes smoother.

Generalized Regression Neural Network Prediction

Input: O;//Original Data Set $\{Data_{Traffic\quad Flow}\}$

Output: PVF;//Prediction Vehicle Flow $\{y_{Vel}\}$

Begin

Train = O.getData(0.7, randomSeed);

Vali = O.getData(0.15, randomSeed);

Test = O.getData(0.15, randomSeed);

GRNN = createModel();

GRNN1 = GRNN.train(Train);

$b_{\rho 0}$ = getWeightVector(ρ_0);

$Layer_0(x)$ = generateFunction($b_{\rho 0}$);

$b_{\rho 1}$ = getWeightVector(ρ_1);

$Layer_1(x)$ = generateFunction($b_{\rho 1}$);

i = 1;
tempValue = 0;
while($MSE_{Layer(i-1)(x)} < MSE_{Layer(i)(x)}$)
i++;

$b_{\rho i}$ = getWeightVector(ρ_i);

$Layer_i(x)$ = generateFunction($b_{\rho i}$, tempValue);

$b_{\rho i0}$ = getWeightVector(ρ_{i0});

$Neural_{00}(x)$ = generateFunction($b_{\rho i0}$);

$b_{\rho i1}$ = getWeightVector(ρ_{i1});

$Neural_{01}(x)$ = generateFunction($b_{\rho i1}$);

j = 1;
while ($MSE_{Neural(i)(j-1)(x)} < MSE_{Neural(i)(j)(x)}$)
j++;

$b_{\rho ij}$ = getWeightVector(ρ_{ij});

$Neural_{ij}(x)$ = generateFunction($b_{\rho ij}$, tempValue);

tempValue = $[y_{Vel}]_{\rho i-1}$;

GRNN2 = GRNN1.validate(Vali);
End

4 Conclusion

Here we have taken the real-time time series data which belong to the urban traffic tunnel in Wuhan. The training sample is the sample for these two model and next 288 data points as test samples to show the efficiency of the prediction method. The test-data results of our prediction would be compared with the actual data which are shown in the Fig. 4.

In this paper we reviewed several measures of Intelligent Transportation System models which are in order to solve the complicated data aggregation and its prediction work. Our main goal was to show that the better time series prediction models for the Intelligent Transportation System which can perform efficiently in this task. At end, from the test sample we could find that the Autoregressive integrated moving average (ARIMA) model works better than Generalized Regression Neural Network (GRNN) at the prediction model for short time periods.

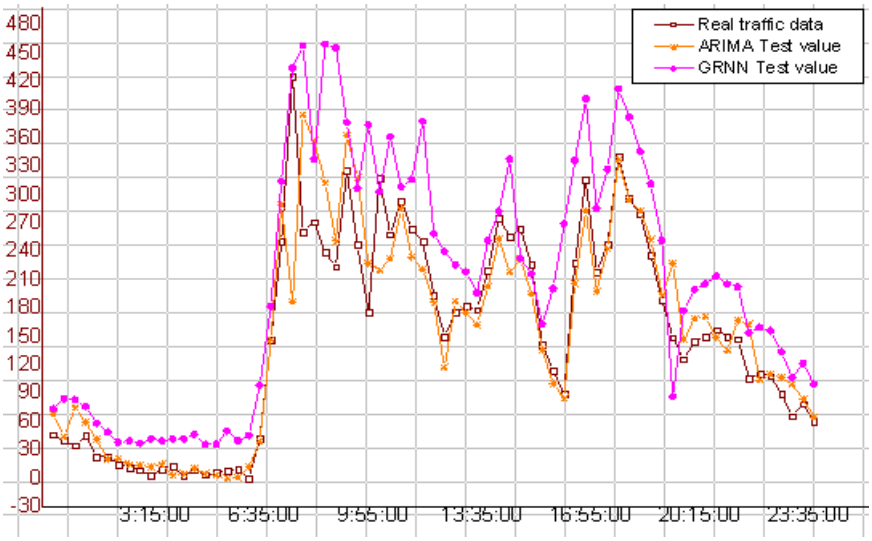


Fig. 4. Result

References

1. Beresford, A.R., Bacon, J.: Intelligent Transportation System. IEEE Pervasive Computing 5(4), 63–67 (2006)
2. Bunn, D.W., Karakatsani, N.: Forecasting electricity prices, London Business School Working Paper (2003)
3. Chen, Y., Dong, G., Han, J., Wah, B.W., Wang, J.: Multi-dimensional regression analysis of time-series data streams. In: VLDB 2002 Proceedings of the 28th International Conference on Very Large Data Bases, pp. 323–334 (2002)
4. Zadeh, L.A.: The roles of soft computing and fuzzy logic in the conception, design and deployment of intelligent systems. In: Proceedings of the 6th IEEE International Conference on Fuzzy Systems, Barcelona, Spain (1997)
5. Gaber, M.M., Zaslavsky, A., Krishnaswamy, S.: Mining Data Streams. A Review.ACM SIGMOD Record Homepage Archive 34(2), 18–26 (2005)
6. Nicolaisen, J.D., Richter Jr., C.W., Sheblé, G.B.: Price signal analysis for competitive electric generation companies. In: Proceedings of the International Conference on Electric Utility Deregulation and Restructuring and Power Technologies, London, UK, pp. 66–71 (2000)
7. Kohzadi, N., Boyd, M.S., Kermanshahi, B.: A comparison of Artificial neural Network and Time Series Models for forecasting commodity prices. Neural Computing 10(2), 169–181
8. Zhang, G.P.: Time series forecasting using a hybrid ARIMA and NN model. Neural Computing, 159–175 (2003)
9. Saini, L.M., Soni, M.K.: Artificial neuralnetwork based peak load forecasting using Levenberg-Marquardt and quasi-Newton methods. IEEE Proc. -Gener. Transm. 149(5), 578–584 (2002)

10. Hagan, M.T., Menhaj, M.B.: Training feedforward networks with the Marquardt algorithm. IEEE Trans. Neural Network. 5(6), 989–993 (1994)
11. Fahrmair, M., Spanfelner, B.: Security and privacy rights management for mobile and ubiquitous computing. In: Workshop on UbiComp Privacy, Tokyo, Japan, September 11 (2005)
12. Huang, E., Antoniou, C., Wen, Y., Ben-Akiva, M.: Real-Time Multi-Sensor Multi-Source Network Data Fusion Using Dynamic Traffic Assignment Models. In: Intelligent Transportation Systems, pp. 1–6. ITSC (2009)
13. Kang, Y., Lee, H., Chun, K., Song, J.: Classification of Privacy Enhancing Technologies on Life-cycle of Information. In: Proceeding of The International Conference on Emerging Security Information, Systems, and Technologies, pp. 66–70 (2007)
14. Breeden, J.L.: Modeling data with multiple time dimensions. Computational Statistics and Data Analysis 51(9), 4761–4785 (2007)
15. Tseng, F.-M.: Fuzzy seasonal ARIMA model for forecasting. Fuzzy Sets Systems 126, 367–376 (2002)
16. Jang, J.S.: Predicting chaotic time series with fuzzy if-then rules. In: IEEE International Conference on Fuzzy Systems, San Francisco, USA, pp. 1079–1084 (1993)
17. Liao, Kao, W.-H., Fan, Y., Ming, C.: Data aggregation in wireless sensor networks using ant colony algorithm. Journal of Network and Computer Applications 31(4), 387–401 (2008)
18. Cooper, J., James, A.: Challenges for database management in the internet of things. Source. IETE Technical Review 26(5), 320–329 (2009)
19. Ritchie, K.M.S.: Comparison of Traditional and Neural Classifiers for Pavement Crack Detection. ASCE 120(4), 552–569 (1994)
20. Breeden, J.L.: Modeling data with multiple time dimensions. Computational Statistics and Data Analysis 51(9), 4761–4785

A Model-Driven Interactive System

Hao Wu and Qing-yi Hua

School of Information Science and Technology Northwest University,
710069 Xi'an, China
{wh,huaqy}@nwu.edu.cn

Abstract. Model-driven interactive system development mainly solves the user interface development problems of the diverse and heterogeneous client devices in the new computing model such as pervasive computing and cloud computing. This paper analyzed the problems that existed in this development approach, and put forward an architecture pattern ALV+. This pattern was integrated with CAMELOAN reference framework. So a novel model-driven interactive system development approach was designed. This approach divided the user interface into several models: abstract interface, concrete interface, device interface and event description, which made the presentation and transformation of the models simple, and decreased the redundancy among the models. Then a case study was used to verify the effectiveness of the approach, and confirmed that this method could solve the problems of the model-driven approach to some extent.

Keywords: new computing model, model-driven, architecture pattern, CAMELOAN reference framework.

1 Introduction

Model-driven interactive system development is considered to come to the third stage of its evolution process [1]. In this stage it mainly solves the user interface development problems of the diverse and heterogeneous clients in the new computing model such as pervasive computing and cloud computing. These development methods have obtained some achievements, but there are some problems in these methods: 1) A weakness of model-driven development approaches is that it seldom uses the technology of the mature architecture pattern, so it is difficult to apply the research results of architecture pattern in recent 30 years, so as to bring so many problems about the design, development, test, evaluation, deployment and evolvement of systems [2]. 2) In the model-driven development approaches, the presentation and transformation of models are complex and difficult to master. These approaches produce the limited user interface that the development tools are supported [3]. 3) In quite a few model-driven system development methods, the models are divided into interactive models, presentation models, platform models and so on. Although the complex sorts can definitely describe the properties of use interface from different views, there is redundancy that is difficult to erase among many models. The redundancy leads to the

Y. Yang, M. Ma, and B. Liu (Eds.): ICICA 2013, Part II, CCIS 392, pp. 430–439, 2013.

cumbersome use interface generation process. 4) Model-driven development approach is a step wise refinement approach to address the static appearance of user interface, but there are weaknesses in solving the dynamic behavior and control problems, because the response and disposal of events and the data transformation between user interface and the application computation are often closely interrelated with final interface, resolving these problems at the abstract level often leads to interweave the user interface and the application computation.

This paper puts forward a model-driven development approach for interactive system, which supports the auto-production of user interface for multi-devices, and aims to do some exploratory research for the aforementioned problems.

2 Related Works

2.1 Interactive System Architecture Pattern ALV

Interactive system architecture pattern ALV shows as Figure 1 [4]. In this pattern, shared abstract denotes the application computation of system, view denotes the user interface of system, and link links the abstract with view by the constrains-based ways. ALV pattern has a good effect in supporting CSCW. The front devices in CSCW environment resemble client devices in new computation mode: there is a share application service in back end. But in ALV-based CSCW environment, front devices are fat client that have the strong interface presentation and computation power. They have the same interactive modalities; all are the standard configuration of PC. Whereas in the new computation mode, the presentation and computation power of front devices have the remarkable differences. They may be fat client, but they may be the thin clients that have the weak presentation and computation power and have the different modalities. So ALV pattern couldn't apply to the new computation mode directly, it needs some expansions and improvements.

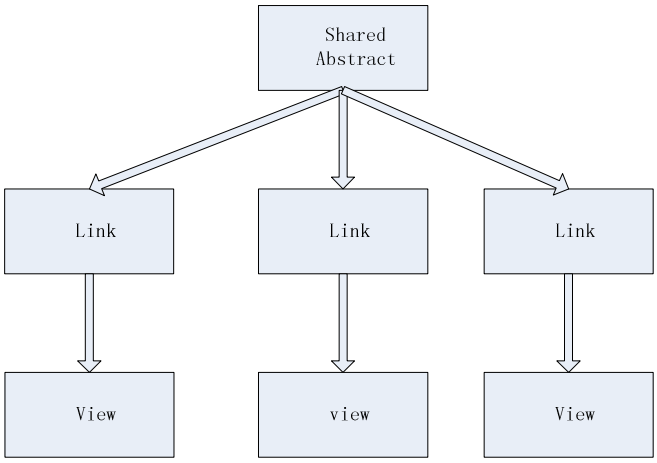


Fig. 1. This Figure shows architecture pattern ALV

2.2 CAMELEON Reference Framework

CAMELEON reference framework [5] shows as figure 2. In this framework, the development of interactive system begins with building concept and task models. Then the framework produces the abstract models from these models. The abstract models transforms into the concrete models by the models mapping. In the last, the final user interface is produced by the concrete models. CAMELEON reference framework doesn't give the specific models; it just refers to the four stages of the life periods of the user interface development. Many a model-driven user interface development works were based this framework [1] [6]. This framework is a top-down development mode for the new system development. The problems that this framework exists have been explained in the preceding part of the text. This paper presents a method that doesn't preclude this approach, but it only limits to the building of static user interface. For the dynamic parts of user interface, it adopts the down-top design ways, it makes a clear separation between user interface and application computation.

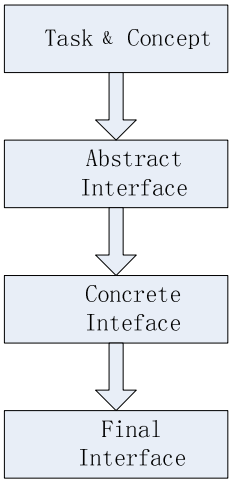


Fig. 2. This Figure shows the CAMELOAN reference framework

3 Architecture Pattern ALV+

For using ALV pattern in the new computation mode, it has to be extended and evolved. This paper names the pattern extended as ALV+ to illuminate it was extended by the ALV, showing as figure 3. Shared abstract still denotes the application computation, *view* denotes the user interface of the client also. The difference with ALV mainly lies in that the function of the *link* doesn't link the abstract and the view by the constrains-based manner no more, and the transverse lines between the links is added to support the dynamic switch user interface between the devices.

In the new computation mode, the client devices that have the different computation, presentation power and interactive modalities access the same service.

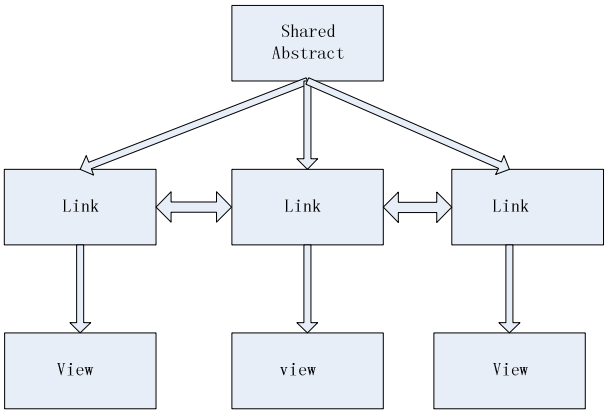


Fig. 3. This figure shows the ALV+ pattern

The problems that ALV+ solves are that the same application service needs to develop the different user interface for these devices that have the large differences, and the user interface may be switched between the devices. So, the function of the link unit is modified as having the function of "the application service transformer". Like that the transformer transforms the high voltage electrical current that is carried from power plant into different low voltage electrical current according to the devices' needs, the application transformer (link) reifies the abstract interface that is passed by application service with the concrete interface that exists link unit for the different client devices, so the final interfaces for these devices are built. At the same time, while the user interface of one device need to switch to the user interface of another device at running time, it saves the states data of the former, and transmits to the latter to dynamically produce corresponding user interface(show as the transverse line of figure 3). While deploying the components of ALV+, the abstract, link and view can locate at the different devices like ALV, they link together by the network. Also according to reality needs, the abstract and link, or link and view could be located at the same machine.

4 Model-Driven User Interface Development Approach

Our model-driven user interface development approach shows as figure 4. In this figure, application service sends the abstract interface to the link unit by needs. Abstract interface is the logic description of the final interface (a window or a HTML page) that is independent of platform and program language. After the link unit receives this abstract interface, it finds out the concrete interface that reifies this abstract interface in link unit. The concrete interface is dependent on the platform and program language. In the C/S architecture, after abstract interface combines with concrete interface, device interface is produced. Device interface is the accurate description of the final interface. Device interface is distributed to view unit. It could make network transform easy. View unit automatically generates final interface by the device interface. In the B/s architecture, after abstract interface combines with concrete interface, the HTML page is automatically generated. The events that the final interfaces give birth to are defined by the event description. Event description.

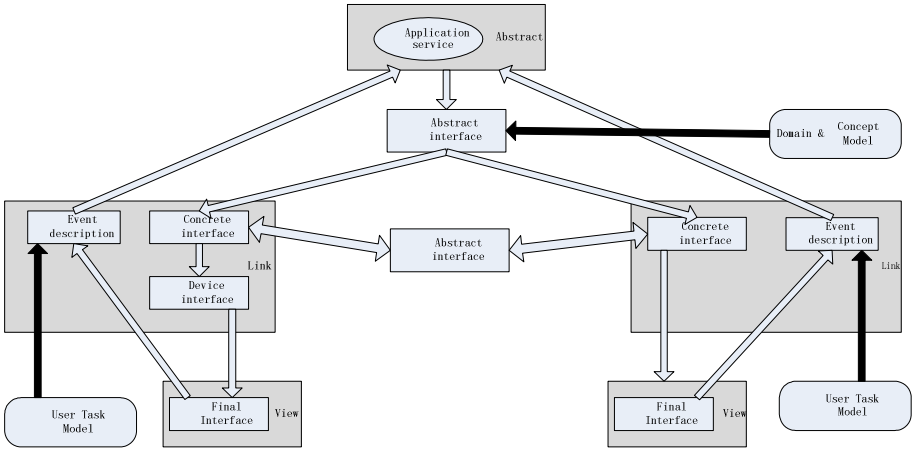


Fig. 4. This figure shows model-driven user interface development approach

Application service. After these data are handled by application service, they may return a new abstract interface to the link unit. While the interface needs to switch between the devices, the link unit of the source device collects state data and generates the abstract interface. The abstract interface is sent to the link unit of the target device. The target device then transforms the abstract interface into the final interface in the target device.

This approach uses the ALV+ architecture pattern: the application service in the top of the figure is the shared abstract of the ALV+. The concrete interface, device interface and event description in the middle of the figure is the link of the ALV+. The final interface in the bottom of the figure is the view of the ALV+. The abstract interface between the two concrete interfaces is the carrier of the data that the link levels of the two devices need to transform while the interface of the one device needs to switch to the other device at run time. In this approach ALV+ works as the organization of the system components and the guidance of the components deployment. This is the embodiment that architecture pattern in the model-driven interactive system development exert influence.

This approach also uses the CAMELOAN reference framework: the approach includes the task, conception model, abstract model, concrete model, and final interface. The differences with the conventional CAMELOAN reference framework use are: 1) In this approach two components, device interface and event description, are introduced by the needs. 2) This method combines the top-down development way with down-top development way. Task and conception model leads to abstract interface. The foundation that the abstract interface is produced is the needs of the interaction between application service and final interface [7-9]. This is a top-down ways. The task model guides the design of the event description. The event description could clearly define its contents after the concrete interface confirms the component type that produces the event. This is a down-top ways. In this process, although the final interface doesn't clearly present, it should be assumed in the

design's brain while designing the concrete interface. 3) CAMELAON reference framework gives a development process [10], this approach adopts its basic conception. But our approach is a flow. This flow mainly uses the undertone arrows to link together in the figure. The fuscous arrows show the leading relation.

This approach has these advantages: 1) because the most application computation works have finished by the application service, and the final interface just has the function of data transform that the application computation needs, the requirement for computation power of devices is furthest lowered, the difference among client devices computation power in multi-device environment is shielded [11]. 2) Because introduction of interactive requirement leads to the abstract interface model, every abstract interface corresponds to a frame/page interface, which makes the presentation of the interface straightforward. At the same time, the concrete interface straight depicts the abstract interface, there needn't complex model transform arithmetic. The express of the abstract interface and the concrete interface doesn't have redundancy. 3) The abstract interface is independent of the platform and the program language, in the event description of the different devices the data format is uniform, so the application service is independent of the specific platforms and languages. The works that the system supports new devices are simple, the only need is to add the new link unit to the system [12]. 4) The programmers that have different duty write the different parts of the interface description, which makes the work assignment clear. For example, application programmers write abstract interface, interface designers write concrete interface. 5) The building of concrete interface could aim at the different presentation power and modality of devices, so the requirements of multi-device environment are meet.

The description of the abstract interface, the concrete interface, the event description and the device interface uses XML language to express.

4.1 Abstract Interface

Abstract interface is the logic level of the interface, which has the independency on platforms and languages. It gives out the type of interaction (ToI), the content of the interaction, and the correlative attributes of interaction, and so on. By analyzing the existing services and correlative papers, we define eight kinds of ToI: input, output, select, modify, create, destroy, start and end. In the work of this paper, we assume that these ToIs could cover all the possible interaction between application services and user interface.

Abstract interface is usually composed by the forest which makes of groups and ToIs. The child nodes of the group may be groups or ToIs, but ToI doesn't have child nodes no more. Group shows the construction and organization relations between ToIs. A ToI denotes a interaction behavior that could happen between application service and user interface, which presents one or more interface components while being implemented.

4.2 Concrete Interface

Concrete interface is the physical level of interface, which mainly presents the static appearance that is special to some type of device interface. It is closely correlative to the interface language (Such as Java, html).

Concrete interface describes the ToIs and groups of the abstract interface. In concrete interface, the description for the elements of abstract interface divides into three levels: group, type and name. The description for the group would work on all the nodes that belongs the tree or sub-tree whose root node is the group; The description for the type makes the ToI that is the type have same characters; The description for the name aims to the ToI that has the given name.

Concrete interface could be sorted as basal and specific interface. Basal concrete interface includes the default description for every type of ToI, which could be reused by several services. Specific concrete interface is designed for the specific demands of one or more service, which has high priority in the interface combination. If a ToI or a group of a abstract interface doesn't have name mapping, then basal concrete interface is used to achieve the specification for them.

From the definition of the abstract interface and the concrete interface and their interrelation we can see that abstract interface and concrete interface don't have complex model presentation, mapping and transformation arithmetic, but the description ability doesn't be lowered.

4.3 Device Interface

Device interface only appears in the application of C/S architecture, which is the whole description of final interface, and combined by the abstract interface and concrete interface. But the combination isn't the simple combination. We throw off the redundant components of the interface description in the combination process. Device interface is automatic produced by the generation engine. The generation engine checks the solecism and semantic error of the description languages. Device connection unit would transfer the device interface to the client, which is more simple and convenient than transferring the source codes or binary codes of the interface.

4.4 Event Description

Event description describes the dynamic sides of interface, which describes the communication and control between interface and service. Event description describes the data format that needs by the semantic computation that interface transfers to the application service. The consistency of interface feedback data format could assure the device independence of the application service, improve the dialog independence of interface and application service. The absolute separation of interface and application service is often unpractical. For the data handle that needn't the interaction between interface and application, we provide a callback function lib. While interface is running, the functions in the callback function lib are called so as to fasten the speed of the interface feedback. For example, the local date function is called so as to acquire the client time.

In the event description, the description of the events base on the events that could be supported by the implement language, such as the events in the Java language.

By the method of the event handle we can see that in this method the application computation and interface presentation and the communication and control between the two are clearly defined. The description of the event handle and the final interface handle for the event has the one by one corresponding relation. So the user interface automatic generation process is predigested ulteriorly.

Fig. 5. This figure shows smart phone version user interface

5 Case Studies

To verify that this approach can effectively support multi-device user interface development, we used this approach to build calendar service software that can add, edit, query, and show calendar. The core functions of the service and the link units located in a windows server respectively. The service communicated with the link units by the socket. We respectively developed concrete interface and event description for PC version and mobile phone version interface which used Java Swing and HTML version interface. One can access the service by the PCs and smart phones that support Java language. Figure 5, figure 6 and figure 7 are respectively the screenshot of the calendar presentation function for Java phone version, PC version and HTML version, which illuminate that the same abstract interface could realize different interface for different devices and languages. By the practically developing to verify, this approach could automatically produce the elaborate user interface for the different devices, and it is easy to study and use.

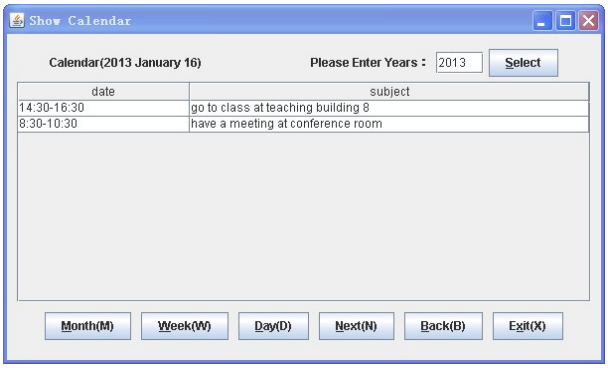


Fig. 6. This figure shows the Java Swing version user interface

6 Conclusions

This paper puts forward a model-driven interactive system development approach. This approach has these advantages: 1) the classic model-driven method CAMELOAN reference framework combined with the architecture pattern, which not only make use of all advantages of user interface development that model-driven

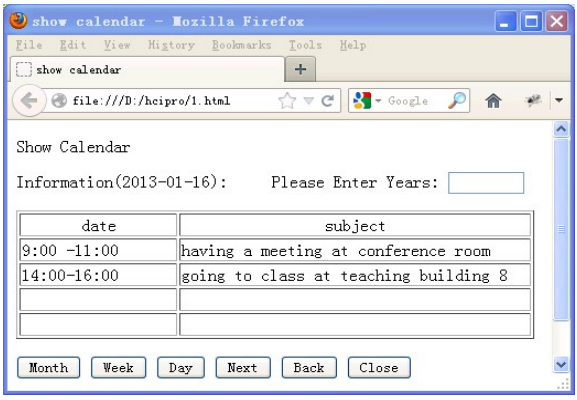


Fig. 7. This figure shows the HTML version user interface

Approach brings, but also the use of the architecture pattern makes the organization structure of system function components clear, deployment convenient, and is convenient to system test and evolution at the later stage. 2) In this approach the presentation and transformation of the models are straightforward and suit for all kinds of devices user interface development. 3) In this approach there isn't redundancy among the models. 4) This approach used top-down and down-top development way, which make static interface development a step wise refinement process, and in the dynamic parts of the interface application computation could better separate with the presentation of the interface. 5) This approach supports the dynamic switch interface between devices. The next work is to make this approach support

more devices and languages, and to get the experience that use the approach for the large application system, so we could improve this approach.

Acknowledgments. The research is supported by the National Natural Science Foundation of China (Granted No. 61272286) and the Specialized Research Fund for the Doctoral Program of Higher Education of China (Granted No. 20126101110006).

References

1. Paterno, F., Santoro, C., Spano, L.D.: MARIA: A universal, declarative, multiple abstraction-level language for srvice-oriented applications in ubiquitous environments. ACM Transactions on Computer-Human Interaction 16(4), 19–30 (2009)
2. López-Sanz, M., Vara, J.M., Marcos, E., et al.: A Model-Driven Approach to Weave Architectural Styles into Service-Oriented Architectures. In: Proceeding MoSE+DQS 2009, pp. 53–60. ACM, New York (2011)
3. Vanderdonckt, J.: A MDA-Compliant Environment for Developing User Interfaces of Information Systems. In: Pastor, Ó., Falcão e Cunha, J. (eds.) CAiSE 2005. LNCS, vol. 3520, pp. 16–31. Springer, Heidelberg (2005)
4. Ralph, D.H.: The abstraction—Hnk—view paradigm: using constraints to connect user interface to applications. In: Proceedings of the Conference on Human Factors in Computing Systems, California, pp. 335–342 (1992)
5. Calvary, G., et al.: A Unifying Reference Framework for multi-target user interfaces. Interacting with Computers 15(3), 289–308 (2003)
6. Saleh, E., Kamel, A.M.R., Fahmy, A.: A model driven engineering design approach for developing multi-platform user interfaces. WSEAS Transactions on Computers 9(5), 536–545 (2010)
7. Lorenz, A.: Research Directions for the Application of MVC in Ambient Computing Environments. In: Workshop on Pattern-Driven Engineering of Interactive Computing Systems (PEICS), pp. 336–348. ACM, New York (2010)
8. Ahmed, S., Ashraf, G.: Model-based user interface engineering with design patterns. The Journal of Systems and Software 80(8), 1408–1422 (2007)
9. Vara, J.M., Marcos, E.: A framework for model-driven development of information systems: Technical decisions and lessons learned. Journal of Systems and Software, 2368–2384 (2012)
10. Ameedeen, M.A., Bordbar, B., Anane, R.: Model interoperability via Model Driven Development. Journal of Computer and System Sciences 77(2), 332–347 (2011)
11. Izquierdo, A.L.C., Jouault, F., Cabot, J., Molina, J.G.: API2MoL: Automating the building of bridges between APIs and Model-Driven Engineering. Information and Software Technology 54(3), 257–273 (2012)
12. Agner, L.T.W., Soares, I.W., Stadzisz, P.C., Simão, J.M.: A Brazilian survey on UML and model-driven practices for embedded software development. Journal of Systems and Software 86(4), 997–1005 (2012)

WAP Mobile Teaching System Based on Campus Network

Fang Chen[1] and Zhengjie Li[2]

[1] Academy of Fine Arts, Huanggang Normal University, 438000, Huanggang
[2] School of Foreign Studies, Huanggang Normal University, 438000, Huanggang

Abstract. The mobile teaching system of WAP teaching website based on digital wireless campus network, its design is primarily on teaching of WAP site development and construction of hardware environments based on WAP technology. Thesis start from current students e-learning habits, elaborated on WAP Mobile teaching systems concepts and technologies, through the analysis of the mobile learning system software design of the main teaching site WAP model, designed for the digital wireless systems hardware environment of the campus to achieve digital wireless campus network based on WAP Mobile teaching system construction of the teaching website.

Keywords: campus network, WAP, WAP teaching website, mobile teaching system.

1 Introduction

This applies when we are working on teaching Web site teaching mode of network teaching resources in time, WAP technology we saw the rise of another outline of the teaching model. Unlimited fast and convenient WAP to mobile terminals, wireless networks are not subject to geographical constraints, as opposed to current computer, free movement of mobile terminal compact, easy to carry [1-6]. 3G use of communications technology brought more good development prospects to the WAP technology: cheaper communication charges, faster communication speed, rich content resources [7-11]. WAP technology will also be a mobile world "new Internet". WAP technology momentum, so that we can have unlimited vision to a mobile teaching system: students sitting on the subway and holding teaching of WAP Mobile phone, cell phone keyboard, should be able to finish today's homework of teachers [12-16]. WAP technology brings wireless, teaching anywhere, anytime. As long as the signal coverage, teaching through a WAP Terminal. This model of mobile learning system teaching reform to our future with new vision [17].

2 An Overview of the WAP Mobile Teaching System

2.1 The Website of WAP

Connecting mobile devices with WAP is wireless communication protocols of the Internet, WAP terminals are run in a wireless network. WAP site is run on a WAP site

Y. Yang, M. Ma, and B. Liu (Eds.): ICICA 2013, Part II, CCIS 392, pp. 440–449, 2013.

server that is running on a WAP server in the Internet [18-22]. That is, the WAP terminals and server-side facilities are located in two different areas, namely areas of wireless and wired Internet. When connecting wireless wired Internet and WAP, gateway, gateway will encode the WAP client requests Internet recognition signals that are sent to the server-side. Then there will be server-side signal decoded into a WAP client can recognize signal, sent to the WAP client. WAP network structure consists of three components, the WAP gateway, WAP client and WAP content server, integral in these three areas [23, 24]. Which WAP gateway protocols "translation", is a bridge linking two different mobile networks and the World Wide Web; WAP content server stores a lot of information, to provide WAP Terminal user to access, query, browse, and so on.

2.2 The Definition of Mobile Education System

Mobile teaching means in mobile computing devices at any time, any place with the help of teaching. Mobile education system in teaching website based on WAP, WAP teaching Web site acts as a information of organizers and processing, clearing house for teaching and as a mobile system, play a teaching role of the platform. Teachers through the network teaching platform, using a range of teaching strategies have the purpose to spread knowledge to the learners [25-30].

Based on WAP teaching website Shang of moved type teaching system is such of a teaching system: to WAP wireless environment for support of, with WAP terminal equipment through wireless network connection Internet, access Internet Shang of through WAP teaching website, in WAP teaching website of network managers and network teachers of teaching policy and learning organization of control Xia, students through knowledge browse, and real-time Exchange, and online quiz, and online answering, and message feedback, way and teachers for remote interactive, to in any time, and Anywhere a mobile networked teaching system of teaching. Due to this teaching system through learner reports available for developers to advance the development of online teaching resources for implementation of the teaching process, so we also call this mobile learning system teaching system based on connection of mobile browsing. By WAP wireless communications technology enables wireless access to the Internet, allow learners to browse through online information, for teaching purposes.

Mobile education system teaching website based on WAP, it integrates the instructional design, software engineering, development, learning theory, information theory, and other ideological theory, also contains a teaching website, multimedia teaching courseware, teaching equipment and mobile terminals, is the organic combination of theory and teaching resources. Developer basis theory guidance through technology achieve, completed has WAP teaching website of development, building has based on WAP teaching website Shang of moved type teaching system of information center, acts as a with teaching platform of role; network teachers according to teaching of need, select suitable of teaching policy, through WAP teaching website this platform, issued teaching task, implementation network teaching, and received teaching in the various problem of feedback, according to reality timely amendment teaching; Learning based on learning theory by teaching online browse WAP sites for learning and reflect problems in the learning and homework in a timely manner. Figure 1 reflects a teaching website based on WAP Mobile learning system of structural drawings.

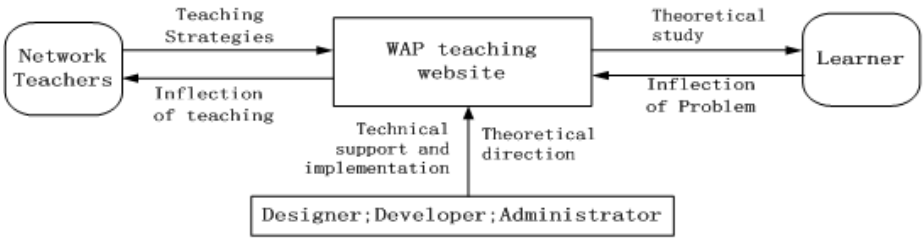


Fig. 1. WAP mibile teaching system of teaching website structure

3 WAP Mobile Characteristics of the Teaching System

On a teaching website based on WAP Mobile teaching system, whose main characteristics are: mobility, interaction of teaching content, teaching activities implementation of digital and multimedia.

3.1 Teaching Activities Mobility

Mobile teaching system very important characteristic is mobility, through mobile devices is no longer allow organizers of teaching and learning is limited in front of the computer desk, gives learners a learning experience anywhere, anytime.

3.2 Interactive Teaching Content

Temporal and spatial separation between teacher and students and two-way communication between teachers and students is one of the characteristics of Internet teaching, mobile education system using mobile communication technology and mobile devices, making exchanges between teachers and students and between life and is no longer a problem, it provides management and learning support services with a good range of interactive channels and ways.

3.3 The Implementation of Digital and Media

Mobile learning is an e-learning based on mobile devices. With e-learning is similar, it is also in the e-learning environment, use of digital learning resources for e-learning, it renders the learning material can be text, images, and video/audio, and other multimedia information form. Implementations of mobile learning is a number, it also provides all the characteristics of network teaching system of General, across time and space, wide object, resource-rich, multimedia, open, interact, diversity, diversity, the roles of teacher and student learning style teaching story of diversity, sharing, collaboration, autonomy, education information updated quickly. In addition to these characteristics, system of wireless ' mobile teaching bondage, free mobile, anytime, anywhere learning, put small, flexible and other characteristics. With the popularity of 3G network coverage and 3G mobile phone, WAP-based mobile education system will better reflect its advantage.

4 Mobile Software Environment Design of the Teaching System

Moving in instructional systems development, teaching Web site development is its core content. An excellent teaching website, is a combination of teaching and technical perfection. In the development of a teaching Web site, in addition to technical factors, must also be advanced guidance teaching design concepts. In teaching Web site development of WAP, should also comply with this principle. We call this Web development model called the ID-SE model, the so-called ID-SE model is instructional design (Instructional Design) and software engineering (Software Engineering) combine, guide the teaching of mobile Web design, the flow chart as shown in Figure 2.

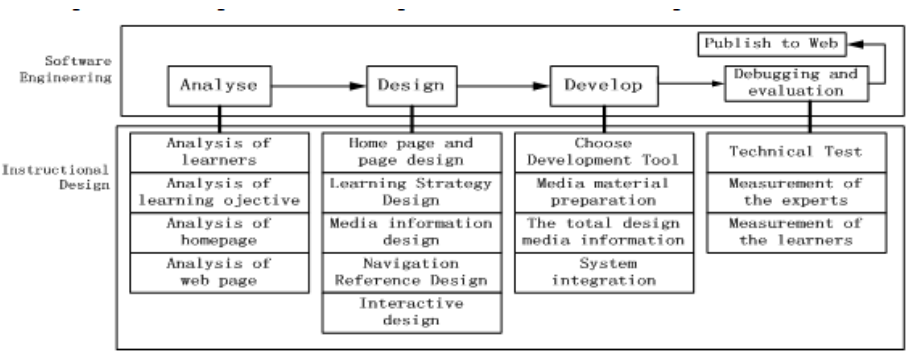


Fig. 2. ID-SE model website design filw chart

Mobile teaching system built primarily around theories and technology systems integration, from a technical point of view, when WAP teaching website development is completed, do we have to do is surround Web site construction of hardware and software support services. WAP system mainly by services and WAP access, WAP gateway server. As shown in Figure 3 WAP system application infrastructure model.

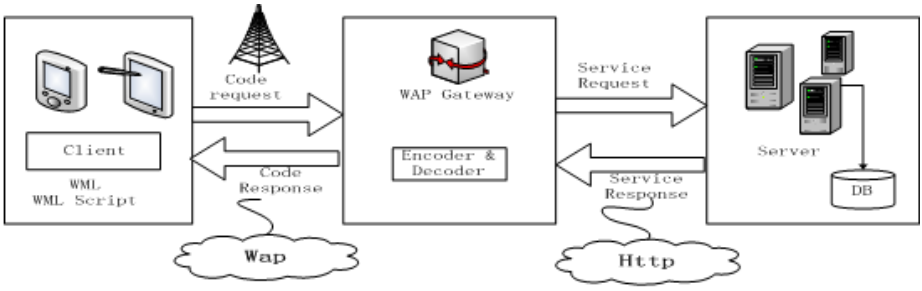


Fig. 3. The basic structure model of WAP application system

4.1 The Realization of the WAP Access Services

WAP access should include WAP terminals receiving device, supporting WAP wireless network environments. WAP Mobile devices on wireless networks, achieving the purpose of access network services. In at this time we need of equipment has moved terminal equipment phone, wireless network, to China moved communications company of GPRS wireless network for cases, specific set way following: in turn select "service", and "set", and "connection set", and "modified service set", and "transmission way set", in "GPRS access points" Servicers name and IP address, "landed type" select automatically, "user name" and "password" option default for empty.

4.2 WAP Gateway Serve

Complete WAP micro-browser and WAP gateway server transfer protocol conversion, transfer text formatting language. WAP gateway receives requests for WAP Mobile phones, through the HTTP protocol to the WAP server, WAP server responds to the request, and returns the appropriate WML Web pages WAP gateway, returned by the WAP gateway to WAP Mobile phone that made the request. WAP gateway construction costs are very expensive, we use China Mobile GPRS wireless network, provided by the company, you can pay using the WAP gateway.

4.3 Construction of WAP Content Server

WAP gateway using the HTTP protocol to access WAP WML file servers, WAP server therefore does not require a dedicated hardware, software and Web-compatible. Using WML description language for the page, rather than HTML. Following to "Windows Server 2003+IIS" for cases build WAP server, built a site folder first you to created a independent Yu you of main site of new of site folder, then in IIS platform created a new of MIME type first select WAP site of property, through "control panel-> management tools->IIS (Internet Server)" open Internet information service (IIS) Management Manager, then in IIS Service in the Administrator window, click after the host computer name, press the right mouse button, click on the "Properties" item, select "Internet information services", select "MIME mapping" "Edit" button again to add a new file type, put the following types of WML files in:

Table 1. Wml document extension and the corresponding content type

The relevant extension	Content type
wml	Text/vnd. wap. wml
wmlc	Application/vnd. wap. wmle
wmls	Text/vnd. wap. wml script
wmlsc	Application/vnd. wap. wml script
wbmp	Image/vnd. wap. wbmp

After setting is complete, restart the IIS service, and WML Web pages so that you can be placed on your own Web site, you can provide WAP Web page service.

After you complete the system after the development and construction of hardware and software services, a mobile teaching system is completed. Network teaching of teachers as long as a reasonable use of WAP Web site platform, play freedom of wireless, mobile learning, advantages of anytime, anywhere learning, instructional design to be reasonable, good steering students to use this platform to learn, teaching will be able to achieve good results.

5 Conclusions

The thesis according to the technical features of WAP, WAP teaching Web development based on architecture teaching system of teaching website based on WAP Mobile framework, introduced digital wireless campus network based on WAP Mobile learning system teaching website development solutions, respectively, on teaching Web site development, mobile WAP hardware environment to build the teaching system design and planning. In the broad prospects of development of WAP technology, WAP site on the teaching for our future educational reform brings new vision, mobile teaching under conditions of maturing technology, the system will pop up, break the limits of space, time, place of mobile teaching mode in occupying a place in the field of education.

Mobile learning research has attracted attention of many experts and scholars at home and abroad, on the theory of mobile learning, technology, applications made extensive and intensive studies have been conducted in areas such as language learning a useful attempt, but did not like the mobile learning and eLearning the same way as traditional learning into our learning activities. Mobile learning is not widely popular for many reasons, the hardware aspects of the network speed, the performance of mobile phones and other mobile devices, the software has also failed to provide a large number of courses suitable for carrying out mobile learning, mobile learning mode of exploration and promotion of research is not enough caused widespread popularity of mobile learning not important reason.

As communication networks and the development of mobile terminals, mobile learning hardware to carry out the preliminary conditions have been met, in particular the development and popularization of 3G for mobile learning system operation provides a better network platform. Needless to say, in mobile learning is necessary and urgent, the application of mobile learning system to provide users with mobile learning services, with wireless communication network and cheap mobile phones and other mobile devices break the constraints of time, virtually anywhere, anytime for college students, adult education students, learners in remote areas to provide educational resources. Mobile learning as a new approach to learning, mobile learning related technology research currently is in a stage of rapid development in mobile learning has important practical significance and broad prospect.

Mobile learning is into people's work and life, is gradually changing the way we learn, the development of mobile communication technology is gradually lifted people's work, life and study time and place and for people anywhere in the mobile

environment work, life and learning provides the technical basis. Therefore, distance learning and mobile learning system mobile environment design of the study has important theoretical and practical significance.

WAP can support the vast majority of currently used wireless device, including mobile phones, FLEX pagers, two-way radio communication equipment. These devices relative to the desktop personal computer is concerned, CPU features is weak, small memory, power supply constrained wireless environment, the display is small, limited input capabilities. On the transmission network, WAP can also support a variety of current mobile networks, such as GSM, CDMA, PHS, etc., it can support the future third- generation mobile communication systems. But the relative bandwidth of a wired network using the Internet, wireless network bandwidth resources are always limited. Taking into account the above limitations and disadvantages, WAP fully draw on the idea of the Internet, and make certain modifications and simplified. That is, applications and web content using a standard data format, using the PC browser software used on similar micro-browser, the application of standard communication mode browsing.

WAP standard defines an application environment, allowing designers to develop device-independent user interface, and then use WMLScript (WML script) a WAP programming language, the logic embedded in the executable to the mobile terminal. Thus, the actual operation of the mobile terminal, a mini- browser, it is very much like a PC using IE or Netscape browsers.

WAP (Wireless Communication Protocol) is a digital mobile telephone, Internet or other personal digital assistant (PDA), a computer for communication between applications open global standard. The birth of this standard is the result of efforts of members of the WAP Forum, WAP Forum in June 1997, by Nokia, Ericsson, Motorola and wireless Planet (Unwired Planet) got together and formed a. The goal is to WAP WAP this technique, it can be a lot of information and a wide variety of Internet business into mobile phones, PALM among other wireless terminals. Wherever you go, whenever and as long as you need information, you can open your WAP phone and enjoy the endless online information or online resources. Such as: General news, weather, stock market dynamics, business reports, the current exchange rate. E-commerce, online banking will be implemented one by one. You can also get sports results anytime, anecdotes and humorous entertainment stories add life to taste, you can also take advantage of online booking function, the living arrangements methodically.

Mobile learning is a new form of learning. The goal is to achieve any time, any place, any way teaching and learning activities. Compared with other forms of learning, mobile learning with learning convenience, individualized instruction, interactive rich, context -related and other characteristics. Mobile learning has become the application of modern educational technology in a new direction, how to make effective use of mobile technology in education is becoming a hot research.

The main work is to:

(A) analysis of the theoretical basis of mobile learning, mobile learning technology base and forms of learning.

(2) analysis and design of a mobile learning system architecture, design, user management, course management, operations and test management, interaction, teaching management, eLearning access and adaptive function modules.

(3) Design of the adaptive system improvement program. Demand from the learner 's perspective on the content, the adaptive configuration system is divided into three parts: First, the device-independent part ; Second, adaptive equipment parts; three are user definable and intelligent analysis section.

(4) The system uses ASP + WML + Access development technology, to achieve a WAP-based mobile learning system, user management, course management, teaching management, message -based interaction and answering functions, practice shows that with good availability.

The design of this mobile learning system placed on a test run in a virtual host, until now reached more than 20,000 page views times. In carrying out mobile learning assisted learning experiment, subjects students staffing a WAP-enabled mobile phones, there are more than 20 kinds of phone models, all open GPRS monthly subscription, students can carry anywhere via mobile phone preparation, review, answering, educational administration and other learning activities.

Application of the system for teaching has brought many benefits, mainly reflected in: Learning convenience is improved. Via mobile phone anytime, anywhere to teaching information, understanding, learning and knowledge answering. Learning activities are no longer confined to the classroom or fixed location, virtually anywhere, anytime learning. After a period of probation, browser-based access to carry out specific forms of mobile learning is still groping in the use of effects to be used for a longer period of time and then test.

The developed system deficiencies and future research mainly in the following aspects:

(A) the current system development work completed only part of the various functions and stay in practice, testing further improved.

(2) adaptive function directly affects the user interaction with the system, this article only proposed and designed adaptive system improvement program, also requires further study and implementation.

(3) mobile learning system moving in the direction of development of intelligent, adaptive intelligent systems, intelligent answering is an important direction for future research.

(4) mobile learning environment "emotional missing", etc. is also a key problem to be solved.

In addition, mobile learning involves computer technology, educational technology and psychology, curriculum and teaching theory and other disciplines and technical areas, to further improve the system also needs to strengthen the theoretical and technical learning.

Mobile learning hardware conditions are gradually improving, theories and technologies are mature, the convenience of mobile learning to learn, learning personalization, rich social interaction, situational relevance, connectivity and other features of the existing mode of teaching welcome to the new reform opportunities. However, mobile learning system involves a wide range of technologies, international standards and more learner adaptability issues for mobile learning applications challenges. For educators and technology developers future challenges will be to find ways to ensure that this new learning is highly adapted to the environment, personal, collaborative and long-term, that is truly learner- centered learning.

6 Foundation Item

This research was financially supported by the Hubei Provincial Department Of Education science and technology research projects (Grant NO. 04ba161), the Hubei Provincial Department Of Education science and technology research projects (Grant NO. Q20132910) and the Hubei Provincial Department Of Education humanities and social science research projects (Grant NO. 2012G195).

References

1. Zuo, C., Fei, C., Hu, C., Guo, J.: Mobile education concepts and key technologies. China Distance Education (September 2005)
2. Xu: happened ; thesis aims ; WAP technical support for mobile learning system research and exploration. Computer Development & Applications (August 2010)
3. Liu, H., Xiaobai: JAVA-based educational management system design and implementation. Computer Knowledge and Technology 31 (2010)
4. Feng, Z., Zhang, H.: WAP2.0 the Senate based performance management system design and implementation. Fujian Computer (March 2010)
5. Even Runjiang ; role-based access control in educational management system application. Fujian Computer (September 2010)
6. Meiya: Senate Colleges and Universities system function and design analysis. Science and Technology Information, 31 (2008)
7. Yan, Z.: China was healthy; Web-based Teaching Management System Design and Implementation. Technology of West China 27 (2010)
8. Meng, Z.: Based on JSP / WML for WAP database application implementation. Modern Computer (Professional Edition) (June 2004)
9. Yue, X., Yun: WAP-based mobile campus network design. Computer Systems & Applications (April 2006)
10. Jie, J., Yao, S., Mei, T.: Wireless application protocol layer of security research and implementation. Zhejiang University of Technology (March 2001)
11. Bin, W.: Based.NET / WAP distributed college open laboratory Reservation Management System Design and Implementation. China Ocean University (2010)
12. Mei, G.: MVC-based Web application framework for research and application of Hohai University (2004)
13. Li, L.: NET-based road quality management system development study. Wuhan University (2004)
14. Hua, G.: calculated based on third party mobile commerce security protocols. Fuzhou University (2005)
15. Undertaking the; Parts Data Model Base Management System Research and Implementation of Zhejiang University (2005)
16. Li, Z.: Mobile payment platform unified access dial -E research and improve the Taiyuan University of Technology (2005)
17. Feng, Y.: Mobile e-commerce implementation techniques comparative study Wuhan University (2005)
18. Liu, H.: Mobile police information system overall solution of Shandong University (2006)
19. Yan, M.: WEB- commerce system based on mobile research and practice. Jilin University (2006)

20. Chen: ECC and restrictive blind signature based mobile electronic cash protocol. Yanshan University (2007)
21. Wu, X., Fan, C.: General Packet Radio Service application and promotion. In: 11th Cross-Strait Information Management Strategy Symposium (2005)
22. Tanghe, Feng, L., Zhang: Based on WAP technology Passenger motor car segment overhaul mobile messaging solutions. In: Railway Passenger Dedicated Line Construction Technology Conference Proceedings (2005)
23. Long, Y., Lei, D.: CDMA20001X mobile network advantages to achieve the government e-government system. In: Inner Mongolia Institute of Communications 2005 Annual Conference Proceedings (2005)
24. Fen, H., Lian, L.: Phone mobile meteorological integrated information service system. In: IT development and application in the field of meteorology Proceedings (2005)
25. Chun, Z., Yang, Y., Wuning: Rabbit3000 embedded platform based MMS sending. In: 2006 Communication Theory and Technology Progress - Eleventh National Youth Communication Conference Proceedings (2006)
26. Xu, Yang, Y.: Mobile commerce security Platform. In: 2006 Communication Theory and Technology Progress - Eleventh National Youth Communication Conference Proceedings (2006)
27. Hua, J., Jie, L., Junliang, C., Xue, Z.: Domestic WAP site usability studies. In: Second Joint Conference on Harmonious Human Machine Environment (HHME2006) - The 2nd China Conference on Human-Computer Interaction (CHCI 2006) Papers set (2006)
28. Lili, S., Chen: Mobile phone content service system design and implementation. In: 2006 National Open Distributed and Parallel Computing Conference Proceedings (2006)
29. Ren, Z.: WAP business chain and business model analysis. In: Communications Development Strategy and Management Innovation Conference Proceedings (2006)
30. Chun, Z., Yang, Y., Wuning: GPRS- MMS protocol module design and implementation. In: 25th Chinese Control Conference (two volumes) (2006)

A Robust Fault Detection Approach in Network Control Systems with Short Time Delay

Zijian Dong and Yongguang Ma

North China Electric Power University, Baoding 071000, China

Abstract. For the complexity of network, networked control systems (NCS) have higher level of demands for safety and reliability than general systems. Therefore, the research on fault detection of networked control system has great theoretical and applied value. The networked control systems can be modeled using the Matlab/Simulink. There a state observer is designed, and then a simulation is conducted in allusion to the network's time-delay problem. Its influence upon the system's performance shows good.

Keywords: networked control systems (NCS), fault detection and diagnosis (FDD), simulation, state observer.

1 Introduction

As the network control system space more widely distributed, the structure more complex, so the control system performance requirements are also increasing. NCS event of failure, may cause huge casualties and property losses, and even lead to catastrophic accidents [1-4]. Therefore, to improve the stability of NCS, reliability and safety, the fault detection of NCS is very important [5-8]. Currently, NCS's research focuses on the mathematical model, stability analysis and network transmission performance analysis and other aspects of the system, the failure detection and diagnosis research is relatively small, so the study of this subject has a positive significance [9-12].

Networked Control Systems (NCS) is the control system actuators, sensors and controllers are connected together via a network of distributed control system. Internet-based network control system on the one hand can be achieved between the network control system with the Internet and network direct communication between the control system, on the other hand can reduce the variety of connection cables, reduce costs and ease of maintenance [13-16]. In the control system, the introduction of the communication network is not the purpose is to improve the system performance, but that the control system breaks limitations on space, thereby reducing the complexity of the control system, centralized and efficient implementation of operations [17-20]. At the same time, the intervention of the network, the inevitable existence of the control loop network delay, which is often cause system instability or adverse performance reasons, how to effectively overcome the impact of network delay, designed to meet certain performance metrics of the extension when the network controller, network control system the main problems to be solved. Thus the delay of networked control systems research has important theoretical and practical significance. This paper gives a

Y. Yang, M. Ma, and B. Liu (Eds.): ICICA 2013, Part II, CCIS 392, pp. 450–459, 2013.

proof of the delay and network control systems and random delay of networked control systems l_2-l_ ∞ performance criteria to ensure that the control output peak is less than a certain value. And the first proposed this system L_1 performance criteria. Focuses on the delay of networked control systems l_2-l_ ∞ robust control, random delay of networked control systems l_2-l_ ∞ filtering and robust L_1 control problems. First, the network control system for a delay to study a model of this system, the application of Lyapunov stability theory and LMI (Linear Matrix Inequality, abbreviated LMI) technique is proposed and proved the robustness l_2-l_ ∞ performance criteria. And so designed as to allow the l_2-l_ ∞ sufficient condition for the existence of controllers. Secondly, the use of the Bernoulli distributed random variables, linear matrix inequalities; a full- order filter exists a sufficient condition that the filtering system is mean square exponentially stable and has given l_2-l_ ∞ performance constraints. Again, study the delay of networked control systems L_1 control problem. Establish a system L_1 performance criteria, the performance criteria can ensure that the system is asymptotically stable and has a worst case peak - peak performance constraints. The L_1 performance analysis, design, robust L_1 controller [21]. Finally, in l_2-l_ ∞ state feedback controller and L_1 state feedback controller contains a sufficient condition for existence of inverse constrained LMIs form, through the cone complementarity linearization (CCL) method to convert it to LMIs constrained by the minimization problem this allows the use of standard mathematical software to solve the controller parameters. Numerical simulation using the proposed method.

2 The FDD with Short Time Delay

Short latency network control system fault diagnosis research has made some achievements [1]. At present, the research is generally assumed that the system exists through the sensor to the controller delay, or the presence of the controller to the actuator delay, and delay may be considered a fixed constant, random, or a time uncertain amount, the use of state estimation, the output estimation method, the fault detection or diagnosis.

Literature [2] uses the method of the parity equation with a short delay network control system in fault detection research. Literature [3] In the asynchronous network control system established under delay fault detection model is constructed asynchronous delay system state observer, and construct the frequency domain faults detecting residual generator to achieve fault detection. Literature [4] for a short time-delay network control system, first construct a short delay network control system of measurement redundancy and parity equation of full decoupling parity vector-existing conditions, to achieve fault diagnosis.

3 The Robust Fault Detection Approach

3.1 The FDD Approach

Fault diagnosis (FDD) is the use of information about the operation of the system and the existing knowledge, information processing, and ultimately get the system

operational status and fault conditions of the process. Troubleshooting tasks can be divided into the following aspects: fault modeling, fault detection, fault isolation, fault identification, fault evaluation and decision-making.

Overall, the fault diagnosis method is broadly divided into three categories: model-based approach, signal-based methods, and knowledge-based approach. Model-based approach is first developed, it generally requires more accurate system model. Signal-based approach to system model accuracy requirements are not particularly high, but to use the signal model to deal with the problem. Knowledge-based systems is accompanied by the emergence of increasingly complex, it is especially suitable for difficult to obtain accurate mathematical model of the system case.

3.2 Observer-Based Fault Detection System of NCS

Let the equation of state of the controlled object is

$$\begin{cases} \dot{x}(t) = Ax(t) + Bu(t) \\ y(t) = Cx(t) \end{cases} \tag{1}$$

Network control system set up continuous-time system controlled object state space model is

$$\begin{cases} \dot{x}(t) = Ax(t) + Bu(t) + B_d d(t) + B_{fa} f_a(t) \\ y(t) = Cx(t) + D_d d(t) + D_{fs} f_s(t) \end{cases} \tag{2}$$

Observer-based NCS fault detection system shown in Figure 1.

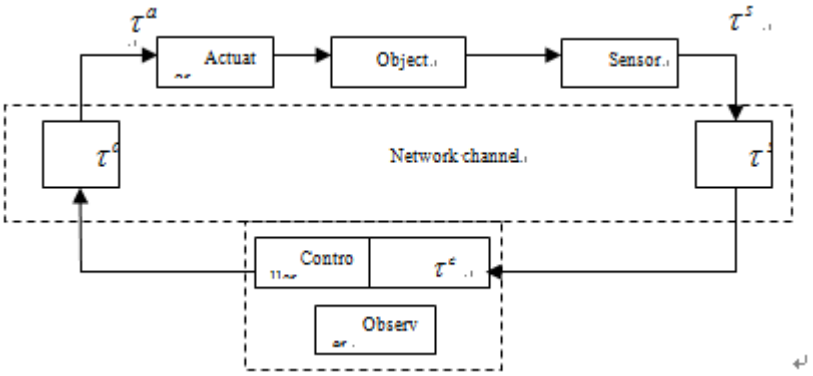


Fig. 1. Observer-based NCS fault detection system

4 The Modeling of Uncertainty Short Delay NCS

Let the linear time-invariant equations of state, such as the controlled object shows

$$\begin{cases} \dot{x}(t) = Ax(t) + Bv(t) + B_w w(t) \\ y(t) = Cx(t) \end{cases} \tag{3}$$

From Figure 2, the network control system discrete model of generalized controlled object.

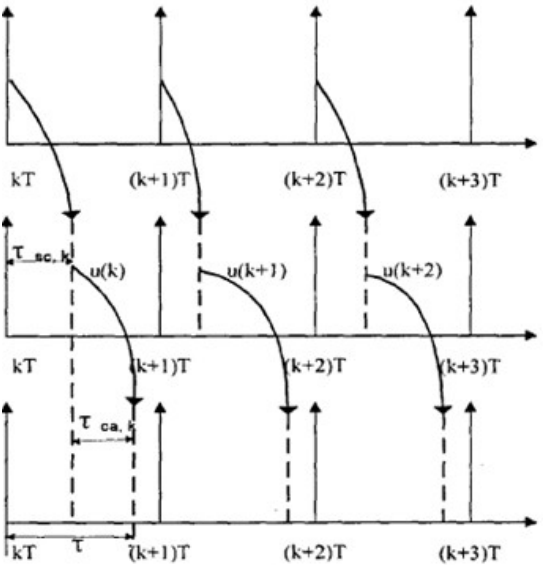


Fig. 2. A short timing delay NCS diagram

$$x_{k+1} = Gx_k + H_0(\tau_k)u_k + H_1(\tau_k)u_{k-1} + H_w w_k \tag{4}$$

Conclusion 3.1 network control system controlled objects such as formula (3) shows, the controlled object contains the network generalized discrete-time model (4) can be equivalent to

$$\begin{aligned} x_{k+1} &= Gx_k + (H_0 + DF(\tau_k)E)u_k + (H_1 - DF(\tau_k)E)u_{k-1} + H_w w_k \\ y_k &= Cx_k \end{aligned} \tag{5}$$

Where G, H_w, H_0, H_1, D, E, As a constant matrix, The uncertainty matrix $F(\tau_k)$ satisfy $F^T(\tau_k)F(\tau_k) \le I$.

5 The Short Delay Fault Detection Methods of NCS

Theorem 4.1 If there exist symmetric positive definite matrix and the matrix, as well as constant $\delta > 0$ satisfying the following matrix inequalities

$$\begin{bmatrix} -X & Y^T & X(\overline{E}\overline{K})^T \\ Y & -X+\delta\overline{D}\overline{D}^T & 0 \\ \overline{E}\overline{K}X & 0 & -\delta I \end{bmatrix} < 0 \tag{6}$$

Observer error system is asymptotically stable.

The observer error system stability is guaranteed after the fault detection residuals if $\varepsilon_k = He_k$ so, H is the output weight matrix, then by selecting the appropriate fault detection threshold $\overline{\varepsilon}$, can determine whether the system is malfunctioning. Judgment criteria are as follows:

$$\begin{cases} \|\varepsilon_k\| \le \overline{\varepsilon} \\ \|\varepsilon_k\| > \overline{\varepsilon} \end{cases} \tag{7}$$

Where, $\|\varepsilon_k\| = \sqrt{\varepsilon_k^T \varepsilon_k}$ is euclidean vector norm.

6 Robustness and Sensitivity Analysis

Threshold $\overline{\varepsilon}$ directly affects the robustness and fault detection sensitivity factor. In order to improve the sensitivity of fault detection can be used as unknowns, including by solving LMI Constrained minimization problem, to obtain γ the optimal value is γ^* , This value is to determine the threshold size factors. In this way, the fault threshold is selected

$$\overline{\varepsilon} = \gamma^* \sigma_m \tag{8}$$

It satisfacties

$$\begin{bmatrix} -P+H^T H & 0 & G^T P-(G^{-d^{sc}})^T C^T Q \\ * & -\gamma_1 I & P \\ * & * & -P \end{bmatrix} < 0 \tag{9}$$

The solution is γ_1^* , then $\gamma^* = \sqrt{\gamma_1^*}$ is the Robustness of H_∞ Fault observer optimize performance.

The problem is a linear inequality constraints and linear objective function convex optimization problem that can be applied to solve LMI toolbox.

7 Simulation Algorithm

Let the controlled object as network control system as follows:

$$\begin{cases} \dot{x}(t) = \begin{bmatrix} -1 & 0.2 \\ 0 & -1.5 \end{bmatrix} x(t) + \begin{bmatrix} 1.5 & 0 \\ 0 & 1 \end{bmatrix} v(t) + w(t) \\ y(t) = \begin{bmatrix} 1 & 1 \end{bmatrix} x(t) \end{cases} \tag{10}$$

Let sampling period $T = 0.1s$, The network information transmission total delay $\tau_k \in [0,0.1]$, And is a time-varying uncertain. Controller node and actuator nodes uses event-driven. Observer designed to ensure the stability of the error system, and can effectively detect the failure.

Depending on the system for continuous sampling period controlled object model discretization, and in accordance with Section of the method can be discretized controlled object model

$$\begin{aligned} x_{k+1} &= Gx_k + (H_0 + DFE)u_k + (H_1 - DFE)u_{k-1} + w_k \\ y_k &= Cx_k \end{aligned} \tag{11}$$

Solving the coefficients formula as follows:

By the formula (11), $G = e^{AT} = \begin{bmatrix} 0.9048 & 0.0177 \\ 0 & 0.8607 \end{bmatrix}$, the Eigenvalues of A are -1, -1.5. And can be diagonalized, respectively corresponding eigenvectors $\begin{bmatrix} 1 & 0 \end{bmatrix}^T$, $\begin{bmatrix} -0.3714 & 0.9285 \end{bmatrix}^T$, $\Lambda = \begin{bmatrix} 1 & -0.3714 \\ 0 & 0.9285 \end{bmatrix}$.

From the (5), $H_0 = \Lambda diag(-\frac{1}{\lambda_1}, \cdots, -\frac{1}{\lambda_n})\Lambda^{-1}B = \begin{bmatrix} 1.5 & 0.1333 \\ 0 & 0.6667 \end{bmatrix}$

$$D = \Lambda diag(\frac{1}{\lambda_1} e^{\lambda_1 \alpha_1}, \cdots, \frac{1}{\lambda_n} e^{\lambda_1 \alpha_n}) = \begin{bmatrix} -1 & 0.2476 \\ 0 & -0.6190 \end{bmatrix}$$

$$F = diag(e^{\lambda_1 (T - \tau_k - \alpha_1)}, \cdots e^{\lambda_n (T - \tau_k - \alpha_n)}) = \begin{bmatrix} e^{-(0.1 - \tau_k)} & 0 \\ 0 & e^{-1.5(0.1 - \tau_k)} \end{bmatrix}$$

$$E = \Lambda^{-1}B = \begin{bmatrix} 1.5 & 0.4 \\ 0 & 1.077 \end{bmatrix}$$

From the (6),

$$H_1 = \Lambda diag(\frac{1}{\lambda_1}e^{\lambda_1 T},\cdots,\frac{1}{\lambda_n}e^{\lambda_n T})\Lambda^{-1}B = \begin{bmatrix} -1.3573 & -0.1324 \\ 0 & -0.5738 \end{bmatrix}$$

If the original system controller gain matrix

$$K = \begin{bmatrix} -0.0368 & -0.0035 \\ -0.0035 & -0.0454 \end{bmatrix}$$

According to Corollary, the application of the LMI toolbox in MATLAB numerical solution can be obtained:

Observer gain matrix

$$L = \begin{bmatrix} 0.4612 & 0.4304 \end{bmatrix}^T$$

8 Simulation Result

When the system is normal, the observer status of the original system x_1, x_2 traces are shown in Figure 1 and Figure 2. When the system is uncertain the delay of the state observer estimation still better able to track the actual state of the system.

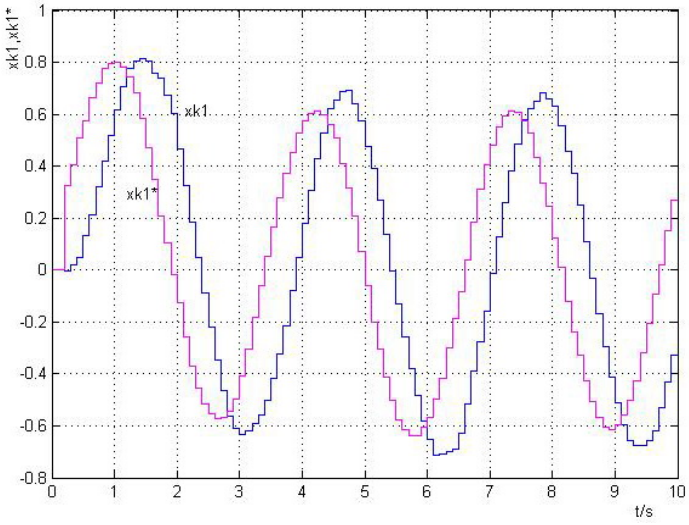


Fig. 3. The state results of x_1

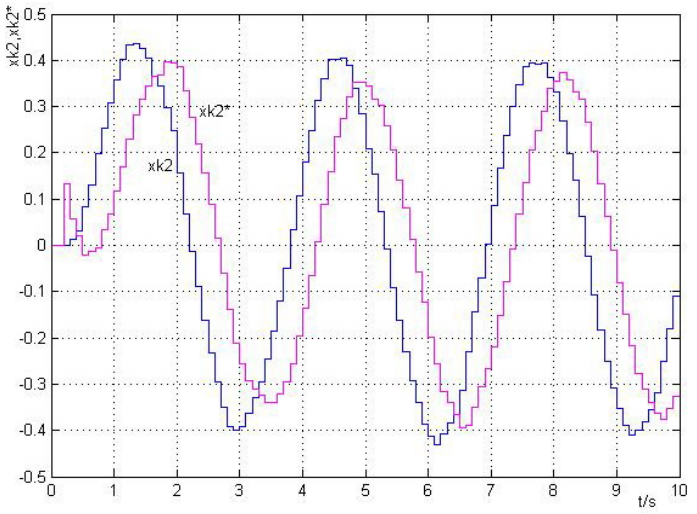


Fig. 4. The state results of x_2

It is assumed that the system failure occurs, when $t = 5s$ the unit step mutations, so that the observer output weighting matrix, then the fault detection results are shown in Figure 3(The Fault Detection results of x_1), .4(The Fault Detection results of x_2) and .5(The Fault Detection results of Output Residuals). From the above chart we can see that when $t = 5s$ a failure occurs, the fault observer can quickly detect the fault occurs, an indication of the design of state observer is valid.

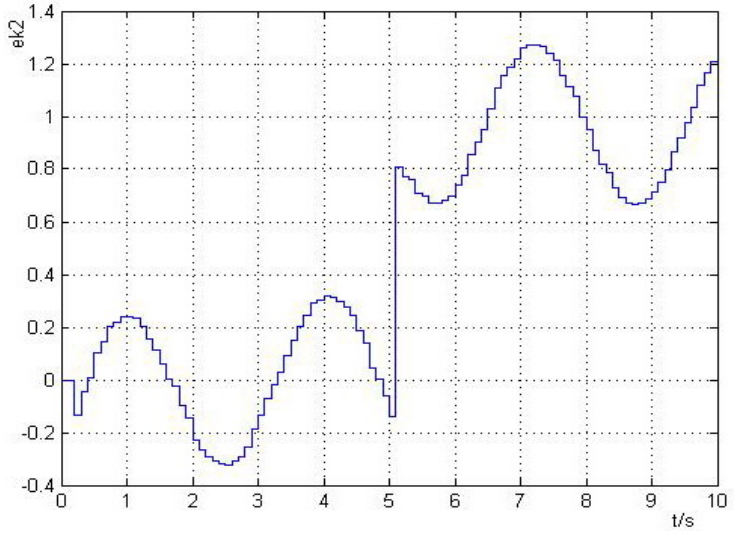


Fig. 5. The Fault Detection results of x_1

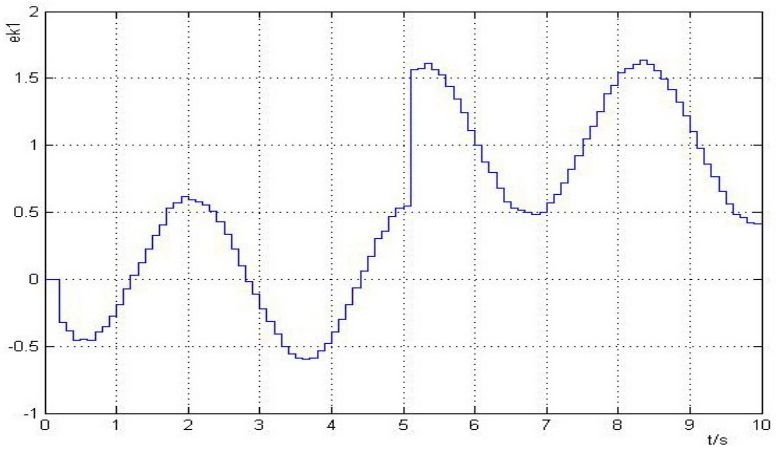


Fig. 6. The Fault Detection results of x_2

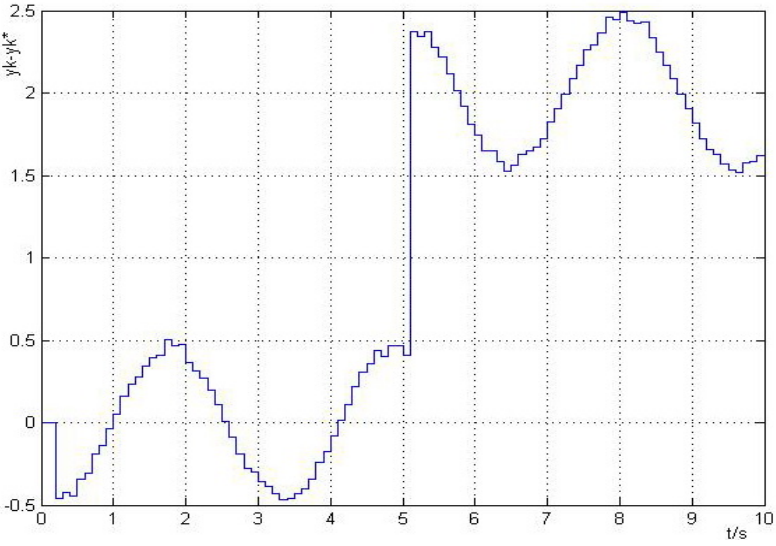


Fig. 7. The Fault Detection results of Output Residuals

9 Conclusions

This paper discusses the short-latency network control system in the absence of packet loss in case of fault detection. First, assume for the time-driven sensors, controllers, actuators for the event-driven model of the system are analyzed and described; the establishment of parametric uncertainties include discrete equations. Then, through the design of state observer to consider the system's fault detection. Lyapunov stability theorem obtained fault observer system stability conditions. Simulation results verify the effectiveness of the method.

References

1. Na, L., Zhang, L., Han migration: Having a saturated nonlinear constrained network control system Robust H_ ∞ optimal control. Northeastern University (Natural Science) 05 (2008)
2. Zhang, L., Chen, D., Liao, L., Wu, X.Y.: Casting aluminum electromagnetic rapid design of networked control systems. Industrial Control Computer 01 (2006)
3. Dai, J.: Delay system approach of networked control systems. Computer Engineering and Applications 20 (2009)
4. ROCKETS, Wu, X.-G.: Web-based control system for Underwater Vehicle Guidance System. Computer Measurement & Control 11 (2005)
5. in the training, Chen, H., Wang, Y.: Closed-loop delay characteristics based on Markov network control system. Control Theory & Applications 02 (2002)
6. Xu, B.-G., Xu, Y., Zhou, X.: Interconnected Systems with Delays Decentralized Stabilization : Linear Matrix Inequality Approach (English). Control Theory & Applications 03 (2002)
7. Guo, Y., Shao-Yuan: Networked control system H_ ∞ state feedback controller design. Control Theory & Applications 03 (2008)
8. Wang, W., Bin, L., Cai, F., Yang, F.-W.: Random delay of networked control systems H_ ∞ Output Feedback Controller Design. Control Theory & Applications 05 (2008)
9. Xie, C., Weili: A class of multi- packet transmission networked control systems H_ ∞ Robust Control. Control Theory & Applications 02 (2010)
10. Rotary L., Yong, Wu: Delay -based H_ ∞ filter design and its application in the network. Control Theory & Applications 04 (2010)
11. Zheng, F., Bing, H.: Networked Control System Research. Anhui Electrical Engineering Professional Technique College 03 (2006)
12. Ming, Y., Li, Q., Wang, C., Xiang, L., Yuan, Z.: Based on State Observer for Networked Control Systems H_ ∞ Control Analysis. Ordnance Technology 07 (2008)
13. Liu, L., Tong, C.-N.: Networked control systems a unified Markov jump model. Beijing University of Technology 11 (2007)
14. The original Jumei, R.Q.: The original security level; Networked Control Systems Control Method. Beijing Technology and Business University (Natural Science) 05 (2005)
15. Zhang, L., Liu, Y.: ROCKETS PROJECTILES unmanned underwater vehicle real-time system design. Ship Engineering, 03 (2009)
16. Bing, K., Zhao, H., Ren, L., Yan, D.: Networked Control Systems Modeling Methods. Journal of Jilin University (Information Science Edition), 01 (2006)
17. Zhi, R.: Pan significantly Gansu dynamic output feedback control system for real-time network stability analysis. Ship Power Technology 09 (2011)
18. Wang, Z.: Method based on LMI for networked control systems l_2-l_ ∞ Control. Journal of Yangtze University (Natural Science Edition) Polytechnic volume, 02 (2010)
19. Wu, C., Zhe, X., He, J., Xu, H.L.: Networked Control Systems Modeling and Control. Electronics Optics & Control 09 (2009)
20. Tang, H., Zhang, A., Shu- in: A large class of time-varying uncertainty interconnections Robust decentralized adaptive control system Electric Machines and Control 01 (2006)
21. Jay, Ning, Shuqing: Neuron non- model control system stability analysis. In: China Control Conference (1996)

LBS and Mobile Internet-Based Monitoring System Architecture

Kun Liu[1,2,*], Kun Ma[1,2], and Wei Liu[1,2]

[1] School of Information Science and Engineering,
University of Jinan, Jinan, Shandong, China 250022
[2] Shandong Provincial Key Laboratory of Network Based Intelligent Computing,
University of Jinan, Jinan, Shandong, China 250022

Abstract. With the improvement of phone performance and enhancement of the quality of LBS services, mobile Internet-based surveillance is possible. We propose an LBS and mobile Internet-based position monitoring system architecture, which under the premise of ensuring information security can achieve remote management and monitoring on the host. The architecture is based on the Android platform that can provide a host of real-time location, location reporting and queries, time control, the application of applied statistics and constraints, intercept phone messages and other remote monitoring services. Host establishes the binding relationship with the controller, then you can remotely control terminal for the real-time monitoring. Control client can view the controller via LBS location and the surrounding environment, and adjust as needed monitoring methods. Meanwhile, the architecture provides a Web client query feature, with which the controller receives commands interacting with the server; the remote host can query information through the Web page, so as to achieve uniform treatment and location recording purposes.

Keywords: Android, LBS, Monitoring System.

1 Introduction

With the rapid development of mobile Internet technology, a lot of mobile terminals will have the "access" to network [1-4]. A variety of applications in the mobile Internet also profoundly change people's lives in the information age, and location-based services (Location Based Services, LBS) and mobile network application system are "embedded" in people's daily lives. We can monitor the situation at home or any other location of the mobile terminal condition at any time via their mobile phones [5]. For now, LBS technology is relatively mature technology, as cloud computing is the third wave of information technology through the LBS and cloud computing technology, we can achieve a complete mobile Internet technology-based monitoring system [6, 7].

As a smart phone mobile phones become mainstream products terminal and gradually replace the traditional phones, LBS technology has very good prospects for development. With the popularity of smart phones, mobile application service is unprecedented developed, which greatly enriches people's way of life, but also

* Corresponding author.

Y. Yang, M. Ma, and B. Liu (Eds.): ICICA 2013, Part II, CCIS 392, pp. 460–468, 2013.

provides users with a convenient information service and location-based services have gained popularity possibility [8-12]. For the above, this paper, based on the Android system, design and implement a mobile Internet monitoring system architecture based on LBS and cloud platform.

2 A System Architecture Design

2.1 Overall System Architecture

The system should achieve three application subsystems, and each system is a client. The first part aims to design the Android desktop application side with the capability of an implementation of the real-time location, location reporting, remote command response, time control, application information, statistics, application usage constraints and telephone message interception. This is the main part of the system, which is being monitored. The second part aims to design a monitoring client to interact with the server, to receive process and respond to the host's request. The third party designs control software for the host to convenient remote monitoring and management to reduce the difficulty of the operation, and the overall system architecture is shown in Figure 1:

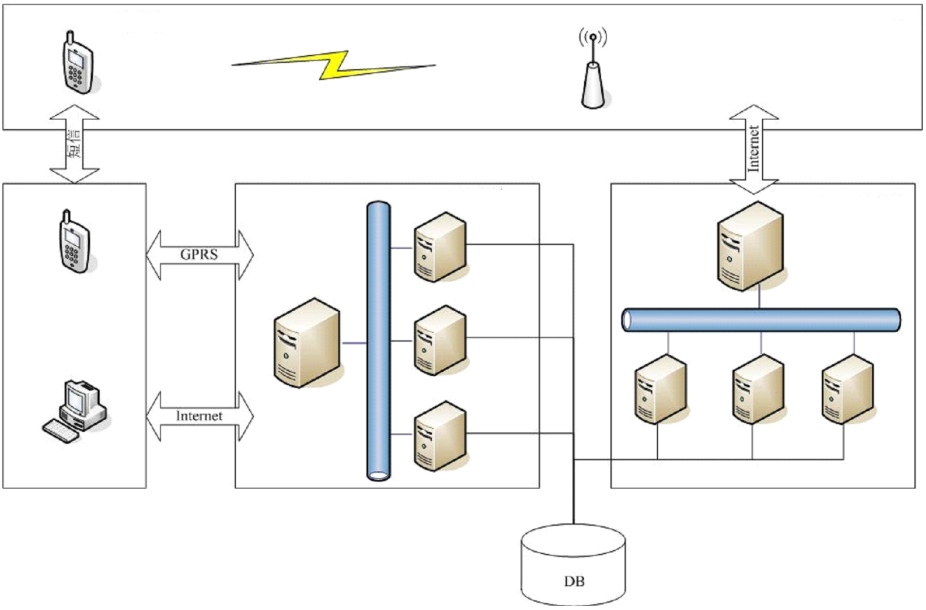


Fig. 1. Overall architecture based on the location of LBS and mobile Internet monitoring system

2.2 The Monitored Client Functions

Through SMS interception, it achieves a set of practical, complete remote instruction system, based on TCP / IP protocol communicates with the server through Socket way, using a local database cache failed messages;

It uses a Baidu positioning API to implement GPS, WI-Fi, GPRS environment of the mobile phone positioning, and uses the Alarm mechanisms to ensure the phone's normal operation of positioning in sleep mode;

It uses timing detection means of the Service and Handler to inspect and correct the operational status of the system;

It designs a reasonable time limit system to facilitate remote host restriction on the use of mobile applications;

It designs a family numbering system and desktop shortcuts plug, and within the time limit intercept non-family number calls;

Through the local database and FTP, it achieves application statistics and reporting, and adopt encapsulate data in JSON format, increasing portability;

2.3 Server-Side Functionality

Based on TCP / IP protocol, it communicates with the monitored through Socket-side and respond to the request of monitoring client;

Record all uploaded data being monitored.

Functionality of the controller:

The host can bind with the monitored side, and send remote commands after the success;

Through the host controller, it sends commands to see the location, application usage statistics;

3 System Architecture Analysis Functions

System mainly consists of hosts, the controlled, Socket server and Web server side, and the function of each part is shown in Figure 2:

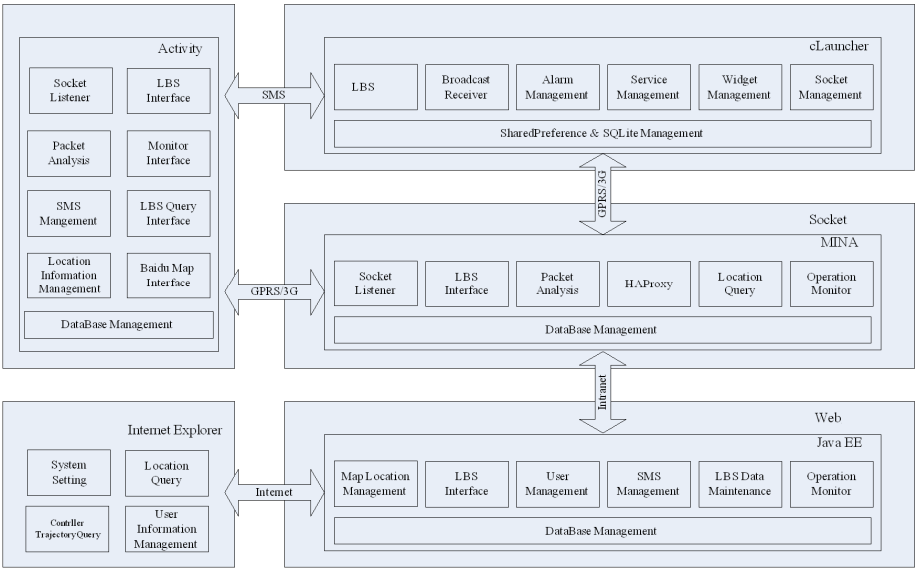


Fig. 2. Position monitoring system function chart based on LBS and mobile Internet

LBS Interface: Using Baidu LBS interfaces to easily implement phone positioning in WI-Fi, GPS, and GPRS network. Callback interface provides developers with the current location of the mobile phone details, and developers can be personalized based on these data processing.

Launcher: applications on Android are divided into many varieties, such as desktop applications, general applications, lock screen applications and so on. Launcher application is the default when you start your phone system, if there are multiple it will prompt you to select. Common applications are not automatically displayed. Android provides an interface to obtain an ordinary application, thus the task of indicating the application entry falls on the Launcher. Android comes with Launcher which provides the basic functions needed by a desktop application, such as sliding around, add a desktop shortcut, the application list display and other basic functions. Therefore, we can develop customized desktop applications on the basis of the Launcher.

Broadcast Receiver: The broadcast mechanism in Android phones which used for passing messages between different applications. Many functions of the phone are achieved like this, for example, switch machines, receive text messages, incoming call, etc., the monitoring system is inevitable to use this feature in large scale. Broadcasting can be used in two ways, one is through the register and its subclass to implement a receiving system provides good news. Such as registering a receive SMS broadcast, when the phone receives a broadcast, it firstly runs methods written by us for processing, and then decides which to perform certain operations or intercept messages. The same approach also applies to telephone interception and switch machine messages. Another is to declare it a broadcast which can be achieved in the context of two different deliveries of messages. Further, the broadcasting system can use the Internet as a data carrier body Intent, so you can easily transfer data.

Alarm Management: Sometimes, a Service background operation does not meet our needs. For example, we hope when we do not use the phone, the system can also perform some operations for us at the regular time. The service does not work this time, because we no longer use the phone for some time, a Linux kernel will put the Android system into hibernation in order to save resources. In this state, Android system will only retain some of the necessary process, and stop other operations. Based on this model, if you want to achieve a background operation, the easiest way is to implement a timing clock, and the alarm in Android is implemented based on this approach. When Android phone in hibernation, Alarm remains valid. When the specified time arrives, the clock program can send a broadcast, which is we can accept broadcast and do the appropriate action.

SQLite: It is a lightweight database on Android, with basic data storage capacity that can be used to store some application data. In addition, many of the data in the system are also stored in this database.

Service: As the name implies, it is the service component in Android phones. When the phone screen is not stuck in the current application, Service still can perform certain actions in the background, which is one way to achieve the usage of a cell phone monitoring software. The application program can start particular Service components through Start Service method, and then run a timer in the background detection program in order to achieve timing detection function.

Shared Preference: It is the most simple data storage methods in Android. The system creates a storage location for each application that can be used to store some configuration information to key on the way to save.

App Widget: Desktop plug-ins in Android, which can be placed directly on a screen desktop application, and quick dial controls need this.

Socket: Communication between the server and the phone using Socket transmission through a particular protocol encapsulation over binary data.

4 System Technical Architecture Analysis

The host and the controlled systems are developed based on the Android platform ADT, as Socket server uses Apache Mina framework, and Web server uses the lightweight Java EE framework.

The front-end of Socket server uses Apache Mina framework, which is a multi-threaded non-blocking Java Socket component, loading a lot of long connections to meet the client's interaction with the mobile; business process layer uses the Spring framework to manage the business logic Bean; persistence layer uses Hibernate framework to interact with the database. In Web server technology architecture, the system is divided into DBMS, OP, BP, RP and UI. The OP provides persistence services for entity in the system; BP achieves specific business operations, and is responsible for business process management system; the control later plays an adaptive role in the BP and the UP and is responsible for exception handling and diary; the UI achieves customer interaction.

5 Business Process Design

To ensure the information security, the host needs to go through the appropriate certification process, and then obtains monitoring data of the controller.

The entire system process is divided into three phases, namely: binding phase, phase and binding service lift phase. The main business processes are as follows:

The host remote sends SMS through "bind" request command to the controlled.

The controlled deals with local host settings to determine whether to allow binding, and then return the binding results to the host.

Establish accounts for the controlled and the host, and store monitoring data.

Create a Web account for the host to and store monitoring data for the monitoring client.

The host receives and processes SMS messages, and stores the results locally.

The host send request through text message for the controlled.

The controlled deals with local settings, check the location information, package the location information and request instructions and sent them to the server.

The server receives and parses the package instructions, calls the network interface, accesses to data, and saves the local database.

The server packages the data and returns it to the host.

Receive and parse the data, and edit the message returned host.

The host receives and processes SMS messages via the corresponding interface display.

The host sends "Bind release" command to the host via SMS.

The controlled process local settings, lifts and remotes bindings, unbinds results and returns them to the host.

Delete account associated between the host and the remote, and change the monitoring data into historical data.

Remove the remote Web account, and change the monitoring data into historical data.

The host receives and processes SMS messages, and stores the results locally.

6 System Implementation

6.1 Remote Management Command

Management Directive is a command used by the host to send commands to the controlled through text messages. Only the host needs to send a simple character or text to the controller; you can view the monitoring client location, modify the time setting, and modify family number and so on.

The controlled side decides whether to respond binding instructions through intercepting the received messages, query letters numbers and bind numbers.

The controlled, after receiving instructions, would do different treatments which can be divided into two types: one is the direct handling and response; the other is that they request processing firstly and response to the control after obtaining the results.

The remote management commands include: "binding", "registration", "location", "map", "unbind" and so on. In which binding instructions can direct response registration, location, maps and other formatting commands.

6.2 Timing Positioning

First, the timing position reporting feature in this system uses Android-based Baidu map development interface. Baidu map provides developers with a range of LBS interfaces, with separated positioning and mapping interface, increasing developer flexibility of use. The system selects only positioning interface provided by Baidu maps. Using Baidu positioning interfaces on the Android platform simply needs to import the appropriate library file, and then implement a callback listener following the instructions in the help documentation, and set the parameters before the start, such as location update interval, the return data content, to receive the current location information in circulating within the set time.

Secondly, the system requires timing reporting their position as long as the phone is on, but the phone screen is not always on. When the user stops operating after a period of time, Android system will automatically enter the sleep state. Therefore, the system uses the high-level process component Alarm Management coming to the Android

platform. Alarm Management can also run some operations when the system is in dormancy, and the monitoring system obtains its instance through the Android system static method and registers a loop Alarm. So every once in a while, the alarm will send a Broadcast Receiver message to tell us that it's time to update the position. We perform a location update our block in previously registered Broadcast Receiver's on Receive method. In this way, monitoring system achieves a regular position acquisition function.

Finally, after getting to the location information, we should package it in accordance with the provisions of good server communication protocol encapsulation, and the encapsulated system has got a string of binary data. Then the corresponding port connected to the server to establish Socket communication connection to the server before sending the encapsulated binary data string.

6.3 Application Usage Restrictions

The control system is required to make certain restrictions for the controlled applications to use. The system is divided into the controlled mode and normal mode. In normal mode, you can use the phone's full functionality, and set binding register numbers binding after the time limit.

Before coming into the alleged mode, you must select this system as the default desktop before the application of the system. Desktop applications in the Android system are distinguished from the general applications based on a marker. After the phone starts, if there is no default desktop, it will list all desktop applications to allow users to choose, and the common application is to start from the desktop.

The system provides application interfaces using the time limit set for the master in normal mode. The master can easily set up the mode and limit of the controlled in the normal mode. In controlled mode, if in limit time, in addition to desktop applications, other applications are not allowed to be used. The system default time of day is divided into four time periods, each of which can be set individually, and the time also can be set by day.

6.4 SMS and Call Blocking

In the remote command, we have introduced a way of intercepting SMS, which intercepts and filters the received SMS by registering a broadcast. In addition to blocking the host's command, the system also intercepts non-family number for the controller.

Call blocking. Get the Telephony Manager through Android system statically; call it a listen method to monitor our custom Phone State Listener. As a call outages or incomes, the on-call State-Changed method will be called. Its parameter contains three states: CALL_STATE_IDLE (hang up), CALL_STATE_OFFHOOK (answer), and CALL_STATE_RINGING (Bell). For a non-family member, the monitoring system will automatically call their shield.

6.5 Application Usage Statistics

The application usage statistics are a more difficult part, and this module is only doing its simple implementation. I.e., statistic the frequency of use, stores the number in a database, and then regularly uploads it to the FTP server. This is achieved as follows:

When you start other applications, in fact, you start an activity of the application. Then the activity displaying in the system (i.e. Launcher Desktop) will call its life cycle functions on-Stop. While monitoring the system, we will obtain the current top application through the tools Activity-Helper (the Android system is running an Activity-Task, then we can learn the running application program through getting the top of the stack Activity). At this point, we start the application and its state statistics, and store it in the local database.

6.6 Family Numbers

Monitoring system provides users with desktop plug-ins with a family number speed dial. SOS Widget Provider is a subclass of App Widget Provider, which corresponds to a desktop plug. On the plug-in board, there are five family number buttons such as the top, bottom, left, right, and center. When you click a button, it will trigger the five different broadcasting.

7 Conclusion

The issue, with the current development trend of mobile applications, according to the needs of the mobile device monitoring, designs and develops a monitoring client monitoring system based on the Android platform, achieving the functions of a remote command, timed positioning, messaging and telephone interception, application usage statistics, family number of settings and other functions for mobile application development.

The thesis analyzes shortcomings and development status of the existing project host monitoring software, summarize their functions, and improve some of the features, enhancing its usability and providing a reference for subsequent similar software.

The topic, based on the Launcher, expands and modifies the source code to reduce the workload, providing developers with experience. At the same time, it does a lot of research on the Launcher desktop mechanism, learns a lot of them and gained a lot of experience in the future of the Android application development.

The project, based on the Android platform, employs a large number of Android components and mechanisms, including the use of basic knowledge to cover almost the majority of the knowledge required for application development, so the design of this system is also a very good understanding and learning Android development opportunity.

The subject achieved a viable monitoring client monitoring system based on the Android platform, although whose function should be further improved and enhanced, but it basically meets the needs of existing applications and can solve some of the everyday problems for people.

In achieving results at the same time, because of the technology and energy constraints, there are also some shortcomings and deficiencies in the application of statistical areas which need further improvement and upgrading.

Acknowledgement. This article is one of the research results of Shandong Province Colleges Science and Technology Project, number J12LN44.

References

1. Chen, X., Chaoyi, Z.: Android mobile Positioning System Design and Development Oriented Children. Electronic World 5(12), 137–138 (2012)
2. Qiu, M.F., He, Z.: Map Matching Study Based on Mobile Phone Base Station Location Data. Traffic Information and Safety 4(29), 52–57 (2011)
3. Chen, S.: Mobile Monitoring System Design and Application Analysis Based on the Android Platform. Software Magazine 33(10), 52–54 (2012)
4. Lai, C., Long, X., Li, J., Wu, L.: Mobile Background Monitoring Technique Based on the Android Platform. Computer Knowledge and Technology 6(33), 9472–9474 (2010)
5. Zhang, P., Zhang, X.: Monitoring System Design and Implementation of Vulnerable Groups Based on the Android Platform. Hunan University 33(12), 79–82 (2012)
6. Liu, Y., Liu, A., Xie, T., Song, Y.: Personnel Positioning and Monitoring System Design and Implementation Based on the Android Platform. Zhengzhou Institute of Light Industry 27(6), 17–20 (2012)
7. Zhong, Y., Wang, J., Hu, J., Zhao, Z.: Android Kernel Transplantation and the Application in the Remote Monitoring. Chongqing University of Technology 25(1), 102–106 (2011)
8. Ren, Q.: LBS Applications Research Based on Android. East China Normal University (2012)
9. Sheng, Z.: LBS Applications System Development Technology Research Based on Android Platform. Yunnan University (2012)
10. Dai, M.: Mobile Positioning Regional Service Design and Application Based on the Android Platform. Kunming University of Science (2012)
11. Xie, C., Zhou, J., Li, S., Jia, L.: Mobile Monitoring System of Transmitting Station. IERI Procedia (3), 81-86 (2012)
12. Kristian, Y., Armanto, H., Frans, M.: Utilizing GPS and SMS for Tracking and Security Lock Application on Android Based Phone. Procedia - Social and Behavioral Science (57), 299–305 (2012)

Model Design of Digital Instructional System Based on AR Technology

Riji Yu[1,2], Cunchen Tang[1], and Shushan Hu[1]

[1] College of International Software, Wuhan University, Wuhan City, Hubei Province, 430079, China
[2] Animation Department, College of Arts, Hubei University, Wuhan City, Hubei Province, 430062, China

Abstract. With the rapid development of such disciplines and areas as computer vision, computer graphics, Internet, communication network and the software and hardware of computers, researches on the theory of AR technology and on the commercial application have been undertaken with great achievements in recent years, hence successful application of AR technology to such domains as military affairs, education and entertainment. This thesis shall begin with the model of animation characters, research the feasibility of application of AR technology to instructional system of design of animation model, design the model of digital instructional system of model design based upon AR technology. Through experimental tests and researches, it also analyzes the advantages of that system in teaching practices of model designing, argument relevant key technology in the system and provides experimental results of the working system.

Keywords: augmented reality, model design, tracking and three-dimensional registration.

1 Digital Teaching of Model Design Based upon AR Technology

AR technology (the abbreviation of Augmented Reality) is the main focus of global researches in recent years. Augmented reality takes on such a picture: when a virtual object is "added" to actual reality, such "adding" refers not to simple superposition but to seamless fusion of real world and virtual objects [1-5]. For example, figure 1 demonstrates the effect of application of augmented reality. Office table, the phone and surrounding wall belong to actual environment yet the lamp and two chairs are virtual three-dimensional objects [2, 6].

It is different from immersion experience general virtual reality brings to the user. Milgram, an American scholar researching on augmented reality once provided a continuum from a complete actual world to a complete virtual world, as figure 2 demonstrates [7-11].

Y. Yang, M. Ma, and B. Liu (Eds.): ICICA 2013, Part II, CCIS 392, pp. 469–478, 2013.

Fig. 1. The effect of application of augmented reality

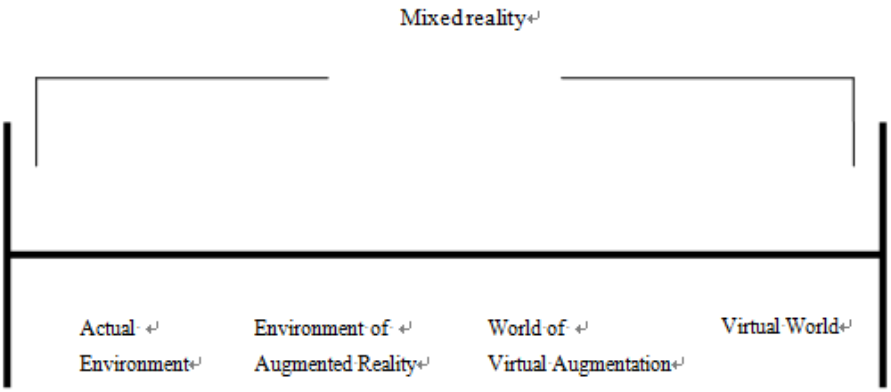


Fig. 2. The continual unity of actual world-- virtual world

Lester Madden, a British researcher mentioned in his monograph of augmented reality, that augmented reality shall include technology of five aspects [4, 12-16], namely, first, connective technology between actual world and digital pictures; second, interactive technology of virtual objects in real time; third, tracking technology in real time; four, identification technology for pictures and objects; five, providing digital materials and data in real time. Ronald Azuma, a scientist in HRL Laboratory, America, once also defined three features of AR [17-21], that is, the connection between actual objects and virtual ones in real world, interaction in real time and three-dimensional registration of virtual objects. In order to realize the basic application of AR technology thus users could finally experience the perfect fusion between actual world and virtual objects, the core key technology is tracking and

registration technology, which refers mainly to the technology of the achievement of position information of camera in actual world and that of placing exactly virtual objects in actual world. The AR application system must detect in real time features and relevant locations of marketed objects in actual scene, and identify the location and interior parameters after the conversion of coordinate systems of different types, thus to obtain exact information of cameras, which is called tracking process.

2 The Design of Animation Model

The design of animation model is an extremely important part in instructional process. For instance, figure 3 demonstrates the example of role design.

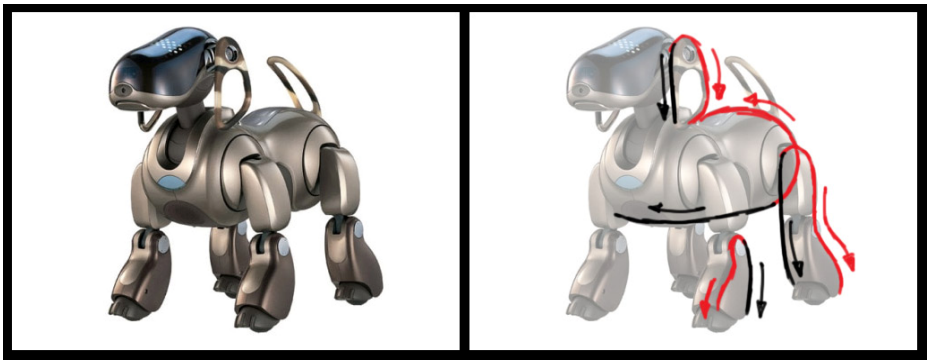


Fig. 3. The example of design of role model

There are at least several advantages of the application of AR technology to the teaching of the design of animation model as follows:

First, there are a large number of images of model design. Role model materials could be digitally dealt with in large quantities, and kept in the bank of virtual objects, thus to demonstrate model images of a great variety to the most degree and stimulate creative inspiration of students.

Second, one could sketch digitally those roles. The role modeling materials may be concrete and may be not in the bank of virtual objects, and whether model objects are animated or not, they could be sketched quite conveniently.

Third, one could observe dynamic models directly and clearly. One could control action of roles in real time with the use of interaction of AR technology in real time, and have command of key morphological characters of original animation immediately.

Four, one could observe the structure of role model directly and clearly. Different morphological structures of role models could be demonstrated conveniently and arbitrarily from different angles and perspectives, from the whole to parts and details and from external to internal structure.

Five, one could observe directly and clearly different photos of roles from different shooting scales. With the employment of interactivity of AR technology in real time and through the movement of marked objects in AR systems, one could dynamically

simulate different role imagines in screen from camera lens of different distances, which could be seen directly and clearly.

Six, instructional cost could saved in extremely large quantities. In the teaching process of sketching of real objects it takes a large sum of money to purchase sketched objects. In the digital instructional system employing AR technology, all role models are digital virtual objects, and the technology of computer graphics is used to make the sketched models almost real and also to be capable of designing digital sketched objects actual life could not provide.

3 The Design of the Establishment of System

3.1 Modules of the Entire System Structure

The function of the design of this system lies in that, in the teaching process of the design of animation models the system could demonstrate from many angles and shooting scales a large quantity of creative and inspiring modeling images and such morphological features as point, line, face and solid of the model, which facilitates interpretation and explanation of teachers and digital sketching of students [22-26].

3.2 Working Procedures of the System

According to modules of systematic structure demonstrated by figure 5, the design of digital instructional system of animation models based upon AR technology includes generally the following five basic working procedures. First, the system collects modules with video stream of marked objects of role models, thus to capture the images of real world by cameras. Second, through video stream of actual scenes collected in real time, one could extract and identify the features of the marked objects, and after the camera tracts modules and virtual role models register modules, the system calculate and identify the interior parameters and exterior parameters of the camera and location information of the marked objects in real world. Third, according to the obtained parameters of cameras, modules are plotted with virtual role models and one could invoke corresponding role models from digital model bank of role models, and plot the module in real time. Four, modules are compounded from virtual role models and actual scenes, hence seamless fusion of the two. Five, in order to demonstrate the output of audio stream of augmented reality, this system shall use projection devices with large screens as the output terminal of demonstration, revolve virtual role models by direction keys of the left and the right at keyboards, and amplify and shrink the images of virtual role models by the key “+” and “-”, hence interaction of human and machines.

4 Key Technology of the System

4.1 Tracking and Three-Dimensional Registration

Considering that the system use the system of augmented reality based upon PC platform, and that camera of high definition tends to be placed at set position, methods

of tracking and registration of computer vision based upon the recognition of marked objects are used, thus to improve tracking function and robustness of the system. In order to place virtual role models exactly into the three-dimensional space of real world, namely, to identify the three-dimensional locations and directions of virtual role models in real world, the essence is to identify the three-dimensional coordinates of any point of virtual role models in real world. Therefore, the process of three-dimensional registration is involved with calculation of how coordinate system of the world, coordinate system of cameras and coordinate system of screens convert mutually, as figure 4 demonstrates.

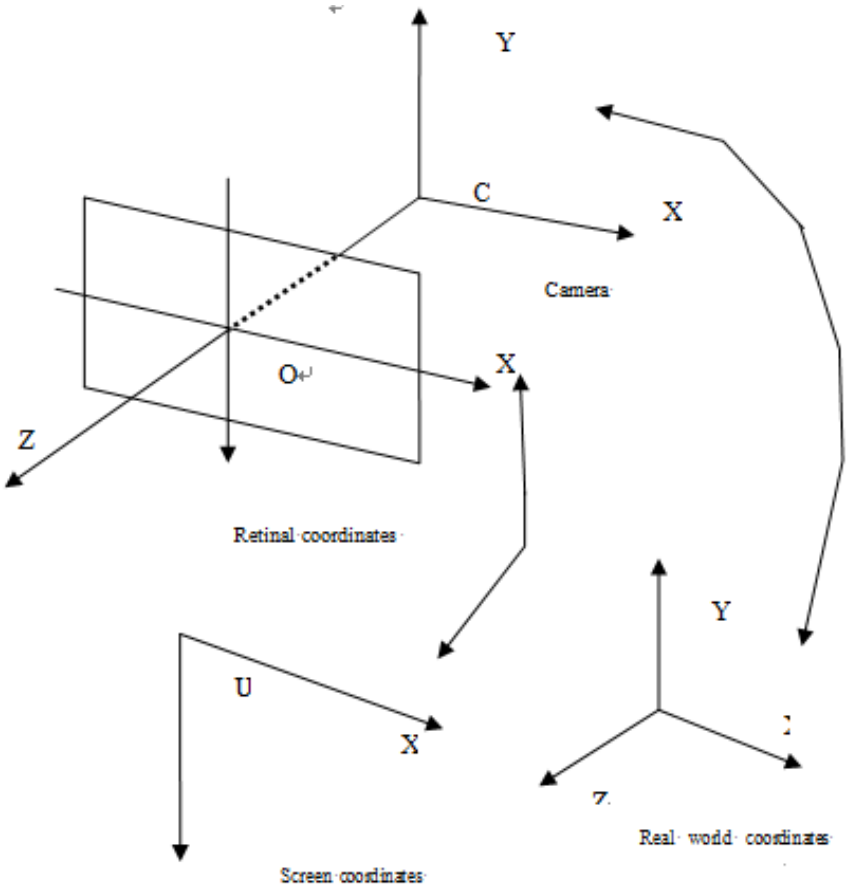


Fig. 4. The conversion of different coordinates in the system

4.2 How to Calculate the Conversion of Coordinates

The coordinates of any point in three-dimensional space of real world are supposed to be $W(X_w, Y_w, Z_w)$, which is $X_w = (X_w, Y_w, Z_w, 1)^T$ with the equation. The

coordinates of the point in coordinate system of computer screen are supposed to be $X_u = (X_u, Y_u, 1)^T$. Therefore, the system shall conduct the following conversion of coordinate systems in the process of tracking and registration.

The conversion of coordinate system of real world ($WX_wY_wZ_w$) and coordinate system of cameras ($CX_cY_cZ_c$) is demonstrated as Equation (1) [6].

$$\begin{pmatrix} X_C \\ Y_C \\ Z_C \\ 1 \end{pmatrix} \quad R_{WC}T_{WC} \begin{pmatrix} X_W \\ Y_W \\ Z_W \\ 1 \end{pmatrix} \tag{1}$$

Where: R_{wc}—rotation matrix of $CX_cY_cZ_c$ relative to $WX_wY_wZ_w$;

T_{wc}—translation quantity of three-dimensional space of $CX_cY_cZ_c$ relative to $WX_wY_wZ_w$. Both represent the position and pose of coordinate system of cameras relative to coordinate system of real world.

The conversion of coordinate system of real world and coordinate system of screens is demonstrated as Equation (2) [6].

$$\begin{pmatrix} X_u \\ Y_u \\ Z_u \\ 1 \end{pmatrix} \quad KR_{WC}T_{WC} \begin{pmatrix} X_W \\ Y_W \\ Z_W \\ 1 \end{pmatrix} \tag{2}$$

Where: K—interior parameters of camera, including focal length, the pixel size on the horizontal and vertical imaging plane and so forth. The system mainly calculate exterior parameters of cameras, namely, just calculate the two matrices of R_{wc} and T_{wc}.

$$T_{wc} = \begin{pmatrix} 1 & 0 & 0 & V_x \\ 0 & 1 & 0 & V_y \\ 0 & 0 & 1 & V_z \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{3}$$

Where: V_x, V_y, V_z represent respectively translation vector of coordinate system of camera relative to coordinate system of real world.

α, β and γ are supposed to represent respectively the rotation angles of the X axis, the Y axis and the Z axis of coordinate system of camera relative to coordinate system of real world, thus, the components on the X axis, the Y axis and the Z axis represent respectively as follows:

$$R_x(\alpha) = \begin{pmatrix} \cos\alpha & -\sin\alpha & 0 & 0 \\ \sin\alpha & \cos\alpha & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{4}$$

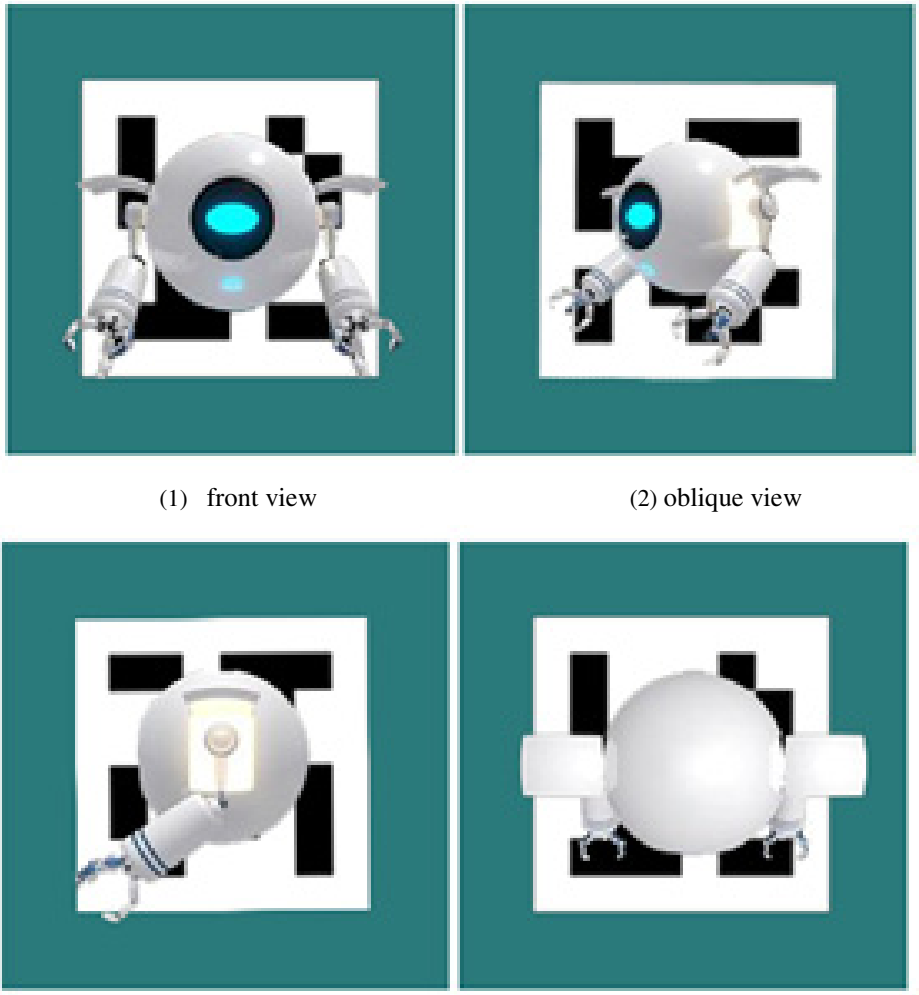
$$R_y(\beta) = \begin{pmatrix} \cos\beta & 0 & \sin\beta & 0 \\ 0 & 1 & 0 & 0 \\ -\sin\beta & 0 & \cos\beta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{5}$$

$$R_z(\gamma) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\gamma & -\sin\gamma & 0 \\ 0 & \sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{6}$$

With the equations of (4), (5) and (6), $R_{wc} = R_x(\alpha)\ R_y(\beta)\ R_z(\gamma)$.

4.3 The Achievement and Experimental Effects of the System

The basic developmental device of the system is ARtoolkit2.71.3, the bank of graphs adopts OpenGL, and the establishment of digital model bank of role models uses 3DS-MAX9. Figure 5 demonstrates the operational effects of the system on the platform of Windows XP. Through the change of location and perspectives of the marked objects of role model, the system renders and plots in real time the effects of digital role models from different angles, as figure 5 demonstrates. Compared with traditional teaching of design of animation models, teaching practices in this system have noticeable advantage and could improve the amount of instructional information and teach structural and morphological features of different role models from different angles and shooting scales, which is beyond the comparison of traditional instructional ways and methods.



(1) front view (2) oblique view

(3) lateral view (4) top view

Fig. 5. Experimental effect of the system

5 Conclusion and Prospects

The thesis probes into the digital instructional system of animation role models based upon AR technology, which is just a bold endeavor and experiment. There are many problems needing further researches and resolution in the operation of the system, such as the problem of shadow of virtual role models and the technology of light source detection in AR system, which are the directions of further research with efforts.

Acknowledgments. The Science and Technology Research Plan Project Tacking a Critical Point in Wuhan (2013060501010151).

References

1. Carmigniani, J., Furht, B., Anisetti, M., Ceravolo, P., Damiani, E., Ivkovic, M.: Augmented reality technologies, systems and applications. Springer Science Business Media, LLC (2010)
2. Azuma, R.T., Hughes Research Laboratories: A Survey of Augmented Reality. Tele-operators and Virtual Environments 6 (4), 355–385 (1997)
3. Milgram, P., Kishino, F.: A Taxonomy of Mixed Reality Visual Displays. IEICE Trans. Information Systems E77-D(12) (1994)
4. Madden, L.: Professional Augmented Reality Browsers for Smartphones. Wiley Publishing, and Inc. (2011)
5. Bennett, D.: Character Design Principles, http://www.disneyanimation.com/viza627/chardev/charDesign.html
6. Hammad, A.: Concordia University, Canada, Distributed Augmented Reality for visualizing, Collaborative Construction Tasks, pp. 171–183. Springer Science Business Media B.V. (2009)
7. Venkatesh, V.: Determinants of Perceived Ease of Use: Integrating Control, Intrinsic Motivation, and Emotion into the Technology Acceptance Model. Information Systems Research 11(4), 342–365 (2000)
8. Vessey, I., Glass, R.: Strong vs. weak approaches to systems development. Communications of the ACM 41(4), 99–102 (1998)
9. Walls, J.G., Widmeyer, G.R., El Sawy, O.A.: Building an Information System Design Theory for Vigilant EIS. Information Systems Research 3(1), 36–59 (1992)
10. Wand, Y., Weber, R.: On the Deep Structure of Information Systems. Information Systems Journal (5), 203–233 (1995)
11. Wand, Y., Weber, R.: On the Ontological Expressiveness of Information Systems Design Analysis and Design Grammars. Journal of Information Systems 3(3), 217–237 (1993)
12. Weber, R.: Editor's Comments: Still Desperately Seeking the IT Artifact. MIS Quarterly 27(2), iii–xi (2003)
13. Weber, R.: Ontological Foundations of Information Systems, Coopers & Lybrand, Brisbane, Australia (1997)
14. Weber, R.: Toward a Theory of Artifacts: A Paradigmatic Base for Information Systems Research. Journal of Information Systems 1(2), 3–19 (1987)
15. WfMC. Workflow Standard–Interoperability Wf-XML Binding. Document Number WFMC-TC- 1023, Version 1.0, Workflow Management Coalition (2000)
16. Nunamaker Jr., J.F., Chen, M., Purdin, T.D.M.: Systems development in information systems research. Journal of Management Information Systems 7(3), 89–106 (1990/1991)
17. Nunamaker, J.F., Dennis, A.R., Valacich, J.S., Vogel, D., George, J.F.: Electronic meeting systems. Communications of the ACM 34(7), 40–61 (1991)
18. Orlikowski, W.J.: Using Technology and Constituting Structures: A Practice Lens for Studying Technology in Organizations. Organization Science 11(4), 404–428 (2000)
19. Orlikowski, W.J., Barley, S.R.: Technology and Institutions: What Can Research on Information Technology and Research on Organizations Learn From Each Other? MIS Quarterly 25(2), 145–165 (2001)

20. Orlikowski, W.J., Suzanne Iacono, C.: Research Commentary: Desperately Seeking the "IT" in IT Research–A Call to Theorizing the IT Artifact. Information Systems Research 12(2), 121–134 (2001)
21. Orman, L.V.: Electronic Markets, Hierarchies, Hubs, and Intermediaries. Information Systems Frontiers 4(2), 213–228 (2002)
22. Pahl, G., Beitz, W.: Engineering Design: A Systematic Approach. Springer, London (1996)
23. Parsons, J., Wand, Y.: Emancipating instances from the tyranny of classes in information modeling. ACM Transactions on Database Systems (TODS) 25(2), 228–268 (2000)
24. Petroski, H.: Invention by design: how engineers get from thought to thing. Harvard University Press, Cambridge (1996)
25. Purao, S., Storey, V.C., Han, T.: Improving Analysis Pattern Reuse in Conceptual Design: Augmenting Automated Processes with Supervised Learning. Information Systems Research 14(3), 269–290 (2003)
26. Rittel, H.J., Webber, M.M.: Planning Problems Are Wicked Problems. In: Cross, N. (ed.) Developments in Design Methodology. John Wiley & Sons, New York (1984)

Programming Error Based on Ellipse Variable Programming

Chen Xiaohong, Ye Jun, and Ling Xunfeng

Zhejiang Institute of Mechanical & Electrical Engneering, Hangzhou, 310053, China
{60188357,15771041,152589782}@qq.com

Abstract. To reduce variable programming errors, improve products NC processing precision, based on ellipse, researched the programming approximate error characteristics, when taking θ, x, y as its argument, found that the approximate error distributions and the maximum approximate errors are clearly different; By the analysis of error characteristics, proposes a theory that selecting appropriate independent variables and using multi-variable segmentation programming to reduce programming errors.

Keywords: variable programming, programming error, ellipse.

1 Introduction

Parts from NC processing must be exist error, as programming error, control error, feed error, locating error and tool setting error, which programming error is related with approximation error, interpolation error, rounding error, the approximation error is caused by approximatting the part outline curve with approximation method. Because variable programming uses approximation theory, in fact, it is implemented by the methods of " segment generation of curve "or "arc generation of curve ", as the result of that, there will be approximate error, This article introduces the different programming errors caused by different arguments, [1-3] theory was advanced to reduce errors by selecting appropriate independent variables and using multi-variable segmentation programming.The following analysis description is based on ellipse [4, 5].

2 Programming Error Study

2.1 Basic Study Methods

In the case of processing ellipse $\frac{x^2}{40^2}+\frac{y^2}{30^2}=1$, analyze programming approximate errors originated from the arguments of the centrifugal angle θ, coordinates x, y, respectively, Using " segment generation of curve " approximation method, First, to establish right-handed Cartesian rectangular coordinates system for programming, origin by ellipsoid center, long shaft of this ellipse is on the X-axis, the short axis is

Y. Yang, M. Ma, and B. Liu (Eds.): ICICA 2013, Part II, CCIS 392, pp. 479–487, 2013.

on the Y-axis, for graphics' symmetry, only the parts on the first quadrant is analylzed, second, the argument in first quadrant is divided equally into 18 value, the segment $A_nA_{n+1}$by connecting a neighbor nodal points [4, 5] instead the elliptic curve segment A_nNA_{n+1}, for example showed as fig1, finally,select the length MN as the maximum error for analyzing, where, 1 $/\!/ A_nA_{n+1}$, and tangent to the ellipse to the point N, MN $\perp A_nA_{n+1}$, point M is the pedal, as shown in Figure 1.

3 Programming Error Study Based on the Arguments *θ,x,y* Respectively

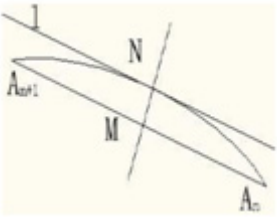


Fig. 1. Approximate error analysis calculation

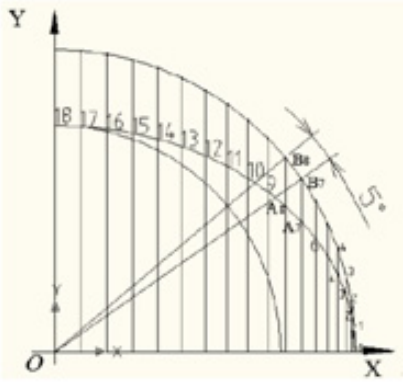


Fig. 2. Approximate error calculation based on θ

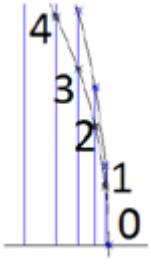


Fig. 3. Local amplification

3.1 Programming Approximate Error Analysis Based on θ Argument

See Figure 2, establishing Cartesian programming coordinate system, select the centrifugal angle θ as argument, initial angle 0°, θ loop step 90°/18=5°, That is, from 0 °, the nodes on the ellipse are obtained each increase 5 °, to program by the method that chord length insteads arc length, calculating the maximum approximation errors of each insteading, the results are shown in Table 1.

Table 1. Analyzing data based on the argument of centrifugal angle $\theta(\Delta\theta=5°)$

Nodal point	A_0	A_1	A_2	A_3	A_4	A_5	A_6	A_7	A_8	A_9
Centrifugal angleθ/[°]	0	5	10	15	20	25	30	35	40	45
Node coordi-nate x_A/[mm]	40	39.848	39.392	38.637	37.588	36.252	34.641	32.766	30.642	28.284
Node coordi-nate y_A/[mm]	0	2.615	5.209	7.765	10.261	12.679	15	17.207	19.284	21.213
chord length L/[mm]		2.619	2.634	2.664	2.708	2.762	2.826	2.896	2.970	3.046
tangency point x_N/[mm]		39.963	39.657	39.052	38.142	36.955	35.480	33.736	31.734	29.491
tangency point y_N/[mm]		1.309	3.916	6.493	9.021	11.481	13.853	16.119	18.263	20.268
foot point x_M/[mm]		39.924	39.621	39.016	38.115	36.924	35.452	33.709	31.711	29.470
foot point y_M/[mm]		1.306	3.909	6.483	9.007	11.4631	13.832	16.097	18.239	20.240
Error value ΔMN/[mm]		0.0380	0.0378	0.0374	0.0368	0.0361	0.0353	0.0344	0.0335	0.0327

Nodal point	A_{10}	A_{11}	A_{12}	A_{13}	A_{14}	A_{15}	A_{16}	A_{17}	A_{18}
Centrifugal angleθ/[°]	50	55	60	65	70	75	80	85	90
Node coordi-nate x_A/[mm]	25.712	22.943	20	19.905	13.681	10.353	6.946	3.486	0
Node coordi-nate y_A/[mm]	22.981	24.575	25.981	27.189	28.191	29.978	29.544	29.886	30
chord length L/[mm]	3.122	3.194	3.262	3.323	3.376	3.420	3.454	3.477	3.488
tangency point x_N/[mm]	27.024	24.351	21.492	18.470	15.307	12.028	8.658	5.221	1.745
tangency point y_N/[mm]	22.118	23.801	25.302	26.610	27.716	28.611	29.289	29.743	29.971
foot point x_M/[mm]	27.006	24.335	21.479	18.459	15.299	12.022	8.653	5.218	1.744
foot point y_M/[mm]	22.092	23.774	25.274	26.582	27.688	28.583	29.260	29.715	29.943
Error value ΔMN/[mm]	0.0319	0.0312	0.0305	0.0300	0.0295	0.0291	0.0289	0.0287	0.0286

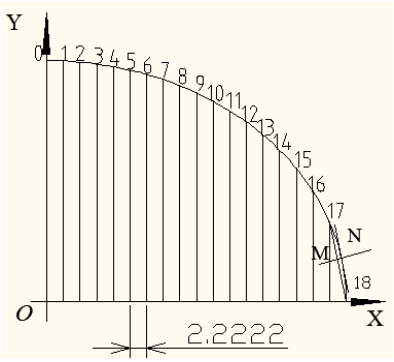


Fig. 4. Approximate error calculation based on *x* argument

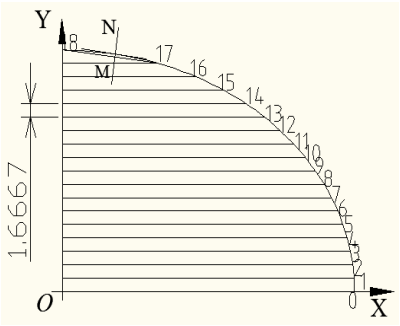


Fig. 5. Approximate error calculation based on *y* argument

3.2 Programming Approximate Error Analysis Based on x Argument

See Figure 4, establishing Cartesian programming coordinate system, select the coordinate *x* as argument, initial value 0, *x* loop step 40/18 = 2.2222, That is, from 0, the nodes on the ellipse are obtained each increase 2.2222, to program by the method that chord length insteads arc length, calculating the maximum approximation errors of each insteading, the results are shown in Table 2.

Table 2. Analyzing data based on the argument of coordinate *x* (adjacent node Δx=40/18= 2.2222)

Nodal point	A0	A1	A2	A3	A4	A5	A6	A7	A8	A9
Node coordi -nate xA/[mm]	0	2.222	4.444	6.667	8.889	11.111	13.333	15.556	17.778	20
Node coordi -nate yA/[mm]	30	29.954	29.814	29.580	29.250	28.819	28.284	27.638	26.874	25.981
chord length L/[mm]		2.223	2.227	2.234	2.247	2.264	2.286	2.314	2.350	2.395
tangency point xN/[mm]		1.112	3.335	5.558	7.781	10.004	12.227	14.451	16.674	18.898

Table 2. (*Continued.*)

tangency point yN/[mm]	29.988	29.896	29.709	29.427	29.047	28.564	27.974	27.269	26.441
foot point xM/[mm]	39.924	39.621	39.016	38.115	36.924	35.452	33.709	31.711	29.470
foot point yM/[mm]	1.306	3.909	6.483	9.007	11.4631	13.832	16.097	18.239	20.240
Error value ΔMN/[mm]	0.0380	0.0378	0.0374	0.0368	0.0361	0.0353	0.0344	0.0335	0.0327
Nodal point	A10	A11	A12	A13	A14	A15	A16	A17	A18
Node coordi -nate xA/[mm]	22.222	24.444	26.667	28.889	31.111	33.333	35.556	37.778	40
Node coordi -nate yA/[mm]	24.944	23.746	22.361	20.750	18.856	16.583	13.744	9.860	0
chord length L/[mm]	2.452	2.525	2.619	2.745	2.920	3.179	3.606	4.474	10.107
tangency point xN/[mm]	21.122	23.347	25.572	27.798	30.027	32.258	34.496	36.758	39.441
tangency point yN/[mm]	25.476	24.360	23.069	21.571	19.820	17.739	15.186	11.831	5
foot point xM/[mm]	27.006	24.335	21.479	18.459	15.299	12.022	8.653	5.218	1.744
foot point yM/[mm]	22.092	23.774	25.274	26.582	27.688	28.583	29.260	29.715	29.943
Error value ΔMN/[mm]	0.0319	0.0312	0.0305	0.0300	0.0295	0.0291	0.0289	0.0287	0.0286

3.3 Programming Approximate Error Analysis Based on Y Argument

See Figure 5, establishing Cartesian programming coordinate system, select the coordinate y as argument, initial value 0, y loop step is 30/18 = 1.6667, That is, from 0, the nodes on the ellipse are obtained each increase 1.6667, to program by the method that chord length insteads arc length, calculating the maximum approximation errors of each insteading, the results are shown in Table3.

Table 3. Analyzing data based on the argument of coordinate y (adjacent node Δy=30/18 =1.6667)

Nodal point	A_0	A_1	A_2	A_3	$\mathbf{A}_4$	A_5	A_6	A_7	A_8	A_9
Node coordi -nate y_A/[mm]	0	1.667	3.333	5	6.667	8.333	10	11.667	13.333	15
Node coordi -nate x_A/[mm]	40	39.938	39.752	39.440	39.000	38.426	37.852	36.851	35.832	34.64
chord length L/[mm]		1.668	1.6778	1.696	1.724	1.763	1.813	1.876	1.954	2.049

Table 3. (*Continued.*)

tangency point x_N/[mm]	39.985	39.861	39.612	39.236	38.729	38.085	37.298	36.359	35.254
tangency point y_N/[mm]	0.834	2.501	4.168	5.836	7.503	9.171	10.838	12.506	14.174
foot point x_M/[mm]	39.969	39.845	39.597	39.221	38.714	38.070	37.283	36.344	35.239
foot point y_M/[mm]	0.833	2.499	4.165	5.832	7.498	9.164	10.830	12.497	14.163
Error value ΔMN/[mm]	0.015	0.016	0.016	0.016	0.016	0.016	0.017	0.018	0.018
Nodal point	A_{10}	A_{11}	A_{12}	A_{13}	A_{14}	A_{15}	A_{16}	**A_{17}**	**A_{18}**
Node coordi -nate y_A/[mm]	16.667	18.333	20	21.667	23.333	25	26.667	28.333	30
Node coordi -nate x_A/[mm]	33.259	33.662	29.814	27.666	25.142	22.111	18.325	13.147	0
chord length L/[mm]	2.16509	2.309	2.488	2.719	3.025	3.459	4.136	5.440	13.252
tangency point x_N/[mm]	33.968	32.480	30.758	28.762	26.427	23.652	20.248	15.775	6.667
tangency point y_N/[mm]	15.842	17.510	19.179	20.849	22.520	24.193	25.872	27.568	29.580
foot point x_M/[mm]	33.953	32.465	30.743	28.746	26.411	23.635	20.229	15.751	6.614
foot point y_M/[mm]	15.829	17.496	19.162	20.829	22.495	24.162	25.828	27.495	29.162
Error value ΔMN/[mm]	0.019	0.021	0.023	0.026	0.030	0.036	0.048	0.077	0.422

4 Results of Error Study

Analysis Tables 1~3 and Figure 6, can conclude the followings.

4.1 Maximum Programming Approximate Errors

Same elliptic curves, same number of nodes, the maximum programming approximation error is less relatively when the argument is θ and not x, y, the error value is even larger than 10 times when x,y are arguments, Secondly, the approximation error value is less when argument is y (short axis) and not x (long axis).

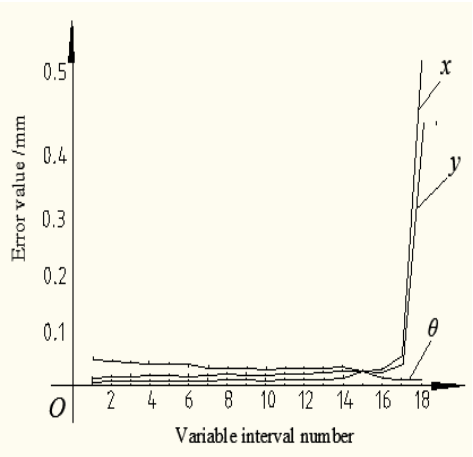


Fig. 6. Approximate error distribution comparison for three argument

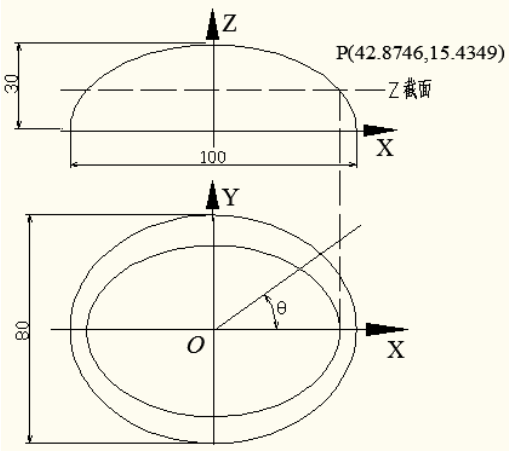


Fig. 7. Ellipsoid programming schematic diagram

4.2 Programming Approximate Error Equality

Same elliptic curves, same number of nodes, and arguments are *x, y*, θ, the average error values are 0.054, 0.047, 0.033, respectively. while the difference between maximum error and averge error are 0.500, 0.375, 0.005 respectively. And the difference between maximum and minimum error margin are 0.542, 0.407, 0.009 respectively, approximation error vary greatly. Therefore, when θ is argument, approximation error value is less, error uniformity is better.

4.3 Programming Approximate Error Distributions

Analyze the approximate error distribution carefully when *x* and *y* is the argument, most value of the error is less, and less than that when θ is argument, and has a

relatively small change. only a small number of intervals, the error changes rapidly, eventually becomes a catastrophe, therefore, if programmed with the well error interval rationally, programming approximate error will be significantly reduced.

5 Study Significance and Its Application

The analysis as above shows that, In the variable programming of CNC machining, the programming approximate maximum error differs widely because of selecting different arguments for the same product, Therefore, a reasonable selection of programming argument, can reduces programming approximate error, and can improves precision products, At the same time, according to the distribution of approximation error, in the majority of interval, the distribution of approximation error is more stable and uniform regardless of the choice of which argument , only small part of the interval has mutation , So take advantage of well interval , Use multiple arguments to program separately, avoid error mutation area effectively, can significantly improve the accuracy of the programming. and has practical guidance for variable programming. A case as following can illuminate.

To machining a half ellipsoid with CNC Milling machine, and the equation for the ellipsoid is $\frac{x^2}{50^2}+\frac{y^2}{40^2}+\frac{z^2}{30^2}=1$, establish Cartesian programming coordinate system as Figure 7, In order to highlight this programming method, the programming is simplified, cutter is looked as a tool without size, there is no the tool radius compensation, only arrange the finishing process, put down the tool in the lower left corne, and process from the down to top hierarchically, z is argument in down part, x is argument in the top , and centrifugal angle θ is argument when process the ellipse hierarchically, the cutting direction is clockwise, basic program is as fllowing.

O1; (bivariate segmented programming)

```
G54 G90 G40G17 G0 Z100;
M3S1000;
X-65 Y-55;
Z2 M08;
#13=0; (Z variable, initial value 0)
N13#11=50*SQRT[1-#13*#13/[30*30]]; (long half- axle a' of
ellipse in horizontal section)
#12=40*SQRT[1-#13*#13/[30*30]];  (the medium half- axle
b'  in the section)
G1Z[#13]F20; (move cutter along Z direction)
X-#11 F100; (move cutter along X direction)
Y0; (cut into along Y direction)
#4=180; (track by ellipse in the top section, the
centrifugal angle   as argument, initial angle is 0 °)
N3#1=#11*COS[#4]; (x coordinate of  nodal  points  in the
section)
#2=#12*SIN[#4]; (y coordinate of nodal  points in the
section)
```

```
G1X#1Y#2F100; (track by chord)
#4=#4-1;  (reduce argument Z, step0.1)
IF[#4GE-180]GOTO3; (condition of tracking by ellipse)
#13=#13+0.1; (add argument z, step0.1)
IF[#13LE15.4349]GOTO13;  (the loop condition for z
argument)
#21=42.8746; (X42.8746 is coordinate of intersection)
N11#23=30*SQRT[1-#21*#21/[50*50]]; (z value in the
section)
#22=40*SQRT[1-#23*#23/[30*30]]; (the mediu half- axle b'
in the section)
G1Z[#23]F20;
X-#21 F100;
Y0;
#4=180;
N5#2=#21*COS[#4];
#3=#22*SIN[#4];
G1X#2Y#3F100;
#4=#4-1;
IF[#4GE-180]GOTO5;
#21=#21-0.1; (reduce argument x, step0.1)
IF[#21GE0]GOTO11;(the loop condition for x argument)
G0 Z100M09;
X0Y200;
M30;
```

6 Summary

Variable programming is widely used in CNC machining, because the program is concise, easy to test, save machining time. The theory on selecting reasonable argument and segmentation programming with multi-arguments to improve precision has practical significance in the field of CNC, and it provides a good idea to improve programming precision in variable programming.

References

1. Valentino, J.V., Goldenberg, J.: Introduction to Computer Numerical Control, vol. 12(5), pp. 46–52. Courier Kendallville Inc., Kendallville (2008)
2. Smid, P.: CNC Programming Handbook: A Comprehehnsive Guilde to Practical CNC Programming, vol. 18(4), pp. 68–77. Industrial Press (2009)
3. Chen, X., Meng, Q.: Analysis of Cutter Radius Compensation. International Journal of Plant Engineering and Management 22(3), 5–12 (2009)
4. Feng, Z.: CNC macro programming methods, techniques and case, vol. 7(5), pp. 22–28. China Machine Press, Beijing (2007)
5. Liu, H.: Quality comparison of ellipsoid surface machined by equal-distance and bi- equal-distance method. Mechanical Research and Application 14(4), 12–17 (2009)

Library Management System Based on Recommendation System

Fu Jia[1] and Yan Shi[2]

[1] Library of Mudanjiang Medical University, Mudanjiang, Heilongjiang, China, 157011
[2] Information Technology Center of Mudanjiang Medical University, Mudanjiang, Heilongjiang, China, 157011

Abstract. With wide application of management system, information data grows rapidly. On one hand, people have a large number of information resources. On the other hand, the time cost and difficulty of people finding the proper information increases. In recent years, recommendation system attracts the attention of many researchers including management science, informatics and physics. From the application of recommendation system to library management system, the paper analyzes the key technology, makes application research on book recommendation system based on content filtering and collaborative filtering, and proposes collaborative filtering algorithm which is an improved recommendation algorithm. From the perspective of system application and design, the paper designs the structure, function module and user interface of book recommendation system.

Keywords: recommendation system, book recommendation system, recommendation algorithm, content filtering.

1 Introduction

At present, most libraries have installed management information system, which realizes digital management of major business processes. Book management system can be used to work on these data such as counting the quantity and turnover, query of readers borrowing books and retrieval of books. Although the functions provide convenience for managers and readers, improve efficiency and reduce the intensity, they still lack of deep mining of important data, full development of data value and further expansion of services. These data includes abundant and important information such as hobbies and interests of readers, relevance of readers and similarity of books. Book recommendation system is based on massive data to unearth valuable best books for readers through specific algorithms. After introducing basic principles, system modeling and key technologies of system, the paper tries to use the design of recommendation system to mine deep information of data for realizing recommendation functions.

Y. Yang, M. Ma, and B. Liu (Eds.): ICICA 2013, Part II, CCIS 392, pp. 488–497, 2013.

1.1 Collaborative Filtering Algorithm Based on Classification

1.1.1 Initial Score of Books

As there are hundreds of thousands of books, it is impossible for every book to be evaluated by readers. Meanwhile, many new books enter database everyday, and these books can't achieve readers' evaluation, which causes data sparse. In order to solve the problem, combined with the practical work in the library, the paper takes initial scoring method.

The new books need to be classified and received data entry by cataloguing staff before entering the system, in which cataloguing staff record the information including the classification, author and summary of the books [1]. In order to make new books receive evaluation before circulation, the paper adds a data processing program, initial score of new books, which means that an original recommended value is assigned to new books according to summary, author and preface of new books. It needs to be completed by cataloguing staff or subject librarian with knowledge. And the new books which have been evaluated can be provided timely for readers to solve the problem of data sparse.

Before implementing recommendation system, a large amount of new books which aren't evaluated and have been saved in the system can be assigned the score by the system subjectively. According to collection turnover and circulation frequency of books, the system figures out a recommendation score, as shown in Formula 1.

$$p=v0+n/10+t/20 \quad \text{(Formula 1)}$$

In the formula, v0 means the cardinality of score and can be set as -0.5, n means the circulation times in recent two years. When $n>10$, $n/10$ takes 1. t means the means borrowing time, and the unit is day. When $t>20$, $t/20$ takes 1.

The above two steps can ensure that all recommended books can achieve an initial score before implementing recommendation system.

1.2 Classification of Books and Readers

Chinese Library Classification is used to classify the recommended books. And the classification is different from that in recording work. There is no need to make detailed category analysis, and it only needs to bring it into the corresponding secondary category according to the classification number of books. And one thing needing to be point out is that when a book is classified, it not only belongs to one category, but also belongs to two classifications or more categories [2].

The classifications of these books are judged according to the books which have been evaluated by readers. Table 1 is the classification of readers. From the table, we can judge the category that readers belong to.

Table 2 is the information which is extracted from borrowing data. And from the books which have been evaluated by readers and their favorite books, we can judge the classification of readers.

Table 1. Classification of readers

Classification	B1	C2	C3	TU	TN	TP	R3	Q2	Q5
a									
b									
c									
d									
e									
f									
g									
h									

Table 2. Classification of books based on evaluation of readers

	Evaluated books	Favorite books	Category of favorite books	Classification of readers
Reader 1	a,b,d,g	a,b	F0,F2	F0,F2
Reader 2				
Reader 3				
Reader 4				
Reader 5				
Reader n				

2 Impact Factors of Reader Similarity

2.1 Readership

In general, the recommendation ability of undergraduates, graduate students, lectures, associate professors and professors on some professional books improves gradually, the reason for which is that the research time and research level on some professions improve gradually. Therefore, the first impact factor w1 of reader similarity is represented by readership.

Table 3. Weighting efficient of readership

	Undergraduate	graduate students	Lectures	Associate professors	Professors
w1	k1	k2	k3	k4	1

w1 means weighting efficient of readership, and k1,k2,k3 and k4 increases gradually. And weighting efficient of professor is 1. The influence and recommendation of professors is the greatest.

2.2 Amount of Evaluated Books

When the readers read large number of books in some field and evaluate them, it means that the time of readers studying the field increases, mastering comprehensive knowledge, and improving EXP. And the evaluation has great influence. Therefore, the amount of the evaluated books is used as the second impact factor of reader similarity, w2.

$$w2=\begin{cases} nu/A & (nu<A) \\ 1 & (nu\geq A) \end{cases} \tag{2}$$

2.3 Accuracy of Evaluation

After a book is borrowed by many readers and receive lots of evaluation, it has a mean score which is the nearest to the real score of the book. When the score of a reader on a book is close to means score, it means that the greater the accuracy of the reader evaluating the book, the greater the recommendation capacity. Therefore, accuracy of evaluation is used as the third impact factor of reader similarity, w3.

$$w_3=\frac{\sum_{i\in B}\left(1-\frac{\left|v_i-\overline{v}_i\right|}{Max-Min}\right)}{n} \tag{3}$$

In the formula, i means a book that has been evaluated by readers, B means the set of books that have been evaluated by readers, vi indicates that the score of readers on book I, $\overline{v}_i$ means the mean score that i receives, Max and Min mean the maximum score and minimum score that i receives, and n means the amount of books that readers have evaluated.

In conclusion, the factors influencing reader similarity can be expressed by the following formula

$$w''=(\alpha 1w1+\alpha 2w2)\times w3 \tag{4}$$

In the formula, α1+α2=1。

2.4 Algorithm Process

2.4.1 Classification of Books and Readers

According to the classification methods of Chinese Library Classification, secondary category is used as different classifications to partition the books into the category to which they belong. And the books belonging to cross-discipline subject are classed into the same category.

And the readers can be classified according to the score of readers. The books which have higher score of readers and more evaluation represent the interests of readers, and the readers can be classified into the corresponding category. And the classifications which have lower score of readers and less evaluation doesn't mean the interests of readers, the readers doesn't belong to these categories, and the books belonging to these categories can't be recommended. Readers always focus on the books in some fields because of the restrictions of profession, interest and time, so the classifications of the same readers need to be 2-3 types. Too many classifications not only need large calculation, but also make recommendation accuracy low.

After determining the category to which target reader belongs to, the readers of the category are analyzed. And the readers whose scores are less than the threshold are shielded, and effective readers are retained.

2.4.2 Calculation of Similarity

For collaborative filtering algorithms based on classifications, the relevance of readers of the category is great, and the evaluation of readers on some books is equivalent to the sum of the evaluation of readers on the books,

$$v_{a,k} = \sum_{j \in ck} v_{a,j} \tag{5}$$

In the formula, va,k means the evaluation of reader a on a category ck, j means a certain kind of books of a category ck, and va,j means the evaluation of reader a on the book j.

If a book belongs to many categories, the evaluation on the categories should be weighted, as shown in formula 6.

$$v_{a,k} = \sum_{j \in ck} p_{j,k} v_{a,j} \tag{6}$$

In the formula, pj,k means the probability of book j belonging to the category ck, and it meets the requirement, $\sum_{k \in C} p_{j,k} = 1$ which means that the sum of probability that book j belongs to the categories should be 1.

Pearson correlation coefficient algorithm is used to calculate the similarity among readers. The similarity between reader a and reader b is calculated as shown in formula 6.

$$w_{a,b} = \frac{\sum_{i=1}^{M} (v_{a,i} - \overline{v}_a)(v_{b,i} - \overline{v}_b)}{\sqrt{\sum_{i=1}^{M} (v_{a,i} - \overline{v}_a)^2 \sum_{i=1}^{M} (v_{b,i} - \overline{v}_b)^2}}$$

In the formula, M means the books that reader a and reader b evaluate, $\overline{v}_a, \overline{v}_b$ mean the means score of the books that reader a and reader b evaluate. And the nearest neighbor set U of reader a can be determined according to the similarity.

2.4.3 Generation Recommendation Results

In the category of the nearest neighbor set of reader a belonging to, the possible score of the books on reader a is predicted. According to formula 5, in view of the impact factors of reader similarity, the paper proposes the following formula to predict scores.

$$p_{a,k} = \overline{v}_a + \frac{\sum_{b \in U} (v_{b,j} - \overline{v}_b) \times w'_{a,b} \times w''_{a,b}}{\sum_{b \in U} w'_{a,b}} \tag{7}$$

In the formula, pa,k means the prediction score of reader a on the book k, $\overline{v}_a$, $\overline{v}_b$ mean the mean score of reader a and b on the books, U means the nearest neighbor set of a, $w'_{a,b}$ indicates correction value of similarity for reader a and b, and $w''_{a,b}$ means the impact factors of a and b similarity.

After getting the prediction score pa,k of reader a on the book k, we can get the prediction score of reader a on the book j by using formula 8.

$$p_{a,j} = \sum_{k \in R} wk \times p_{a,k} \tag{8}$$

In the formula, wk means the probability of book j belonging to category k, and R means the set of categories to which j belongs to. From the calculation, we can get the books which have the greatest prediction score for recommendation.

3 System Architecture Design of Book Recommendation System

Book recommendation system includes three layers, data warehouse layer, data process layer and user display layer. The structure chart is shown in Figure 1.

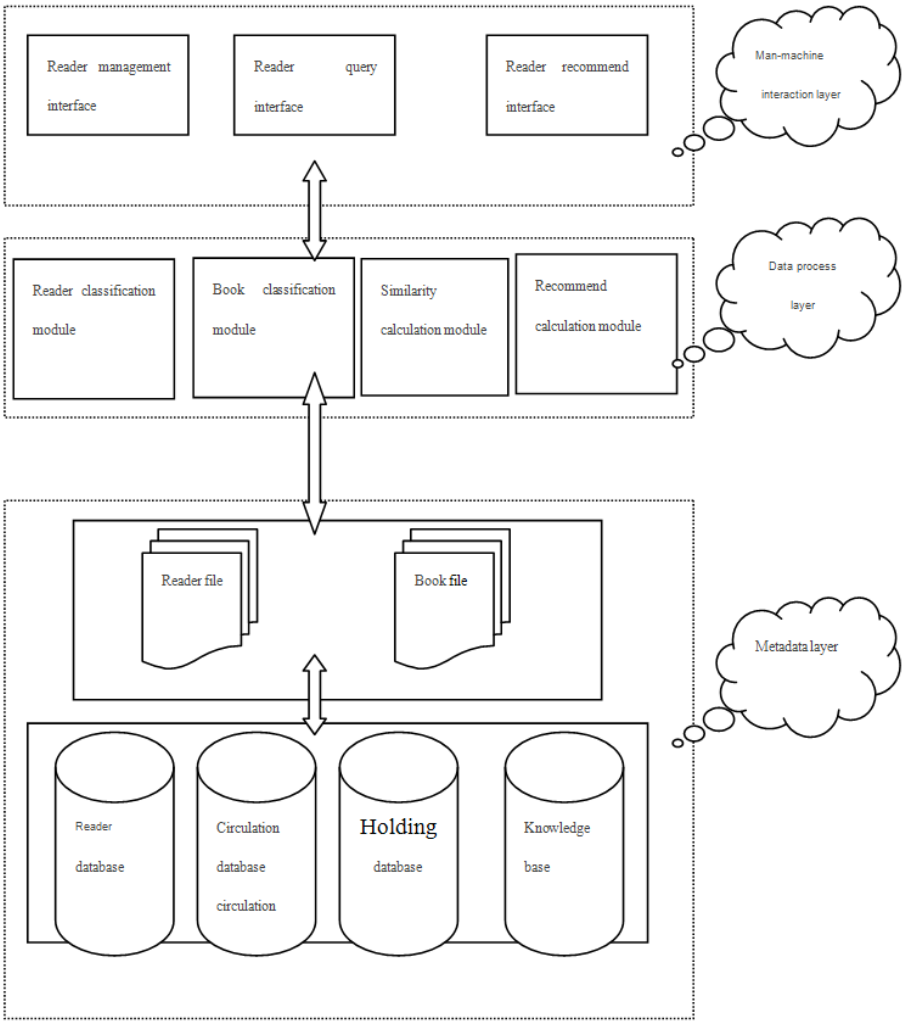


Fig. 1. System architecture of book recommendation system [3]

The first layer, data warehouse layer, belongs to metadata layer and includes reader database, circulation database, holding database and knowledge base. The layer preprocesses metadata including all kinds of information, and gets the data meeting the requirements which forms reader file and book file to be passed to the second layer.

The second layer is the core of the recommendation system and is the key of recommendation algorithm and propulsion technology. In the layer, the metadata in the first layer receives similarity calculation, the target data is filtered out, and the results are passed to the third layer to be showed t o the readers.

The third layer, user displayer layer, belongs to man-machine interaction layer. The requirements of readers are described in the interface, for example, the machine transforms query terms and interests description of readers into machine language to be passed to the data processing layer. Meanwhile, the machine outputs the processing results in user display layer.

3.1 Function Module Design of Book Recommendation System

In the practice of programming technique, the paper proposes the method of modular design. In the program, modules are the set of programs including data specification and executable statement. The set is the code block which is packaged together according to a certain rules for realizing the functions, and it is also called function module. In form, modules are some separate units which can be combined, separated and updated [4]. And each module completes an interdependent function.

The process of designing book recommendation system function module should obey the rule of high cohesion and low coupling to design interactive interface with independent and friendly functions, intelligent data analysis and the function module with processing capacity.

According to the analysis on the process of recommendation services, the paper divides book recommendation system into reader classification module, book classification module, similarity calculation module and recommendation calculation module.

3.2 Interface Design of Book Recommendation System

Throughout history, we can find that the original common tab styles and the later headlines are various. With the improvement of utilization of Google and Baidu network, more and more users, especially young people, like simple and clear retrieval interface style. Many network libraries correct the home pages according to the preference of readers and simple style. National Science Library in Chinese Academy of Sciences is taken as an example.

It is obvious that new homepage of National Science Library in Chinese Academy of Science has unified and simple interface style. And the readers can get straight to the point. So designing the interface of book recommendation system should strive for simple, friendly and individualized style.

3.2.1 Homepage of Book Recommendation System

Homepage of book recommendation system mainly includes retrieval field, new book recommendation column and popular book column [5]. The readers input key words in retrieval column to inquire books, magazines and database. New book recommendation column recommends new books for readers, and popular book column recommends the books which are borrowed most for readers. The recommended books face all readers

and have no individualized recommendation function. After logging in, the readers click correlation recommendation or collaborative recommendation menu, and enter correlation recommendation interface and collaborative recommendation interface, which can achieve the corresponding individualized recommendation.

3.2.2 Correlation Recommendation Interface

In correlation recommendation interface, after the readers retrieve Focus on AutoCAD2008 Mechanical Drawing, the system not only shows call number, author and editor of the book, but also displays a recommendation list beneath the interface. The system uses filtering technology based on content, and provides a recommendation list aiming at the property of the retrieved book and the interests of readers. The recommendation list only provides autograph, author and editor, but the readers can click title links to inquire detailed content. Besides receiving basic information including inquiring books, the readers can consult recommendation list to borrow the relevant books.

3.2.3 Collaborative Recommendation Interface

After the readers choose the function of collaborative recommendation, book recommendation system applies collaborative filtering algorithm to determine the nearest neighbor set of readers according to the profession and interests of readers, and calculates the recommendation book of the readers according to the evaluation. Through the comparison, we can find that collaborative filtering can recommend new interesting target for readers.

4 Conclusion

From the perspective of the application and design of book recommendation system, the paper analyzes the structure design of recommendation system, function modules and reader interface of the system, which establishes the foundation for further development on book recommendation system.

With wide application of information network, information data increases rapidly. On one hand, people have a large number of information resources. On the other hand, the time cost and difficulty of people finding the proper information increases. As Engels said, once the society has the requirement of technology, the need can promote the science more than ten universities. The individualized requirements of people in information field will promote rapid development of recommendation system. When recommendation system is widely applied to real network, people can enjoy humanized recommendation service brought by recommendation system.

Acknowledgement. Foundation item: 2013 annual scientific project of Education Department of Heilongjiang province (12532397).

References

1. Ma, Y.: Study on an adaptive collaborative filtering book recommendation system. Journal of Information 5, 105-106+109 (2008)
2. Zhu, Q., Hao, H., Liu, J., Jiao, P.: Application of DARVM to book recommendation system. Library and Information Service 15, 78–81+87 (2012)
3. Kun, D.: Study on book recommendation system based on collaborative filtering algorithm. Modern Library and Information Technology 11, 44–47 (2011)
4. Ma, D.: Construction of book reputation in university libraries. Library Works and Researches 3, 119–121 (2012)
5. Jing, M., Yu, Y.: Application of collaborative book recommendation model based on borrowing time score. Library and Information Service 3, 117–120 (2012)

Enterprise After-Sales Management System Based on J2EE

Yang Hongxia[1], Mu Ping[2], and Sun Yuqing[2]

[1] City College of Xi'an Jiao Tong University, Xi'an710021, China
[2] Third oil production plant of Daqing Oilfield, Heilongjiang Province, Daqing, China

Abstract. With the globalization of market and economy, more and more enterprises realize that it is more and more important to improve the competiveness of enterprise and after-sales service system. Traditional after-sales service system only focuses on the business process of the system and ignores a lot of important information in various after-sales service data. Data mining is a Knowledge Discovery in Database. It can help to find out some potential and valuable information and knowledge from daily experiences in favor of decision-making. The paper combines the design thought of data mining and J2EE architecture technology to analyze the application of data mining to after-sales service.

Keywords: data mining, J2EE, MVC, after-sales service system, decision tree.

1 Introduction

In fact, as competitive advantage of enterprises comes from high evaluation of consumer on products or service. And service quality has become the new focus of trade competition in the world [1]. So how to combine information management means and effectively improve service quality has become one of the hottest topics in recent years. In order to improve operational efficiency and promote the scientificalness of decisions, after-sales departments of enterprises establish their own new after-sales management system or reincorporate and form their own new after-sales information management system based on the existing information system [2, 3].

As an international brand, JVC enjoys a good reputation in china and in the world. Electronic products of JVC Corporation are popular with people and occupy a certain position in the market. JVC products spread all over the country. JVC Corporation has many Chinese users. At present, JVC Corporation has 100 after-sales service maintenance stations in China. In order to make after-sales service maintenance stations in China work well, the corporation needs to establish complete and specified informationized after-sales service management system [4, 5]. The system is after-sales service system developed by JVC (China) Trading Company Ltd which includes basic functions such as input, statistics, analysis and query. As after-sales networks of enterprises spread widely and data needs centralized processing, the system adopts B/S mode, which not only supports the process of remote business data, but also guarantees the safety of data. The system makes general management on all after-sales networks of

Y. Yang, M. Ma, and B. Liu (Eds.): ICICA 2013, Part II, CCIS 392, pp. 498–507, 2013.

the company in the country and after-sales service departments of the department. According to different businesses, after-sales service departments and every regional manager correspondingly authorize the after-sales networks and departments. Each department and network visit after-sales service system through internet. The system needs to be confirmed by the user log in. The system should have good maintainability, expandability and portability, which meets the requirements of the development in the future.

2 Enterprise After-Sales Service System Uses J2EE Architecture

Generally speaking, enterprise after-sales service system has the following characteristics [6].

(1) Decentralized application

Resource management faces every basic maintenance network, so the client application must realize zero-install and zero-maintenance. The client of resource management system should use the browser and combine Applet/JSP technique to realize decentralized query and entry.

(2) Flexible extension

The solution of resource management system should be a software platform with abundant content. It makes full use of the existing hardware and software information. Besides providing provisioning function, it should have great configurability and customization ability, which not only supports individualized requirements of different users, but also adapts to changing management models through extension mechanism.

(3) Seamless integration

Resource management system is one part of the enterprises information management system. In order to realize the maximum integration of the system, the integration between it and other systems must be realized by standard methods or interfaces, which can really establish dynamic and frictionless system integration.

(4) System based on open and standard architecture has long-term applicability, extensibility and interoperability.

The designed enterprise after-sales system adopts advanced J2EE architecture and distributed application model to transform and optimize the original system. The program, database and other components of the system are concentrated on the server side. Except for the browser, the client side doesn't need any component. The data and documents which need to be inquired by the user are from the dame data source, which ensures the timeliness and integrity of the data. The system provides the integration framework of the immediate layer for satisfying the requirements of high availability, high reliability and extensibility. J2EE provides unified development platform, which not only reduces the cost and complexity of developing multi-tier applications, but also provides strong support for the application integration. Good wizard supports package and deploy application, which not only enhances security mechanism, but also improves the performance. As time goes on, the requirements of the user are changing. When business treatment logic changes, it only needs to modify the program on the server side, which not only makes the development and release process easy, but also has no influence on the user.

3 System Requirements

3.1 Network Requirements and Characteristics

The requirements of after-sales service data center network in JVC Trading Co, Ltd are as follows [7].

(1) It demands to establish reliable, safe, advanced, manageable, scalable and stable network system of the whole JVC after-sales service, for supporting stable operation of the business system.

(2) There are about 100 maintenance stations all over the country in which 30 core maintenance networks not only need to connect with the headquarters in Shanghai for data transmission, but also need to consider extensibility of network system.

(3) The characteristics of wide-area network of JVC after-sales service system are that there are a lot of maintenance stations and concurrent User.

(4) There is a lot transmission data in each maintenance station, there is less download data and the data is focused on the headquarters in Shanghai.

(5) Privileged mobile users need to visit the business system in the headquarters.

3.2 Business Process Flow and Functional Requirements

After-sales service maintenance process means the process that the user takes the broken machine to after-sales maintenance network specified by the manufacture. The process is one of the important and commonest processes in after-sales service process.

The general functional requirements which need to be realized by the system include:

(1) Effectively, rapidly and accurately processing the circulation and transmission of the orders that various after-sales service need, which provides the record for the transmission of machines and parts, and realizes paperless office.

(2) Effectively, rapidly and accurately providing service process and strengthening effective management on service. After-sales service information transmission not only has unblocked processing channels, but also should be timely.

(3) Subjectively fulfilling promises, eliminating the problems, reducing external quality loss and improving the lifetime of the equipments and customer satisfaction.

(4) Mastering the condition of consumer complaints, realizing the disadvantages of products and service, retrieving the market impact, properly handling and eliminating consumer complaints.

(5) Measuring and analyzing customer satisfaction and evaluating effectiveness of quality management system. It can improve the product quality and service work.

3.3 Interface with Other Internal Software Systems

Besides internal data flowing, the designed after-sales service system needs to exchange data with other application systems, so the integrity, accuracy and consistency of data are important indexes of data quality. The information base of the system must guarantee the integrity, accuracy and consistency of the data.

(1)System interface list

The system needs to establish interfaces with part library management system used by JVC Trading Co, Ltd. And the interfaces run through the business process and supports data exchange. After-service system needs to make use of some resources of the part library management system of JVC to serve for after-sales service system. The interfaces needing to be completed include part library interface, packing list data interface and type list interface.

(2) Introduction instructions

The timing backend import mode is used to import the data in the corresponding data sheet of after-sales service system from the original part data library.

3.4 Requirements of System Expandability

With the development of system application and after-sales service management business, more subsystems may increase, so the system should be extensible in data layer, exchange layer and application service layer, and the system can realize increasing data source nodes, the content of application service, roles and objects by configuration. The extensibility of the system is manifested as follows.

(1)Extensibility of data layer

The system can be compatible with multiple types of data formats, and supports the mainstream operation system and database system. The system can support custom data exchange format which can make data exchange with the customer. And the customer can directly send electric data to the system form the information system.

(2) Extensibility of exchange layer

The system can configure data exchange nodes, increase new nodes, increase and decrease the exchanged data item and define the exchanged standard.

(3) Extensibility of application service layer

Application service layer demands to have the characteristics of realizing platform and customizability. The system can customize user management, authority management, organization management and menu management. The process of developing application system should be simplified to improve the operation efficiency and increase the extensibility of application.

3.5 Overall Performance Requirements

According to the nature of business in after-sales service management system, the whole system platform must have enough ability of processing data. And the server should have more than30% of processing allowance in the peak time, for improving impact resistance of the system and avoiding performance deterioration. Overall performance requirements of the system are as follows.

(1)It should have good portability and supports various operating system platform, database platform and programming language.

(2)It needs to have rigorous and perfect audit system, which not only ensures the safety and auditability of information exchange, but also ensures that data transmission is safe, rapid and accurate.

(3) It provides automated information exchange and has good expansibility, which fully considers the requirements of information content and business growth in the future.

(4) Providing a unified and specified access interface for the existing business system of each department and reserving information security support platform to guarantee the manageability of the system.

(5) Multiple data exchange platform systems can be cascaded, have great elasticity and expansibility, and support distributed structure, which provides basis for system expansion in the future.

(6) Supporting the colony of multiple servers, ensuring the elasticity of the system and supporting lots of concurrent users.

4 Framework Design

The process of framework design needs to determine the application framework of the system, establish a common web framework based on J2EE which not only should satisfy various functions of the user service, but also should comprehensively consider all technical limitation such as database connection, user interface, communication, safety and equipments.

Generally speaking, the Network service cycle of web application based on J2EE can be divided into four stages, acquiring the client request, executing business logic, choosing the page to jump and generating optional page. Figure1 is a detailed introduction on the process of the whole response of each HTTP request (POST or GET). (Filter request is optional. Service is used to write filter interception request and provides unified service for each request.

Map request to HTML action has a realized request service corresponding to each request.

Execute HTML action means that the program executes the business logic which is executed by some requests in network layer.

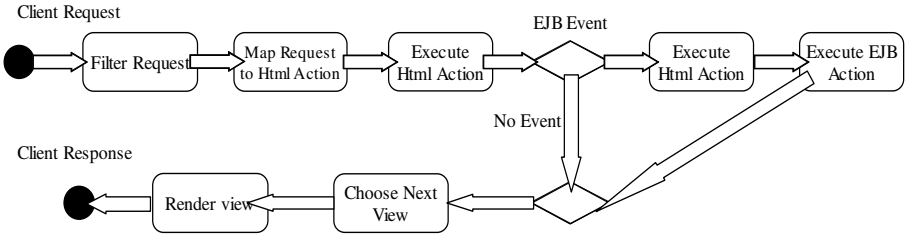


Fig. 1. Response process of system request

Choose Next view means to choose the next view to be realized based on the existing page, HTML action results and other optional state.

Render view means to render a page and returns it to the client user. HTML action may be returned to a serial EJB event which encapsulates a request and parameter in EJB layer.

Map event to EJB action means that EJB event includes the name of EJB action.

Execute EJB action means that the program sees EJB event as a parameter to execute EJB operation.

5 Database Design

As the business of the company is frequent and the data size is great, we need to use large relational database management tools -Oracle9i database system to manage the database. And Oracle9i is used to establish and maintain database. The system needs an effective data organizing form and management mechanism. The database design should satisfy the requirement that the data between each management module should be shared and transmitted smoothly. And after-sales service database with marketing and service support as the theme should be established.

5.1 Naming Rules of Data Sheet

(1) The common words are abbreviated. In the naming process, the words are abbreviated according to the semantics. As Oracle database can change the field name into the capital or the lower, the upper and lower line needs to be added. For example, DataManage, DM is data management, Dtl means detail and the detailed list of data management is defined as DM_Dt.

(2) If the sheet or the name of the filed only has one word, abbreviation should not be used, and it suggests using complete words. For example, Material, the name of material table is Material and the field code is Ma_ID.

(3) Master-detail sheet depending on relation. The suffix Main of master sheet can be not written.

(4) Data type is the field of text, and the suffix of type can be omitted.

5.2 Major Data Sheets

The system establishes JTS database under oracle9i and sets up more than 100 data sheets in which basic data modules and after-sales business management of the core part have 26 pieces of sheets.

Basic data sheets include:

DM_PART
DM_MACHINE
DM _GUARANTEE
DM_SHIP
DM _GOODSTYPE
DM_SYMPTOM
DM_DEAL

DM_SYMPTOM_DEAL
DM_PRESELLMACHINE
BIL_SUBMITSERVICE
BIL_SUBMITSERVICELIST
BIL_SERVICE
BIL_SERVICEOUT
BIL_SERVICEPART
BIL_SERVICEPARTOUT
BIL_FREESERVICE
BIL_FREESERVICELIST
TRANFERSERVICE
BIL_PARTORDER
BIL_ORDERLIST
BIL_PAYMENT
BIL_CONCTRACT

Data dictionary describes the storage structure information of data sheet in database. The data dictionary of data sheets used for after-sales service of the system is shown from Table 5-1 to Table 5-6.

Table 1. BIL_FREESERVICE

Field name	Type	If it is empty	Illustration
BIL_FREESEQ	Nvarchar2(17)	N	PK
BIL_JCCADM	Number(1)	N	
BIL_DUTIER	Nvarchar2(8)	N	
BIL_RESPONSER	Nvarchar2(8)	N	
BIL_FIXPRO	Number(1)	N	
… …	… …	… …	… …

Table 2. BIL_FREESERVICELIST

Field name	Type	If it is empty	Illustration
BIL_LISTESEQ	Nvarchar2(20)	N	PK
BIL_NO	Number(2)	N	
BIL_MAC_DE	Nvarchar2(15)	N	
BIL_MAC_NO	Nvarchar2(16)	N	
BIL_CONTENT	Nvarchar2	N	
… …	… …	… …	… …

Table 3. BIL_SERVICE

Field name	Type	If it is empty	Keys
BIL_COMPUTERCODE	Nvarchar2(17)	N	PK
BIL_SITENAME	Nvarchar2(30)	N	
BIL_PRINTCODE	Nvarchar2(12)	N	F
BIL_GOODSTYPE	Nvarchar2(2)	N	
BIL_MACTYPE	Nvarchar2(15)	N	
BIL_MACNO	Nvarchar2(15)	N	
BIL_SERVICETYPE	Number(1)	N	
BIL_FEETYPE	Number(1)	N	
BIL_FIXTYPE	Number(1)	N	
… …	… …	… …	… …

Table 4. BIL_SUBMITSERVICE

Field name	Type	If it is empty	Keys
BIL_SER_SEQ	Nvarchar2(20)	N	PK
BIL_SER_NO	Nvarchar2(16)	N	
BIL_COMPUTERCODE	Nvarchar2(17)	N	F
BIL_SUB_TIMES	Number(2)	N	
BIL_PAY_NO	Nvarchar2(19)	N	F
IMPORTTIME	Date	Y	
USERCODE	Nvarchar2(5)	N	

Table 5. BIL_PAYMENT

Field name	Type	If it is empty	Illustration
BIL_PAYMENT_NO	Nvarchar2(19)	N	PK
BIL_CLIENT_CODE	Nvarchar2(15)	N	
IMPORTTIME	Date	Y	
USERCODE	Nvarchar2(5)	N	

Table 6. BIL_SERVICEPART

Field name	Type	If it is empty	Illustration
BIL_PAERSEQ	Nvarchar2(20)	N	PK
BIL_COMPUTERCODE	Nvarchar2(17)	N	F
BIL_PARTSCODE	Nvarchar2(15)	N	F
BIL_PARTSNAME	Nvarchar2(60)	N	
BIL_PRICE	Number(2)	N	
BIL_PARTSNUM	Number(2)	N	

5.3 Realization of System Configuration

Server configuration: As described in the previous chapter, the application server of the system uses WebSphere. IBM WebSphere application server also contains IBM HTTP server. It includes Servlet run-time engine, high-performance database connection program, and provides application services suc has pre-linking, session and state management. WebSphere also supports XML file structure. And there is high-speed database coupling connection pool.

The major configurations are shown as follows. The port is set as 9080, action scope is localhost/server1, high-speed database coupling connection pool of JDBC is configured, Oracle JDBC Thin Driver is selected from the programs of JDBC, classes12.jar and nls_charset11.jar of ORACLEare added, JNDI name is configured in data source and database name, server name and data source name are input. For example, the instance name of ORACLE database which is established on server with IP as 192.168.1.25 is server, the configuration data source is jdbc:oracle: thin:@192.168.1.25:1521@server, the database name is server and the server name is 192.168.1.25.

Configuration of web.xml file: web.xml uses the form of xml file to describe the controller objects of the system. It can define the initialized parameters of the system, realize url information required by the customer, concrete mapping relationship on the server side and configure EJB quote. The following marks are added to web.xml.

<servlet>
<servlet-name>Svl_System</servlet-name>
<display-name>Svl_System</display-name>
<servlet-class>com.pantech.src.develop.system.Svl_System</servlet-class>
<init-param>
<param-name>application</para-name>
</servlet>

The servlet object points to one servlet controller of the system and appoints the initialized parameters.

<servlet-mapping>
<servlet-name> Svl_System </servlet-name>
<url-pattern>/ Svl_System </url-pattern>
</servlet-mapping>

servlet- mapping object realizes url information mapping relationship on the server side.

<welcomee-file-list>
<welcome-file>Login.jsp</welcome-file>
<welcome-file>default.jsp</welcome-file>
</welcome-file-list>

welcome-file-list object configures the initialize view of system operation.

<filter>
<filter-name>SetCharacterEncodingFilter</filter-name>
<display-name>SetCharacterEncodingFilter</display-name>
<filter-class>com.pantech.
common.filter.SetCharacterEncodingFilter</filter-class>
</filter>

filter object configures system filter, and SetCharacterEncodingFilter is used for character encoding.

6 Conclusion

With the globalization of market and economy, more and more enterprises realize that it is more and more important to improve the competiveness of enterprise and after-sales service system. Management information system is an important part of enterprise informationization. As market economy develops rapidly, consumers are paying attention to after-sales service of products. As a part of sales management of management information system, the function of after-sales service system is more and more important. Based on expounding classification ID3 decision tree algorithm, the paper establishes credit analysis decision support application model, which provides a reference for developing other data mining application systems.

References

1. Guisen, Z.: Analysis on status of enterprise information construction, http://www.jsetc.gov.cn
2. Wei, Q., Rao, Y.: Sales management system based on B/S structure. Hebei Industrial Science and Technology 20(1), 6–9 (2003)
3. Lin, Z.: Construction of enterprise information system facing the user. Research and Development of Computer 34(9), 657–661 (1997)
4. Li, Z., Yan, Z.: Development of C/S, B/S, three layers of structure theory. Journal of Henan Normal University 30(3), 24–27 (2002)
5. Li, H., Yin, Z., Jin, T.: Study on Application and development of design mode to J2EE system. Computer Engineering 31(14), 228–230 (2005)
6. Han, Y.: Application of computer to after-sales service management. Automotive Technology (3), 29–30 (1996)
7. Liu, W., Luo, Y.: Instruction manual of J2EE development. Electronic Industry Press (2004)

Bounds for the Norms of Symmetric Toeplitz Matrices in Information Theory

Zhaolin Jiang and Ziwu Jiang

Department of Mathematics, Linyi University, Linyi 276005, P.R. China
jzh1208@sina.com, zwjiang@gmail.com

Abstract. In recent years, there have been several papers on the norms of some special matrices. However, the bounds of symmetric Toeplitz matrix is still not clear, which is an important issue for related research area. In this paper, focusing on the theoretical analysis and mathematical proof, we find out some bounds for a symmetric Toeplitz matrix which is widely used in information theory with the first row $(x_0, x_1, \cdots, x_{n-1})$ in the case of $x_i = P_i$ and $x_i = Q_i$, where P_i and Q_i are Pell number and Pell-Lucas number respectively.

Keywords: Information Theory, Symmetric Toeplitz matrix, Norm, Pell number, Pell-Lucas number.

1 Introduction

In recent years, there have been several papers on the norms of some special matrices. However, the bounds of row symmetric Toeplitz matrix is still not clear, which is an important issue for related research area. Symmetric Toeplitz band matrices occur in many mathematical as well as scientific and engineering investigations. As an example, pentadiagonal Toeplitz matrices arise from 2nd and 4th order differential equations with various boundary conditions. In recent years, a number of fast algorithms for computing the determinant of a pentadiagonal Toeplitz matrix were developed [2, 3].

In the case of nonstructured matrices, this problem is commonly solved by reducing the matrix to a tridiagonal form and applying algorithms that compute the eigensystem of the tridiagonal form. In the case of structured matrices, these algorithms have a clear drawback: the tridiagonalization destroys the structure of the matrix. An array of antennas mounted on vehicles, ships, aircraft, satellites, and base stations is playing more and more important role in peoples daily life. However, many conventional subspace-based algorithms need explicitly know the noise characteristics of the array of sensors, which is regarded as being difficult for some practical applications. As to different certain conditions or different application backgrounds, the inverse eigenvalue problems have various formulations, and the structured inverse eigenvalue problems can be considered as the problems with the special structure. Many results on the inverse eigenvalue problems with Jacobi matrices and symmetric band matrices have been established. Recently, some other matrices with special structure such as Toeplitz matrix and

Y. Yang, M. Ma, and B. Liu (Eds.): ICICA 2013, Part II, CCIS 392, pp. 508–517, 2013.

Hankle matrix often occur in signal processing, system identification and image processing.

The Pell and Pell-Lucas sequences P_n and Q_n are defined by the recurrence relations

$$P_0 = 0,\ P_1 = 1,\ P_n = 2P_{n-1} + P_{n-2} \quad \text{for} \quad n \geq 2, \tag{1}$$

and

$$Q_0 = 2,\ Q_1 = 2,\ Q_n = 2Q_{n-1} + Q_{n-2} \quad \text{for} \quad n \geq 2. \tag{2}$$

If we start from $n = 0$, then Pell and Pell-Lucas sequences are given by

$$\begin{array}{cc} n & 0\,1\,2\,3\,4\,5\,6\,7\cdots \\ P_n & 0\,1\,2\,5\,12\,29\,70\,169\cdots. \\ Q_n & 2\,2\,6\,14\,34\,82\,198\,478\cdots \end{array}$$

In [1], their Binet forms are given by

$$P_n = \frac{1}{2\sqrt{2}}\left[\left(1+\sqrt{2}\right)^n - \cos(\pi n)\left(1+\sqrt{2}\right)^{-n}\right],$$
$$Q_n = \left[\left(1+\sqrt{2}\right)^n + \cos(\pi n)\left(1+\sqrt{2}\right)^{-n}\right].$$

Now we introduce some useful properties of the Pell and Pell-Lucas numbers as follows:

Lemma 1 ([2, 3]). *Let P_n be the n-th Pell number and Q_n be the n-th Pell-Lucas number, then we have*

1. $Q_n = 2P_n + 2P_{n-1}$;
2. $Q_nQ_{n-1} = 8P_nP_{n-1} + 4(-1)^{n-1}$;
3. $\displaystyle\sum_{i=0}^{n} P_i = \frac{P_{n+1} + P_n - 1}{2}$;
4. $\displaystyle\sum_{i=0}^{n} Q_i = \frac{Q_{n+1} + Q_n}{2}$;
5. $\displaystyle\sum_{i=0}^{n} P_i^2 = \frac{P_{n+1}P_n}{2}$;
6. $\displaystyle\sum_{i=0}^{n} Q_i^2 = 2(P_{n+1}^2 - P_n^2 + 1) = \frac{Q_{n+1}Q_n}{2} + 2.$

Lemma 2. *For $n \geq 1$, then we have the following recursion formulas*

1. $\displaystyle\sum_{i=0}^{n} P_iP_{i+1} = \frac{(P_{n+1} + P_n)^2 - 1}{4}$
$= \dfrac{1}{16}Q_{n+1}^2 - \dfrac{1}{4}$;

2. $\sum_{i=0}^{n} Q_i Q_{i+1} = 4P_{n+1}^2$
$= \dfrac{Q_{n+1}^2}{2} + 2(-1)^n.$

Proof. According to the properties of Pell and Pell-Lucas numbers as well as Lemma 1, it is not difficult to get the results.

In recent years, there have been several papers on the norms of some special matrices[3–12]. In [3], the authors give upper and lower bounds for the spectral norms of matrices $A = C_r(F_{k,0}, F_{k,1}, \cdots, F_{k,n-1})$ and $B = C_r(L_{k,0}, L_{k,1}, \cdots, L_{k,n-1})$, where $\{F_{k,n}\}_{n\in N}$ and $\{L_{k,n}\}_{n\in N}$ are k-Fibonacci and k-Lucas sequences respectively, and they also give the bounds for the spectral norms of Kronecker and Hadamard products of these matrices $A = C_r\left(F_0^{(k,h)}, F_1^{(k,h)}, \cdots, F_{n-1}^{(k,h)}\right)$ and $B = C_r\left(L_0^{(k,h)}, L_1^{(k,h)}, \cdots, L_{n-1}^{(k,h)}\right)$, where $F_n^{(k,h)}$ and $L_n^{(k,h)}$ are (k,h)-Fibonacci and (k,h)-Lucas numbers respectively[4]. Akbulak [5] found upper and lower bounds for the spectral norms of Toeplitz matrices such that $a_{ij} \equiv F_{i-j}$ and $b_{i-j} \equiv L_{i-j}$. Solak and Bozkurt [7] had found out upper and lower bounds for the spectral norms of Cauchy-Toeplitz and Cauchy-Hankel matrices in the forms $T_n = \left[\frac{1}{a+(i-j)b}\right]_{i,j=1}^{n}$, $H_n = \left[\frac{1}{a+(i+j)b}\right]_{i,j=1}^{n}$. Solak [9, 10] had defined $A = [a_{ij}]$ and $B = [b_{ij}]$ as $n \times n$ circulant matrices, where $a_{ij} \equiv F_{(\mathrm{mod}(j-i,n))}$ and $b_{ij} \equiv L_{(\mathrm{mod}(j-i,n))}$, then he has given some bounds for the A and B matrices concerned with the spectral and Euclidean norms.

Since different intervals can be processed in parallel, it becomes possible to greatly reduce the execution time by using parallel computational resources like clusters of computers or multicore computers. From practical point of view, pentadiagonal matrices frequently arise from boundary value problems involving fourth order derivatives and a fast computational formula for the determinants are required to test efficiently the existence of unique solutions of the PDEs. These problem arises in many applications, such as the discretization of partial differential equations (PDEs) in 2D or 3D by finite difference or finite element approximation. In the general case (non-structured matrices), these algorithms are used when memory space limits are reached. They are also used if only a few eigenvalues/eigenvectors are required, maybe the largest ones or the smallest ones. This approach is also preferred in the structured matrix field since the structure of the matrix is not destroyed in the process. Another useful property of the symmetry of Toeplitz matrices allows us to extract symmetric and skew-symmetric eigenvalues at the same time. Once the linear system is solved the process goes on in parallel. Among these methods, the covariance differencing approach is notably worthy and has been proposed for bearing estimation of desired signals embedded in colored noise. The noise component is eliminated by forming the difference of the original and transformed covariance matrices with an assumption that the noise covariance matrix is a Hermitian symmetric Toeplitz matrix.

In this paper, we consider one kinds of special matrices defined as follow.

Definition 1. *Let $X = \{x_0, x_1 \ldots, x_{n-1}\}$ is a given sequence, an $n \times n$ matrix T_n is called a symmetric Toeplitz matrix if it is of the form*

$$T_n = \begin{pmatrix} x_0 & x_1 & x_2 & \cdots & \cdots & x_{n-1} \\ x_1 & x_0 & x_1 & \ddots & & \vdots \\ x_2 & x_1 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & x_1 & x_2 \\ \vdots & & \ddots & x_1 & x_0 & x_1 \\ x_{n-1} & \cdots & \cdots & x_2 & x_1 & x_0 \end{pmatrix}. \tag{3}$$

Here X is called the input sequence, and obviously, $T(i, j) = x_{|i-j|}$ for each $i, j = 1, \ldots, n$. The symmetric Toeplitz matrix has been an object of great interest in mathematics for a long time(see e.g. [13–15]). The sample autocovariance matrix in time series analysis also has the symmetric Toeplitz structure.

As is well known, the inverse eigenvalue problems for matrices is just to confirm the elements of matrices under certain conditions due to the information of matrices of eigenvalues or eigenvectors. The problems are not only one of the hottest topics in numerical linear algebra, but also have found successful applications in a variety of areas such as discrete simulation, structure design, structure finite element model updating, system identification, nuclear spectroscopy of particle physics .i.e. These researches are significant in the study of theories and applications. Such factors indicate that the number of eigenvalues in each subinterval must not be too large, and that the width of the subinterval must be controlled as well, since very large subintervals with eigenvalues in the extremes would slow down the convergence. The suitable maximum number of eigenvalues per subinterval depends on these two factors, but it also depends on the matrix spectrum, the solvers and the hardware architecture.

In this study, we define symmetric Toeplitz matrices with input sequences $\{P_0, P_1, \cdots, P_{n-1}\}$ and $\{Q_0, Q_1, \cdots, Q_{n-1}\}$ respectively, and then we give lower and upper bounds for the spectral norms of these matrices. A process uses cores inside the node by means of light processes, that is, threads. The model does not make restrictions about the number of processes that can be mapped in a single scheme. There are several factors that must be taken into account when determining the subintervals. It can be observed that it takes a few Lanczos iterations until the eigenvalues start to be extracted.

Now we start with some preliminaries. Let A be any $n \times n$ matrix. The ℓ_p norms of the matrix A are defined as

$$\|A\|_p = \left(\sum_{i=1}^{m} \sum_{j=1}^{n} |a_{ij}|^p \right)^{\frac{1}{p}} \qquad (1 \le p < \infty).$$

The well-known Frobenius (or Euclidean) norm of matrix A is

$$\|A\|_F = \left(\sum_{i=1}^{m} \sum_{j=1}^{n} |a_{ij}|^2 \right)^{\frac{1}{2}},$$

and also the spectral norm of matrix A is

$$\|A\|_2 = \sqrt{\max_{1 \le i \le n} \lambda_i(A^H A)},$$

where A is $m \times n$ and A^H is the conjugate transpose of matrix A. The following inequality holds:

$$\frac{1}{\sqrt{n}} \|A\|_F \le \|A\|_2 \le \|A\|_F. \tag{4}$$

Let $A = [a_{ij}]$ and $B = [b_{ij}]$ be $n \times n$ matrices. The Hadamard product of A and B is defined by $A \circ B = [a_{ij} b_{ij}]$. If $\|\cdot\|$ is any norm on $n \times m$ matrices, then [9]

$$\|A \circ B\| \le \|A\| \cdot \|B\|.$$

Let A and B be arbitrary $n \times m$ matrices. Kronecker product of A and B is given to be

$$A \otimes B = \begin{bmatrix} a_{11}B & \cdots & a_{1m}B \\ \vdots & & \vdots \\ a_{n1}B & \cdots & a_{nm}B \end{bmatrix},$$

then [9]

$$\|A \otimes B\|_F = \|A\|_F \|B\|_F.$$

2 Main Results

Theorem 1. *Let $n \times n$ matrix A be as $A = [a_{ij}]$ such that $a_{ij} = P_{|i-j|}$, then*

$$\sqrt{\frac{(P_{n-1} + P_n)^2 - 1}{n}} \le \|A\|_2 \le P_{n-1} + P_n - 1,$$

where $\|\cdot\|_2$ is the spectral norm and P_n denote the n-th Pell number.

Proof. The matrix A is of the form

$$A = \begin{pmatrix} P_0 & P_1 & P_2 & \cdots & \cdots & P_{n-1} \\ P_1 & P_0 & P_1 & \ddots & & \vdots \\ P_2 & P_1 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & P_1 & P_2 \\ \vdots & & \ddots & P_1 & P_0 & P_1 \\ P_{n-1} & \cdots & \cdots & P_2 & P_1 & P_0 \end{pmatrix}, \tag{5}$$

From the definition of Frobenius norm, we can get

$$\|A\|_F^2 = nP_0^2 + 2\sum_{i=1}^{n-1}\sum_{j=1}^{i} P_j^2.$$

According to Lemma 1 and Lemma 2, we can get

$$\|A\|_F^2 = nP_0^2 + 2\sum_{i=1}^{n-1}\sum_{j=1}^{i} P_j^2 = \sum_{i=1}^{n-1} P_i P_{i+1} = \frac{(P_{n-1}+P_n)^2 - 1}{4}.$$

Once an appropriate number of eigenvalues has been decided, the main interval containing all the desired eigenvalues is divided into a reasonable number of subintervals. We know that, $\frac{1}{\sqrt{n}}\|A\|_F \le \|A\|_2 \le \|A\|_F$ from equivalent norms where $\|\cdot\|_F$ is the Euclidean norm, so

$$\sqrt{\frac{(P_{n-1}+P_n)^2 - 1}{n}} \le \|A\|_2.$$

On the other hand,

$$M_1 = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \ddots & 0 & 0 \\ 0 & 0 & 0 & \ddots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{pmatrix},$$

and

$$M_2 = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & 0 \\ 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ddots & 0 & 0 \\ 0 & 0 & 0 & \cdots & 1 & 0 \end{pmatrix},$$

then

$$A = P_0 E_n + \sum_{i=1}^{n-1} P_i M_1^i + \sum_{i=1}^{n-1} P_i M_2^i.$$

We can get

$$\begin{aligned} \|A\|_2 &= \left\| P_0 E_n + \sum_{i=1}^{n-1} P_i M_1^i + \sum_{i=1}^{n-1} P_i M_2^i \right\|_2 \\ &\le \sum_{i=1}^{n-1} P_i \|M_1\|_2^i + \sum_{i=1}^{n-1} P_i \|M_2\|_2^i. \end{aligned} \tag{6}$$

Because

$$M_1^H M_1 = \begin{pmatrix} 0 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix},$$

and

$$M_2^H M_2 = \begin{pmatrix} 1 & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 1 & 0 \\ 0 & \cdots & 0 & 0 \end{pmatrix},$$

then

$$\|M_1\|_2 = \|M_2\|_2 = 1.$$

It is clear that $\|E_n\|_2 = 1$, so we can get the other result as follows

$$\begin{aligned} \|A\|_2 &\le P_0\|E_n\|_2 + \sum_{i=1}^{n-1} P_i \|M_1\|_2^i + \sum_{i=1}^{n-1} P_i \|M_2\|_2^i \\ &= P_0 + 2\sum_{i=1}^{n-1} P_i = P_{n-1} + P_n - 1. \end{aligned}$$

Thus, the proof is completed.

Theorem 2. *Let $n \times n$ matrix B be as $B = [b_{ij}]$ such that $b_{ij} = Q_{|i-j|}$, then*

$$\sqrt{\frac{Q_n^2 - 4(-1)^n}{2n}} \le \|B\|_2 \le Q_{n-1} + Q_n + 2,$$

where $\|\cdot\|_2$ is the spectral norm and Q_n denote the n-th Pell-Lucas number.

Proof. The matrix B is of the form

$$B = \begin{pmatrix} Q_0 & Q_1 & Q_2 & \cdots & \cdots & Q_{n-1} \\ Q_1 & Q_0 & Q_1 & \ddots & & \vdots \\ Q_2 & Q_1 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & Q_1 & Q_2 \\ \vdots & & \ddots & Q_1 & Q_0 & Q_1 \\ Q_{n-1} & \cdots & \cdots & Q_2 & Q_1 & Q_0 \end{pmatrix}. \tag{7}$$

From the definition of Frobenius norm, we have

$$\|B\|_F^2 = nQ_0^2 + 2\sum_{i=1}^{n-1}\sum_{j=1}^{i} Q_j^2.$$

According to Lemma 1 and Lemma 2, we can get

$$\|B\|_F^2 = nQ_0^2 + 2\sum_{i=1}^{n-1}\sum_{j=1}^{i} Q_j^2 = 4n + 2\sum_{i=1}^{n-1}\left(\frac{Q_iQ_{i+1}}{2} - 2\right)$$
$$= 4 + \sum_{i=1}^{n-1} Q_iQ_{i+1} = \frac{Q_n^2 - 4(-1)^n}{2}.$$

We know that, $\frac{1}{\sqrt{n}}\|B\|_F \le \|B\|_2 \le \|B\|_F$ from equivalent norms where $\|\cdot\|_F$ is the Euclidean norm, so

$$\sqrt{\frac{Q_n^2 - 4(-1)^n}{2n}} \le \|B\|_2.$$

On the other hand, according to M_1 and M_2 defied in Theorem1, we can get

$$B = Q_0E_n + \sum_{i=1}^{n-1} Q_iM_1^i + \sum_{i=1}^{n-1} Q_iM_2^i.$$

Because

$$\|E_n\|_2 = \|M_1\|_2 = \|M_2\|_2 = 1,$$

the other result can be obtained as follows

$$\|B\|_2 \le Q_0\|E_n\|_2 + \sum_{i=1}^{n-1} Q_i\|M_1\|_2^i + \sum_{i=1}^{n-1} Q_i\|M_2\|_2^i$$
$$= 2 + 2\sum_{i=1}^{n-1} Q_i = Q_{n-1} + Q_n + 2.$$

Thus, the proof is completed.

Corollary 1. *Let $n \times n$ matrix $A = [a_{ij}]$ and $B = [b_{ij}]$ such that $a_{ij} = P_{|i-j|}$, $b_{ij} = Q_{|i-j|}$, then the spectral norm of Hadamard product of A and B is valid the inequality*

$$\|A \circ B\|_2 \le (P_{n-1} + P_n - 1) \times (Q_{n-1} + Q_n + 2).$$

Proof. Since $\|A \circ B\|_2 \le \|A\|_2\|B\|_2$, the proof is trivial by Theorems 1 and 2.

Corollary 2. *Supposed that matrices $A = [a_{ij}]$ and $B = [b_{ij}]$ such that $a_{ij} = P_{|i-j|}$, $b_{ij} = Q_{|i-j|}$, then the Frobenius norm of Kronecker product of A and B is*

$$\|A \otimes B\|_F = \frac{\sqrt{2[(P_{n-1} + P_n)^2 - 1] \times [Q_n^2 - 4(-1)^n]}}{4}.$$

Proof. Since $\|A \otimes B\|_F = \|A\|_F\|B\|_F$, the proof is trivial by Theorems 1 and 2.

Acknowledgments. This work was supported by the NSFC (Grant Nos. 11301252).

References

1. Kocer, E.G., Tuglu, N., Stakhov, A.: On the m-extension of the Fibonacci and Lucas p-numbers. Chaos. Soliton Fract. 40, 1890–1906 (2009)
2. Bicknell, M.: A primer on the Pell sequence and related sequences. Fib. Quart. 13, 345–349 (1975)
3. Shen, S.Q., Cen, J.M.: On the spectral norms of r-circulant Matrices with the k-Fibonacci and k-Lucas numbers. Int. J. Contemp. Math. Sciences 5, 569–578 (2011)
4. Shen, S.Q., Cen, J.M.: On the norms of circulant matrices with the (k, h)-Fibonacci and (k, h)-Lucas numbers. Int. J. Contemp. Math. Sciences 6, 887–894 (2011)
5. Akbulak, M., Bozkurt, D.: On the norms of Toeplitz matrices involving Fibonacci and Lucas numbers. Hacet. J. Math. Stat. 37, 89–95 (2008)
6. Solak, S., Bozkurt, D.: Some bounds on l_p matrix and l_p operator norms of almost circulant, Cauchy-Toeplitz and Cauchy-Hankel matrices. Math. Comput. Appl. 7, 211–218 (2002)
7. Solak, S., Bozkurt, D.: On the spectral norms of Cauchy-Toeplitz and Cauchy-Hankel matrices. Appl. Math. Comput. 140, 231–238 (2003)
8. Solak, S., Bozkurt, D.: A note on bound for norms of Cauchy-Hankel matrices. Numer. Linear Algebra Appl. 10, 377–382 (2003)
9. Solak, S.: On the norms of circulant matrices with the Fibonacci and Lucas numbers. Appl. Math. Comput. 160, 125–132 (2005)
10. Solak, S.: Erratum to 'On the norms of circulant matrices with the Fibonacci and Lucas numbers. Appl. Math. Comput. 160, 125–132 (2005), Appl. Math. Comput. 190, 1855–1856 (2007)
11. Mathias, R.: The spectral norm of a nonnegative matrix. Linear Algebra Appl. 139, 269–284 (1990)
12. Yalçiner, A.: Spectral norms of some special circulant matrices. Int. J. Contemp. Math. Sciences 3, 1733–1738 (2008)
13. Grenander, U., Szegö, G.: Toeplitz forms and their applications. California Monographs in Mathematical Sciences. University of California Press, Berkeley–Los Angeles (1958)
14. Jiang, Z.L., Zhou, Z.X.: Circulant Matrices. Chengdu Technology University Publishing Company, Chengdu (1999)
15. Basak, A., Bose, A.: Limiting spectral distributions of some band matrices. Period. Math. Hungar. 63, 113–150 (2011)
16. Visick, G.: A quantitative version of the observation that the Hadamard product is a principal submatrix of the Kronecker product. Linear Algebra Appl. 304, 45–68 (2000)
17. Ammar, G.S., Gragg, W.B.: Superfast solution of real positive definite Toeplitz systems. SIAM J. Matrix Anal. Appl. 9, 61–76 (1988)
18. Brent, R.P., Gustavson, F.G., Yun, D.Y.: Fast solution of Toeplitz systems of equations and computations of Pad approximations. J. Math.Phys 1, 259–295 (1980)
19. Makhoul, J.: On the eigenvectors of symmetric Toeplitz matrices. IEEE Trans. Acoust., Speech, Signal Processing 11, 868–872 (1990)

20. Willsky, A.S.: Relationships between digital signal processing and control and estimation theory. Proc. IEEE 66, 996–1017 (1988)
21. Chan, R.H., Ng, M.K.: Conjugate Gradient Methods for Toeplitz Systems. SIAM Review 38(3), 427–482 (1996)
22. Mackens, W., Voss, H.: The Minimum Eigenvalue of a Symmetric Positive-Definite Toeplitz Matrix and Rational Hermitian Interpolation. SIAM Journal on Matrix Analysis and Applications 18(3), 521–534 (1997)
23. Capizzano, S.S.: Toeplitz Preconditioners Constructed from Linear Approximation Processes. SIAM Journal on Matrix Analysis and Applications 20(2), 446–465 (1998)
24. Lin, F.-R., Ng, M.K., Chan, R.H.: Preconditioners for Wiener–Hopf Equations with High-Order Quadrature Rules. SIAM Journal on Numerical Analysis 34(4), 1418–1431 (1997)
25. Bose, A., Sen, A.: Spectral norm of random large dimensional noncentral Toeplitz and Hankel matrices. Electron. Comm. Probab. 12, 29–35 (2007)
26. Bryc, W., Dembo, A., Jiang, T.: Spectral measure of large random Hankel, Markov and Toeplitz matrices. Ann. Probab. 34(1), 1–38 (2006)
27. Hammond, C., Miller, S.J.: Distribution of eigenvalues for the ensemble of real symmetric Toeplitz matrices. J. Theoret. Probab. 18(3), 537–566 (2005)

New Path Following Algorithm for Convex Programming

Xiaona Fan and Dianxuan Gong

College of Science, Nanjing University of Posts and Telecommunications, Nanjing, Jiangsu, 210046, P.R. China
kjpaadsd@126.com
College of Sciences, Hebei United University, Tangshan, 063009, P.R. China

Abstract. To study a path following algorithm for tracing a constraint shifting combined homogony method for convex nonlinear programming. The method is based on the technique of β -cone neighbourhood which ensures the obtained iterative points are in the interior of the constraint shifting feasible set. To establish the global linear convergence of the algorithm the numerical experiments indicate that the algorithm is efficient.

Keywords: Convex Nonlinear Programming, Constraint Shifting Combined Homogony, Path Following Algorithm, Global Convergence.

1 Introduction

In this paper, we consider the following convex nonlinear programming (CNLP) problem

$$\min f(x) \quad \text{s.t.} x \in \Omega, \tag{1}$$

Where $\Omega = \{x \in R^n : g_i(x) \le 0, i = 1,2,\cdots,m\}$ is the feasible set of CNLP and its strictly feasible set is $\Omega^0 = \{x \in R^n : g_i(x) < 0, i = 1,2,\cdots,m\}$. $f(x), g_i(x)$ are all sufficiently smooth convex functions. The recent development of the CNLP can refer to $[1,4,5,6,7,8,10]$. In this paper, we consider to solve CNLP by a homogony method.

$x \in R^n$ Is a solution of the CNLP problem (1) if and only if there exists a $y \in R^m$ satisfying the Karush-Kuhn-Tucher condition, i.e.

$$\nabla f(x) + \nabla g(x) y = 0, \quad Yg(x) = 0, g(x) \le 0, y \ge 0, \tag{2}$$

Where $Y = \mathrm{diag}(y)$

In 2005, Shang and Yu [9] proposed a new homogony method, called constraint shifting combined homogony (CSCH for short) method for convex nonlinear programming. It is "warm started" and does not require the initial point being an interior point of the feasible set. They constructed the homogony equation as follows:

$$H(x,y,\mu) = \begin{pmatrix} (1-\mu)[\nabla f(x) + \nabla g(x) y] + \mu(x - x^0) \\ Y[g(x) - \mu\Theta(g(x^0) + e)] - \mu Y^0[g(x^0) - \Theta(g(x^0) + e)] \end{pmatrix} \tag{3}$$

Y. Yang, M. Ma, and B. Liu (Eds.): ICICA 2013, Part II, CCIS 392, pp. 518–526, 2013.

Where $Y = \mathrm{diag}(y)$, $Y^0 = \mathrm{diag}(y^0)$, $(x^0, y^0) \in R^n \times R^m_{++}$, $e = (1, \cdots, 1)^T \in R^m$, $\Theta = \mathrm{diag}(\theta_1, \cdots, \theta_m)$, that

$$\theta_i = \begin{cases} 0, & g_i(x^0) < 0 \\ 1, & g_i(x^0) \geq 0. \end{cases} \tag{4}$$

Denote the constrained set of CSCH:

$$\Omega(\mu) = \{x \mid g_i(x) - \mu\theta_i(g_i(x^0) + 1) \leq 0, i = 1, \cdots, m\}, \mu \in [0,1] \tag{5}$$

$$\Omega(\mu)^0 = \{x \mid g_i(x) - \mu\theta_i(g_i(x^0) + 1) < 0, i = 1, \cdots, m\}, \mu \in [0,1] \tag{6}$$

$$\partial\Omega(\mu) = \Omega(\mu) \setminus \Omega(\mu)^0, \mu \in [0,1] \tag{7}$$

$$I(x, \mu) = \{i \in \{1, \cdots, m\} \mid g_i(x) - \mu\theta_i(g_i(x^0) + 1) = 0, \mu \in [0,1]\} \tag{8}$$

When $\mu = 0$, constraint shifting combined homogony equation becomes is equivalent to the K-K-T system of CNLP.

In [9], under the following two conjectures, existence and convergence of the homogony path is proved.

Conjecture1. $\Omega(\mu)^0$ Is nonempty

Conjecture2. There exists $\hat{x} \in \Omega(\mu)^0$ such that

$$X = \{x \in \Omega(1) \mid (x - \hat{x})^T \nabla f(x) < 0\} \tag{9}$$

Is a bounded set, that is to say, there exists a constant $M > 0$, such that for all $x \in X$ we have $\|x\| \leq M$.

2 Algorithm

Motivated by some stronger theoretical results obtained by many algorithms, we prefer to view μ as an independent variable rather than a parameter. We construct the function as follows:

$$G(w, \mu) = \begin{pmatrix} H(w, \mu) \\ \mu \end{pmatrix} \tag{10}$$

Note that $G(w, \mu) = 0$ if and only if $w = (x, y)$ solves (1.2).

Now, we give the definition of the β -cone neighborhood. We call the set

$$C = \{(w, \mu) : H(w, \mu) = 0, \mu \in (0,1]\} \tag{11}$$

The central path which is to be followed we next define a β-cone neighborhood around the central path

$$N(\beta,\mu)=\{(w,\mu):\|H(w,\mu)\|\le\beta\mu,\mu\in(0,1]\}, \tag{12}$$

Where $\beta>0$ is called the width of the neighborhood.

Lemma1. If $\beta_0=\min_{i\in 1:m}|\,y_i^0[g_i(x^0)-\theta_i(g_i(x^0)+1)]\,|$ and $\beta\in(0,\beta_0)$, we have

$$N(\beta,\mu)\subset\Omega(\mu)^0\times R_{++}^m\times(0,1]. \tag{13}$$

Proof. By the definition of the β-cone neighborhood, we know when $\|H(w,\mu)\|\le\beta\mu$, the inequality $|(H(w,\mu))_i|\le\beta\mu$ holds, for $i=n+1,n+2,\cdots,n+m$, that is to say,

$$\begin{aligned}&\mu y_i^0[g_i(x^0)-\theta_i(g_i(x^0)+1)]-\beta\mu\le y_i[g_i(x)-\mu\theta_i(g_i(x^0)+1)]\\&\le\mu y_i^0[g_i(x^0)-\theta_i(g_i(x^0)+1)]+\beta\mu.\end{aligned} \tag{14}$$

For $\beta\in(0,\beta_0)$, in terms of $w^0\in R^n\times R_{++}^m$ and $\beta_0\le -y_i^0[g_i(x^0)-\theta_i(g_i(x^0)+1)]$, $i=1,2,\cdots,m$, we have

$$\begin{aligned}&y_i[g_i(x)-\mu\theta_i(g_i(x^0)+1)]\\&<(y_i^0[g_i(x^0)-\theta_i(g_i(x^0)+1)]+\beta_0)\mu\le 0,i=1,2,\cdots,m.\end{aligned} \tag{15}$$

Hence,

$$N(\beta,\mu)\subset\Omega(\mu)^0\times R_{++}^m\times(0,1]. \tag{16}$$

Algorithm (the predictor-corrector algorithm for the solution curve)

Step 0 (Initialization)

Set $k:=0,\mu_0:=1$. Take the width of β-cone neighborhood $\beta\in(0,\beta_0)$, $\alpha\in(0,1),\delta\in(0,1)$.

Step 1 (Termination Criterion)

If $\mu_k=0$, stop, and $w^k:=(x^k,y^k)$ solves (2)

Step 2 (Computation of the Newton Direction)

Compute the solution (w^k,μ_k) of the linear system

$$G(w^k,\mu_k)+\nabla G(w^k,\mu_k)^T\begin{pmatrix}w^k\\ \mu_k\end{pmatrix}=\begin{pmatrix}0\\ (1-\alpha)\mu_k\end{pmatrix} \tag{18}$$

Step 3 (Backtracking Line Search)

Let λ_k be the maximum of the values $1,\delta,\delta^2,\cdots$ such that

$$\begin{aligned}&\|H(w^k+\lambda_k w^k,(1-\alpha\lambda_k)\mu_k)\|\\&\le(1-\alpha\lambda_k)\beta\mu_k.\end{aligned} \tag{19}$$

Set

$$w^{k+1} := w^k + \lambda_k w^k \tag{20}$$

$$\mu_{k+1} := (1-\alpha\lambda_k)\mu_k \tag{21}$$

$$k := k+1 \tag{22}$$

and return to step 1.

Proposition1. If $f, g_i, i = 1,2,\cdots,m$ are all twice continuously differentiable and convex functions, for $w^0 \in R^n \times R^m_{++}$ and $(w,\mu) \in \Omega(\mu)^0 \times R^m_{++} \times (0,1]$, $\nabla G(w,\mu)$ is a nonsingular matrix

Proof By the homogony equation (6), we have

$$\frac{\partial H(w,\mu)}{\partial w} = \begin{pmatrix} Q(w) & (1-\mu)\nabla g(x) \\ Y\nabla g(x)^T & P(x) \end{pmatrix}, \tag{23}$$

Where $Q(w) = \mu I + (1-\mu)(\nabla^2 f(x) + \sum_{i=1}^m y_i \nabla^2 g_i(x))$ and $P(x) = diag(g_i(x) - \mu\theta_i(g_i(x^0)+1))$. Noticing that $P(x)$ is a nonsingular matrix, we have

$$\begin{aligned} |\frac{\partial H(w,\mu)}{\partial w}| &= \begin{vmatrix} \overline{Q}(w) & 0 \\ Y\nabla g(x)^T & P(x) \end{vmatrix} \\ &= |\overline{Q}(w)|\,|P(x)|, \end{aligned} \tag{24}$$

Where $\overline{Q}(w) = Q(w) - (1-\mu)\nabla g(x) P(x)^{-1} Y \nabla g(x)^T$.

Since $\mu \in (0,1]$, $Q(w)$ is a positive definite matrix for all $w \in \Omega(\mu)$ and $g(x) < 0, y > 0$, we have $\overline{Q}(w)$ is a positive definite matrix. Therefore, $|\frac{\partial H(w,\mu)}{\partial w}| \neq 0$, i.e., $\partial H(w,\mu)/\partial w$ is a nonsingular matrix

Based on the above results and by simple computation, we obtain

$$\nabla G(w,\mu) = \begin{pmatrix} Q(w) & (1-\mu)\nabla g(x) & x - x^0 - \nabla f(x) - \nabla g(x) y \\ Y\nabla g(x)^T & P(x) & -Y^0 g(x^0) \\ 0 & 0 & 1 \end{pmatrix}. \tag{26}$$

Hence $|\nabla G(w,\mu)| = \begin{vmatrix} Q(w) & (1-\mu)\nabla g(x) \\ Y\nabla g(x)^T & P(x) \end{vmatrix} = |\frac{\partial H(w,\mu)}{\partial w}| \neq 0$, which shows that $\nabla G(w,\mu)$ is a nonsingular matrix.

3 Global Linear Convergence

Before the global linear convergence analysis, we discuss the property of the mapping $G(w,\mu)$.

Lemma 2. Assume that $G(w,\mu)$ is defined by the equation (8) given the bound convex set $\mathrm{M} \subset R^{n+m}$, for $\forall \omega \in \mathrm{M}$ and $\mu \in (0,1]$, there exists a positive constant $C>0$ such that

$$\left\| \frac{\partial^2 G_i(w,\mu)}{\partial (w,\mu)^2} \right\| \le C, i = 1,2,\cdots,n+m+1. \tag{27}$$

Proof. Since $\dfrac{\partial^2 G_i(w,\mu)}{\partial (w,\mu)^2}$ regards with w,μ, when w belong to a bounded set and $\mu \in (0,1]$, there exists a constant $C>0$ such that

$$\left\| \frac{\partial^2 G_i(w,\mu)}{\partial (w,\mu)^2} \right\| \le C, i = 1,2,\cdots,n+m+1. \tag{28}$$

This completes the proof of the Lemma.

We are now in position to state and prove the global linear convergence result for the algorithm described in the preceding section. Assume that the algorithm does not terminate finitely. In addition, we make the following conjecture on the infinite sequence $\{(w^k,\mu_k)\}$ generated by the algorithm.

Conjecture3. Given $\mu \in (0,1]$ and the point $w \in N(\beta,\mu)$, there exists $M>0$ such that

$$\left\| \nabla G(w,\mu)^{-1} \right\| \le M \tag{30}$$

Suppose that Conjecture 3 holds for the infinite sequence $\{(w^k,\mu_k)\}$ generated by the algorithm. Then

(i) For $k = 0,1,2,\cdots$,

$$(w^k,\mu_k) \in N(\beta,\mu_k), \tag{31}$$

$$(1-\alpha\lambda_{k-1})\cdots(1-\alpha\lambda_0) = \mu_k. \tag{32}$$

(ii) For all $k \ge 0$, $\lambda_k \ge \hat{\lambda} = \delta\bar{\lambda}$, where $\bar{\lambda} = \min\{1, \dfrac{2(1-\alpha)\beta}{C\sqrt{n+m}M^2(\beta+\alpha)^2}\}$.

Therefore, μ_k converges to 0 at a globally linear rate.

(iii) The sequence $\{\omega_k\}$ converges to a solution of (6), *i.e.* $\{x^k\}$ converges to a solution of the CNLP problem.

Proof (i) we can establish (31) and (32) by induction on k .

(ii) Noting (11) and after some computations, we have

$$\begin{pmatrix} \Delta\omega^k \\ \Delta\mu_k \end{pmatrix} = \nabla G(\omega^k, \mu_k)^{-1} \begin{pmatrix} -H(\omega^k, \mu_k) \\ -\alpha\mu_k \end{pmatrix}, \tag{33}$$

Then

$$\left\| \begin{pmatrix} \Delta\omega^k \\ \Delta\mu_k \end{pmatrix} \right\| \le M(\|H(\omega^k, \mu_k)\| + \alpha\mu_k) \le M(\beta + \alpha)\mu_k . \tag{34}$$

On the other hand, recalling the relation between the mappings G and H , we have

$$\begin{aligned}
&\begin{pmatrix} H(\omega^k + \lambda\Delta\omega^k, (1-\alpha\lambda)\mu_k) \\ 0 \end{pmatrix}_i \\
&= \left(G(\omega^k + \lambda\Delta\omega^k, (1-\alpha\lambda)\mu_k) - \begin{pmatrix} 0 \\ (1-\alpha\lambda)\mu_k \end{pmatrix} \right)_i \\
&= G_i(\omega^k, \mu_k) + \lambda \frac{\partial G_i(\omega^k, \mu_k)}{\partial(\omega^k, \mu_k)} \begin{pmatrix} \Delta\omega^k \\ \Delta\mu_k \end{pmatrix} - \begin{pmatrix} 0 \\ (1-\alpha\lambda)\mu_k \end{pmatrix}_i \\
&+ \frac{1}{2}\lambda^2 \begin{pmatrix} \Delta\omega^k \\ \Delta\mu_k \end{pmatrix}^T \frac{\partial^2 G_i(\tilde{\omega}^k, \tilde{\mu}_k)}{\partial(\omega^k, \mu_k)^2} \begin{pmatrix} \Delta\omega^k \\ \Delta\mu_k \end{pmatrix} \\
&= (1-\lambda) G_i(\omega^k, \mu_k) + \lambda \begin{pmatrix} 0 \\ (1-\alpha)\mu_k \end{pmatrix}_i - \begin{pmatrix} 0 \\ (1-\alpha\lambda)\mu_k \end{pmatrix}_i \\
&+ \frac{1}{2}\lambda^2 \begin{pmatrix} \Delta\omega^k \\ \Delta\mu_k \end{pmatrix}^T \frac{\partial^2 G_i(\tilde{\omega}^k, \tilde{\mu}_k)}{\partial(\omega^k, \mu_k)^2} \begin{pmatrix} \Delta\omega^k \\ \Delta\mu_k \end{pmatrix} \\
&= (1-\lambda) \begin{pmatrix} H(\omega^k, \mu_k) \\ 0 \end{pmatrix}_i + \frac{1}{2}\lambda^2 \begin{pmatrix} \Delta\omega^k \\ \Delta\mu_k \end{pmatrix}^T \frac{\partial^2 G_i(\tilde{\omega}^k, \tilde{\mu}_k)}{\partial(\omega^k, \mu_k)^2} \begin{pmatrix} \Delta\omega^k \\ \Delta\mu_k \end{pmatrix}
\end{aligned} \tag{36}$$

Where $(\tilde{\omega}^k, \tilde{\mu}_k) = (\omega^k + \lambda\theta_i\Delta\omega^k, u_k + \lambda\theta_i\Delta u^k), \theta_i \in (0,1)$,And the second equality follows from (5). Let

$$A = \left(\frac{\partial^2 G_1}{\partial(\omega^k, u_k)^2}, \cdots, \frac{\partial^2 G_{n+m}}{\partial(\omega^k, u_k)^2} \right) \tag{37}$$

Then $\left\|A(\tilde{\omega}^k, \tilde{\mu}_k)\right\| \le \sqrt{n+m}C$ From the deduction given above, we have

$$\begin{aligned}
&\left\|H(\omega^k + \lambda\Delta\omega^k, (1-\alpha\lambda)\mu_k)\right\| \\
&\le (1-\lambda)\left\|H(\omega^k, \mu_k)\right\| + \frac{1}{2}\lambda^2\left\|A(\tilde{\omega}^k, \tilde{\mu}_k)\right\|\left\|(\Delta\omega^k, \Delta\mu_k)\right\|^2 \\
&\le (1-\lambda)\beta\mu_k + \frac{\sqrt{n+m}}{2}\lambda^2 CM^2(\beta+\alpha)^2\mu_k{}^2 \\
&\le (1-\lambda)\beta\mu_k + \frac{\sqrt{n+m}}{2}\lambda^2 CM^2(\beta+\alpha)^2\mu_k .
\end{aligned} \tag{38}$$

It is easy to verify that

$$(1-\lambda)\beta\mu_k + \frac{\sqrt{n+m}}{2}\lambda^2 CM^2(\beta+\alpha)^2\mu_k \le (1-\alpha\lambda)\beta\mu_k . \tag{39}$$

Whenever $\lambda \le \dfrac{2(1-\alpha)\beta}{C\sqrt{n+m}M^2(\beta+\alpha)^2}$

Therefore, taking

$$\bar{\lambda} \le \min\left\{1, \frac{2(1-\alpha)\beta}{C\sqrt{n+m}M^2(\beta+\alpha)^2}\right\} \tag{40}$$

We have $\lambda_k \ge \hat{\lambda} = \delta^2\bar{\lambda}$, which combines with (32) implying $u_k \le (1-\alpha\hat{\lambda})^k u_0 = (1-\alpha\hat{\lambda})^k$, for all k sufficient large. Thus $\{u^k\}$ converges globally linearly to zero.

(iii) Let $(\Delta\omega^k, \Delta\mu_k)$ be chosen to satisfy the Newton equation (11), then,

$$\begin{aligned}
&\left\|\begin{pmatrix}\Delta\omega^{k+1}\\ \Delta\mu_{k+1}\end{pmatrix} - \begin{pmatrix}\Delta\omega^{k}\\ \Delta\mu_{k}\end{pmatrix}\right\| = \lambda_k\left\|\begin{pmatrix}\Delta\omega^{k}\\ \Delta\mu_{k}\end{pmatrix}\right\| \le \left\|\begin{pmatrix}\Delta\omega^{k}\\ \Delta\mu_{k}\end{pmatrix}\right\| \\
&\le M(\beta+\alpha)\mu_k \\
&\le M(\beta+\alpha)(1-\alpha\hat{\lambda})^k
\end{aligned} \tag{41}$$

Therefore, $\{\omega^k\}$ is a Cauchy sequence and converges to a point ω^* .It follows from $(\omega^k, \mu_k) \in N(\beta, \mu_k)$ that $\omega^* \in \Omega \times R^m$. Hence, ω^* is a solution of the homogony equation (3) and the point x^* solves the CNLP.

4 Numerical Experiments

We have compared our algorithm with the algorithm of [2] for several examples. The algorithms were implemented in Mat lab. For all test problems, we choose the accuracy parameter $\varepsilon = 10^{-6}$ for termination criterion of the Step 1, $u_0 = 1.0$, $\alpha = 0.7$, $\delta = 0.5$, $\beta = \beta_0$ in our algorithm and in the algorithm of [2], we take the accuracy

parameters $\varepsilon_1 = 10^{-4}$, $\varepsilon_2 = 10^{-3}$, $\varepsilon_3 = 10^{-6}$ and $h_0 = 0.3$. All of the initial points are chosen as the exterior of the feasible region. IT denotes the number of iterations, ACPU denotes the average total cost time for solving the problem among the ten runs and u^* denotes the value when the algorithm terminates.

Example 4.1 $\min f(x), s.t. g_i(x) \le 0, i = 1,2,\cdots,4,$

Where $f(x), g_i(x), i = 1,2,\cdots,4$ are as follows:

$$f(x) = (x_1 - 2)^2 + (x_2 - 4)^2 \tag{42}$$

$$g_1(x) = -x_1 \tag{43}$$

$$g_2(x) = -1 - x_2 \tag{44}$$

$$g_3(x) = x_1 + x_2 - 3 \tag{45}$$

$$g_4(x) = (x_1 - 1)^2 + x_2^2 - 4 \tag{46}$$

In this example, we take the initial exterior point $x_0 = (3,3)^T$. The experiment results see Table 1.

Example 4.2 $\min f(x), s.t. g_i(x) \ge 0, i = 1,2,3,4,$

Where $f(x), g_i(x), i = 1,2,3,4$ are as follows:

$$f(x) = x_1^2 + x_2^2 - 2x_1 + x_1x_2 + 1 \tag{47}$$

$$g_1(x) = 3x_1 + 5x_2 - 4 \tag{48}$$

$$g_2(x) = 3x_1^2 + x_2^2 - 2x_1x_2 - 8x_2 + 10 \tag{49}$$

$$g_3(x) = x_1 \tag{50}$$

$$g_4(x) = x_2 \quad g_4(x) = x_2. \tag{51}$$

In this example, we take the initial exterior point $x_0 = (-1,0)^T$ .The experiment results see Table 2

Table 1. The numerical results of Example 4.1

Method	ACPU	IT	x^*	$f(x^*)$	u^*
New	0.0297	19	$(1.0000, 2.0000)^T$	5.0000	$3.2792*10^{-7}$
Old	0.1484	130	$(1.0000, 2.0000)^T$	5.0000	$4.6585*10^{-7}$

Table 2. The numerical results of Example 4.2

method	ACPU	IT	x^*	$f(x^*)$	u^*
new	0.0937	51	$(1.3333, 2.0000)^T$	0.1111	$6.6567*10^{-7}$
old	0.2139	194	$(1.3333, 2.0000)^T$	0.1111	$3.0931*10^{-7}$

References

1. Chandrasekaran, V., Sanghavi, S., Parrilo, P.A., Willsky, A.S.: Rank-Sparsity Incoherence for Matrix Decomposition. SIAM Journal on Optimization 21(2), 146–151 (2011)
2. Lin, Z.H., Yu, B., Feng, G.C.: A combined homotopy interior point method for convex nonlinear programming. Applied Mathematics and Computation 84, 193–211 (1997)
3. McCormick, G.P.: The projective SUMT method for convex programming. Mathematics of Operations Research 14, 203–223 (1989)
4. Nedic, A.: Asynchronous Broadcast-Based Convex Optimization over a Network. IEEE Transactions on Automatic Control 56(6), 1337–1351 (2011)
5. Nedic, A.: Random Algorithms for Convex Minimization Problems. To appear in Mathematical Programming, Special issue in honor of Paul Tseng 129, 225–253 (2011)
6. Nedic, A., Ozdaglar, A.: Convergence Rate for Consensus with Delays. Journal of Global Optimization 47(3), 437–456 (2010)
7. Nedic, A., Bertsekas, D.P.: The Effect of Deterministic Noise in Subgradient Methods. Mathematical Programming 125(1), 75–99 (2010)
8. Ram, S.S., Nedic, A., Veeravalli, V.V.: Distributed Stochastic Subgradient Projection Algorithms for Convex Optimization. Journal of Optimization Theory and Applications 147(3), 516–545 (2010)
9. Shang, Y.F., Yu, B.: A Boundary Moving Combined Homogony Method for convex nonlinear programming (Chinese). Numerical Mathematics: A Journal of Chinese Universities 27, 311–315 (2005)
10. Srivastava, K., Nedic, A.: Distributed Asynchronous Constrained Stochastic Optimization. IEEE Journal of Selected Topics in Signal Processing 5(4), 772–790 (2011)

Improved Randomized Algorithm for Square Root Extraction

Lihua Liu, Zhengjun Cao, and Qian Sha

Department of Mathematics, Shanghai Maritime University, China
caoszhj@shu.edu.cn

Abstract. Square root extraction plays a key role in cryptosystems based on elliptic curves. Motwani and Raghavan had proposed an algorithm for square root extraction over finite field F_p, where p is an odd prime. It is a randomized algorithm with expected running time $O(\text{len}(p)^4)$. Its complexity relies mainly on the loops calling Euclid's algorithm for polynomials over F_p. In this paper, we propose an improvement of it. The new algorithm calls a subroutine for computing a Legendre symbol. Since the running time of computing a Legendre symbol is much less than that of Euclid's algorithm for polynomials over F_p, the new algorithm is more efficient. It only takes time $O(\text{len}(p)^3)$. We also compare the new algorithm with those algorithms for square root extraction over finite fields.

Keywords: Quadratic Residue, Legendre Symbol, Randomized Algorithm.

1 Introduction

Square root extraction plays a key role in cryptosystems based on elliptic curves. The typical applications of root extraction are point compression in elliptic curves and operation of hashing onto elliptic curves [1, 2].

Consider a degree 2 polynomial $f(X)$ over a finite field F_p, where p is an odd prime. Without loss of generality, we assume that the polynomial is manic, i.e., the leading coefficient is 1. We also assume that the polynomial is not irreducible, which means that it has roots over F_p and can be factored into linear terms as follows: $f(X) = X^2 + \alpha X + \beta = (X - a)(X - b)$ Here $\alpha, \beta \in F_p$ are the coefficients, and $a, b \in F_p$ are the roots of the polynomial. We make the simplifying assumption that the two roots are distinct; otherwise, if a is the only root, it must be the case that $\alpha \equiv -2a \pmod p$ and $\beta \equiv a^2 \pmod p$ [3-5]. These equations can be easily checked and would yield the desired root. Furthermore, we assume that neither root is 0 otherwise, the polynomial is easily factored. Notice that the problem of finding square roots of a quadratic residue β is the special case where the polynomial is $f(X) = X^2 - \beta$.

Y. Yang, M. Ma, and B. Liu (Eds.): ICICA 2013, Part II, CCIS 392, pp. 527–535, 2013.

There are two typical methods to solve the problem; the Tonelli-Shanks method [6, 7] and the Cipolla-Lehmer method [8, 9]. The Tonelli-Shanks algorithm takes time $O(\text{len}(p)^3 + \alpha^2 \text{len}(p)^2)$, where $2^\alpha s = p-1$, s is odd. At worst (if almost all of $p-1$ is a power of 2), this is $O(\text{len}(p)^4)$. It was extended to the general r-th roots extraction by Adleman et al [10-12].

The Cipolla-Lehmer algorithm does not depend on the exponent of 2 in $p-1$. Its complexity can be reduced to $O(\text{len}(p)^3)$. Motwani and Raghavan [13-15] had described an algorithm for square root extraction over finite field F_p. It does not depend on the exponent of 2 in $p-1$, either. It is a randomized algorithm with polynomial expected running time $O(\text{len}(p)^4)$. Its complexity relies mainly on the loops calling Euclid's algorithm for polynomials over F_p.

Our contribution in this paper, we propose an improvement of Motwani-Raghavan algorithm. The new algorithm calls the subroutine for computing a Legendre symbol. Since the running time of computing a Legendre symbol is much less than that of Euclid's algorithm for polynomials over F_p, the new algorithm is more efficient. It only takes time $O(\text{len}(p)^3)$.

2 Preliminary

Let $Z_n = \{0,1,\cdots,n-1\}$ be the set of all numbers smaller than n, $Z_n^* = \{x | 1 \le x \le n \text{and} \gcd(x,n) = 1\}$ be the set of numbers in Z_n that are coprime to n.

Definition 1. A residue $a \in Z_n^*$ is said to be a quadratic residue if there exists some $x \in Z_n^*$ such that $x^2 \equiv a (\text{mod} n)$. If a is not a quadratic residue, then it is referred to as a quadratic non-residue.

Theorem 2 (Euler's Criterion): For prime p, an element $a \in Z_p^*$ is a quadratic residue if and only if $a^{\frac{p-1}{2}} \equiv 1 (\text{mod} p)$.

Definition 3 (Legendre Symbol): For any prime p and $a \in Z_p^*$, we define the Legendre symbol $\left[\frac{a}{p}\right] = 1$, if a is a quadratic residue (mod p), or -1, if a is a quadratic non-residue (mod p)

The following theorem summarizes the essential properties of the Legendre symbol [9].

Theorem 4 let p be an odd prime, and let $a, b \in Z$. Then we have (1) $\left[\frac{a}{p}\right] \equiv a^{\frac{(p-1)}{2}} (\text{mod} p)$; in particular, $\left[\frac{-1}{p}\right] = (-1)^{\frac{(p-1)}{2}}$; (2) $\left[\frac{a}{p}\right]\left[\frac{b}{p}\right] = \left[\frac{ab}{p}\right]$; (3)

$a \equiv b(\mathrm{mod}p)$ implies $\left[\frac{a}{p}\right]=\left[\frac{b}{p}\right]$; (4) $\left[\frac{2}{p}\right]=(-1)^{\frac{(p^2-1)}{8}}$; (5)(Law of quadratic reciprocity) if p is an odd prime, then $\left[\frac{q}{p}\right]\left[\frac{p}{q}\right]=(-1)^{\frac{(p-1)}{2}\cdot\frac{(q-1)}{2}}$.

For an integer a, we define $\mathrm{len}(a)$ to be the number of bits in the binary representation of $|a|$. Given $a \in Z_n$ and a non-negative integer e, the repeated-squaring algorithm computes $a^e(\mathrm{mod}n)$ using just $O(\mathrm{len}(e))$ multiplications in Z_n, thus taking time $O(\mathrm{len}(e)\mathrm{len}(n)^2)$. Therefore, we have the following result.

Proposition 5 [13] for an odd prime p, we can test whether an integer a is a quadratic residue modulo p by either performing the exponentiation $a^{\frac{(p-1)}{2}} \pmod p$ or by computing the Legendre symbol $\left[\frac{a}{p}\right]$. Assume that $0 \le a < p$. Using a standard repeated squaring algorithm, the former method takes time $O(\mathrm{len}(p)^3)$, while using Euclidean-like algorithm, the latter method takes time $O(\mathrm{len}(p)^2)$.

Clearly, to test whether an integer a is a quadratic residue modulo p, the method to compute the Legendre symbol $\left[\frac{a}{p}\right]$ is to be preferred.

Let R be a ring. Let us define the length of a polynomial $f(X) \in R[X]$, denoted $\mathrm{len}(f)$, to be the length of its coefficient vector. Analogous to algorithms for modular integer arithmetic, we can also do arithmetic in the residue class ring $R[X]/(f)$, where $f \in R[X]$ is a polynomial of $\deg(f)>0$ whose leaden coefficientlc (f) is a unit.

Proposition 6 [7]. Let $R[X]/(f)$ be a residue class ring, where $f \in R[X]$ is a polynomial of $\deg(f)>0$ whose leading coefficient lc (f) is a unit. Given $g \in R[X]/(f)$ and a non-negative exponent e, using repeated-squaring algorithm we can compute g^e taking $O(\mathrm{len}(e)\deg(f)^2)$ operations in R.

Remark 7 using a standard representation for F_p, each operation in F_p takes time $O\left(len(p)^2\right)$

3 Algorithm-1

Motwani and Raghavan [10] had proposed an algorithm for computing square roots over finite field F_p. The algorithm is based on the following proposition.

Proposition 8 an element $t \in Z_p^*$ is a quadratic residue modulo an odd prime p if and only if $X-t$ is a factor of the polynomial $X^{\frac{p-1}{2}}-1$.

This proposition follows from Euler's Criterion, since $X-t$ is a factor if and only if t is a root of the polynomial $X^{\frac{p-1}{2}}-1$.

Suppose that the roots a and b of $f(X)$ are such that $\left[\frac{a}{p}\right]=1$ and $\left[\frac{a}{p}\right]=1$ We have $(X-a)\left|X^{\frac{p-1}{2}}-1\ (X-b)\right.\dagger X^{\frac{p-1}{2}}-1$. Thus $\gcd\left(f(X), X^{\frac{p-1}{2}}-1\right)=(X-a)$ since a given polynomial $f(X)$ should have roots with same Legendre symbols, we can not directly apply the above method to factor it. Consider r chosen uniformly at random from F_p. Define the polynomial $f_r(X)=f(X-r)=(X-a-r)(X-b-r)$

Clearly, the roots of $f_r(X)$ are $a+r$ and $b+r$, which are both uniformly distributed over F_p. Moreover, by the following lemma it is easy to find the two roots are random and independent.

Denote the event that $\left[\frac{a}{p}\right]\left[\frac{b}{p}\right]=-1$ by $E_p(a,b)$, where $a,b\in F_p^*$, we have:

Lemma 9 [10]. Let $a,b\in F_p^*$ and $a\neq b$. For r chosen uniformly at random from F_p^*, the random variables $A=a+r \pmod p$ and $B=b+r \pmod p$ satisfy $\Pr\left[E_P(A,B)\right]=\frac{1}{2}-O\left(\frac{1}{P}\right)$

Input: Odd prime p and a non-irreducible, monic, square-free, degree 2 polynomial $f(x)=X^2+\alpha X+\beta(\mathrm{mod}p)$

Output: The roots a and b of $f(X)$ over F_p

1-1 choose r uniformly at random from F_p^*

1-2 compute $g(X)=X^2+\alpha' X+\beta'$ such that $g(X)=f(X-r)$, as follows $\alpha'\leftarrow\alpha-2r(\mathrm{mod}p)$ $\beta'\leftarrow\beta-\alpha r+r^2(\mathrm{mod}p)$.

1-3 if $\beta'=0$ then return $a=-r$ and $b=-r-\alpha'$

1-4 compute $h(X)=\gcd(g(X), X^{\frac{p-1}{2}}-1)$ using Euclid's algorithm (for polynomials)

1-5 if $h(X)=g(X)$ or $h(X)=1$ then go to Step 1

1-6 let $h(X)=X-c$ and compute $A\leftarrow c, B\leftarrow-\alpha'-A$

1-7 return $a=A-r$ and $b=B-r$

By the above Lemma, Algorithm-1 succeeds in each loop with probability at least 1/2. In each loop, the cost is mainly dominated by that of the Euclid's algorithm for polynomials over F_p. By Proposition 6 and Remark 7, we know each loop takes time $O\left(len(p)^3\right)$. If we do this for more than $O(1)len(p)$ different randomly chosen r, then with probability $>1-\left(\frac{1}{p}\right)^{O(1)}$ at least one of them will give a root of $f(X)$. Thus, the expected time of this algorithm is $O\left(len(p)^4\right)$

4 Improvement of Algorithm-1

By Table 1, it is easy to find that the gcd trick in each loop is to be used to check that $\left[\frac{a+r}{p}\right]\left[\frac{b+r}{p}\right]=-1$ but this condition can be more easily verified. Based on this observation, we now present an improvement of Algorithm-1.

Denote the event that $\left[\frac{s}{p}\right]=-1$ by $\zeta_p(s)$, where $s\in F_p^*$, we have:

Lemma 10. Suppose that $f(X)=X^2+\alpha X+\beta=(X-a)(X-b)$. Here $\alpha,\beta\in F_p$ are the coefficients, and $\alpha,\beta\in F_p$ are the roots of the polynomial. Define the polynomial $f_r(X)=f(X-r)=(X-a-r)(X-b-r)$ where r chosen uniformly at random from F_p^*. The random variable $\beta'=\beta-\alpha r+r^2 \pmod p$ satisfies $\Pr\left[\zeta_p(\beta')\right]=\frac{1}{2}-O\left(\frac{1}{P}\right)$

Proof Since $f(X)=X^2+\alpha X+\beta=(X-a)(X-b)$ and $f(X)=f(X-r)=(X-a-r)(X-b-r)=X^2+(\alpha-2r)X+(\beta-\alpha r+r^2)$ we have $(a+r)(b+r)=\beta-\alpha r+r^2 \pmod p$. Thus $\left[\frac{a+r}{p}\right]\left[\frac{b+r}{p}\right]=-1\Leftrightarrow\left[\frac{\beta-\alpha r+r^2}{p}\right]=-1$, $Ep(a+r,b+r)\Leftrightarrow\zeta_p(\beta-\alpha r+r^2)$

By Lemma 9, we have $\Pr\left[\zeta_p(\beta')\right]=\frac{1}{2}-O\left(\frac{1}{P}\right)$

By Lemma 10, the gcd trick in each loop in the original algorithm can be reasonably replaced by the trick to check that $\left[\frac{\beta-\alpha r+r^2}{p}\right]=-1$

Table 2: Algorithm-2

Input: Odd prime p and a non-irreducible, monic, square-free, degree 2 polynomial $f(x)=X^2+\alpha X+\beta(\mathrm{mod}p)$

Output: The roots a and b of $f(X)$ over F_p.

2-1 chooses r uniformly at random from F_p^*

2-2 compute $g(X)=X^2+\alpha' X+\beta'$ such that $g(X)=f(X-r)$, as follows $\alpha'\leftarrow\alpha-2r(\mathrm{mod}p)$, $\beta'\leftarrow\beta-\alpha r+r^2(\mathrm{mod}p)$

2-3 if $\beta'=0$ then return $a=-r$ and $b=-r-\alpha'$

2-4 compute $[\frac{\beta'}{p}]$ using Euclidean-like algorithm if $[\frac{\beta'}{p}]=1$ then go to Step 1

2-5 compute $h(X)=\gcd(g(X),X^{\frac{p-1}{2}}-1)$ using Euclid's algorithm (for polynomials)

2-6 let $h(X) = X - c$ and compute $A \leftarrow c, B \leftarrow -\alpha' - A$

2-7 return $a = A - r$ and $b = B - r$.

Apparently, the improved algorithm succeeds in each loop with probability at least 1/2. In each loop, the cost is mainly dominated by that of computing the Legendre Symbol $\left[\frac{\beta - \alpha r + r^2}{p}\right]$. By Proposition 5, we know each loop takes time $O\left(len(p)^2\right)$. Thus, the improvement only takes expected time $O\left(len(p)^3\right)$.

Table 1. Comparison between Algorithm-1 and Algorithm-2

	Algorithm-1	Algorithm-2
computation	$\gcd(f(X-r), X^{\frac{p-1}{2}} - 1)$	$\left[\frac{\beta - \alpha r + r^2}{p}\right]$
time	$O(\text{len}(p)^3)$	$O(\text{len}(p)^2)$
total time	$O(\text{len}(p)^4)$	$O(\text{len}(p)^3)$

It should be observed that the original algorithm can be directly used over any finite field F_p, while the improvement cannot be used in such case. We now make a comparison between the new algorithm and the original as follows.

5 Comparisons between Algorithms for Square Root Extraction

5.1 Tonelli-Shanks Algorithm

The Tonelli-Shanks algorithm is based on the following facts. Write $p-1$ in the form $2^t \cdot s$, where s is odd. Given a quadratic residue β and a quadratic nonresidue r, we have $\left(\beta^s\right)^{2^{t-1}} \equiv 1 \pmod{p}$, $\left(r^s\right)^{2^{t-1}} \equiv -1 \pmod{p}$. If $t = 1$, then $\left(\beta^s\right)^{2^{t-1}} \equiv 1 \pmod{p}$. Since $(2, s) = 1$, one can obtain the least nonnegative integer t' such that $s \mid 2t' - 1$. Hence, we have $\left(\beta^s\right)^{2^{t'-1}} \equiv 1 \pmod{p}$. Therefore, $\left(\beta^s\right)^2 \equiv \beta \pmod{p}$ It means that $\beta^{t'}$ is a square root of β with respect to the modular p. In this case, it only takes time $O\left(len(t')len(p)^2\right)$.

If $t \geq 2$, then $\left(\beta^s\right)^{2^{t-2}} \pmod{p} \in \{1, -1\}$. Take $k_1 = 0$ or 1 such that $\left(\beta^s\right)^{2^{t-2}} \left(r^s\right)^{2^{t-1} \cdot k_1} \equiv 1 \pmod{p}$ Since $\left(\beta^s\right)^{2^{t-3}} \left(r^s\right)^{2^{t-1} \cdot k_1} \pmod{p} \in \{1, -1\}$, take Take $k_2 = 0$ or 1 such that $\left(\beta^s\right)^{2^{t-3}} \left(r^s\right)^{2^{t-2} \cdot k_1} \left(r^s\right)^{2^{t-1} \cdot k_2} \equiv 1 \pmod{p}$ Likewise, we obtain

$k_3,\ldots,k_{t-1}\in\{0,1\}$ such that $\left(\beta^s\right)\left(r^s\right)^{2k_1+2^2k_2+\cdots+2^{t-1}k_{t-1}}\equiv 1 \pmod p$)Thus, we have $\left(\beta^{\frac{s+1}{2}}\right)^2\left(\left(r^s\right)^{k_1+2k_2+\cdots+2^{t-2}k_{t-1}}\right)^2\equiv\beta \pmod p$)

Input: Odd prime p and $X^2\equiv\beta(\mathrm{mod}p)$.

Output: A square root of β .

Step 1: choose r uniformly at random from F_p^*, compute $[\frac{r}{p}]$ using Euclidean-like algorithm.

Step 2: if $[\frac{r}{p}]=1$ then go to Step 1.

Step 3: compute t,s such that $p-1=2^t s$, where s is odd, and compute $a\leftarrow r^s, b\leftarrow\beta^s,\ \ h\leftarrow 1$

Step 4: for $i=1$ to $t-1$ compute $d=b^{2^{t-i-1}}$ if $d\neq 1$, then $h\leftarrow h\cdot a$, $b\leftarrow b\cdot a^2$ $a\leftarrow a^2$ end for

Step 5: return $\beta^{\frac{s+1}{2}}\cdot h$

We now discuss the time estimate for this algorithm. To find a quadratic no residue r, it takes expected time $O\left(len(p)^3\right)$. To compute $b^{2^{t-i-1}} \pmod p$, it takes time $O\left((t-i-1)len(p)^2\right)$. Since $1+2+\cdots+(t+1)=\frac{t(t-1)}{2}$, the loop takes time $O\left(t^2len(p)^2\right)$. Thus, the final estimate is $O\left(len(p)^3+t^2len(p)^2\right)$. At worst (if almost all of $p-1$ is a power of 2), this is $O\left(len(p)^4\right)$. Notice that the procedure to find a quadratic no residue r can be done previously. In such case, Tonelli-Shanks algorithm takes time $O\left(t^2len(p)^2\right)$.

5.2 Cipolla-Lehmer Algorithm

Given a quadratic residue β with respect to the modular p, where p is an odd prime. Choose r uniformly at random from F_p^* until you find r such that $r^2-\beta$ is a quadratic no residue. Let α denote the element $\sqrt{r^2-\beta}$ in the quadratic extension F_{p^2}. Then compute $h=(r+\alpha)^{\frac{p+1}{2}}$. Since $\alpha^p=\left(\alpha^2\right)^{\frac{p-1}{2}}\alpha=\left(r^2-\beta\right)^{\frac{p-1}{2}}\alpha=-\alpha$ we have $h^2=(r+\alpha)^{p+1}=\left(r+\alpha^p\right)(r+\alpha)=(r-\alpha)(r+\alpha)=\beta$.

Table 2. Cipolla-Lehmer algorithm

Input: Odd prime p and $X^2 \equiv \beta(\mathrm{mod}p)$. Output: A square root of β.
Step 1: choose r uniformly at random from F_p^*, and compute $[\frac{r^2-\beta}{p}]$ using Euclidean-like algorithm
Step 2: if $[\frac{r^2-\beta}{p}]=1$ then go to Step 1
Step 3: compute $f(x) \leftarrow x^2 - 2rx + \beta$
Step 4: Return $x^{\frac{p+1}{2}}(\mathrm{mod}f(x))$

To find r such that $\left[\frac{r^2-\beta}{p}\right]=-1$, it takes expected time $O\left(len(p)^3\right)$ to compute $x^{\frac{p+1}{2}} \pmod{f(x)}$, it takes time $O\left(len(p)^3\right)$ Thus, the algorithm takes expected time $O\left(len(p)^3\right)$.

6 Further Discussions

We now make a comparison between those algorithms for square root extraction. See the table 6. Notice that the trick of computing a Legendre symbol can not be generalized to any finite field F_q.

The Tonelli-Shanks algorithm and Cipolla-Lehmer algorithm have to check $r^{\frac{p-1}{2}} \equiv -1$ or $\left(r^2-\beta\right)^{\frac{p-1}{2}} \equiv -1$ over F_p. This takes time $O\left(len(p)^3\right)$. Since we do this for more than $O(1)len(p)$ different randomly chosen r, the expected time for them in such case is $O\left(len(p)\right)^4$ so does the Algorithm-1. There is no any apparent difference between these algorithms.

By the way, the author [14] had recently presented a novel idea to compute square roots over finite fields, without being given any quadratic no residue.

Table 3. Comparison

	Tonelli-Shanks	Cipolla-Lehmer	Algorithm-2
computation	$\left[\frac{r}{p}\right], b^{2^{r-i-1}}$	$\left[\frac{r^2-\beta}{p}\right]$	$\left[\frac{\beta-\alpha r+r^2}{p}\right]$
root	$\beta^{\frac{q+1}{2}} \cdot h$	$x^{\frac{p+1}{2}}(\mathrm{mod}f(x))$	constant term of $\gcd\left(g(X), X^{\frac{p-1}{2}}-1\right)$
total time	$O(\mathrm{len}(p)^3 + t^2\mathrm{len}(p)^2)$	$O(\mathrm{len}(p)^3)$	$O(\mathrm{len}(p)^3)$

7 Conclusions

In this paper, we propose an improvement of one algorithm for square root extraction over finite field F_p. The original algorithm takes time $O\left(len(p)^4\right)$, while the new algorithm only takes time $O\left(len(p)^3\right)$. We also compare those algorithms for square root extraction over finite fields.

Acknowledgements. This work is supported by the National Natural Science Foundation of China (Project 11171205), and the Key Disciplines of Shanghai Municipality (S30104).

References

1. Adleman, L., Manders, K., Miller, G.: On taking roots in finite fields. In: Proc. 18th IEEE Symposium on Foundations on Computer Science (FOCS), vol. 13, pp. 75–177 (1977)
2. Boneh, D., Boyen, X., Shacham, H.: Short Group Signatures. In: Franklin, M. (ed.) CRYPTO 2004. LNCS, vol. 3152, pp. 41–55. Springer, Heidelberg (2004)
3. Boneh, D., Franklin, M.: Identity-based Encryption from the Weil Pairing. SIAM J. Computing 32(3), 586–615 (2003)
4. Barreto, P., Voloch, J.: Efficient Computation of Roots in Finite Fields. Designs, Codes and Cryptography 39, 275–280 (2006)
5. Cao, Z., Fan, X.: Extension of Barreto-Voloch Root Extraction Method. In: Qing, S., Susilo, W., Wang, G., Liu, D. (eds.) ICICS 2011. LNCS, vol. 7043, pp. 184–189. Springer, Heidelberg (2011)
6. Cipolla, M.: Un metodo per la risolutione della congruenza di secondo grado, Rendiconto dell'Accademia Scienze Fisiche e Matematiche, Napoli, Ser 3, pp. 154–163 (1903)
7. Gathen, J., Gerhard, J.: Modern computer algebra, 2nd edn., vol. 16, pp. 224–228. Cambridge University Press (2003)
8. Lehmer, D.: Computer technology applied to the theory of numbers. Studies in Number Theory, vol. 15, pp. 117–151. Pretice-Hall, Englewood Cliffs (1969)
9. Lidl, R., Niederreiter, H.: Introduction to finite fields and their applications, pp. 16–19. Cambridge University Press (1986)
10. Motwani, R., Raghavan, P.: Randomized Algorithms, vol. 11, pp. 26–31. Cambridge University Press (1995)
11. Shanks, D.: Five number-theoretic algorithms. In: Proc. 2nd Manitoba Conf. Numer. Math., Manitoba, Canada, vol. 11, pp. 51–70 (1972)
12. Smart, N.: An Identity Based Authenticated Key Agreement Protocol Based on the Weil Pairing. Electronics Letters 38, 630–632 (2002)
13. Shoup, V.: A computational introduction to number theory and algebra., vol. 18, pp. 147–151. Cambridge University Press (2005)
14. Sze, T.: On taking square roots without quadratic no residues over finite fields. Mathematics of Computation 80, 1797–1811 (2011)
15. Tonelli, A.: Bemerkung über die Auflosung quadratischer Congruenzen. Göttinger Nachrichten 18, 344–346 (1891)

Solutions of Ill-Posed Linear Equations

Yamian Peng, Jincai Chang, and Yan Yan

College of Science, Hebei United University Tangshan 063009 Hebei P.R. China
pymz806@163.com

Abstract. Linear system of equations is been used more and more widely in social life. Most people use the estimated value for a variety of computing that will cause a lot of errors. Familiar with a variety of ill-posed linear equations solution can make us grasp the algorithm and make the error reduce to the minimum in practice, thereby increasing the accuracy to reduce unnecessary trouble. The $Ax = b$ calculation solution equivalent to solve the $(A + E)x = b$ perturbation equations of floating point error analysis results shown. We choose algorithm to make the $\| E \|$ as small as possible. In order to simplify the calculation, the perturbation matrix generally desirable as the simplest rank one type, this paper discusses the problem and gives a feasible algorithm.

Keywords: ill-posed problems, Ill-posed Linear Equations, Successive Adjustment Elimination Method.

1 Introduction

In practical application, we need to carry out a variety of measurement, a measurement will have error, for linear equations, even if the error is very small, the calculated solution would be a lot worse real solutions, the error equation (i.e. disturbed equation) is called the ill-posed linear equations, the disturbance after equations known as the ill-posed linear equations [1].

In the equation $AX = b$, if the matrix A and (or) the right side of the small changes caused by b , equation $AX = b$ solution of tremendous change, then the equation group for" ill-posed" equations [2]. Matrix A (relative to the solution of equations or inverse.) known as the" sick" matrix, otherwise known equations as" normal" equations, A is known as the" normal" matrix. On the" small change" and" great changes" are relative terms, there is no quantitative criteria; "the matrix" concept. (1) Must be clearly "sick" is the nature of the matrix itself characteristic, and the computational tools and the calculation method of independent. But in practical computation" ill-posed" degree is by using a computational tool and method of calculation show. (2) Must be clearly "sick" is relative to what. For the solution of equations (or inverse) is ill-posed matrix (the solution of equations in terms of), for finding the eigenvalues is not necessarily pathological, and vice versa [3].

Y. Yang, M. Ma, and B. Liu (Eds.): ICICA 2013, Part II, CCIS 392, pp. 536–545, 2013.

The conventional methods, such as the selection of the main element method for ill-conditioned problems almost incapable of action. Equations of the morbid problems seriously affect the accuracy of the results and reliability; people have been exploring the effective solution [4]. Now many methods, such as iterative correction method, projection method, recursive method, ordinary differential equation method and so on, from the algorithm is simple, effective on look, a lot of not ideal, such as some of the effectiveness of the algorithm is better but more complicated, some algorithms are actually only pathological degree not too serious, smaller the problem, as the scale of the problem increases, morbid degree increased, the precision of the solutions will be very obvious or even decreased dramatically [5].

Simple iterative method is adopted for all values of a value to calculate, by repeated calculation can gradually closer to the true solution, has high precision, calculation speed in general, but for ill-conditioned linear equations, a simple iterative method and error, also can cause very big error.

2 Successive Elimination of Adjustment

Similar to the vector norm definition, can define $n \times n$ order matrix of the norm.

Definition: Set in $\|\cdot\|$ phalanx of variable n order real value function, and meet the conditions as follows:

$\|A\| \geq 0$, and $\|A\| = 0$ when, and only when $A = 0$;

For any real number α, all have $\|\alpha A\| = |\alpha| \|A\|$;

For any two n orders square A, B there'll be $\|A + B\| \leq \|A\| + \|B\|$;

$\|AB\| \leq \|A\| \|B\|$ (Compatibility condition);

It says $\|A\|$ matrix of the norm for A.

Establish: $A = (a_{ij})$ is n order square formation and common matrix norm have:

(1) 1- norm $\|A\|_1 = \max\limits_{1 \leq j \leq n} \sum\limits_{i=1}^{n} |a_{ij}|$;

(2) ∞ - norm $\|A\|_\infty = \max\limits_{1 \leq i \leq n} \sum\limits_{j=1}^{n} |a_{ij}|$;

(3) 2-norm $\|A\|_2 = \sqrt{\lambda_{\max}(A^T A)}$, among them $\lambda_{\max}(A^T A)$ is the maximum eigenvalue of the $A^T A$.

Example 1: With $A = \begin{bmatrix} 1 & -3 \\ -1 & 2 \end{bmatrix}$ matrix, and the norm $\|A\|_p$. $p = 1, 2, \infty$.

Solution: $\|A\|_\infty = \max\{1 + |-3|, |-1| + 2\} = 4$, $\|A\|_1 = \max\{1 + |-1|, |-3| + 2\} = 5$, $A^T A = \begin{bmatrix} 1 & -1 \\ -3 & 2 \end{bmatrix} \begin{bmatrix} 1 & -3 \\ -1 & 2 \end{bmatrix} = \begin{bmatrix} 2 & -5 \\ -5 & 13 \end{bmatrix}$. Characteristic equations are

$\left|\lambda I - A^T A\right| = \begin{vmatrix} \lambda - 2 & 5 \\ 5 & \lambda - 13 \end{vmatrix} = 0$. Characteristic root is $\lambda_1 = \frac{15 + \sqrt{221}}{2}$, $\lambda_2 = \frac{15 - \sqrt{221}}{2}$. So $\|A\|_2 = \sqrt{\frac{15 + \sqrt{221}}{2}} \approx 3.864$.

Theorem set y is equations of the solution as follows:

$$(A + \sum_{i=1}^{k} bt^{(i)\perp}) y = b \tag{1}$$

$$AX = b \tag{2}$$

Solution can be expressed as

$$X = \frac{1}{1 - \sum_{i=1}^{k} y^{\perp} t^{(i)}} y \tag{3}$$

$$u^{(0)} V^{(0)\perp} = (b^{(0)} + \lambda_0 e_1) \cdot \frac{1}{\lambda_0} (e_1^{\perp} + \frac{\lambda_0 - b_1^{(0)}}{b_{i_0}^{(0)}} e_{i_0}^{\perp})$$

$$= \frac{1}{\lambda_0} (\begin{bmatrix} b_1^{(0)} \\ b_2^{(0)} \\ \vdots \\ b_{i_0}^{(0)} \\ \vdots \\ b_n^{(0)} \end{bmatrix} + \lambda_0 \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ \vdots \\ 0 \end{bmatrix}) \cdot ([1 \quad 0 \quad \cdots \quad 0 \quad \cdots \quad 0] + \frac{\lambda_0 - b_1^{(0)}}{b_{i_0}^{(0)}} [0 \quad 0 \quad \cdots \quad 1 \quad \cdots \quad 0])$$

$$= \frac{1}{\lambda_0} \begin{bmatrix} b_1^{(0)} + \lambda_0 \\ b_2^{(0)} \\ \vdots \\ b_{i_0}^{(0)} \\ \vdots \\ b_n^{(0)} \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & \cdots & \frac{\lambda_0 - b_1^{(0)}}{b_{i_0}^{(0)}} & \cdots & 0 \end{bmatrix}$$

$$= \begin{bmatrix} 1+\frac{b_1^{(0)}}{\lambda_0} & 0 & \cdots & 0 & \frac{(b_1^{(0)}+\lambda_0)(\lambda_0-b_1^{(0)})}{\lambda_0 b_{i_0}^{(0)}} & 0 & \cdots & 0 \\ \frac{b_2^{0}}{\lambda_0} & 0 & \cdots & 0 & \frac{b_2^{(0)}(\lambda_0-b_1^{(0)})}{\lambda_0 b_{i_0}^{(0)}} & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{b_{i_0-1}^{0}}{\lambda_0} & 0 & \cdots & 0 & \frac{b_{i_0-1}^{(0)}(\lambda_0-b_1^{(0)})}{\lambda_0 b_{i_0}^{(0)}} & 0 & \cdots & 0 \\ \frac{b_{i_0}^{0}}{\lambda_0} & 0 & \cdots & 0 & 1-\frac{b_1^{(0)}}{\lambda_0} & 0 & \cdots & 0 \\ \frac{b_{i_0+1}^{0}}{\lambda_0} & 0 & \cdots & 0 & \frac{b_{i_0+1}^{(0)}(\lambda_0-b_1^{(0)})}{\lambda_0 b_{i_0}^{(0)}} & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{b_n^{0}}{\lambda_0} & 0 & \cdots & 0 & \frac{b_n^{(0)}(\lambda_0-b_1^{(0)})}{\lambda_0 b_{i_0}^{(0)}} & 0 & \cdots & 0 \end{bmatrix}$$

Direct will (1.3) and (1.2) the very easy to validate this theorem. Successive elimination adjustment introduces below process of calculation. In order to narrative convenient, remember $A^0 = A$, $b^0 = b$, get $A^0 X^0 = b^0$, Among them A^0 is singular matrix. Set $b^0 \neq 0$, and $b^0 \neq a_0 e_1$ (a_0 is a nonzero constant), If you can find a certain i_0, making them $| b_{i_0}^{(0)} | = \| b^{(0)} \|_\infty = \max | b_i^{(0)} | (1 \le i \le n)$. Can set

$\lambda_0 = sign(b_i^{(0)}) \| b^{(0)} \|_2$ ($sign(x) = \begin{cases} 1 & x > 0 \\ 0 & x = 0 \\ -1 & x < 0 \end{cases}$).

$u^{(0)} = b^{(0)} + \lambda_0 e_1$, $v^{(0)} = \frac{1}{\lambda_0}(e_1 + \frac{\lambda_0 - b_1^{(0)}}{b_{i_0}^{(0)}} e_{i_0})$. Design matrix $M^0 = I - u^{(0)} V^{(0)\perp}$

$$M^0=\begin{bmatrix} -\frac{b_1^{(0)}}{\lambda_0} & & & -\frac{(b_1^{(0)}+\lambda_0)(\lambda_0-b_1^{(0)})}{b_{i_0}\lambda_0} & \\ -\frac{b_2^{(0)}}{\lambda_0} & 1 & & -\frac{b_2^{(0)}(\lambda_0-b_1^{(0)})}{b_{i_0}\lambda_0} & \\ \vdots & & \ddots & \vdots & \\ -\frac{b_{i_0-1}^{(0)}}{\lambda_0} & & 1 & -\frac{b_{i_0-1}^{(0)}(\lambda_0-b_1^{(0)})}{b_{i_0}\lambda_0} & \\ -\frac{b_{i_0}^{(0)}}{\lambda_0} & & & \frac{b_1^{(0)}}{\lambda_0} & \\ -\frac{b_{i_0+1}^{(0)}}{\lambda_0} & & & -\frac{b_{i_0+1}^{(0)}(\lambda_0-b_1^{(0)})}{b_{i_0}\lambda_0} & \ddots \\ \vdots & & & \vdots & 1 \\ -\frac{b_n^{(0)}}{\lambda_0} & & & -\frac{b_n^{(0)}(\lambda_0-b_1^{(0)})}{b_{i_0}\lambda_0} & 1 \end{bmatrix}$$

Can get $M^0 b^{(0)}=(-\lambda_0,0,\cdots,0)^{\perp}$, and M^0 has the following shape

$$\begin{bmatrix} -\frac{b_1^0}{\lambda 0} & & & * & \\ * & 1 & & * & \\ \vdots & & \ddots & \vdots & \\ * & & 1 & * & \\ * & & & \frac{b_1^0}{\lambda_0} & \\ \vdots & & & \vdots & \ddots \\ * & & & * & 1 \\ * & & & * & 1 \end{bmatrix} \tag{4}$$

Type on both sides by the singular left matrix M^0, and keep $\overline{A^0}=M^0A^0$, $C^{(0)}=M^0b^{(0)}$, can get $\overline{A^0}X^{(0)}=C^{(0)}$.

Select adjust vector is $t^{(0i)}=-\frac{\overline{a_{1i}^{(0)}}}{c_1^{(0)}}e_i$, $i=1,2,...,n-1$

$$t^{(0t)}=sign(\overline{a_{1n}^{(0)}})e_n \tag{5}$$

Will (8) type deformation solution

$$(\overline{A^0} + \sum_{i=1}^{n} C^{(0)} t^{(0i)\perp}) y^{(0)} = C^{(0)} \tag{6}$$

Written for

$$B^0 = \overline{A^0} + \sum_{i=1}^{n} C^{(0)} t^{(0i)\perp} \tag{7}$$

B^0 has a shape $\begin{bmatrix} 0 & \cdots & b_{1n}^{(0)} \\ \cdots & \cdots & \cdots \\ B^0 & \cdots & b^{(0)} \end{bmatrix}$ from $B^0 y^{(0)} = C^{(0)}$ get $y_n^{(0)} = -\dfrac{\left\| b^{(0)} \right\|_2}{b_{1n}^{(0)}}$, $y_1^{(0)}, \ldots, y_{n-1}^{(0)}$ can be solved from a low of the equation.

Remember $A_1 X^{(1)} = b^{(1)}$, including $A_1 = \overline{B^0}$, $b^{(1)} = \dfrac{\left\| b^{(0)} \right\|_2}{b_{1n}^{(0)}} \overline{b^{(0)}}$, then and repeatable (4) to (7) discussion. The general discussion, the various angles in standard 0 to k can. So after n-1 the elimination process, can will (4) into a order equations. Back to the generation process is very easy to give:

Hypothesis have already worked out $x^{(k+1)}$, then make $y_i^{(k)} = x_i^{(k+1)}$ $(i = 1, 2, \cdots, n-k-1)$ and $y_{n-k}^{(k)}$ in the elimination process that is already find out, so they get vector $y^{(k)}$, further can be calculated $x^{(k)}$ by

$$x^{(k)} = \frac{1}{1 - \sum_{i=1}^{n-k} t^{(ki)\perp} y^{(k)}} y^{(k)}, k = n-1, n-2, ..., 0 \tag{8}$$

3 Short Solutions of Linear Equations

Examples 2: The set $A : R^n \to R^n$ is a linear operator, namely $Ax = b$ is a linear equations, we know that when condition number $Cond(A) >> 1$, the solution of linear equations is discontinuous to rely on data b ,namely this kind problem is not stable ill-posed problem. Study the following 5 order Pascal matrix of linear equations

$$\begin{bmatrix} 0.5 & 0.5 & 0.5 & 0.5 & 0.5 \\ 0.5 & 1 & 1.5 & 2 & 2.5 \\ 0.5 & 1.5 & 3 & 5 & 7.5 \\ 0.5 & 2 & 5 & 10 & 17.5 \\ 0.5 & 2.5 & 7.5 & 17.5 & 35 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \\ 6.5 \\ 12 \\ 20 \end{bmatrix} \tag{9}$$

The solution is $x = [0, 1, 0, 1, 0]^T$. If the item has a right equation tiny disturbance to

$$\begin{bmatrix} 0.5 & 0.5 & 0.5 & 0.5 & 0.5 \\ 0.5 & 1 & 1.5 & 2 & 2.5 \\ 0.5 & 1.5 & 3 & 5 & 7.5 \\ 0.5 & 2 & 5 & 10 & 17.5 \\ 0.5 & 2.5 & 7.5 & 17.5 & 35 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 1 \\ 2.93 \\ 6.53 \\ 12 \\ 20.03 \end{bmatrix} \tag{10}$$

Use simple iterative method to solve the equations of the disturbance after:

$$A = \begin{bmatrix} 0.5 & 0.5 & 0.5 & 0.5 & 0.5 \\ 0.5 & 1 & 1.5 & 2 & 2.5 \\ 0.5 & 1.5 & 3 & 5 & 7.5 \\ 0.5 & 2 & 5 & 10 & 17.5 \\ 0.5 & 2.5 & 7.5 & 17.5 & 35 \end{bmatrix},$$

$$b = \begin{bmatrix} 1 \\ 2.93 \\ 6.53 \\ 12 \\ 20.03 \end{bmatrix},$$

$$D^{-1} = \begin{bmatrix} 2 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{3} & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{10} & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{35} \end{bmatrix}$$

$$B = I - D^{-1}A = \begin{bmatrix} 0 & -1 & -1 & -1 & -1 \\ -0.5 & 0 & -1.5 & -2 & -2.5 \\ -\frac{1}{6} & -0.5 & 0 & -\frac{5}{3} & -2.5 \\ -0.05 & -0.2 & -0.5 & 0 & -1.75 \\ -\frac{1}{70} & -\frac{1}{14} & -\frac{3}{14} & -0.5 & 0 \end{bmatrix},$$

$$g = D^{-1}b = \begin{bmatrix} 2 \\ 2.93 \\ 2.1767 \\ 1.2 \\ 0.5723 \end{bmatrix} \tag{11}$$

Set $x^{(0)} = (0,0,0)^T$, $x^{(1)} = Bx^{(0)} + g = (2, 2.93, 2.1767, 1.2, 0.5723)^T$.

$$x^{(2)} = Bx^{(1)} + g = \begin{bmatrix} 0 & -1 & -1 & -1 & -1 \\ -0.5 & 0 & -1.5 & -2 & -2.5 \\ -\frac{1}{6} & -0.5 & 0 & -\frac{5}{3} & -2.5 \\ -0.05 & -0.2 & -0.5 & 0 & -1.75 \\ -\frac{1}{70} & -\frac{1}{14} & -\frac{3}{14} & -0.5 & 0 \end{bmatrix} \begin{bmatrix} 2 \\ 2.93 \\ 2.1767 \\ 1.2 \\ 0.5723 \end{bmatrix} + \begin{bmatrix} 2 \\ 2.93 \\ 2.1767 \\ 1.2 \\ 0.5723 \end{bmatrix} \tag{12}$$

$$= \begin{bmatrix} -4.879 \\ -5.1658 \\ -3.0524 \\ -1.57588 \\ -0.73199 \end{bmatrix}$$

So go on, Such as Table 1.

Table 1. The iterative results

k	$x_1^{(k)}$	$x_2^{(k)}$	$x_3^{(k)}$	$x_4^{(k)}$	$x_5^{(k)}$
1	2	2.93	2.1767	1.2	0.5723
2	-4.879	-5.1658	-3.0524	-1.57588	-0.73199
3	12.52606	14.92981	10.02921	5.2842915	2.453006
4	-30.6963	-35.0779	-22.3155	-11.7196	-5.46432
5	76.57747	88.85159	58.02531	30.47076	14.15811
6	-189.505	-218.733	-141.192	-74.1885	-34.5375
7	470.6516	544.1919	353.1195	185.4586	86.25305
8	-1167.02	-1348.62	-873.093	-458.674	-213.420
9	2895.811	3346.978	2169.001	1139.308	530.0025

Can be seen from the table with simple iterative method with the calculation result of true solution more and more far, obviously, for sick linear equation to calculate with simple iteration method is feasible.

Examples 3: Solving $\begin{pmatrix} 0.78 & 0.563 \\ 0.913 & 0.659 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}=\begin{pmatrix} 0.149 \\ 0.819 \end{pmatrix}$ the exact solution of the equation is $X=(-362906,502783)^T$, the inverse of the coefficient matrix is $A^{-1}=\begin{pmatrix} 659000 & -563000 \\ -913000 & 780000 \end{pmatrix}$.

Visible, the equations is sick, but because the exact solution of the order of magnitude of quite big, can use successive adjustment elimination method for solving.

Solution: set $A^0=A=\begin{bmatrix} 0.78 & 0.563 \\ 0.913 & 0.659 \end{bmatrix}$, $b^0=b=\begin{bmatrix} 0.149 \\ 0.819 \end{bmatrix}$,

$$v^0=\frac{1}{\lambda_0}(e_1+\frac{\lambda_0-b_1^{(0)}}{b_{i_0}^{(0)}}e_{i_0})=\frac{1}{0.692962}(\begin{bmatrix}1\\0\end{bmatrix}+\frac{0.692962-0.149}{0.819})\begin{bmatrix}0\\1\end{bmatrix}\approx\begin{bmatrix}1.443080573\\0.958462752\end{bmatrix}$$

$$\overline{A^0}=M^0A^0=\begin{bmatrix}-0.215019005 & -0.806989215\\ -1.181882989 & 0.215019006\end{bmatrix}\begin{bmatrix}0.78 & 0.563\\ 0.913 & 0.659\end{bmatrix}\approx\begin{bmatrix}-0.904495977 & -0.6528616\\ -0.725556379 & -0.523702598\end{bmatrix}$$

$$C^{(0)}=M^0b^{(0)}=\begin{bmatrix}-\lambda_0\\0\end{bmatrix}=\begin{bmatrix}-0.692962\\0\end{bmatrix},\ \overline{A^0}X^{(0)}=C^{(0)}.$$

$$t^{(01)}=-\frac{\overline{a_{11}^{(0)}}}{c_1^{(0)}}e_1=\begin{bmatrix}-1.305260573\\0\end{bmatrix},\ t^{(02)}=sign(\overline{a_{12}^{(0)}})e_2=\begin{bmatrix}0\\-1\end{bmatrix}.$$

$$1-\sum_{i=1}^{n}t^{(0i)\perp}y^{(0)}=1-\left[-1.305260573,-1\right]\begin{bmatrix}12.47309639\\-17.2806755\end{bmatrix}=-0.00003456\cdot$$

Table 2. The use different algorithms

Algorithm	x_1	x_2	Relative error
Complete primary elimination method	-363002, 1953	502916, 2721	2.65/10000
Orthogonal method	-362804, 4617	502640, 6658	2.81/10000
Adjust the successive elimination	-362804, 3172	502725, 3261	1.15/100000
Exact solutions	-3629069	502783	

So according to calculation method calculation can get close to continue in the exact solution of the results.

In addition the use different algorithms are given respectively for the calculation of the corresponding solution and the relative error. Results such as Table 2.

From the example can be seen, the adjustment with successive elimination method is used to solve the income and exact solutions of the solution of the deviation from the few, the relative error is small, it is a kind of ideal solution.

4 Summary

The use of simple iterative method to prove the pathological linear equations tiny disturbance caused the calculation result and the real value is far, thus trying to explain that the find a good solution of linear equations of the necessity of pathological methods. Then use the successive elimination method to adjust the ill-posed linear equations are solved, with an example proves the accuracy of the algorithm.

Acknowledgement. This paper is supported by the Natural Scientific Fund Project in Hebei Province (A2011209019) & (A2012209043). Hebei United University science study imbrues item (z201019).

References

1. Liu, J.: Ill-posed problems of regularization and application, vol. 4(3), pp. 12–15. Science Press, Beijing (2005)
2. Xiao, T., Yu, S., Wang, Y.: Numerical solution of inverse problems, vol. 4(1), pp. 3–9. Science Press, Beijing (2003)
3. Burden, R.L.: Numerical Analysis, 7th edn., vol. 20(4), pp. 21–26. Higher education press (2001)
4. Bjorck, A.: Least Squares Methods. In: Ciarlet, P., Lionseds, J.L. (eds.) Handbook of Numerical Analysis, Vol 1, vol. 8(4), pp. 45–49. North Holland Amsterdam (1990)
5. Ting, L., Cheng, J.: Numerical methods, vol. 3(5), pp. 19–23. Beijing University of Technology Press, Beijing (2005)

Soution of Convection-Diffusion Equations

Yamian Peng, Chunfeng Liu, and Linan Shi

College of Science, Hebei United University Tangshan 063009 Hebei P.R. China
py7m806@163.com

Abstract. Partial differential equations are an important part of mathematics in science and its numerical solution occupies an important position in the numerical analysis. Partial differential equations are closely related to human life and it has important research value. At present, people have studied its solutions in depths and achieved a lot of valuable results. The current solution is the finite element method and finite different method. The convection-diffusion equation is more closely related to human activities, especially complex physical processes. The behavior of many parameters in flow phenomena follows the convection-diffusion equation, such as momentum and heat. The convection-diffusion equation is also used to describe the diffusion process in environmental science, such as the pollutant transport in the atmosphere, oceans, lakes, rivers or groundwater. The research of the convection-diffusion equation is of great importance. Partial differential equation theory has important applications in the solution of the convection-diffusion equation. This chapter mainly talks about the application of the finite difference method in the solution of the convection-diffusion equation.

Keywords: Partial Differential Equations, Differential Format, Convection-Diffusion Equation, Finite Element Method.

1 Introduction

At present, the numerical solution of partial differential equations is mainly two categories: the finite difference method and the finite element method. The advantage of the finite element method is that the region boundaries are more flexible and the results are more accurate. But it also has disadvantages [1-3]. For example, it often requires us to solve the large banded sparse matrix and its computation and storage volume requirements are more difficult to achieve [4-7]. It is more difficult to achieve the implicit scheme and the workload in the preparation of the computer programs is larger [8]. As a result, these disadvantages hindered the further development and application of the finite element method [9, 10].

The finite difference method, as a traditional numerical method to solve partial differential equations, has achieved great success in the recent years [11, 12]. Researchers have achieved much good research result. The main idea of the finite difference method is to use a linear combination of discrete function value to substitute the derivative in order to achieve related difference format in differential equations [13-16]. By solving differential equations, we achieve the approximation of the solution of the differential equations. Our main purpose is to reduce the error and improve the accuracy by improving differential format [17, 18].

Y. Yang, M. Ma, and B. Liu (Eds.): ICICA 2013, Part II, CCIS 392, pp. 546–555, 2013.

2 The Finite Differential Method

Here is the initial value problem for hyperbolic equations and parabolic equations. Its solving region is $D_1 = \{(x,t) \mid -\infty < x < \infty, t \geq 0\}$.

We can draw two clusters of parallel lines in the upper half $x-t$ plane and have the upper half plane into a rectangular grid [19-22]. These lines can be called the grid lines and these intersections can be called the grid points or nodes. The lines paralleling the t axis are equidistant. We set the distance with $\Delta x > 0$ or h which is called space step. The lines paralleling x axis are not equidistant according to concrete problems. For simplicity, we assume that they are also equidistant. We set the distance with $\Delta t > 0$ sometimes τ . They are called time step. The two cluster grid lines can be written as such: $x = x_i = j\Delta x = jh, j = 0, \pm 1, \pm 2...$; $t = t_n = n\Delta t = n\tau, n = 0, 1, 2, \ldots$.

Mesh nodes (x_j, t_n) are sometimes abbreviated as (j,n).

Here is the initial boundary value problem for hyperbolic equations and parabolic equations. We assume the solving region is $D_1 = \{(x,t) \mid 0 < x < l, t \geq 0\}$.

The regional grid is constituted by straight lines paralleling t axis and x axis. They are $x = x_j, j = 0, 1, ..., J$, $t = t_n, n = 0, 1, 2, \ldots$.

In it, $x_i = j\Delta x = jh, \Delta x = h = \frac{l}{J}; t_n = n\Delta t = w\tau$.

Example3: Here is boundary value problem of elliptic equations. The solving region is a bounded domain D on the plane $x - y$ .Its boundary Γ is piecewise smooth curve. Take the steps which are along the x axis and the y axis and make two clusters of straight lines paralleling them.

$x = x_i = i\Delta x, i = 0, \pm 1, \pm 2, \ldots$, $y = y_i = j\Delta y, j = 0, \pm 1, \pm 2, \ldots$.

If the distance of two nodes along the x or y axis is only one step, the two nodes can be called two adjacent nodes. If a node's all four adjacent nodes belong to $D \cup \Gamma$, this node can be called an internal node. If a node's four adjacent does not belong to $D \cup \Gamma$, this node can be called boundary node.

A variety of finite difference methods for solving partial differential equations, using the series expansion method is the most commonly used method.

From the initial value of the convection equation:

$$\frac{\partial u}{\partial t} + a\frac{\partial u}{\partial x} = 0, x \in R, t > 0 \tag{1}$$

$$u(x,0) = g(x), x \in R \tag{2}$$

And the diffusion equation initial value problem.

$$\frac{\partial u}{\partial t} = a\frac{\partial^2 u}{\partial x^2}, x \in R, t > 0 \tag{3}$$

$$u(x,0)=g(x), x\in R \tag{4}$$

For discussion, assume that the $u(x,t)$ initial value problem of partial differential equations, the solution *Taylor* is sufficiently smooth progression commence there.

$$\begin{cases} \dfrac{u(x_j,t_{n+1})-u(x_j,t_n)}{\tau}=[\dfrac{\partial u}{\partial t}]_j^n+O(\tau) \\ \dfrac{u(x_j,t_{n+1})-u(x_j,t_{n-1})}{2\tau}=[\dfrac{\partial u}{\partial t}]_j^n+O(\tau^2) \\ \dfrac{u(x_{j+1},t_n)-u(x_j,t_n)}{h}=[\dfrac{\partial u}{\partial x}]_j^n+O(h) \\ \dfrac{u(x_j,t_n)-u(x_{j-1},t_n)}{h}=[\dfrac{\partial u}{\partial x}]_j^n+O(h) \\ \dfrac{u(x_{j+1},t_n)-u(x_{j-1},t_n)}{2h}=[\dfrac{\partial u}{\partial x}]_j^n+O(h^2) \end{cases} \tag{5}$$

$$\frac{u(x_{j+1},t_n)-2u(x_j,t_n)+u(x_{j-1},t_n)}{h^2}=[\frac{\partial^2 u}{\partial x^2}]_j^n+O(h^2) \tag{6}$$

Use type 1 and type 3 in (1.5):

$$\frac{u(x_j.t_{n+1})-u(x_j.t_n)}{\tau}+a\frac{u(x_{j+1}.t_n)-u(x_j.t_n)}{h}=[\frac{\partial u}{\partial t}+a\frac{\partial u}{\partial x}]_j^n+o(\tau+h)$$

If $u(x,t)$ you meet the smooth solution of partial differential equations (1.1).

$[\dfrac{\partial u}{\partial t}+a\dfrac{\partial u}{\partial x}]_j^n=0$ This can be seen, the partial differential equation (x_j,t_n) can be approximated in the Department with the following equation instead.

$$\frac{u_j^{n+1}-u_j^n}{\tau}+a\frac{u_{j+1}^n-u_j^n}{h}=0\ ,\ j=0,\pm1,\pm2,...,n=0,1,2,... \tag{7}$$

Where u_j^n is an $u(x_j,t_n)$ approximation? Finite difference equation (1.7) called the approximation of differential equations or difference equations can be rewritten into the form of easy calculation (1): $u_j^{n+1}=u_j^n-a\lambda(u_{j+1}^n-u_j^n)$.

Among them, $\lambda=\dfrac{\tau}{h}$ is known as grid.

Differential equation (1.7) coupled with the discrete form of the initial conditions (1.2).

$$u_j^0=\varphi_j\ ,\ j=0,\pm1,... \tag{8}$$

Time layer can advance, calculate the value of the layers, the differential equation (1.7) and (1.8) together constitute a differential format, advancing the first time layer to layer the first time, the formula (1.7) provides a direct calculation u_j^{n+1} of the expression of point by point, saying (1.7) the explicit form. The (5) and the type one and four style, you can get another differential equation (1.1) approximation of differential equations.

$$\frac{u_j^{n+1}+u_j^n}{\tau}+a\frac{u_j^n-u_{j-1}^n}{h}=0 \tag{9}$$

The first type of the Central and the fifth type (1.5), you can get another differential equation (1.1) approximation of differential equations.

$$\frac{u_j^{n+1}-u_j^n}{\tau}+a\frac{u_{j+1}^n-u_{j-1}^n}{2h}=0 \tag{10}$$

$$u_j^{n+1}=u_j^n-\frac{a\lambda}{2}(u_{j+1}^n-u_{j-1}^n) \tag{11}$$

Equation (1.10) is called the central difference scheme, (1.7) and (1.9) called eccentric differential format. Diffusion equation differential format using the same method can be constructed approximation.

$$\frac{u_j^{n+1}-u_j^n}{\tau}+a\frac{u_{j+1}^n-2u_j^n+u_{j-1}^n}{h^2}=0\ ,\ j=0,\pm1,\pm2,...,n=0,1,2,... \tag{12}$$

Consider diffusion equation (3) by integrating this equation, first of all selected points region, and located in the plane *x-t*, the integral region:

$D=\{(x,t)\mid x_j-\frac{h}{2}\le x\le x_j+\frac{h}{2},t_n\le t\le t_{n+1}\}$, internalizes $\iint_D\frac{\partial u}{\partial t}dxdt=\iint_D a\frac{\partial^2 u}{\partial t^2}dxdt$

Direct quartered can be

$$\int_{x_j-\frac{h}{2}}^{x_j+\frac{h}{2}}[u(t_n+\tau,x)-u(t_n,x)]dx=\int_{t_n}^{t_{n+1}}[\frac{\partial u}{\partial x}(t,x_j+\frac{h}{2})-\frac{\partial u}{\partial x}(t,x_j-\frac{h}{2})]dt$$

Numerical integration

$$[u(t_n+\tau,x_j)-u(t_n,x_j)]h\approx a[\frac{\partial u}{\partial x}(t_n,x_j+\frac{h}{2})-\frac{\partial u}{\partial x}(t_n,x_j-\frac{h}{2})]\tau \tag{13}$$

The resulting $\frac{u_j^{n+1}-u_j^n}{\tau}+a\frac{u_{j+1}^n-2u_j^n+u_{j-1}^n}{h^2}=0$. That is (1.11), the integral method is also called the finite volume method.

The previous structure of differential format u_j^{n+1} are explicit, in each time level t_{n+1} can be independently worth the time layer, but $\frac{u(x_i,t_n)-u(x_i,t_{n-1})}{\tau}=[\frac{\partial u}{\partial t}]_j^n+o(\tau)$ not always the case, if adopted and style (1.6), you can get another of the diffusion equation differential format (1.3).

$$\frac{u_j^n-u_j^n}{\tau}-a\frac{u_{j+1}^n-2u_j^n+u_{j-1}^n}{h^2}=0 \tag{14}$$

The finite difference scheme contains more than one node in the new time level; this finite difference scheme called the implicit scheme, most of the implicit scheme is suitable for solving initial boundary value problem of differential equations or to satisfy the cycle conditions of initial value problem.

$\frac{\partial u}{\partial t}=a\frac{\partial^2 u}{\partial x^2},0<x<t,t>0$, $u(x,0)=g(x),0<x<t$, $u(0,t)=u(l,t)=0,t>0$, in it $a>0$.

Diffusion equation (14) approximation with differential format, the initial conditions (12) with discrete, the discrete boundary conditions $u_0^n=0,n>0$.

$$u_j^n=0,n>0 \tag{15}$$

In it $J=\frac{l}{h}$.

Order $U^n=(u_1^n,u_2^n,...,u_{j-1}^n)^T$ r shall be written as such:

$$AU^n=U^{n-1} \tag{16}$$

In it $A=\begin{bmatrix}1+2a\lambda & -a\lambda & & & \\ -a\lambda & 1+2a\lambda & -a\lambda & & \\ & \ddots & \ddots & \ddots & \\ & & -a\lambda & 1+2a\lambda & -a\lambda \\ & & & -a\lambda & 1+2a\lambda\end{bmatrix}$, A is strictly diagonally dominant.

As a result, (16) Solvability A is a tri-diagonal matrix, can be used to catch up method. It can be seen from above, using an explicit format and effort solving, implicit scheme for solving does not seem to benefit, but the future will see the implicit scheme can be a large time step, so there is a significant benefit.

3 Solutions of Convection Diffusion Equations

For its simple structure, it is most likely think of is a direct discrete time derivative forward difference quotient, the spatial derivatives using the central difference quotient to approximate the differential equations (1.17), the following differential format:

$$\frac{u_j^{n+1}-u_j^n}{\tau}+a\frac{u_{j+1}^n-u_{j-1}^n}{2h}=\varepsilon\frac{u_{j+1}^n-2u_j^n+u_{j-1}^n}{h^2} \tag{18}$$

This is the convection diffusion equation (1.17) center Explicit Difference Scheme. (1.18) can be rewritten as such:

$$u_j^{n+1}=u_j^n-\frac{1}{2}\lambda(u_{j+1}^n-u_{j-1}^n)+\mu(u_{j+1}^n-2u_j^n+u_{j-1}^n) \tag{19}$$

The center of the convection-diffusion equation (1.18) explicitly format the impact on the diffusion effect, reducing the diffusion effect, we can establish the explicit form of the correction center.

Full and smooth solution set $u(x,t)$ to the convection-diffusion equation (1.18), the following equation can be:

$$\frac{\partial^2 u}{\partial t^2}=\varepsilon^2\frac{\partial^4 u}{\partial x^4}-2\varepsilon a\frac{\partial^3 u}{\partial x^3}+a^2\frac{\partial^2 u}{\partial x^2} \tag{20}$$

$$\frac{\partial^3 u}{\partial t^3}=\varepsilon^3\frac{\partial^6 u}{\partial x^6}-3\varepsilon^2 a\frac{\partial^5 u}{\partial x^5}+3\varepsilon a^2\frac{\partial^4 u}{\partial x^4}-a^3\frac{\partial^3 u}{\partial x^3} \tag{21}$$

Use the series *Taylor* to expand the type (1.18) and combination of (1.20) and (1.21):

$$\begin{aligned}
&\frac{u(x_j,t_{n+1})}{\tau}+a\frac{u(x_{j+1},t_n)-u(x_{j-1},t_n)}{2h}-\frac{u(x_{j+1},t_n)-2u(x_j,t_n)+u(x_{j-1},t_n)}{h^2}\\
&=\frac{\partial u}{\partial t}+\frac{\tau}{2}\frac{\partial^2 u}{\partial t^2}+\frac{\tau^2}{6}\frac{\partial^3 u}{\partial t^3}+o(\tau^3)+a\frac{\partial u}{\partial t}+\frac{1}{6}ah^2\frac{\partial^3 u}{\partial t^3}+o(h^4)\\
&-\varepsilon\frac{\partial^2 u}{\partial x^2}-\frac{1}{12}\varepsilon h^2\frac{\partial^4 u}{\partial t^4}+o(h^4)\\
&=\frac{\partial u}{\partial t}+a\frac{\partial u}{\partial t}-\varepsilon\frac{\partial^2 u}{\partial x^2}+\frac{\tau}{2}[\varepsilon^2\frac{\partial^4 u}{\partial x^4}-2\varepsilon a\frac{\partial^3 u}{\partial x^3}+a^2\frac{\partial^2 u}{\partial x^2}]+\\
&\frac{\tau^2}{6}[\varepsilon^3\frac{\partial^6 u}{\partial x^6}-3\varepsilon^2 a\frac{\partial^5 u}{\partial x^5}+3\varepsilon a^2\frac{\partial^4 u}{\partial x^4}-a^3\frac{\partial^6 u}{\partial x^3}]+\frac{a}{6}h^2\frac{\partial^3 u}{\partial x^3}-\frac{\varepsilon}{12}h^2\frac{\partial^4 u}{\partial x^4}+\\
&o(\tau^3+h^4)=\frac{\partial u}{\partial t}+a\frac{\partial u}{\partial x}-(\varepsilon-\frac{a^2}{2}\tau)\frac{\partial^2 u}{\partial x^2}+\frac{1}{6}\varepsilon h(1-6\lambda-\lambda^2)\frac{\partial^3 u}{\partial t^3}+\ldots
\end{aligned} \tag{22}$$

Upwind Difference Scheme to better reflect the case of convection-dominated, but this format is only first order accuracy, in order to better solve the various characteristics of the convection-diffusion equation and fully reflect the differential equations, we need to construct more accurate differential format index hybrid finite difference scheme.

For general convection-diffusion equation to calculate the analytical solution is impossible, in the entire solution region, so we can turn to consider the Can the analytical solution in the local area to solve the answer is yes. The basic idea of the hybrid finite analytic method is: first on the unit of local subdivision boundary conditions to obtain the analytical solution of this unit, and secondly, the use of this analytical deconstruction to create a finite difference scheme. Using the finite difference scheme for the hybrid finite analytic method called hybrid finite difference scheme.

First, the spatial variables x and time variables t equidistant mesh and analyzed in any one subdivision unit $(x,t)\in[x_{j-1},x_{j+1}]\times[t_{n-1},t_{n+1}]$. Equations (1.17) in this unit with freeze coefficient method $\frac{\partial u}{\partial t}$, and even if $\frac{\partial u}{\partial t}=k$, as a constant, this time into a second order k constant coefficient ordinary differential equation (1.17):

$$K+a\frac{du}{dx}=\varepsilon\frac{d^2u}{dx^2} \tag{23}$$

The analytical solution of the derived type (1.23) on $[x_{j-1},x_{j+1}]$ as follows:

$$u=c_1e^{\lambda x}+c_2-\frac{k}{a}x \tag{24}$$

In it, $\lambda=\frac{a}{\varepsilon}$. c_1,c_2 are the undetermined coefficients.

Differential equations satisfy the boundary conditions

$$\begin{cases} u(x_{j-1})=c_1e^{\lambda(j-1)}h+c_2-\frac{k}{a}(j-1)h \\ u(x_{j+1})=c_1e^{\lambda(j+1)}h+c_2-\frac{k}{a}(j+1)h \end{cases} \tag{25}$$

Solve it and achieve that:

$$\begin{cases} c_1=\dfrac{u(x_{j+1})-u(x_{j-1})+2h\dfrac{k}{a}}{e^{\lambda(j-1)h}(e^{2\lambda h}-1)} \\ c_2=u(x_{j-1})-\dfrac{u(x_{j+1})-u(x_{j-1})+2h\dfrac{k}{a}}{e^{2\lambda h}-1}+\dfrac{k}{a}(j+1)h \end{cases} \tag{26}$$

Sub-statute c_1, c_2 into (24):

$$u(x_j) = \frac{u(x_{j+1}) + e^{\lambda h} u(x_{j-1})}{1 + e^{\lambda h}} - k \frac{h}{a} \cdot \frac{e^{\lambda h} - 1}{e^{\lambda h} + 1} \tag{27}$$

De formats it and achieves it:

$$k = \frac{a\{(1 + e^{\lambda h}) u(x_j) - [u(x_{j+1}) + e^{\lambda h} u(x_{j-1})]\}}{h(1 - e^{\lambda h})} \tag{28}$$

Continue to deformity it and achieve it:

$$k = -a \frac{u(x_{j+1}) - u(x_{j-1})}{2h} - \frac{a\lambda}{2} \frac{1 + e^{\lambda h}}{1 - e^{\lambda h}} \frac{u(x_{j+1}) - 2u(x_j) + u(x_{j-1})}{h^2} \tag{29}$$

Time level $(n+1)$ analysis (24), if it is the highest possible accuracy, the difference quotient instead of k using the center, we can construct the following differential format:

$$\frac{u_j^{n+1} - u_j^{n-1}}{2\tau} + a \frac{u_{j+1}^{n+1} - u_{j-1}^{n+1}}{2h} = -\frac{ah}{2} \frac{1 + e^{\lambda h}}{1 - e^{\lambda h}} \frac{u_{j+1}^{n+1} - 2u_j^{n+1} + u_{j-1}^{n+1}}{h^2} \tag{30}$$

Format (1.30) combined with the initial conditions can only get on the even-numbered time value, the value can not be the odd time; this must be considered separately on the odd layer format.

At that time $n = 1$, (1.30) in the following format instead of

$$\frac{u_j^{n+1} - u_j^{n}}{2\tau} + a \frac{u_{j+1}^{n+1} - u_{j-1}^{n+1}}{2h} = -\frac{ah}{2} \frac{1 + e^{\lambda h}}{1 - e^{\lambda h}} \frac{u_{j+1}^{n+1} - 2u_j^{n+1} + u_{j-1}^{n+1}}{h^2} \tag{31}$$

And combine the value of (1.30) and (1.31) all nodes in the entire grid can be obtained.

Building process can be seen from the format, the time derivative using the central difference quotient in the space on the local unit to solve the second order constant coefficient differential equations, boundary conditions, approximate analytical solution. This method combines the finite difference method and analysis, to be called the index hybrid finite analytic method, the format can be called the index hybrid finite analytic format (1.31).

We can verify that the truncation error $o(\tau^2 + h^2)$ in this format, and is absolutely stable.

4 Summary

Convection diffusion equations have been used in several of fields in social life and have important applications. Its solution is not only the practical application's needs, but also an important element in the theory of academic study. As a result, the

research of the convection diffusion equations has great value. For this reason, in recent decades, the convection-diffusion equation theory has gained importance and rapid development and its solution is also changing every day. This paper first introduces the basic knowledge and solutions of the partial differential equations, and then use these to solve some simple problems, and in the end sublimate the above methods and then to solve some complex problems and made a number of high-precision solutions in order to make the result better and more reasonable.

Acknowledgement. This paper is supported by the Natural Scientific Fund Project in Hebei Province (A2011209019) & (A2012209043). Hebei United University science study imbrues item (z201019).

References

1. Liu, J.: Ill-posed problems of regularization and application, vol. 3(2), pp. 12–19. Science Press, Beijing (2005)
2. Xiao, T., Yu, S., Wang, Y.: Numerical solution of inverse problems, vol. 4(5), pp. 78–80. Science Press, Beijing (2003)
3. Burden, R.L.: Numerical Analysis, 7th edn., vol. 8(3), pp. 11–19. Higher education press (2001)
4. Bjorck, A.: Least Squares Methods. In: Ciarlet, P.G., Lions, J.L. (eds.) Handbook of Numerical Analysis, vol. 1(3), pp. 90–99 (1990)
5. Ting, L., Cheng, J.: Numerical methods, vol. 7(11), pp. 110–117. Beijing University of Technology Press, Beijing (2005)
6. Song, X.W.: Order elliptic partial differential operators from over specified boundary data, in "Improperly posed problems". Pitman san Francisco 7(2), 23–31 (1977)
7. Bands, H.T., Kunisch, K.: An approximation theory for nonlinear differential equation with application to identification and control. SIAM, J. Control Optimal. 20(3), 815–849 (1982)
8. Haberman, R.: Applied Partial Differential Equations, vol. 3(6), pp. 10–20. Mechanical industry press (2007)
9. Qiu, J.-M., Shu, C.-W.: Convergence of High Order Finite Volume Weighted Essentially Nonoscillatory Scheme and Discontinuous Galerkin Method for Nonconvex Conservation Laws. SIAM Journal on Scientific Computing 31(1), 584–607 (2008)
10. Persson, P.-O., Peraire, J.: Newton-GMRES Preconditioning for Discontinuous Galerkin Discretizations of the Navier–Stokes Equations. SIAM Journal on Scientific Computing 30 30(6), 2709–2733 (2008)
11. Warburton, T., Hagstrom, T.: Taming the CFL Number for Discontinuous Galerkin Methods on Structured Meshes. SIAM Journal on Numerical Analysis 46(6), 3151–3180 (2008)
12. Pietro, D.A.D., Ern, A., Guermond, J.-L.: Discontinuous Galerkin Methods for Anisotropic Semidefinite Diffusion with Advection. SIAM Journal on Numerical Analysis 46(2), 805–831 (2008)
13. Vergara, C., Zunino, P.: Multiscale Boundary Conditions for Drug Release from Cardiovascular Stents. Multiscale Modeling & Simulation 7(2), 565–588 (2008)
14. Peraire, J., Persson, P.-O.: The Compact Discontinuous Galerkin (CDG) Method for Elliptic Problems. SIAM Journal on Scientific Computing 30(4), 1806–1824 (2008)

15. Girault, V., Sun, S., Wheeler, M.F., Yotov, I.: Coupling Discontinuous Galerkin and Mixed Finite Element Discretizations using Mortar Finite Elements. SIAM Journal on Numerical Analysis 46(2), 949–979 (2008)
16. Ern, A., Guermond, J.-L.: Discontinuous Galerkin Methods for Friedrichs' Systems. Part III. Multifield Theories with Partial Coercivity. SIAM Journal on Numerical Analysis 46(2), 776–804 (2008)
17. Wang, W., Li, X., Shu, C.-W.: The Discontinuous Galerkin Method for the Multiscale Modeling of Dynamics of Crystalline Solids. Multiscale Modeling & Simulation 7(1), 294–320 (2008)
18. Xu, Y., Shu, C.-W.: A Local Discontinuous Galerkin Method for the Camassa–Holm Equation. SIAM Journal on Numerical Analysis 46(4), 1998–2021 (2008)
19. Ern, A., Guermond, J.L.: Discontinuous Galerkin Methods for Friedrichs' Systems. Part II. Second order Elliptic PDEs. SIAM Journal on Numerical Analysis 44(6), 2363–2388 (2006)
20. Kamga, J.B.A., Després, B.: CFL Condition and Boundary Conditions for DGM Approximation of Convection Diffusion. SIAM Journal on Numerical Analysis 44(6), 2245–2269 (2006)
21. Ye, X.: A Discontinuous Finite Volume Method for the Stokes Problems. SIAM Journal on Numerical Analysis 44(1), 183–198 (2006)
22. Buffa, A., Perugia, I.: Discontinuous Galerkin Approximation of the Maxwell Eigenproblem. SIAM Journal on Numerical Analysis 44(5), 2198–2226 (2006)

Product Tolerance Engineering Based on the Mathematical Statistics

Yucang Chao and Weitan Chao

School of Machinery and Electronics Engineering,
Changzhou College of Information Technology Changzhou 213164, Jiangsu Pro China
yucang6303@sina.com

Abstract. Flow process of material production often has a problem that the formulation of the range of products tolerance is either too wide or too narrow, which is because the integrated process is not fully considered. With mathematical statistical method, we can do statistical work about the standard deviation of the products when product is in the stable state. Then choose appropriate comprehensive process capability coefficient, and estimate the actual size distribution of the products. Based on this, we can make products tolerance and solve the problem that makes the range of products tolerance too wide or too narrow. So the products not only can meet customer requirements but also bring its potential to the full. Improve the product quality, and also improve the economic benefit of enterprise.

Keywords: Flow Process of Material Production, Mathematical Statistics. Process Capability Coefficient, Product Tolerance.

1 Introduction

Product tolerance is one of the qualities of the product characteristics, reflects the class of the product and credibility, it is a commitment which the enterprise gives to users. Meanwhile, as far as the enterprise internal, it has a directly affection of the enterprise quality cost and economic benefits [5]. In a flow process of material production, the formulation of the range of products tolerance will too wide or too narrow. The method which makes products tolerance too narrow is based on the previous experience of production, [8, 9] from the view of improving product grade and quality. Then through a stage of the practice, they will find it making for excessive cost which influences the enterprise's economic efficiency. And through the efforts, not only can not ensure the product reach qualified standard, but also has badly hurt worker's working passion. So they relax the product tolerance range through review the product specification again. After a stage, they find the product quality has obviously declined because of the product tolerance id too wide. It makes the product sales and market coverage fell, and the staffs tend to relax. Such repeated instability makes a bad credit in the market.

Y. Yang, M. Ma, and B. Liu (Eds.): ICICA 2013, Part II, CCIS 392, pp. 556–564, 2013.

By this token, it is inevitable that the product tolerance is too wide or too narrow if estimated just by the subjective inference or experience, and which method can accurately establish a reasonable product tolerance? In control of product quality, process capability coefficient is a commonly used method of mathematical statistics. The principle is: when the process is in a stable state, through statistics of sampling product samples, calculate the coefficient and the failure rate of its process capability. If the rate of qualified products of a device can not meet expectations, you must choose a more [3, 6] sophisticated device to product [2-3]. However, when the production is a production of process materials, such as glass fiber mats, paper production; when we have no choice on production equipment, whether we can take advantage of its principles to think about it in turn? That is to say, make statistics of the standard deviation of the equipment produced at the steady state, and then choose the integrated process capability coefficient according to the pass rate of corporate quality objectives to estimate the product tolerance range.

This problem involves in four aspects, which are process capability coefficient and evaluation of process capability; standard deviation σ in production process; tolerance range calculation under the certain [4] process capability coefficient Cp and the relationship between process capability and fraction defective [12].

2 Process Capability Coefficient and Process Capability Evaluation

From Table 1, we can have an intuitional understanding of the process capability coefficient and process capability evaluation.

When the production process is under stable control and the process capability coefficient is equal to 1, how the quality characteristic value distribution to describe the process capability is through the method by six times standard deviation 6σ; namely, within the range of u±3c, products can be guaranteed to meet the quality requirements by the probability of 99.73%. Therefore, process may be deemed to have enough quality assurance. Generally speaking, it may be considered to be poor economy with higher demand such as 8σ [11, 13]. Certainly, in case that the product cost is greatly lower than product price, or the products is more competitive in market, or the systematic factors of quality fluctuation reduce with enriching in manufacture experience, we can make a further comprehensive evaluation of whether adjust the process capability coefficient or not within tolerance range. In a word, we have an actual pass percentage corresponding to a certain process capability coefficient. Then by the statistics of the standard deviation σof the production process, we can work out the products tolerances 6σ [6].

Table 1. The relationship between the process capability coefficient and the product requirements

Process capability coefficient Cp(Cpk)	Fraction defective P (%)	Evaluation of process capability	Disposal
Cp>1.67	P<0.00006	Process capability is much too sufficient The product tolerance is too loose	Products with low production requirements
1.67>Cp>1.33	0.00006<p<0.006	Process capability is much too sufficient The product tolerance is too loose	Produce products with common requirements
1.33>Cp>1.00	0.006<p<0.27	Process capability is up to standard Product tolerance is appropriate	Produce products with high requirements
1.0>Cp>0.67	0.27<p<4.55	Process capability is insufficient Product tolerance is too strict	Produce products with too high requirements
Cp<0.67	P>4.55	Process capability is seriously insufficient Product tolerance is unpractical	Produce products which are hard to be achieved

3 Standard Deviation of Production Process

For what we analyze is the production of process material and each process capability is hard to be calculated by ration, therefore, the process capability involved.

In below is actually the comprehensive process capability.

σ shows the standard deviation of the whole production process, which not only includes the standard capability of finished products, but also contains the standard deviation of unfinished products; hence, we can not work out it. The S represents the standard deviation of finished products (subsample) picked out from this production process. Only if the process is stable, σ can be approximately substitute by S. But there are distinctions between them in concept.

The calculation of standard deviation s of production process, which is illustrated as follows:

Sample one: a batch of mat being continuous produced, and the thickness data of the 100 rolls picking out from that.

Table 2. The thickness data measured from those 100 rolls of mats

0.34	0.32	0.35	0.35	0.36	0.31	0.33	0.31	0.32	0.32
0.32	0.34	0.34	0.32	0.35	0.35	0.34	0.30	0.32	0.31
0.35	0.32	0.33	0.34	0.34	0.36	0.32	0.31	0.30	0.35
0.35	0.33	0.35	0.32	0.35	0.33	0.33	0.32	0.33	0.33
0.33	0.35	0.33	0.34	0.34	0.32	0.31	0.32	0.35	0.32
0.33	0.34	0.34	0.35	0.31	0.32	0.31	0.32	0.32	0.32
0.34	0.33	0.34	0.34	0.36	0.32	0.31	0.30	0.32	0.31
0.33	0.34	0.32	0.32	0.33	0.31	0.30	0.30	0.33	0.31
0.33	0.35	0.36	0.34	0.33	0.33	0.30	0.30	0.30	0.31
0.33	0.33	0.34	0.33	0.35	0.29	0.30	0.30	0.31	0.32

1. Make the frequency distribution table according to the measured data.

Table 3. Frequency distribution table

Group No.	Group interval	Central value xi	Frequency statistics	frequency fi	Simplified central value ui	fi*ui	fi*ui2
1	0.285-0.295	0.29	𝍸	1	-3	-3	9
2	0.295-0.305	0.30	𝍸 𝍸	10	-2	-20	40
3	0.305-0.315	0.31	𝍸 𝍸 𝍸	13	-1	-13	13
4	0.315-0.325	0.32a	𝍸 𝍸 𝍸 𝍸 𝍸	22	0	0	0
5	0.325-0.335	0.33	𝍸 𝍸 𝍸 𝍸	19	1	19	19
6	0.335-0.345	0.34	𝍸 𝍸 𝍸 𝍸	17	2	34	68
7	0.345-0.355	0.35	𝍸 𝍸 𝍸	14	3	42	126
8	0.355-0.365	0.36	𝍸	4	4	16	64
Total				100 Σfi	4 Σui	75 Σfiui	339 Σfiui2

2. Calculation of average value

$$\overline{x} = a + h\frac{\sum fiui}{fi} \tag{1}$$

Example:

$$\overline{x} = 0.32 + 0.01\frac{75}{100} = 0.33$$

3. Calculation of standard deviation s:

$$S = h \cdot \sqrt{\frac{\sum fiui^2}{\sum fi} - \left[\frac{\sum fiui}{\sum fi}\right]^2} \tag{2}$$

Example:

$$S = 0.01 \cdot \sqrt{\frac{339}{100} - \left[\frac{75}{100}\right]^2} = 0.0168$$

4 Calculation of Tolerance

4.1 Calculation of Process Capability Coefficient When the Tolerance Center Coincides with the Size Distribution Center

Calculation formula of process capability coefficient:

$$Cp = T / B = T / 6\sigma \tag{3}$$

In formula:

Cp-process capability coefficient T-tolerance

B-Actual characteristic value distribution range (appropriate production quota)

σ-Standard deviation of process production process (it can be substitute by s)

The calculation of process capability coefficient is related to the stipulation method of quality standard. As quality standards have three stipulation manners of doubleaction, right unidirection and left unidirection, there are three methods to calculate Cp accordingly.

1 When the quality standard is stipulated by the double-way, as it shown in Fig. 1:

$$Cp = (S\mu - Sl) / 6\sigma \tag{4}$$

In formula:

Sμ-Upper limit of the quality standard;

S1-Lower limit of quality standard;

Σ-Overall standard deviation

Sample 2. Use the data of the example 1 to get the tolerance T

$T = Cp \times 6\sigma = 1 \times 6 \times 0.0168 = 0.1008$

The actual sizes distribution range X= Actual size centre

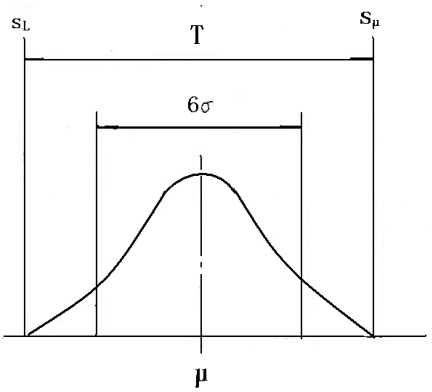


Fig. 1. Doubleaction stipulation of quality standard

2. When the quality standard is stipulated by right unidirection (generally, the tolerances only provide the upper limit standard), as shown in Fig. 2:

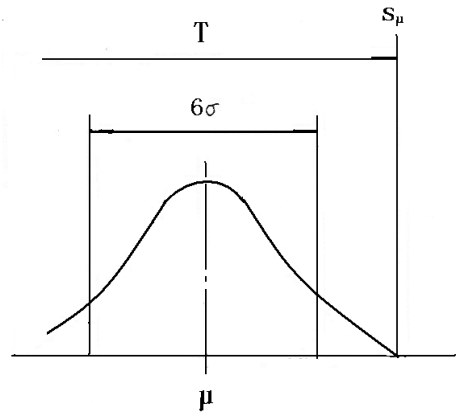


Fig. 2. Right Unidirection Stipulation of Quality Standard

$$Cp = (S\mu - \mu)/3\sigma \tag{5}$$

In formula:

μ-Overall average value.

3 When the quality standard is stipulated by the left unidirection (generally, the tolerances only provide the lower limit standard), as shown in Fig. 3:

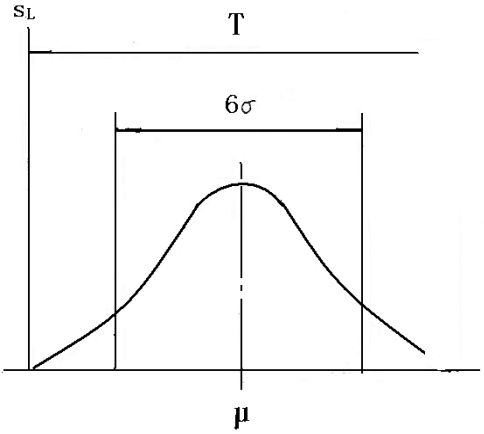


Fig. 3. Left Unidirection Stipulation of Quality Standard

$$Cp = (\mu - Sl) / 3\sigma \tag{6}$$

4.2 Calculation of Process Capability Coefficient When the Tolerance Center doesn't Coincide with the Size Distribution Center

We should modify Cp, and the method is to multiply the modified value k

That is to say: Cpk=(l-K)Cp=(1-K)T/6σ

In formula: k= (1-K)

$$K = \frac{(M - \bar{x})}{T / 2} = \frac{\varepsilon}{T / 2} \tag{7}$$

In the formula:

K-Modified value;

K-Relative offset;

X-Central values of the tolerance range;

E-Distribution canter and absolute offset with tolerance centre=M-x;

T- Tolerance range

We also can use the following formula to indicate process capability coefficient of offset ε;

$Cpk = (T - 2\varepsilon) / 6\sigma$

5 Corresponding Relation between Process Capability Coefficient and Fraction Defective

When quality index distribution obeys to normal distribution and distribution center x and tolerance center position M are co insistent, the certain process capability coefficient value corresponds with certain percent defective.

For example Cp=1, that is T=6σ, known from the normal distribution, now, the normal distribution is0.27%, we can see the corresponding relationship of Cp value and the failure rate (p) from figure 4 and may be proved to be:

$$p = 2\Phi(-3Cp) \tag{8}$$

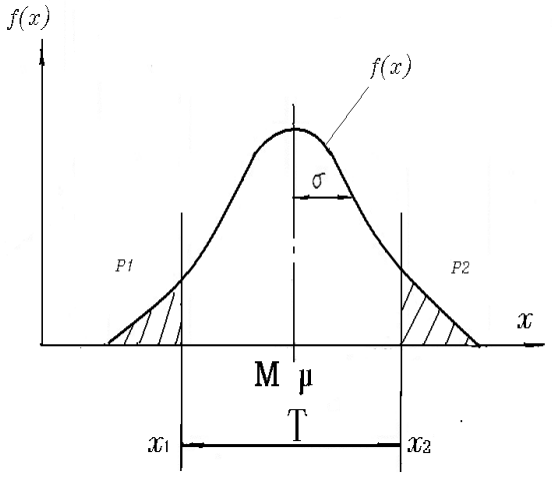


Fig. 4. Relationship between Process Capability Coefficient and Percent Defective

Sample 3. Use the data in Example 1 to calculate percent defective p

$p = 2\Phi(-3Cp)$

When $Cp = 1$

$p = 2\Phi(-3\times 1) = 2\Phi(-3)$

See normal distribution table: $\Phi(-3) = 0.0013$

So, $p = 2\times 0.00135 = 0.0027 = 0.27\%$

6 Conclusion

Cited in this case is the use of tolerance range calculated by process capacity coefficient when tolerance coincides with the center, product quality standards are stipulated by bidirection, as described herein , when the quality standards are the right unidirectional stipulation or left unidirectional stipulation provisions, it can also be calculated by using the formula provided; if customers have special needs, it can also develop tolerance according to the calculation formula of process capability coefficient when tolerance center and the size distribution centers do not coincide. Not enumerate one by one here.

In summary, we can produce an accurate calculation of the tolerance range of products manufactured by the existing equipment through the application of statistical techniques. Otherwise, if the production facility is obsolete and with poor accuracy, it

can only be vain to do high precision product, resulting in the waste of raw materials and loss of enterprise; if production equipment is advanced and with high precision, but the tolerances is given in a very wide range based on the subjective judgment, that have not been able to play a potential role and fail to create the best economic benefit, this is also a waste of resources. Only tolerance range developed by using statistical techniques is in line with the actual product tolerances, which can help corporate leadership make the right decisions, provide expected and accessible objectives for production operators, and sales staff can show the excellent convinced basis of products for customers.

At present, a considerable number of enterprises do not pay attention to the application of statistical techniques. On the one hand, the understanding of statistical techniques knowledge is not enough; on the other hand, they feel it is not economic to take effort on it. Actually, the statistical techniques said above can be mastered by the engineering and technical personnel of junior or secondary technical level; Everything is difficult at the beginning, after confirming of calculation ideas, a general computing operations personnel can do it; and, if use computer for calculation, the calculation like standard deviation can be calculated within a few minutes. We believe that, with the promotion of statistical techniques in the enterprise, it must be able to have a multiplier effect on improving product quality and promoting economic benefit of enterprises.

References

1. Venkateshwara Rao, B.: Uniqueness and convergence of solutions to average run length integral equations for cumulative sum and other control charts. IIE transactions 33, 72–76 (2001)
2. Chao, M.T.: Another Look at the Process Capability Index. Quality and Reliability Engineering International 22, 96–98 (2005)
3. Cui, L.: Application of Tolerance Matching of Probability and statistics 4, 17–20 (2010)
4. Fu, J.C.: Fu on the average runs lengths of quality control schemes using a Markov chain approach. Statistics & Probability Letters 56, 22–28 (2002)
5. Jia, X.: The process capability index analysis of non-normal distribution parameters. Journal of Xidian University 1, 13–15 (2004)
6. Kotz, S., Johnson, N.: Process Capability Indices-A Review. Journal of Quality Technology 34, 66–68 (2002)
7. Ling, Z.: Application of Probability and Mathematical Statistics in the Economic Analysis. Intelligence 11, 34–35 (2010)
8. Marion, R.: Reynolds a general approach to modeling CUSUM charts for a proportion. IIE Transactions 32, 56–59 (2000)
9. Prabhu, S.S., George, C.: Selection of the subgroup size and sampling interval for a CUSUM control chart. IIE Transactions 29, 133–138 (1997)
10. Wang, F.: Quality Evaluation of a manufactured product with multiple characteristics. Quality and Reliability Engineering International 22, 258–262 (2006)
11. Zhang, G.: New Quality Management, pp. 46–51. Higher Education Press, China (2005)
12. Zheng, M.: Mathematical Statistics Notes, pp. 6–11. Fudan University Press, China (2006)
13. Zhou, F.: Commonly Used Mathematical Statistics Methods and Application Examples, pp. 588–590. China Metrology Press, China (2006)

Small Sample Prediction Based on Grey Support Vector Machine

Nie Shaohua

Linyi University, Linyi, Shandong 276000, China

Abstract. The concept of grey system is evolving from the concept of "black box", mainly focusing on objects with clear extension but vague intension. The core part of grey system theory is dynamic modeling, which has organically combined theory and actual situation, so as to solve and instruct actual problem.Support Vector Machine method is determining decision function according to limited sample information and little support vector quality. The counting process is not relevant to space dimensionality, mainly dealing with novel small sample study method of non-linear regression problems. In this paper, through introducing combining grey system theory and support vector machine theory, combining grey prediction model and support vector machine model, the possibility of combining these two has been attained.

Keywords: Grey System, Support Vector Machine, System Modeling.

1 Introduction

Rising in 80's, grey system analyzes incomplete information small samples, to recognize real situation and further grasp operation law of objects. The nature of grey system modeling is finding number with number, exploring system change situation by counting one or more discrete series [1-5].

Support vector machine (SVM) has overcome weaknesses of traditional statistics classification, mainly applied to solve small sample, non-linear problems, with high precision for counting process and counting result which are incomparable with traditional classification method. The grey support vector machine is a kind of machine made by organically integrating grey system theory and support vector machine. This method not only maintains advantages of these two theories, and also has more typical representativeness for dealing with actual problems, and wider application scope [6-10].

Predict the state, enterprises and other organizations to develop policies and programs based primarily on the accuracy of prediction is therefore policy and program development is scientific premise. Forecasting methods are traditional multiple regression, and in recent years developed an artificial neural network prediction, gray prediction. Multiple regression prediction model is simple, easy to use and strong, but difficult to handle high-dimensional, nonlinear mode [11-15]; artificial neural networks, although able to solve the problem of high-dimensional nonlinear prediction,

Y. Yang, M. Ma, and B. Liu (Eds.): ICICA 2013, Part II, CCIS 392, pp. 565–574, 2013.

but it requires a lot of training samples and the generalization ability is not strong so when the available sample is small sample forecast, or to obtain a large number of samples is costly, it will inevitably affect its practicality and economy [16-20]; gray forecast, although with short-term forecasting ability, etc. can be tested, but its long-term forecast ability is poor. Vapnik et al proposed SVM in statistical learning theory developed on the basis of a new kind of machine learning algorithms, is the small sample statistics and forecasts for the best learning theory, support vector machine with perfect mathematical form, intuitive geometric interpretation and good generalization performance, to solve the model selection and less learning, through learning and nonlinear problems, to overcome the slow convergence and easy to fall into local optimal solution and other shortcomings, so support vector machine classification and regression have shown superior performance [21-25].

For small sample data, information processing, many scholars gray system model and support vector machine model to study and achieved certain results [26-30]. Currently, the fusion of these two methods to establish gray support vector machine model, can make up a single (slightly) model is insufficient to achieve good data processing and forecasting effect, has become a very important issue. Firstly, the characteristics of small sample data were studied, a detailed analysis of small sample data modeling and (slightly), specificity, then raise the gray system theory and support vector machine combining ideas to be modeled in the right gray system theory and support vector machine, especially the gray system theory in-depth study, based on the establishment of gray GM (1,1) and support Vector Machine (abbreviated) SVM prediction model on gray GM (1,1) modeling mechanism and error analysis, based on the use of sub- index sequence alignment of the PSO algorithm constructed under the form of background values for parameter optimization, and then combined SVM advantage to (slightly) higher than the number of columns or under conditions similar index series has higher prediction accuracy modeling, simulation examples show that this model can predict long-term due to the complexity of the process industry (multi-variable, nonlinear, time-varying, etc.), a large number of studies have focused on the historical data according to the process, using a variety of non-linear function fitting method to model, however, between the input variables when modeling autocorrelation and cross correlation (abbreviated) type of generalization ability, this paper, gray correlation analysis method as an attribute preprocessor correction based on the attribute relevance of each factor were studied [31].

2 Grey System Theories

2.1 Generation of Grey System

The basic common character of modern control theory and classic control theory is based on correct and precise mathematics model. But under many real situations such as biological system, economic system and social system, it's hardly to attain accurate mathematics models. Chad has proposed fuzzy control theory in 1965.

We have achieved significant breakthrough when we firstly applied mathematics method in fuzzy variable to realize fuzzy prediction and control of system that cannot

form mathematics model. With continuous practicing test in later phase, fuzzy control theory has revealed some drawbacks, such as low information utilization rate, rough control and low precision. Therefore, no matter classic control theory, modern control theory or fuzzy control theory, all study white systems as study object (information is fully confirmed). In fact, in terms of subject for recognition, neither macro-system and micro-system, nor natural social system has complete information. Hence, grey system theory has been created. Grey system is established on grey system theory. Analysis on concept of "grey" made in different perspectives is shown in Table 1.

Table 1. Analysis on concept of "grey" made in different perspectives

Perspective	Black	Grey	White
Information Aspect	Unknown	Incomplete	Complete
Appearance Aspect	Dark	Bright or Dark	Bright
Process Aspect	New	Mixture of New & Old	Old
Nature Aspect	Mixed	Multiple Ingredients	Pure
Method Aspect	Negative	Sublate	Affirmative
Attitude Aspect	Indulging	Tolerant	Strict
Result Aspect	Without Answer	With More Than One Answer	With Only One Answer

2.2 Grey Correlation Analysis

Grey correlation is the foundation of grey system theory, the purpose of grey correlation analysis is analyzing main characters of system based on correlation degree between different factors. The basic thought is determining tightness according to geometrical shape resemblance of curves, and then showing it with correlation level. Grey correlation analysis method has offered solid theoretical foundation for problem analysis, modeling, prediction and decision making.

2.3 Grey GM (1, 1) Model

In the grey theory, the original characteristic data series $X^{(0)}$ is assumes as :

$$X^{(0)} = \left\{ x^{(0)}(1), x^{(0)}(2), \ldots, x^{(0)}(n) \right\} \tag{1}$$

Next, accumulate all original signature sequences, namely, one-time accumulated sequences:

$$x^{(1)}(k) = \{ \sum_{i=1}^{k} x^{(0)}(i), k = 1, 2, \ldots, n \} \tag{2}$$

This sequence has increased regularity of data, more suitable for following study. Later, we can establish grey model as below:

If we set

$$z^{(1)}(k) = 0.5x^{(1)}(k) + 0.5x^{(1)}(k-1) \tag{3}$$

and name $z^{(1)}=\{z^{(1)}(2),z^{(1)}(3),\ldots.z^{(1)}(n)\}$ as the close average value generated sequence of $x^{(0)}$.

Next, we set grey differential equation:

$$x^{(0)}(k) + az^{(1)}(k) = b \tag{4}$$

a and b are uncertain parameters, a is development coefficient, and b is grey action integral.

Formula (4) GM (1, 1) model's whitening equation is:

$$\frac{dx^{(1)}(k)}{dt} + ax^{(1)}(k) = b \tag{5}$$

3 Principle of Support Vector Machine Method

3.1 Linear Regression Support Vector Machine

In the geometrical angle, we can see that the linear regression function:

$$y = f(x) = (w \cdot x) + b \tag{6}$$

is seeking optimal regression hyperplane problem by solving convex quadratic programming. Adopting ε as error function, we have introduced slack variable ξ, and received hyperplane convex quadratic programming problem:

$$Min \frac{1}{2}\|\omega\|^2 + C\sum_i (\xi_j + \xi_j^n) \tag{7}$$

st

$$\begin{aligned} &y_i - (\omega \cdot x_i - b) \le \varepsilon + \xi_j \\ &(\omega \cdot x_i) + b - y_i \le \varepsilon + \xi_j, \quad \xi_j, \xi \ge 0 \end{aligned} \tag{8}$$

Under the condition of KKT, we can receive optimal hyperplane linear regression function as following through Lag range multiplier method:

$$\begin{aligned} &f(x) = (\omega \cdot x) + b \\ &= \sum_{s.v.} (\alpha_i - \alpha_i^*)(x \cdot x_i) + b \end{aligned} \tag{9}$$

With constraint condition, we can receive α_i, α_i^*, b and other optimal hyperplane parameters. If both or one of α_i and α_i^* is 0 or below, namely the training sample (x_i, y_i) of $\alpha_i\alpha_i^*=0$ is existing, we can count and receive vector α_i and α_i^*. Objective function and constraint condition are convex, and this problem has only one minimum overall solution according to optimization theory.

3.2 Analysis about Non-linear Support Vector Machine

When mapping only needs reflection dot product, we can use related kernel function to replace, which is an important step from linear vector machine to non-linear support vector machine. If one sample set is reflected into the characteristic space expanded by Merkcer nuclear's characteristic function

$$\begin{aligned} K(x,y) &= \sum_i \lambda_i \Phi_i(x)\Phi_i(y) \\ &= \left(\Phi(x)\cdot\Phi(y)\right) \end{aligned} \tag{10}$$

through non-linear mapping

$$\Phi(x) = (\sqrt{\lambda_1}\Phi_1(x), \sqrt{\lambda_2}\Phi_1(x), \ldots, \sqrt{\lambda_k}\Phi_k(x)) \tag{11}$$

The regression decision function formula is changed into

$$\begin{aligned} f(x) &= \left(\omega\cdot\Phi(x)\right) + b \\ &= \sum_{i=1}^{l} (\alpha_i - \alpha_i^*)\left(\Phi(x)\cdot\Phi(x_i)\right) + b \end{aligned} \tag{12}$$

According to Mercer theorem, we can receive

$$\begin{aligned} f(x) &= \left(\omega\cdot\Phi(x)\right) + b \\ &= \sum_{i=1}^{l} (\alpha_i - \alpha_i^*) K(x\cdot x_i) + b \end{aligned} \tag{13}$$

by simplifying above function, and then solve support vector, support "strength" and threshold, and thus solve kernel function to count sample space regression value.

3.3 Introduction of Kernel Function

Kernel function is the basis of kernel skill, which is an important constituent part of support vector machine. By introducing kernel function, we not only avoid complicated higher space operational data, complex counting process, but also can handle with high-dimensional problem without knowing characteristic space and non-linear mapping explicit expression.

The definition of kernel function: if X is one subset in R^n, we can name function k (x, x') defined on x.x is kernel. If there is a reflection from X to some Hilbert space F:

$$\vec{\Phi}: X \quad \vec{F} \qquad x \quad \Phi(x) \tag{14}$$

Set

$$k(x, x') = \Phi(x) \cdot \Phi(x) \tag{15}$$

There are three kinds of kernel functions usually applied:

1. Linear kernel function:

$$K(x, y) = x \cdot y \tag{16}$$

2. Multinomial kernel function:

$$K(x, y) = \left[(x \cdot y) + c \right]^d \tag{17}$$

c and d are parameters

3. Gauss kernel function:

$$K(x, y) = exp(-\frac{\| x - y \|^2}{\delta^2}) \; K(x, y) = \left[(x \cdot y) + c \right]^d \tag{18}$$

δ is parameter

4 Applying Grey Support Vector Machine to Predict Small Sample

4.1 Founding Grey Support Vector Machine Prediction Modeling

1) We set time sequence $x^{(0)} = \left(x^{(0)}(1), x^{(0)}(2), \ldots, x^{(0)}(n) \right)$, under three models of SGM(1,1), IGM(1,1) and BGM(1,1), we can receive m values of simulation and n values of prediction.

2) Prediction regularization factor is 1, and radial kernel function sig2 is 0

3) Input value of simulation under three models to count the error under model prediction.

4) Increase value of regularization factor and value of radial kernel function: 1, 0.1, to count sample error again, choose the parameter with smallest error as regularization factory and radial kernel function.

5) Input prediction value fewer than three models, choose three group of n values of prediction in 1) as input values, and then select regularization factor and radial kernel function in Step 2 as parameters to make simulation. The received output n values are values of predication of GMSVM.

4.2 Analysis of Case

According to the models based on original GM (1, 1) model and GMSVM model about financial science and technology investment data of Shandong Province from 1999 to 2002, we have founded prediction model to forecast financial science and technology investment amount in 2003 and 2004, and also made analysis and comparison. As shown in Table 2.

Table 2. Investment in Henan Province from 1999 to 2004 Unit: 100 Million Yuan

Year	1999	2000	2001	2002	2003	2004
Investment	10.51	12.72	14.84	17.89	21.22	26.79

The simulation of SGM model, IGM model and BGM model about investment data from 1999 to 2002 are shown in Table 3:

Table 3. Comparison among Three Models' Values of Simulated Prediction and Actual Values Unit: 100 Million Yuan

Year	Actual Value	SGM Model	IGM Model	BGM Model
		Simulated Value	Simulated Value	Simulated Value
1999	10.51	10.823	10.77	10.784
2000	12.72	12.736	12.694	12.725
2001	14.84	14.986	14.962	15.014
2002	17.89	17.633	17.635	17.716
2003	21.22	20.749	20.786	20.903
2004	26.79	24.415	24.5	24.665

Next, we use simulated results of three models as input factors, actual values as prediction factors, to make support vector machine training, and also make comparison with original GM (I, I) model as shown in Table 4:

Table 4. Comparison among Three Models and GM (I, I) modelModel

Year	Actual Value	Original GM (1, 1) Model	GMSVM Model
		Value of Simulated Prediction	Value of Simulated Prediction
1999	10.51	12.29	10.544
2000	12.72	14.48	12.731
2001	14.84	17.07	14.815
2002	17.89	20.12	17.804
2003	21.22	23.71	20.913
2004	26.79	27.95	25.873

We further count model prediction error and receive Table 4.

Table 5. Count model prediction error

Year	Original GM (1, 1) Model	GMSVM Model
2003	11.7649	-1.4467
2004	4.3446	-3.4229

As shown in Table 5, we can see that error of GMSVM model is clearly smaller than GM (1, 1) model's error, meaning that simulation precision of GMSVM model is higher than GM(1, 1) model.

5 Conclusion

If we combine grey correlation analysis with support vector machine, the prediction precision is higher than regular prediction model. In this paper, study on combining grey system and support vector machine is made and related progress is attained. Because of influence in many aspects, many models still cannot effectively predict.For small sample data, information processing, many scholars gray system model and support vector machine model to study and achieved certain results. Currently, the fusion of these two methods to establish gray support vector machine model, can make up a single (slightly) model is insufficient to achieve good data processing and forecasting effect, has become a very important issue. Firstly, the characteristics of small sample data were studied, a detailed analysis of small sample data modeling and (slightly), specificity, then raise the gray system theory and support vector machine combining ideas to be modeled in the right gray system theory and support vector machine, especially the gray system theory in-depth study, based on the establishment of gray GM (1,1) and support Vector Machine (abbreviated) SVM prediction model on gray GM (1,1) modeling mechanism and error analysis, based on the use of sub- index sequence alignment of the PSO algorithm constructed under the form of background values for parameter optimization, and then combined SVM advantage to (slightly) higher than the number of columns or under conditions similar index series has higher prediction accuracy modeling, simulation examples show that this model can predict long-term due to the complexity of the process industry (multi-variable, nonlinear, time-varying, etc.), a large number of studies have focused on the historical data according to the process, using a variety of non-linear function fitting method to model, however, between the input variables when modeling autocorrelation and cross correlation (abbreviated) type of generalization ability, this paper, gray correlation analysis method as an attribute preprocessor correction based on the attribute relevance of each factor were studied.

References

1. Yu, F., Fengshan, M., Aihua, W., Haijun, Z., Jie, G.: Gray system and neural network combined model in Groundwater Level Prediction. In: Institute of Geology and Geophysics of the 11th (2011) Annual Conference Proceedings (in) (2012)

2. Chun, H.: Grey System Theory in Building Deformation Analysis. In: National Mapping Technology Information Network South sub-network Twenty-fifth Symposium (2011)
3. Qifeng, Z., Lim, P., Open, R., Hung, P.: Riemannian metric based on the training sample class imbalance SVM classification method. In: Twenty-Sixth Chinese Control Conference (2007)
4. Xiaojun, G., Xi, Y., Edmund, Q.: Based on support vector machine multi-class Rotating Machinery Fault Recognition. In: Twenty-Sixth Chinese Control Conference (2007)
5. Zhangshun, W., Shen, Y.-H., Wei, L.: Based on optimized GM (1,1) model of the population prediction methods. In: Twenty-Sixth Chinese Control Conference (2007)
6. Chunjie, Y., Jun, S., Zhang, C.-H.: A new type of battery status line detection and fault prediction algorithm. In: Twenty-Sixth Chinese Control Conference (2007)
7. Liang, L., Xinping, X., Shuhua, M.: Grey particle swarm bilevel linear programming solution. In: Twenty-Seventh Chinese Control Conference (2008)
8. Bin, W., Li, Gong, W.: Based on gray correlation analysis of population and environment in Tianjin. In: Twenty-Seventh Chinese Control Conference (2008)
9. Devon war, Junfeng: Image edge detection algorithm and its application. In: Twenty-Seventh Chinese Control Conference (2008)
10. Rongcheng, L., Devon war: Residuals based model predictive gray fuzzy self-tuning controller design. In: Twenty-Seventh Chinese Control Conference (2008)
11. Qing, L., Guoping, D., Bian, X.-M., Yang, C.L., Paul, Q.l.: Chinese grain production dummy variable model. Anhui Agricultural Sciences, 11 (2005)
12. Feng, T.: kernel-based learning algorithm, 2. Northern Jiaotong University (2003)
13. Huang Shengxiang do themselves, engineering and construction Settlement Prediction equidistant gray model, Geospatial Information, 01 (2004)
14. Licheng, J., Li, Z.: Zhou reached; support vector preselection center distance ratio method, Electronic Technology 03 (2001)
15. Taojie, C.: Gray forecasting model an Extension. Systems Engineering, 4 (1990)
16. Cunzhi, G., Rui, T., Chen, C.: Application of non- equal Interval GM (1,1) model fitting groundwater calculation parameters. Journal of Hohai University (Natural Science), 1 (1999)
17. Wang, G.-S., Yixin, Z.: Support vector machine theory - Statistical Learning Theory. Computer Engineering and Applications, 19 (2001)
18. Yan, B., Si, M.: Grey RBF Neural Network Model Traffic Prediction and Analysis. Computer Engineering and Science, 10 (2008)
19. Qing, L., Licheng, J.: Zhou up. vector-based support vector projection pre-selection. Computer Engineering, 2 (2005)
20. Chen, J., Hong, S., Xian, W.Z.: Gray system modeling data in a transformation method. Jiangsu Polytechnic University (Natural Science), 3 (1999)
21. Xiaojing, F., Bin, Z., Liang, Zhaohui: Based on support vector machine Pipeline Leak Detection Method. In: Sixth National Information Acquisition and Processing Conference Proceedings, vol. (1) (2008)
22. Ruiping, Wenbin, S.: Jiaxing City Precipitation Forecast Tropical Cyclone Research. In: Chinese Meteorological Society 2005 Annual Conference Proceedings (2005)
23. Wang in the text. SVM method quickly revised Application Research. In: Chinese Meteorological Society 2005 Annual Conference Proceedings (2005)
24. Hanzhong, F., Yongyi, C., Qin, C.Y.: LUO Health, Shuangliu Airport Low Visibility Forecast Method.
25. Chinese Meteorological Society 2006 Annual Meeting “aviation weather detection, forecasting, early warning technology progress” Breakout Proceedings (2006)

26. Yang, X.: Support vector machine in Shandong flash geological disasters forecast application experiments. In: Chinese Meteorological Society 2006 Annual Meeting " flash flood monitoring, prediction and evaluation of " Breakout Proceedings (2006)
27. Chinese Meteorological Society 2007 Annual Meeting Proceedings climate credits venue (2007)
28. Yuxia, H., Dongbei, X., Su, P.: SVM method in forest fire prediction. In: Chinese Meteorological Society 2007 Annual Forecast warning and impact assessment techniques venue Proceedings (2007)
29. Zhu, L., Zhu, G.-D., River, S.: SVM method in Urumqi airport runway visual range prediction. In: Chinese Meteorological Society 2007 Annual Forecast warning and impact assessment techniques venue Proceedings (2007)
30. Wanquan, H., Yu'e, L.: Hebei objective sub-county summer rainfall forecasting system. In: Chinese Meteorological Society Annual Conference 2008 Forecast accuracy and Public Weather Services venue Proceedings (2008)
31. Yuqing, M., Yongqing, W., Wang, G.L.: Support vector machine method is applied to the ideal time series prediction. In: Chinese Meteorological Society Annual Conference 2008 climate prediction and forecasting methods venue Proceedings (2008)

Simulation of Liquid Film Flowing Outside the Elliptical Tube

Chunhua Qi, Houjun Feng, Qingchun Lv, and Ke Xu

The Research Institute of Seawater Desalinization and Comprehensive Utilization in Tianjin, State Oceanic Administration, Tianjin City, China.300192

Abstract. The elliptical tube is set as object of the study. From the angles of hydromechanics and heat transfer science, and with technical characteristics of falling film evaporation and seawater desalination of horizontal tubes, the thesis uses theoretical analysis, numerical simulation and experimental study. It researches liquid film flowing outside the elliptical tube, and also heat and mass transfer of the film. Contrastive analysis with circular tube is also conducted. The thesis shall probe into superiority and feasibility when the elliptical tube is applied to falling film evaporation and seawater desalination of horizontal tubes. It conducts exploratory researches for further development of seawater desalination technology in China.

Keywords: elliptical tube, liquid film, simulation, heat transfer, experiment.

1 Introduction

Seawater desalination of low-temperature and multi-effect distillation mostly adopts the technology of falling film evaporation of horizontal tubes. The performance of heat and mass transfer of falling film evaporator of the tube has a direct effect on the cost of seawater desalination. During the distillation of seawater desalination, intensified heat tubes are used, thus saving the amount of heat transfer material and reducing manufacturing cost of devices [1]. Falling film evaporator of the elliptical tube has been developed recently. It has been successfully applied to such industries as air conditioning, refrigeration and chemical industry, achieving a satisfactory energy-saving effect. Many experts and scholars at home and abroad have conducted various researches on heat transfer of the elliptical tube [2-5]. In 2003, Hasan [6] carried out an experimental study on the elliptical tube and circular tube. The study indicated that the performance of heat transfer and flow resistance of the elliptical tube is 1.93～1.96 times that of the circular tube. Chong-yan Chen [7-10] researched boiling heat transfer of falling film outside the horizontal perforated elliptical tube. The researching results indicated that perforated surface of oval cross section could improve prominently heat transfer performance [11-16].

2 Numerical Simulation Research

From the angle of hydromechanics, the thesis uses computational fluid mechanics to build the model of flow of liquid film for elliptical tubes, thus undertaking numerical

Y. Yang, M. Ma, and B. Liu (Eds.): ICICA 2013, Part II, CCIS 392, pp. 575–583, 2013.

simulation. The thesis determines structure and size of the tube where liquid film outside the tube could flow well [17-21].

In the entire computational domain, length (x) ×height (y) ×breadth (z) =18mm×50mm×40mm. The domain could be divided into such six regions as wall region, critical region of wall, inflow region, axial region, effluent region and non-primary region. All use hexahedral mesh. The totality of meshes is 1,420,000. The mesh of computational domain is demonstrated in figure 1.

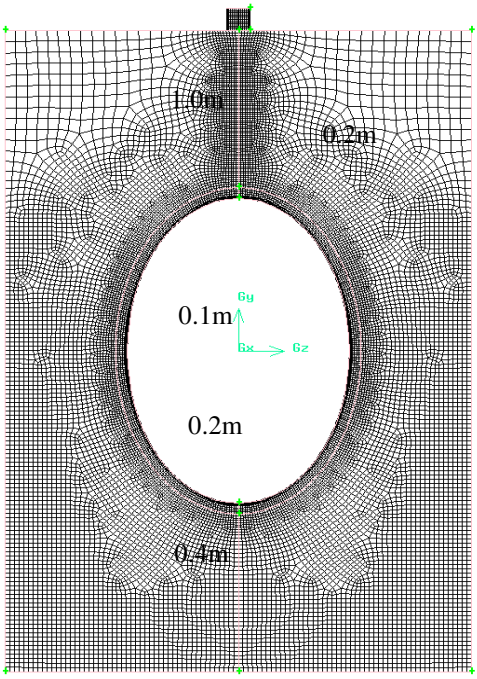


Fig. 1. Figure of division of the mesh

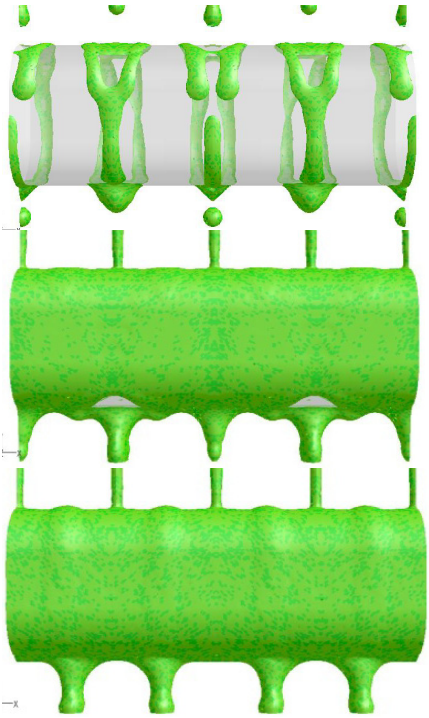
3 Results and Analysis of Numerical Simulation

3.1 Effect of Types of Tubes on the Distribution of Liquid Field

When flow velocity at entrance is 0.2m/s, 0.5m/s and 0.8m/s respectively (and fluid load is 0.037kg/(m·s), 0.093kg/(m·s) and 0.149kg/(m·s) respectively), one calculates numerical simulation of liquid flowing outside the tube for circular tube EP01 and elliptical tube EP03 with the same perimeter of section [22-25].

Table 1. Parameters of tube EP01 and tube EP03

Parameters of models	Shape factors of cross-section E	Parameters of mesh
EP01 long axis of tube wall 25.0 mm, minor axis 25.0 mm	1	1.0 time
EP03 long axis of tube wall 28.5 mm, minor axis 19.0 mm	1.5	1.5 times



V=0.2m/s, 980ms

V=0.5m/s, 840ms

V=0.8m/s, 375ms

Fig. 2(a). The distribution of liquid film with different flow velocities at entrances of circular tube EP01

V=0.2m/s, 645ms

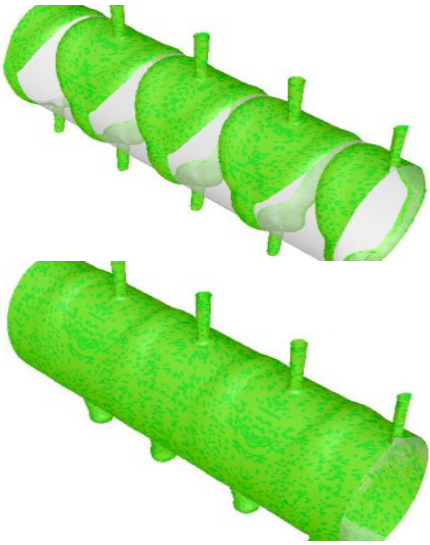


Fig. 2(b). The distribution of liquid film with different flow velocities at entrances of elliptical tube EP03

From figure 2 one may see the whole process. The liquid drop first drips, then spreads, finally falls and breaks away at the top of every tube. It also returns to the original state at the top. Time of breaking away will decrease as flow velocity at entrances increases. In figure 2 (b), when the flow velocity of liquid film of elliptical tubes is 0.2m/s, the liquid will flow discontinuously and like drops. There exist many dry walls at the surface of heat transfer tube. When the flow velocity is 0.5m/s, flow pattern of liquid outside the tube is columnar. Yet the liquid could not completely wet the outer wall of heat transfer tube. There still exist dry regions. When the flow velocity is 0.8m/s, flow pattern of liquid outside the tube is drop-columnar. The outer wall of heat transfer tube is wrapped up with a thin layer of liquid film. When the flow velocity increases to 1.0m/s, flow pattern of liquid outside the tube is columnar. The outer wall of the tube is wrapped up with a thicker layer of liquid film. At the middle region between the two inflows, liquid film will get noticeably thicker and form obvious bulges.

3.2 The Effect of Types of Tubes on Thickness of Liquid Film

The distribution of liquid on the surface of tubes determines thickness of liquid film outside the tube [26-29]. Yet the types of heat transfer tube have a significant effect upon the distribution of liquid outside the tube. When liquid load is 0.149kg/ (m·s), namely, when the flow velocity at entrances is 0.8m/s, according to simulation results, one measures the changes of thickness of liquid film of outer walls of the tube in the directions of circumference and axis of the tube. The simulation results are shown in figure 5, in which θ is the included angle between the line linking the measured point with center of gravity of the tube and the vertical line through the center.

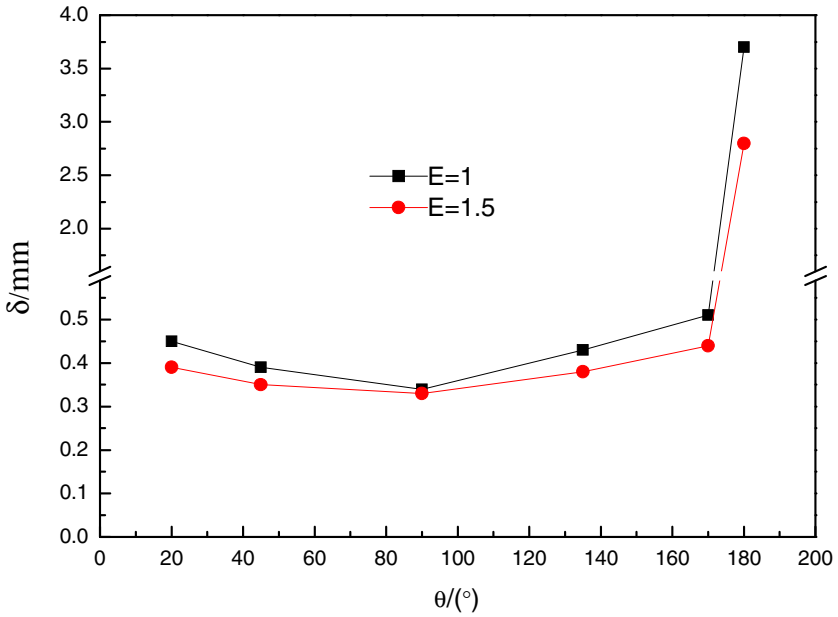


Fig. 3. The distribution of thickness of liquid film outside the tube

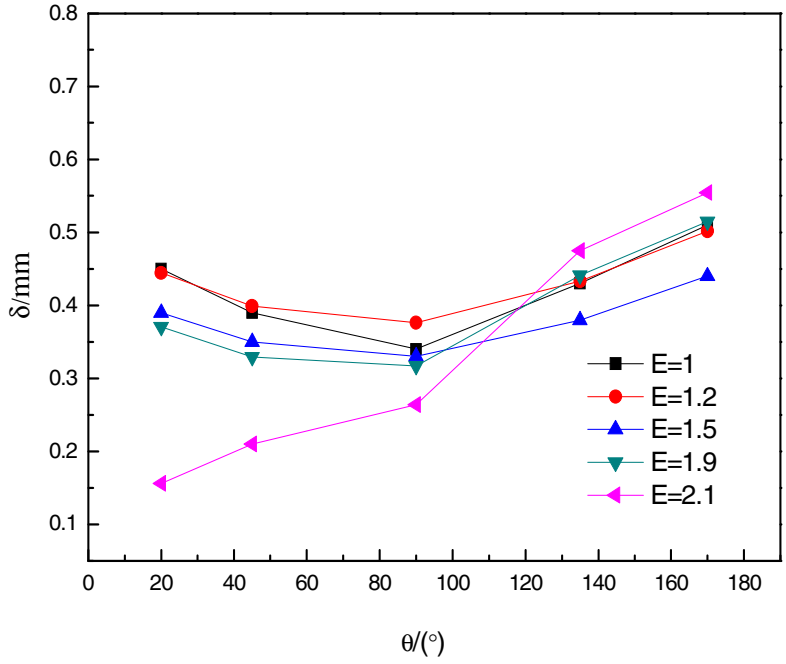


Fig. 4. The distribution of thickness of liquid film outside different elliptical tubes

From figure 3 one may see initial distribution and flowing stage of liquid film at surfaces of heat transfer tube. In most regions, the film at circular tubes is 10~20% thicker than that at elliptical tubes. Yet only at the point where the flow velocity of liquid is the biggest at the circumference (θ =90°), liquid film outside the circular tube is as thick as that outside the elliptical tube. When the film is to break away, its thickness reaches 3.6~3.8mm at the bottom of circular tubes, whereas it also reaches 2.7~2.9mm at the bottom of elliptical tubes. They are respectively 11 and 8 times that of θ =90°. Liquid film is thicker only at the lowest point of elliptical tubes. There exists no concentration of liquid film in other regions. The thickness of liquid film of elliptical tubes at all regions of circumference is smaller than that of circular tubes.

According to figure 4, the bigger the shape factor of cross-section E gets, the smaller the thickness of the film outside elliptical tubes gets. Yet as E gets bigger and bigger, the ill-distribution of the film shall get severer and severer. For example, when E=2.1, at the points of θ=170° and θ=20°, their thicknesses of liquid films outside elliptical tubes are 0.55mm and 0.16mm respectively. The two differ more than twice. The liquid film of E=1.5 is thinner and evener than that of E=1.2.

Simulation results demonstrate that when shape factor of elliptical tubes E=1.5, the liquid film outside the tube is in the best shape. It is also thin and even, which is good for heat transfer.

4 Experimental Studies

4.1 Experimental Devices

Experimental devices refer to the experimental platform of multi-functional distillation, seawater desalination and heat transfer. Such a platform is constituted with evaporimeter, electric boiler, seawater circulating pump, heat-exchange facility, and vacuum pump and metering installation of condensate, and other devices.

4.2 Experimental Results and Analysis

4.2.1 The Effect of Liquid Load on Overall Coefficient of Heat Transfer of Elliptical Tubes

Figure 5 demonstrates change rules of overall coefficient of heat transfer of elliptical tubes when E=1.5 as liquid load changes.

Figure 5 is the change curve. It shows that overall coefficients of heat transfer of elliptical tubes E=1.5 and circular tubes Φ=25 mm change as the liquid loads Γ vary, when difference in temperature of heat transfer is 3℃ and the temperature of evaporation is 70℃. Experimental results demonstrate that heat transfer coefficient of elliptical tube of heat transfer increases 20%~22% that of circular tube. According to figure 6, overall coefficient of heat transfer K first increases and then decreases and again increases as liquid load Γ increases. When K is beyond certain numerical value, the effect of liquid load on overall coefficients of heat transfer lessens. As the load gets bigger, the phenomenon of sputtering of liquid around heat transfer tube gets severer.

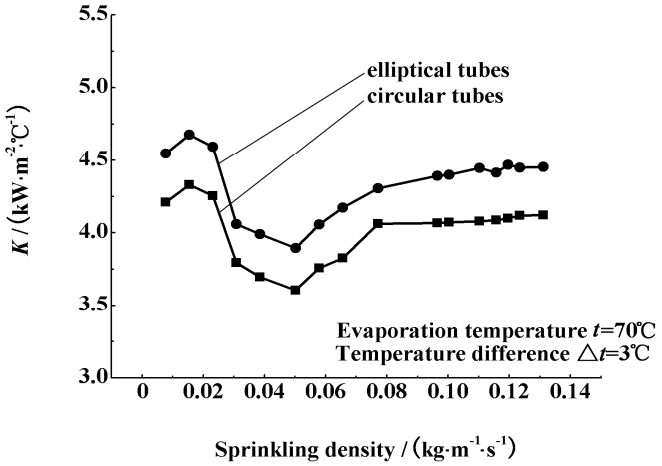


Fig. 5. The effect of liquid load on overall coefficient of heat transfer

4.2.2 The Distribution of Thickness of Liquid Film Outside Elliptical Tubes

The experiment inspects change of thickness of liquid film outside elliptical tubes and circular tubes, when the flow velocity at entrances is 0.8m/s. It is shown in figure 6.

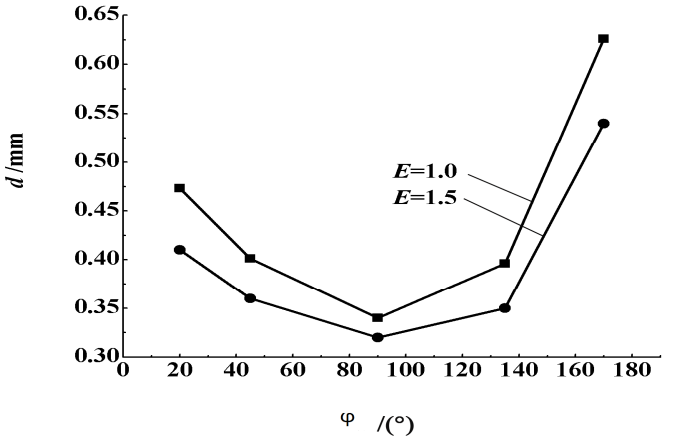


Fig. 6. Contrast of thickness of liquid film outside elliptical tubes and circular tubes

Figure 6 demonstrates the contrast of the experimental results and results of simulated calculation of thickness of liquid film outside elliptical tubes. It is found that the two results basically match. When φ<145°, results of simulated calculation almost match with the measured value. The most error is about 8%. When φ>145°, the measured value is slightly bigger than the calculated one, especially when the measured value of thickness of liquid film of the bottom half of tubes is about 25% more than the calculated value. In experiments, with the phase change process of evaporation, steam molecules exchange themselves constantly between inside and outside at vapor-liquid interface, thus producing steam pressure (reactive force of vapor). Since the

temperature of fluctuated surface is distributed unevenly, the effect of heat capillary force is produced. All these differences shall have some effect upon the stability of surface of liquid film. Basing on boundary-layer theory, Bankoff et al. [8-11] undertook stability analysis of the surface of liquid film under the conditions of evaporation or condensation. They solved the equation O-S under the two conditions. Considering the roles of such factors as reactive force of vapor, heat capillary force, Van der Waals force, quality change, and the effect of un-equilibrium thermodynamics, Bankoff [12] deduced evolution equation of wave with long wavelength of nonlinear and two-dimensional surface of ultrathin liquid film under the conditions of evaporation and condensation at horizontal walls.

5 Conclusion

The thesis first builds a physical model of falling film flowing outside elliptical tubes. It then conducts simulated researches on conditions of liquid flowing outside the tubes. The thesis finds that compared with circular tube, the consumed time of liquid flowing in elliptical tubes is reduced by 40ms. Thickness of liquid film outside elliptical tubes is 10~20% thinner than that outside circular tubes. According to the simulated results, one determines that elliptical tubes of E=1.5 possess some features. For example, liquid film is thin and well-distributed. When liquid flows at their surfaces, it is easier to get columnar, which is good for heat transfer.

The simulated calculation of thickness of liquid film outside elliptical tube basically matches with experimental results. The error is about 8%. It may demonstrate that the model is correct and that simulated calculation is feasible.

One adopts experimental ways to compare the performance of heat transfer of traditional circular tube with that of elliptical tube of E=1.5. The experimental results demonstrate that heat transfer coefficient of elliptical tube of heat transfer increases by 20%~22% compared with circular tubes. It may display that intensified heat transfer of elliptical tubes is obvious.

Acknowledgements. A special fund of Development and Research of Technology of Scientific Research Institutions (No.2013EG149277). A special fund of Operating Expense of Fundamental Scientific Researches of Scientific Research Institutions of Public Welfare (No. K-JBYWF-2013-T1).

References

1. Qi, C.-H., Li, Y., Yu, Y.-J., Feng, H.-J., Zhao, H.-L.: Tests and researches of heat transfer of falling film evaporation and seawater desalination of horizontal tubes. Chemical Industry and Project 28(5), 1–5 (2011)
2. Liu, X.-N., et al.: Numerical simulation of forced-convection heat transfer of staggered oval tube. Refrigeration & Air-Conditioning (1), 16–19 (2004)
3. Qingling, L., Ma, L., Qi, S., et al.: Experimental investigation on heat transfer and f low resistance of an 1.6/1.0 elliptic cylinder in cross-flow. Journal of Qingdao University of Science and Technology 25(1), 54–61 (2004)

4. Li, Q., Qi, S., Zhang, X., et al.: Heat transfer characteristics and flow behaviors of elliptic cylinders in cross-flow. Journal of Qingdao University of Science and Technology 25(5), 434–440 (2004)
5. Saboya, S.M., Saboya, F.E.M.: Experiments on elliptic sections in one-and tow-row arrangements of plate fin and tube heat exchangers. Experimental Thermal and Fluid Science 24, 67–75 (2001)
6. Hasan, A., Siren, K.: Performance investigation of plain and finned tube evaporative cooled heat exchangers. Applied Thermal Engineering 23(3), 3252340 (2003)
7. Chen, Z.-Y., et al.: Research on heat transfer and boiling heat transfer outside horizontal perforated elliptical tubes. Chemical Engineering (8), 35–40 (1990)
8. Bankoff, S.G.: Stability of liquid flow down a heated inclined plane. Int. J. Heat Mass Transfer 14, 377–385 (1971)
9. Bankoff, S.G.: Dynamics and stability of thin heated liquid films. Tran. ASME, J. Heat Transfer 112, 538–546 (1990)
10. Marschall, E., Lee, C.Y.: Stability of condensate flow down a vertical wall. Int. J. Heat Mass Transfer 16, 44–48 (1973)
11. Mintz, Y.: University of California Report AFCRL 690 (1961)
12. Wilkes, J.O.: Ph.D. thesis, University of Michigan (1963)
13. Hellums, J.D., Churchill, S.W.: International Heat Transfer Conference, Boulder, Colorado (1961)
14. Clark, J.A., Barakat, H.Z.: University of Michigan, College of Engineering. Technical Report 1 (1964)
15. Deardorff, J.W.: J. Atmospheric Sci. 21, 419 (1964)
16. Pearson, C.E.: Sperry Rand Research Report SRRC-RR-64-17 (1964)
17. Greenspan, D., Jain, P.C., Manohar, R., Noble, B., Sakurai, A.: University of Wisconsin, Mathematics Research Center. Technical Summary Report 482 (1964)
18. Fromm, E.: Los Alamos Scientific Laboratory Report LA-2910 (1963)
19. Harlow, F.H.: Los Alamos Scientific Laboratory Report LAMS-2452 (1960)
20. Santaló, L.A.: Introduction to Integral Geometry. Hermann & Cie., Paris (1953)
21. Martin, J.C., Moyce, W.J.: Phil. Trans. Roy. Soc (London) A244, 312 (1952)
22. Stoker, J.J.: Water Waves, ch. 10. Interscience Publishers, Inc., New York (1957)
23. Whitham, G.B.: Proc. Roy. Soc (London) A227, 399 (1955)
24. Harlow, F.H., Welch, J.E.: (to be published)
25. Harlow, F.H., Shannon, J.P., Welch, J.E.: Science 149, 1092 (1965) [MEDLINE]
26. Welch, J.E., Harlow, F.H., Shannon, J.P., Daly, B.J.: Los Alamos Scientific Laboratory Report LA-3425 (1965)
27. Spindler, B.: Linear stability of liquid films with interfacial phase change. Int. J. Heat Mass Transfer 25(2), 161–173 (1982)
28. López, P.G., Bankoff, S.G., Miksis, M.J.: Non-isothermal spreading of a thin liquid film on an inclined plane. J. Fluid Mech. 324, 261–286 (1996)
29. Alder, B., Fernbach, S., Rotenberg, M. (eds.): Methods in Computational Physics, vol. 3. Academic Press Inc., New York (1964)

On the Type of Multiple Taylor Series and Random Taylor Series

Wanchun Lu and Youhua Peng

Mathematics Department, Pingxiang College, Pingxiang, China, 337000

Abstract. It is studied that type of entire functions of multiple Taylor series and multiple random Taylor series. The characterization of type of entire functions of multiple Taylor series and multiple random Taylor series has obtained in terms of their Taylor's series coefficients.

Keywords: Polar coordinates, multiple Taylor series, random multiple Taylor series, type.

1 Introduction

Let $z \in C^n$,

$$z = (z_1, z_2, \cdots, z_n) = re^{i\theta} = (r_1 e^{i\theta_1}, r_2 e^{i\theta_2}, \cdots, r_n e^{i\theta_n}),$$

where $r = (r_1, r_2, \cdots, r_n) \in [0,+\infty)^n$, $\theta = (\theta_1, \theta_2, \cdots, \theta_n) \in [0,2\pi]^n$.

Represent $r = (r_1, r_2, \cdots, r_n) \in [0,+\infty)^n$, we have $r = \rho\Phi(\varphi)$, that is

$$r_1 = \rho\Phi_1(\varphi) = \rho\cos\varphi_1,$$

$$r_2 = \rho\Phi_2(\varphi) = \rho\sin\varphi_1\cos\varphi_2,$$

$$r_{n-1} = \rho\Phi_{n-1}(\varphi) = \rho\sin\varphi_1\sin\varphi_2\cdots\sin\varphi_{n-2}\cos\varphi_{n-1},$$

$$r_n = \rho\Phi_n(\varphi) = \rho\sin\varphi_1\sin\varphi_2\cdots\sin\varphi_{n-2}\sin\varphi_{n-1}.$$

where $r = \|z\| = \sqrt{r_1^2 + r_2^2 + \cdots + r_n^2}$, $\Phi = (\Phi_1, \Phi_2, \cdots, \Phi_n)$ and $\varphi = (\varphi_1, \varphi_2, \cdots, \varphi_n) \in [0, \frac{\pi}{2}]^{n-1}$. Hence $z \in C^n$ can be represented by polar coordinates (ρ, φ, θ) in C^n:

$$z = \rho\Phi(\varphi)e^{i\theta} = (\rho\Phi_1(\varphi)e^{i\theta_1}, \rho\Phi_2(\varphi)e^{i\theta_2}, \cdots, \rho\Phi_n(\varphi)e^{i\theta_n}).$$

Y. Yang, M. Ma, and B. Liu (Eds.): ICICA 2013, Part II, CCIS 392, pp. 584–594, 2013.

2 Type of Multiple Taylor Series

Consider an n-tuple Taylor series [1-4]

$$\sum_{\alpha} a_{\alpha} z^{\alpha} = \sum_{\alpha} a_{\alpha_1,\alpha_2,\cdots\alpha_n} z_1^{\alpha_1} z_2^{\alpha_2} \cdots z_n^{\alpha_n} \tag{1}$$

where $\alpha = (\alpha_1, \alpha_2, \cdots, \alpha_n) \in N^n$, Rearrange the terms of the series (1) and represents z by polar coordinates in C^n as follows:

$$\sum_{q=0}^{+\infty} \sum_{|\alpha|=q} a_{\alpha} z^{\alpha} = \sum_{q=0}^{+\infty} P_q(z) = \sum_{q=0}^{+\infty} P_q(\rho \Phi(\varphi) e^{i\theta})$$

$$= \sum_{q=0}^{+\infty} (\sum_{|\alpha|=q} a_{\alpha_1,\alpha_2,\cdots\alpha_n} \Phi_1^{\alpha_1}(\varphi) \Phi_2^{\alpha_2}(\varphi) \cdots \Phi_n^{\alpha_n}(\varphi) e^{i(\alpha_1\theta_1+\alpha_2\theta_2+\cdots\alpha_n\theta_n)}) \rho^q$$

$$= \sum_{q=0}^{+\infty} (\sum_{|\alpha|=q} a_{\alpha} \Phi^{\alpha}(\varphi) e^{i<\alpha,\theta>}) \rho^q$$

$$= \sum_{q=0}^{+\infty} P_q(\Phi(\varphi) e^{i\theta}) \rho^q \tag{2}$$

Where $q \in N$, $|\alpha| = \alpha_1 + \alpha_2 + \cdots + \alpha_n$, and $<\alpha, \theta> = \alpha_1\theta_1 + \alpha_2\theta_2 + \cdots \alpha_n\theta_n$.

Let

$C_q(\varphi, \theta) = | P_q(\Phi(\varphi) e^{i\theta}) |$,

$C_q^{(1)}(\varphi) = \max\{C_q(\varphi, \theta) : \theta \in [0, 2\pi]^n\}$,

$C_q^* = \max\{C_q^{(1)}(\varphi) : \varphi \in [0, \frac{\pi}{2}]^{n-1}\}$,

$B_q(\varphi) = \sum_{|\alpha|=q} | a_{\alpha} | \Phi^{\alpha}(\varphi)$,

$B_q^{(1)}(\varphi) = \max\{| a_{\alpha} | \Phi^{\alpha}(\varphi) : |\alpha| = q\}$,

$B_q^* = \max\{B_q^{(1)}(\varphi) : \varphi \in [0, \frac{\pi}{2}]^{n-1}\}$, $A_q = \max\{| a_{\alpha} | : |\alpha| = q\}$,

then [5-8], we have

Lemma 1[9-11]

$$B_q^{(1)}(\varphi) \le B_q(\varphi) \le (q+1)^n B_q^{(1)}(\varphi), \tag{3}$$

$$B_q^{(1)}(\varphi) \le C_q^{(1)}(\varphi) \le (q+1)^n B_q^{(1)}(\varphi), \tag{4}$$

$$A_q n^{\frac{-q}{2}} \le B_q^* \le C_q^* \le (q+1)B_q^* \le (q+1)A_q, \tag{5}$$

$$\overline{\lim_{q\to+\infty}} \sqrt[q]{C_q(\varphi,\theta)} \le \overline{\lim_{q\to+\infty}} \sqrt[q]{C_q^{(1)}(\varphi)} = \overline{\lim_{q\to+\infty}} \sqrt[q]{B_q^{(1)}(\varphi)} = \overline{\lim_{q\to+\infty}} \sqrt[q]{B_q(\varphi)},$$

$$\frac{1}{\sqrt{2}} \overline{\lim_{q\to+\infty}} \sqrt[q]{A_q} \le \overline{\lim_{q\to+\infty}} \sqrt[q]{C_q^*} = \overline{\lim_{q\to+\infty}} \sqrt[q]{B_q^*} \le \overline{\lim_{q\to+\infty}} \sqrt[q]{A_q}.$$

From [1], the function $f(z)$ defined by (1) is entire in the case $\rho_b = +\infty$, where $\rho_b = (\overline{\lim_{q\to+\infty}} \sqrt[q]{B_q^*})^{-1}$.

Consider the simple Taylor series

$$\psi_1(w,\varphi) = \sum_{q=0}^{+\infty} B_q^{(1)} w^q,$$

$$\psi_2(w) = \sum_{q=0}^{+\infty} B_q^* w^q,$$

$$\psi_3(w,\varphi) = \sum_{q=0}^{+\infty} (q+1)^n B_q^{(1)} w^q,$$

$$\psi_4(w) = \sum_{q=0}^{+\infty} (q+1)^n B_q^* w^q, (w \in C).$$

Let

$$M(\rho,\varphi,\psi_j) = \max\{|\psi_j(w,\varphi)| : |w| = \rho) \ (\varphi \in [0,\frac{\pi}{2}]^{n-1}; j = 1,3),$$

$$M(\rho,\psi_k) = \max\{|\psi_k(w)| : |w| = \rho) \ (k = 2,4),$$

$$m(\rho,\varphi,\psi_1) = \max\{B_q^{(1)}(\varphi)\rho^q : q \in N\},$$

$$m(\rho,\psi_2) = \max\{B_q^* \rho^q : q \in N\},$$

By [1], we have

Lemma2[12, 13]. $m(\rho,\varphi,\psi_1) \le M(\rho,\varphi,f) \le M(\rho,\varphi,\psi_3)$,

$m(\rho,\psi_2) \le M(\rho,f) \le M(\rho,\psi_4)$.

Lemma 3[1]. If the series (1) defines an entire function $f(z)$, then

$$\overline{\lim_{\rho\to\infty}} \frac{\ln\ln M(\rho,\varphi,f)}{\ln\rho} = \mu(\varphi) \Leftrightarrow \overline{\lim_{q\to\infty}} \frac{\ln B_q^{(1)}(\varphi)}{q\ln q} = -\frac{1}{\mu(\varphi)},$$

and

$$\overline{\lim_{\rho\to+\infty}} \frac{\ln\ln M(\rho,f)}{\ln\rho} = \mu \Leftrightarrow \overline{\lim_{q\to+\infty}} \frac{\ln B_q^*}{q\ln q} = \overline{\lim_{q\to+\infty}} \frac{\ln A_q}{q\ln q} = -\frac{1}{\mu}.$$

Where

$$M(\rho,\varphi,f)=\max\{|f(\rho\Phi(\varphi)e^{i\theta})|:\theta\in[0,2\pi]^n\}(\rho\geq 0,\varphi\in[0,\frac{\pi}{2}]^{n-1}),$$

$$M(\rho,f)=\max\{|f(\rho\Phi(\varphi)e^{i\theta})|:\theta\in[0,2\pi]^n,\varphi\in[0,\frac{\pi}{2}]^{n-1}\}(\rho\geq 0).$$

Lemma 4[6]. If entire function $\psi_2(w)$ satisfies

$$\overline{\lim_{\rho\to\infty}}\frac{\log^+\log M(\rho,\psi_2)}{\log\rho}=\mu\quad(0<\mu<\infty),$$

then

$$\overline{\lim_{\rho\to\infty}}\frac{\log^+\log^+ M(\rho,\psi_2)}{\log U(\rho)}=1\Leftrightarrow\overline{\lim_{q\to\infty}}\frac{\log q}{\log U(e^{\frac{-\log B_q^*}{q}})}=1\quad(A_n^*\neq 0),$$

where $U(\rho)=\rho^{\mu}$.

Lemma 5[6]. If the series (1.3) defines an entire function $\psi_2(w)$ and satisfies

$$\overline{\lim_{\rho\to\infty}}\frac{\log^+\log^+ M(\rho,\psi_2)}{\log\rho}=\infty,$$

Then

$$\overline{\lim_{\rho\to\infty}}\frac{\log^+\log^+ M(\rho,\psi_2)}{\log U(\rho)}=1\Leftrightarrow$$

$$\overline{\lim_{q\to\infty}}\frac{\log q}{\log U(e^{\frac{-\log B_q^*}{q}})}=\overline{\lim_{q\to\infty}}\frac{\log q}{\log U(e^{\frac{-\log A_q}{q}})}=1\ (A_n,B_n^*\neq 0),$$

where $U(\rho)=\rho^{\mu(\rho)}$, $\mu(\rho)$ is positive, increasing, differentiable and satisfies

(i) $\lim_{\rho\to\infty}\mu(\rho)=\infty$,

(ii) $\lim_{\rho\to\infty}\frac{\log U(\overline{\rho})}{\log U(\rho)}=1$, $\overline{\rho}=(1+\frac{1}{\log U(\rho)})\rho$,

(iii) $\overline{\lim_{\rho\to\infty}}\frac{\log^+\log^+ M(\rho,f)}{\log U(\rho)}=1$.

Theorem 1. If the series (1) defines an entire function $f(z)$, then

$$\overline{\lim_{\rho\to+\infty}}\frac{\ln M(\rho,\varphi,f)}{\rho^{\mu(\varphi)}}=\sigma(\varphi)\Leftrightarrow\frac{1}{\mu(\varphi)e}\overline{\lim_{q\to\infty}}q\sqrt[q]{(B_q^{(1)}(\varphi))^{\mu(\varphi)}}=\sigma(\varphi)$$

and

$$\overline{\lim_{\rho\to+\infty}}\frac{\ln M(\rho,f)}{\rho^{\mu}}=\sigma \Leftrightarrow \frac{1}{\mu e}\overline{\lim_{q\to\infty}} q\sqrt[q]{(B_q^*)^{\mu}}=\frac{1}{\mu e}\overline{\lim_{q\to\infty}} q\sqrt[q]{A_q{}^{\mu}}=\sigma .$$

Proof By some known results on entire function in C and Lemma 2, we have

$$\overline{\lim_{\rho\to+\infty}}\frac{\ln M(\rho,\varphi,\psi_1)}{\rho^{\mu(\varphi)}}=\overline{\lim_{\rho\to+\infty}}\frac{\ln m(\rho,\varphi,\psi_1)}{\rho^{\mu(\varphi)}}$$

$$\le \overline{\lim_{\rho\to+\infty}}\frac{\ln M(\rho,\varphi,f)}{\rho^{\mu(\varphi)}}\le \overline{\lim_{\rho\to+\infty}}\frac{\ln M(\rho,\varphi,\psi_3)}{\rho^{\mu(\varphi)}} \tag{6}$$

and

$$\overline{\lim_{\rho\to+\infty}}\frac{\ln M(\rho,\psi_2)}{\rho^{\mu}}=\overline{\lim_{\rho\to+\infty}}\frac{\ln m(\rho,\psi_2)}{\rho^{\mu}}$$

$$\le \overline{\lim_{\rho\to+\infty}}\frac{\ln M(\rho,f)}{\rho^{\mu}}\le \overline{\lim_{\rho\to+\infty}}\frac{\ln M(\rho,\psi_4)}{\rho^{\mu}}. \tag{7}$$

For entire functions $\psi_1(w,\varphi)$ and $\psi_3(w,\varphi)$, we have

$$\overline{\lim_{\rho\to+\infty}}\frac{\ln M(\rho,\varphi,\psi_1)}{\rho^{\mu(\varphi)}}=\frac{1}{\mu(\varphi)e}\overline{\lim_{q\to\infty}} q\sqrt[q]{(B_q^{(1)}(\varphi))^{\mu(\varphi)}}$$

$$=\frac{1}{\mu(\varphi)e}\overline{\lim_{q\to+\infty}} q\sqrt[q]{((q+1)^n B_q^{(1)}(\varphi))^{\mu(\varphi)}}=\overline{\lim_{\rho\to+\infty}}\frac{\ln M(\rho,\varphi,\psi_3)}{\rho^{\mu(\varphi)}}.$$

By (6) and Lemma 3, we have

$$\overline{\lim_{\rho\to+\infty}}\frac{\ln M(\rho,\varphi,f)}{\rho^{\mu(\varphi)}}=\sigma(\varphi) \Leftrightarrow \frac{1}{\mu(\varphi)e}\overline{\lim_{q\to\infty}} q\sqrt[q]{(B_q^{(1)}(\varphi))^{\mu(\varphi)}}=\sigma(\varphi).$$

For entire functions $\psi_2(w)$ and $\psi_4(w)$

$$\overline{\lim_{\rho\to+\infty}}\frac{\ln M(\rho,\psi_2)}{\rho^{\mu}}=\frac{1}{\mu e}\overline{\lim_{q\to+\infty}} q\sqrt[q]{(B_q^*)^{\mu}}$$

$$=\frac{1}{\mu e}\overline{\lim_{q\to+\infty}} q\sqrt[q]{((q+1)^n B_q^*)^{\mu}}=\overline{\lim_{\rho\to+\infty}}\frac{\ln M(\rho,\psi_4)}{\rho^{\mu}}.$$

By (5), Lemma 3 and (7), we have

$$\overline{\lim_{\rho\to+\infty}}\frac{\ln M(\rho,f)}{\rho^{\mu}}=\sigma \Leftrightarrow \frac{1}{\mu e}\overline{\lim_{q\to\infty}} q\sqrt[q]{(B_q^*)^{\mu}}=\frac{1}{\mu e}\overline{\lim_{q\to\infty}} q\sqrt[q]{A_q{}^{\mu}}=\sigma .$$

Theorem 2. If the series (1) defines an entire function $f(z)$ and satisfies

$$\overline{\lim_{\rho\to\infty}}\frac{\log^+\log^+ M(\rho,f)}{\log\rho}=\mu \quad (0<\mu<\infty),$$

then

$$\overline{\lim_{\rho\to\infty}}\frac{\log^+\log^+ M(\rho,f)}{\log U(\rho)}=1\Leftrightarrow$$

$$\overline{\lim_{q\to\infty}}\frac{\log q}{\log U(e^{\frac{-\log B_q^*}{q}})}=\overline{\lim_{q\to\infty}}\frac{\log q}{\log U(e^{\frac{-\log A_q}{q}})}=1\ (A_n,B_n^*\neq 0), \tag{8}$$

where $U(\rho)=\rho^{\mu}$.

Proof: First, assume that

$$\overline{\lim_{\rho\to\infty}}\frac{\log^+\log^+ M(\rho,f)}{\log U(\rho)}=1. \tag{9}$$

By Lemma 2, we have

$$\overline{\lim_{\rho\to\infty}}\frac{\log^+\log^+ m(\rho,\psi_2)}{\log U(\rho)}\le 1.$$

By some known results on entire function in $\mathbb{C}$, we have

$$\overline{\lim_{\rho\to\infty}}\frac{\log^+\log^+ M(\rho,\psi_2)}{\log U(\rho)}=\overline{\lim_{\rho\to\infty}}\frac{\log^+\log^+ m(\rho,\psi_2)}{\log U(\rho)}\le 1$$

By Lemma 4, we get

$$\overline{\lim_{q\to\infty}}\frac{\log q}{\log U(e^{\frac{-\log B_q^*}{q}})}\le 1.$$

Assume that $\overline{\lim_{q\to\infty}}\dfrac{\log q}{\log U(e^{\frac{-\log B_q^*}{q}})}<1$, by Lemma 4

$$\overline{\lim_{\rho\to\infty}}\frac{\log^+\log^+ M(\rho,\psi_2)}{\log U(\rho)}<1.$$

Note that

$$\overline{\lim_{\rho\to\infty}}\frac{\log^+\log^+ M(\rho,\psi_2)}{\log U(\rho)}=\overline{\lim_{\rho\to\infty}}\frac{\log^+\log^+ M(\rho,\psi_4)}{\log U(\rho)}$$

and by Lemma 2, we have

$$\overline{\lim_{\rho\to\infty}}\frac{\log^+\log^+ M(\rho,f)}{\log U(\rho)}<1.$$

This contradiction with (9), so we have

$$\overline{\lim_{q\to\infty}}\frac{\log q}{\log U(e^{\frac{-\log B_q^*}{q}})}=1.$$

To prove the reverse process in (8), let $\overline{\lim_{\rho\to\infty}}\frac{\log^+\log^+ M(\rho,f)}{\log U(\rho)}=\alpha<1$. By the above proof we get

$$\overline{\lim_{q\to\infty}}\frac{\log q}{\log U(e^{\frac{-\log B_q^*}{q}})}=\alpha<1.$$

This is contradictory.

Theorem 3 If the series (1) defines an entire function $f(z)$ and satisfies

$$\overline{\lim_{\rho\to\infty}}\frac{\log^+\log^+ M(\rho,f)}{\log\rho}=\infty,$$

Then

$$\overline{\lim_{\rho\to\infty}}\frac{\log^+\log^+ M(\rho,f)}{\log U(\rho)}=1\Leftrightarrow$$

$$\overline{\lim_{q\to\infty}}\frac{\log q}{\log U(e^{\frac{-\log B_q^*}{q}})}=\overline{\lim_{q\to\infty}}\frac{\log q}{\log U(e^{\frac{-\log A_q}{q}})}=1\ (A_n,B_n^*\neq 0),\qquad(10)$$

where $U(\rho)=\rho^{\mu(\rho)}$ from Lemma 5.

Proof: First, assume that

$$\overline{\lim_{\rho\to\infty}}\frac{\log^+\log^+ M(\rho,f)}{\log U(\rho)}=1.\qquad(11)$$

By Lemma 2, we have

$$\overline{\lim_{\rho\to\infty}}\frac{\log^+\log^+ m(\rho,\psi_2)}{\log U(\rho)}\le 1.$$

By some known results on entire function in $\mathbb{C}$, we have

$$\overline{\lim_{\rho\to\infty}}\frac{\log^+\log^+ M(\rho,\psi_2)}{\log U(\rho)}=\overline{\lim_{\rho\to\infty}}\frac{\log^+\log^+ m(\rho,\psi_2)}{\log U(\rho)}\le 1$$

By Lemma 5, we get

$$\overline{\lim_{q\to\infty}}\frac{\log q}{\log U(e^{\frac{-\log B_q^*}{q}})}\le 1.$$

Assume that $\overline{\lim_{q\to\infty}}\dfrac{\log q}{\log U(e^{\frac{-\log B_q^*}{q}})}<1$, by Lemma 5

$$\overline{\lim_{\rho\to\infty}}\frac{\log^+\log^+ M(\rho,\psi_2)}{\log U(\rho)}<1.$$

Note that

$$\overline{\lim_{\rho\to\infty}}\frac{\log^+\log^+ M(\rho,\psi_2)}{\log U(\rho)}=\overline{\lim_{\rho\to\infty}}\frac{\log^+\log^+ M(\rho,\psi_4)}{\log U(\rho)}$$

and by Lemma 2, we have

$$\overline{\lim_{\rho\to\infty}}\frac{\log^+\log^+ M(\rho,f)}{\log U(\rho)}<1.$$

This contradiction with (11), so we have

$$\overline{\lim_{q\to\infty}}\frac{\log q}{\log U(e^{\frac{-\log B_q^*}{q}})}=1.$$

To prove the reverse process in (10), let $\overline{\lim_{\rho\to\infty}}\dfrac{\log^+\log^+ M(\rho,f)}{\log U(\rho)}=\alpha<1$. By the above proof we get

$$\overline{\lim_{q\to\infty}}\frac{\log q}{\log U(e^{\frac{-\log B_q^*}{q}})}=\alpha<1.$$

This is contradictory.

3 Type of Multiple Random Taylor Series

Consider a random n-tuple Taylor series corresponding to (1)

$$\sum_{\alpha} a_{\alpha} z^{\alpha} = \sum_{\alpha} a_{\alpha_1,\alpha_2,\cdots\alpha_n} X_{\alpha_1,\alpha_2,\cdots\alpha_n}(\omega) z_1^{\alpha_1} z_2^{\alpha_2} \cdots z_n^{\alpha_n}, \tag{12}$$

where $\{X(\omega)\}$ in a probability space (Ω, Λ, P), is a sequence of nondegenerate, symmetric and independent complex random variables of the same distribution and satisfying

$$0 < E(|X_m(\omega)|^2) = d^2 < +\infty .$$

Rearrange the terms of the series (12) and represent z by polar in C^n as in (2).

$$\sum_{q=0}^{+\infty} \sum_{|\alpha|=q} a_{\alpha} X_{\alpha}(\omega) z^{\alpha} = \sum_{q=0}^{+\infty} P_q(z,\omega) = \sum_{q=0}^{+\infty} P_q(\rho\Phi(\varphi)e^{i\theta},\omega)$$

$$= \sum_{q=0}^{+\infty} \Big(\sum_{|\alpha|=q} a_{\alpha_1,\alpha_2,\cdots\alpha_n} X_{\alpha_1,\alpha_2,\cdots\alpha_n}(\omega) \Phi_1^{\alpha_1}(\varphi)\Phi_2^{\alpha_2}(\varphi)\cdots\Phi_n^{\alpha_n}(\varphi) e^{i(\alpha_1\theta_1+\alpha_2\theta_2+\cdots\alpha_n\theta_n)} \Big) \rho^q$$

$$= \sum_{q=0}^{+\infty} \Big(\sum_{|\alpha|=q} a_{\alpha} X(\omega) \Phi^{\alpha}(\varphi) e^{i<\alpha,\theta>} \Big) \rho^q$$

$$= \sum_{q=0}^{+\infty} P_q(\Phi(\varphi)e^{i\theta},\omega)\rho^q \tag{13}$$

Let

$$C_q(\varphi,\theta,\omega) = |P_q(\Phi(\varphi)e^{i\theta},\omega)|,$$

$$C_q^{(1)}(\varphi,\omega) = \max\{C_q(\varphi,\theta,\omega) : \theta \in [0,2\pi]^n\},$$

$$C_q^{*}(\omega) = \max\{C_q^{(1)}(\varphi,\omega) : \varphi \in [0,\frac{\pi}{2}]^{n-1}\},$$

$$B_q(\varphi,\omega) = \sum_{|\alpha|=q} |a_{\alpha}||X_{\alpha}(\omega)|\Phi^{\alpha}(\varphi),$$

$$B_q^{(1)}(\varphi,\omega) = \max\{|a_{\alpha}||X_{\alpha}(\omega)|\Phi^{\alpha}(\varphi) : |\alpha| = q\},$$

$$B_q^{*}(\omega) = \max\{B_q^{(1)}(\varphi,\omega) : \varphi \in [0,\frac{\pi}{2}]^{n-1}\},$$

where $\omega \in \Omega$.

Lemma 6[14]. $\overline{\lim\limits_{q\to+\infty}}\sqrt[q]{C_q(\varphi,\theta,\omega)} \le \overline{\lim\limits_{q\to+\infty}}\sqrt[q]{C_q^{(1)}(\varphi,\omega)} = \overline{\lim\limits_{q\to+\infty}}\sqrt[q]{B_q(\varphi,\omega)}$

$= \overline{\lim\limits_{q\to+\infty}}\sqrt[q]{B_q^{(1)}(\varphi,\omega)} = \overline{\lim\limits_{q\to+\infty}}\sqrt[q]{B_q^{(1)}(\varphi)}$ *a.s.*,

and

$\overline{\lim\limits_{q\to+\infty}}\sqrt[q]{C_q^*(\omega)} = \overline{\lim\limits_{q\to+\infty}}\sqrt[q]{B_q^*(\omega)} = \overline{\lim\limits_{q\to+\infty}}\sqrt[q]{B_q^*}$ *a.s.*.

Lemma 7[1]. If the series (12) defines a random entire function $f(z,\omega)$, then

$$\overline{\lim_{\rho\to\infty}}\frac{\ln\ln M(\rho,\varphi,\omega,f)}{\ln\rho} = \mu(\varphi) \quad a.s. \quad \Leftrightarrow \overline{\lim_{q\to\infty}}\frac{\ln B_q^{(1)}(\varphi)}{q\ln q} = -\frac{1}{\mu(\varphi)}$$

and

$$\overline{\lim_{\rho\to+\infty}}\frac{\ln\ln M(\rho,\omega,f)}{\ln\rho} = \mu \quad a.s. \quad \Leftrightarrow \overline{\lim_{q\to+\infty}}\frac{\ln B_q^*}{q\ln q} = \overline{\lim_{q\to+\infty}}\frac{\ln A_q}{q\ln q} = -\frac{1}{\mu},$$

where

$$M(\rho,\varphi,\omega,f) = \max\{|f(\rho\Phi(\varphi)e^{i\theta},\omega)|:\theta\in[0,2\pi]^n\},$$

$$M(\rho,\omega,f) = \max\{|f(\rho\Phi(\varphi)e^{i\theta},\omega)|:\theta\in[0,2\pi]^n\},$$

$$(\rho\ge 0,\varphi\in[0,\frac{\pi}{2}]^{n-1},\omega\in\Omega).$$

Theorem 2. If the series (12) defines a random entire function $f(z,\omega)$, then

$$\overline{\lim_{\rho\to+\infty}}\frac{\ln M(\rho,\varphi,\omega,f)}{\rho^{\mu(\varphi)}} = \sigma(\varphi) \qquad a.s.$$

$$\Leftrightarrow \frac{1}{\mu(\varphi)e}\overline{\lim_{q\to\infty}}q\sqrt[q]{(B_q^{(1)}(\varphi))^{\mu(\varphi)}} = \sigma(\varphi)$$

and

$$\overline{\lim_{\rho\to+\infty}}\frac{\ln M(\rho,\omega,f)}{\rho^{\mu}} = \sigma \quad a.s. \Leftrightarrow \frac{1}{\mu e}\overline{\lim_{q\to\infty}}q\sqrt[q]{(B_q^*)^{\mu}} = \frac{1}{\mu e}\overline{\lim_{q\to\infty}}q\sqrt[q]{A_q{}^{\mu}} = \sigma.$$

Proof. By Theorem 1, we have

$$\overline{\lim_{\rho\to+\infty}}\frac{\ln M(\rho,\varphi,\omega,f)}{\rho^{\mu(\varphi)}} = \sigma(\varphi) \qquad a.s.$$

$$\Leftrightarrow \frac{1}{\mu(\varphi)e}\overline{\lim_{q\to\infty}}q\sqrt[q]{(B_q^{(1)}(\varphi,\omega))^{\mu(\varphi)}} = \sigma(\varphi) \ a.s.$$

and

$$\overline{\lim_{\rho\to+\infty}}\frac{\ln M(\rho,\omega,f)}{\rho^{\mu}}=\sigma \quad a.s. \Leftrightarrow \frac{1}{\mu e}\overline{\lim_{q\to\infty}} q\sqrt[q]{(B_q^*(\omega))^{\mu}}=\sigma \quad a.s.$$

By Lemma 6, we have

$$\overline{\lim_{q\to+\infty}} q\sqrt[q]{(B_q^{(1)}(\varphi,\omega))^{\mu(\varphi)}}=\overline{\lim_{q\to+\infty}} q\sqrt[q]{(B_q^{(1)}(\varphi))^{\mu(\varphi)}} \quad a.s.$$

$$\overline{\lim_{q\to\infty}} q\sqrt[q]{(B_q^*(\omega))^{\mu}}=\overline{\lim_{q\to\infty}} q\sqrt[q]{(B_q^*)^{\mu}}=\overline{\lim_{q\to\infty}} q\sqrt[q]{A_q{}^{\mu}} \quad a.s.$$

So, we obtain the conclusions of the theorem.

References

1. Yu, J.R.: Some properties of multiple Taylor Series and random Taylor Series. J. Acta Mathematica Scientia 26B, 568–576 (2006)
2. Yu, J.R.: Some remarks on holomorphic functions and Taylor Series in Cn. J. Acta Mathematica Scientia 28B, 721–726 (2008)
3. Yu, J.R.: Further applications of polar coordinates in Cn. J. Acta Mathematica Scientia 31B, 1–7 (2011)
4. Valiron, G.: Theory of Integral Functions. Chelsea, New York (1949)
5. Ya, L.B.: Lectures on Entire Functions. American Mathematical Society, New York (1996)
6. Gao, Z.S., Sun, D.C.: On the growth of Taylor series. J. Sys. Sci. & Math. Scis. 14, 73–80 (1994)
7. Kumar, S., Srivastava, G.S.: Maximum term and lower order of entire functions of several complex variables. J. Bulletin of Mathematical Analysis Applications 3, 156–164 (2011)
8. Yu, J.R.: The distribution of values of Taylor series. J. Acta Mathematica Scientia 32B, 1719–1726 (2012)
9. Gopala Krishna, J.: Maximum term of a power series in one and several complex variables. J. Pacific J. Math. 29, 609–621 (1969)
10. Seremeta, M.N.: On the connection between the growth of the maximum modulus of an entire function and the moduli of the coefficients of its power series expansion. J. Amer. Math. Soc. Transl. 288, 291–301 (1970)
11. Shah, S.M.: The maximum term of an entire series. J. Math. Student. 10, 80–82 (1942)
12. Shah, S.M.: The maximum term of an entire series IV. J. Quart. J. Math. Oxford Ser. 12, 112–116 (1950)
13. Sun, D.C.: Taylor series of zero order in unit circle. J. J. of Math. 19, 181–184 (1996)
14. Gao, Z.S., Sun, D.C.: The regular growth of Taylor series. J. Acta Mathematica Scientia. 17, 111–120 (1997)

Fuzzy Number Ranking Based on Combination of Deviation Degree and Centroid

Dan Wang, Guoli Zhang, and Hua Zuo

Department of Mathematics and Physics, North China Electric Power University, Baoding 071003, China,
wang_daan0318@163.com
Department of Mathematics and Physics, North China Electric Power University, Baoding 071003, China
zhanggusoli@ncepu.edu.com
Department of Mathematics and Physics, North China Electric Power University, Baoding 071003, China
zuohuaa_870317@126.com

Abstract. This paper first analyses the disadvantage of two methods of ranking fuzzy numbers based on deviation degree and centroid, then introduces the concept of height weight ratio. After then gives a new method of ranking fuzzy numbers, which can make up the disadvantage of the two methods above. Finally, some illustrative examples are given to verify the effectiveness of the new method.

Keywords: Fuzzy Numbers, Ranking, Deviation Degree, Centroid, Height Weight Ratio.

1 Introduction

Since Zedeh put forward the fuzzy set theory in the 1960s, the theory has been widely used in many fields. In recent yeas, a lot of achievements about the theory and method of ranking fuzzy numbers are obtained. Some scholars studied the problems from different views, and put forward some valuable ranking fuzzy numbers methods. Shuo Yan introduced a revised method for ranking fuzzy numbers using maximizing set and minimization set [1]. Yanjie Wang ranked the fuzzy numbers by possibility degree method, which formed the decision matrix by comparing the fuzzy numbers, and got the ranking vector to rank the fuzzy numbers [2]. Zhongxing Wang introduced a new method for ranking fuzzy numbers by integrating the centroid index [3]. Ali Mahmodi Nejad ranked the fuzzy numbers based on the areas on the left and right sides of fuzzy number [4]. Shanlin Gao considered the possibility degree and α - cut set, and used the distance of the possibility degree and the left ideal points to form a comprehensive priority index [5].

Yager was the first researcher in using the concept of the centroid, and he used horizontal coordinate of the centroid as the priority index. Ying-Ming Wang proposed that the vertical coordinate of the centroid was unreasonable in the method of Chu and Tsao, and got the revised vertical coordinate of the centroid [6]. Then YuJie Wang

Y. Yang, M. Ma, and B. Liu (Eds.): ICICA 2013, Part II, CCIS 392, pp. 595–604, 2013.

found that there existed disadvantage in the method of putting the product of the horizontal coordinate and vertical coordinate as the priority index [7]. Juan Zhao put forward the method of combining the absolute of the horizontal coordinate with the vertical coordinate of the centroid [8]. The greater the priority index, the greater the fuzzy number. But the geometrical significance of this method is not obvious, Zhongguang Fan put forward the method of combining the centroid with its deviation [9-11]. Ruili Xu introduce the concept of deviation degree of fuzzy number based on ideal axis, and getting the priority index by weighting left deviation degree and right deviation degree [12-15]. But in this method, when increasing the number of fuzzy numbers, there is different ranking result because of the change of ideal axis. In order to overcome the disadvantage, Zhongxing Wang added centroid as the stability factor. But, when left deviation degree is zero, this priority index is meaningless [16-20].

This paper introduces the concept of height weight ratio based on combining deviation degree and centroid, and gets a new priority index which can overcome the disadvantage in [21-24] that sort result will different because of the change of ideal axis. We also introduce e to overcome the condition of left deviation degree being zero in [11].

2 Preliminary Knowledge

Definition 1. Let A is a fuzzy number defined on real domain R, with the following membership function :

(1) f_A is piecewise continuous function defined on R to closed interval $[0,\omega]$;

(2) $f_A = 0$, $\forall x \in (-\infty, a_1]$;

(3) $f_A(x)$ is strictly monotone increasing on $[a_i, a_{i+1}], (1 \le i \le m-1)$;

(4) $f_A(x) = \omega$, $\forall x \in [a_m, b_1]$;

(5) $f_A(x)$ is strictly monotone decreasing on $[b_j, b_{j+1}], (1 \le j \le n-1)$;

(6) $f_A(x) = 0$, $\forall x \in [b_n, +\infty)$.

Where $\omega = \max_x f_A(x)$, $f_A(x) = \begin{cases} f_A^{L_i}(x), x \in [a_i, a_{i+1}], 1 \le i \le m-1 \\ f_A^{R_j}(x), x \in [b_j, b_{j+1}], 1 \le j \le n-1 \end{cases}$.

Because $f_A^{L_i}(x) : [a_i, a_{i+1}] \to [0,\omega] (1 \le i \le m-1)$ is continuous increasing and $f_A^{R_i}(x) : [b_j, b_{j+1}] \to [0,\omega] (1 \le j \le n-1)$ is continuous decreasing, the inverse functions of them are existed, denoted by $g_A^{L_i}$ and $g_A^{R_j}$. Then, $g_A^{L_i} : [f_A^{L_{i-1}}(a_i), f_A^{L_i}(a_{i+1})] \to [a_i, a_{i+1}]$ is continuous increasing and $g_A^{R_j} : [f_A^{R_j}(b_{j+1}), f_A^{R_{j-1}}(b_j)] \to [b_j, b_{j+1}]$ is continuous decreasing, and they are integrable.

3 Method of Ranking Fuzzy Numbers Based on Deviation Degree, Centroids and Height Weight Ratio

3.1 Method of Ranking Fuzzy Numbers Based On Deviation Degree and Centroid

Reference [10] gave the method of ranking fuzzy numbers based on deviation degree. On the basis of the concepts left deviation degree and right deviation degree of fuzzy number, a mapping is defined from fuzzy numbers set to real numbers, which makes every fuzzy number have a corresponding real number. Then the compare of fuzzy numbers can be turned into the compare real numbers.

Definition 2 [11]. Let $S=\{A_1,A_2,\cdots,A_T\}$ is a set of fuzzy numbers, and then the concepts left deviation degree and right deviation degree of fuzzy number A_t are defined as follows:

$$
\begin{aligned}
&{}_L(A_t)=\int_0^{f_{A_t}^{L_1}(a_2)}\left[g_{A_t}^{L_1}(y)-x_{\min}\right]dy+\int_{f_{A_t}^{L_1}(a_2)}^{f_{A_t}^{L_2}(a_3)}\left[g_{A_t}^{L_2}(y)-x_{\min}\right]dy+\ldots\\
&+\int_{f_{A_t}^{L_{m-2}}(a_{m-1})}^{\omega_t}\left[g_{A_t}^{L_{m-1}}(y)-x_{\min}\right]dy\\
&I_R(A_t)=\int_0^{f_{A_t}^{R_{n-1}}(b_{n-1})}\left[x_{\max}-g_{A_t}^{R_{n-1}}(y)\right]dy+\int_{f_{A_t}^{R_{n-1}}(b_{n-1})}^{f_{A_t}^{R_{n-2}}(b_{n-2})}\left[x_{\max}-g_{A_t}^{R_{n-2}}(y)\right]dy+\ldots\\
&+\int_{f_{A_t}^{R_1}(b_1)}^{\omega_t}\left[x_{\max}-g_{A_t}^{R_1}(y)\right]dy
\end{aligned}
$$

Where $D_t=\left\{x\middle| f_{A_t}(x)>0\right\}, D=\bigcup_{t=1}^{T}D_t$, $x_{\min}=\inf D$ is said to be left ideal axis, and $x_{\max}=\sup D$ is said to be right ideal axis.

Definition 3. Let $S=\{A_1,A_2,\cdots,A_T\}$ is a set of fuzzy numbers, I is a mapping defined from S to R^1 : $I(A_t)=\alpha I_L(A_t)-(1-\alpha)I_R(A_t)$, $\alpha\in[0,1]$. If $I(A_i)>I(A_j)$ ($I(A_i)=I(A_j)$, $I(A_i)<I(A_j)$), then A_i is called to be larger than (equal to, less than) A_j , and note that as $A_i>A_j$ ($A_i=A_j$, $A_i<A_j$).

From the definitions of $I_L(A_t)$ and $I_R(A_t)$, $I_L(A_t)$ is the area formed by left membership function and left ideal axis, $I_R(A_t)$ is the area formed by right membership function and right ideal axis.

We expect the larger fuzzy number has the properties of a larger distance with left ideal axis, which means a larger left deviation degree, and a smaller distance with right ideal axis, which means a smaller right deviation degree. α is risk attitude of decision maker. Under the same condition, if α is larger, then $\alpha I_L(A_t)$ is larger, $(1-\alpha)I_R(A)$ is smaller. We get the priority index by weighting $I_L(A_t)$ and $I_R(A_t)$. For certain α , a larger fuzzy number can be turned into a larger real

number through mapping I. Then with the compare of real numbers, we can get the order of fuzzy numbers.

The method of ranking fuzzy number relies on the determination of left ideal axis and right ideal axis. When the number of fuzzy numbers changes, left ideal axis and right ideal axis will change, and $I(A_t)$ is different under different ideal axis, which leads to the change of sort result of fuzzy numbers. This is unreasonable.

Example 1. Rank the fuzzy numbers $A_1=(1,2,3;1)$ and $A_2=\left(2,3,4;\frac{4}{5}\right)$ by the above method.

Suppose that left ideal axis is $x_{\min}=1-m$, and right ideal axis is $x_{\max}=4+n$, then we have:

$$I_L(A_1)=\frac{1}{2}+m,\ I_R(A_1)=\frac{3}{2}+n;$$

$$I(A_1)=\alpha(\frac{1}{2}+m)-(1-\alpha)(\frac{3}{2}+n)=(2+m+n)\alpha-(\frac{3}{2}+n);$$

$$I_L(A_2)=\frac{6}{5}+\frac{4}{5}m,\ I_R(A_2)=\frac{2}{5}+\frac{4}{5}n,$$

$$I(A_2)=\alpha(\frac{6}{5}+\frac{4}{5}m)-(1-\alpha)(\frac{2}{5}+\frac{4}{5}n)=(\frac{8}{5}+\frac{4}{5}m+\frac{4}{5}n)\alpha-(\frac{2}{5}+\frac{4}{5}n).$$

Compare the two fuzzy numbers A_1 and A_2:

$$\begin{aligned} I(A_1)-I(A_2)&=(2+m+n)\alpha-(\frac{3}{2}+n)-(\frac{8}{5}+\frac{4}{5}m+\frac{4}{5}n)\alpha+(\frac{2}{5}+\frac{4}{5}n)\\ &=(\frac{2}{5}+\frac{1}{5}m+\frac{1}{5}n)\alpha-(\frac{11}{10}+\frac{1}{5}n) \end{aligned} \tag{1}$$

For certain $\alpha=\frac{1}{2}$, there is $I(A_1)-I(A_2)=-\frac{9}{10}+\frac{1}{10}m-\frac{1}{10}n$.

Let $n=0$, if $m>9$, then $I(A_1)-I(A_2)>0$, $A_1>A_2$. If $m<9$, then $I(A_1)-I(A_2)<0$, $A_1<A_2$.

So there is different sort result by the changing of ideal axis, and this is unreasonable. In order to overcome the disadvantage, reference [11] gave the method of ranking fuzzy numbers, which based on the following definition of centroid.

Definition 4 [11]. The centroid of fuzzy number A_t is defined as follows:

$$M_t=\frac{\int_{-\infty}^{+\infty}xf_{A_t}(x)dx}{\int_{-\infty}^{+\infty}f_{A_t}(x)dx} \tag{2}$$

Based on the concepts of deviation degree, centroid is added as the stability factor, and the priority index is given as follows:

$$R(A_t)=\begin{cases}\dfrac{I_L(A_t)}{\lambda_t+\lambda_{\max}+I_R(A_t)}, M_{\max}\neq M_{\min}\\ \dfrac{I_L(A_t)}{1+I_R(A_t)}, M_{\max}=M_{\min}\end{cases} \tag{3}$$

Where M_t is the centroid of fuzzy number A_t, $M_{\max}=\max\{M_1,M_2,...,M_T\}$, $M_{\min}=\min\{M_1,M_2,...,M_T\}$, $\lambda_{\max}=M_{\max}-M_{\min}$ and $\lambda_t=M_{\max}-M_t$.

But there is also disadvantage in this method. From the priority index we can see that, if the left deviation degree of fuzzy number is zero, then the priority index has no significance, and we can not rank the fuzzy numbers.

Example 2. Rank fuzzy numbers $A_3=(1,1,3;1)$, $A_4=(1,1,3,4;1)$ by the above method.

Because $I_L(A_3)=I_L(A_4)=0$, there is $R(A_3)=R(A_4)$, which implies that A_3 is equal to A_4 in accordance with the above method. But intuitively A_3 is smaller than A_4, so the sort result is unreasonable.

3.2 Method of Ranking Fuzzy Numbers Based on Deviation Degree, Centroid and Height Weight Ratio

Definition 5. Let $S=\{A_1,A_2,\cdots,A_T\}$ is a set of fuzzy numbers, and height weight ratio is defined as follows:

$$\beta_t=\frac{\omega_t}{\omega_{\max}} \tag{4}$$

Where $\omega_t=\max_x f_{A_t}(x)$, $\omega_{\max}=\max\{\omega_1,\omega_2,...,\omega_T\}$.

On the basis of deviation degree and centroid, we introduce the concepts of left relative deviation degree, right relative deviation degree and relative centroid difference.

Definition 6. Let $S=\{A_1,A_2,\cdots,A_T\}$ is a set of fuzzy numbers, left relative deviation degree $J_L(A_t)$, right relative deviation degree $J_R(A_t)$, and relative centroid difference T_t are respectively defined by using formula (5)-(7):

$$J_L(A_t)=\frac{I_L(A_t)}{\omega_{\max}(x_{\max}-x_{\min})} \tag{5}$$

$$J_R(A_t)=\frac{I_R(A_t)}{\omega_{\max}(x_{\max}-x_{\min})} \tag{6}$$

$$T_t = \frac{M_{\max} - M_t}{M_{\max} - M_{\min}} \tag{7}$$

Where $I_L(A_t)$ is left deviation degree, $I_R(A_t)$ is right deviation degree, $x_{\max}$ and $x_{\min}$ are ideal axis, $\omega_t = \max_x f_{A_t}(x)$ and $\omega_{\max} = \max\{\omega_1, \omega_2, ..., \omega_T\}$.

Now, we introduce the new priority index based on deviation degree, centroid and height weight ratio:

$$S(A_t) = \begin{cases} \dfrac{\beta_t e^{J_L(A_t)}}{e^{T_t} + e^{J_R(A_t)}}, M_{\max} \neq M_{\min} \\ \dfrac{\beta_t e^{J_L(A_t)}}{1 + e^{J_R(A_t)}}, M_{\max} = M_{\min} \end{cases} \tag{8}$$

The above priority index reflects that if left deviation degree is larger and right deviation degree is small, then the priority index is larger.

Definition 7. Let A_1 and A_2 are two fuzzy numbers. If $S(A_1) \geq S(A_2)$, we say A_1 is not less than A_2, and note that as $A_1 \geq A_2$. Otherwise we say A_1 is less than A_2, and write that as $A_1 < A_2$.

From the definition above, we have the following properties:

(1) Completeness. $A_1 \geq A_2$, $A_1 \leq A_2$ at least one holds.

(2) Transitivity. If $A_1 \geq A_2$, $A_2 \geq A_3$, there is $A_1 \geq A_3$.

(3) Equivalence. If $A_1 \geq A_2$, $A_1 \leq A_2$, there is $A_1 = A_2$.

Where $A_1 \geq A_2$ means $A_1 > A_2$ or $A_1 = A_2$.

4 Experimental Study

In the following, we will give some examples to illustrate the effectiveness of the methods we introduce.

Example 3. Use the priority index of (8) to rank the fuzzy numbers in example 1.

$M_1 = 2, M_2 = 3$, $\beta_1 = 1, \beta_2 = \frac{4}{5}$, $I_L(A_1) = \frac{1}{2}, I_R(A_1) = \frac{3}{2}, I_L(A_2) = \frac{6}{5}, I_R(A_2) = \frac{2}{5}$,

$S(A_1) = \dfrac{e^{\frac{1}{6}}}{e^1 + e^{\frac{1}{2}}}, S(A_2) = \dfrac{\frac{4}{5}e^{\frac{2}{5}}}{e^0 + e^{\frac{2}{15}}}$, $S(A_1) < S(A_2)$, so there is $A_1 < A_2$.

From the figure 1of fuzzy numbers A_1 and A_2, we can see that the ranking result is reasonable by using the priority index we introduce.

Example 4. Use the priority index of (8) to rank the fuzzy numbers in example 2.

$M_3 = \frac{5}{3}, M_4 = \frac{34}{15}$, $S(A_3) = \frac{e^0}{e^1 + e^{\frac{2}{3}}}$, $S(A_4) = \frac{e^0}{e^0 + e^{\frac{1}{6}}}$, $S(A_3) < S(A_4)$, so there is $A_3 < A_4$. From the Fig. 2 of fuzzy numbers A_3 and A_4 , we can see that the ranking result is reasonable by using the priority index we introduce.

Example 5. Compare fuzzy numbers $A_5 = (1,4,7;1), A_6 = (2,4,5;1)$, $M_5 = 4, M_6 = \frac{11}{3}$,

$I_L(A_5) = \frac{3}{2}, I_R(A_5) = \frac{3}{2}, I_L(A_6) = 2, I_R(A_6) = \frac{5}{2}$, $S(A_5) = \frac{e^{\frac{1}{4}}}{e^1 + e^{\frac{1}{4}}}, S(A_6) = \frac{e^{\frac{1}{3}}}{e^0 + e^{\frac{5}{12}}}$,

$S(A_5) < S(A_6)$, so there is $A_5 < A_6$.

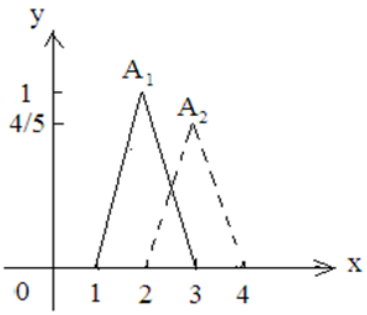


Fig. 1. Membership function image of A_1 and A_2

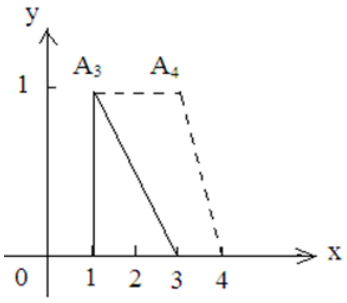


Fig. 2. Membership function image of A_3 and A_4

From the Fig. 3 of fuzzy numbers A_5 and A_6 , we can see that the ranking result is reasonable by using the priority index we introduce.

Example 6. Compare fuzzy numbers $A_7 = (1,2,3;1), A_8 = (4,5,6;\frac{1}{2})$, $M_7 = 2, M_8 = 5$, $\beta_7 = 1, \beta_8 = \frac{1}{2}$, $I_L(A_7) = \frac{1}{2}, I_R(A_7) = \frac{5}{2}, I_L(A_8) = \frac{7}{4}, I_R(A_8) = \frac{1}{4}$, $S(A_7) = \frac{e^{\frac{1}{10}}}{e^1 + e^{\frac{1}{2}}}, S(A_8) = \frac{\frac{1}{2}e^{\frac{7}{20}}}{e^0 + e^{\frac{1}{20}}}$, $S(A_7) < S(A_8)$, so there is $A_7 < A_8$.

From the figure 4 of fuzzy numbers A_7 and A_8, we can see that the ranking result is reasonable by using the priority index we introduce.

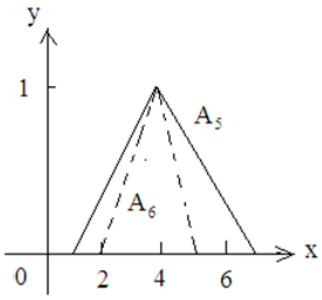


Fig. 3. Membership function image of A_5 and A_6

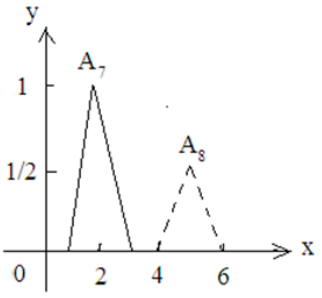


Fig. 4. Membership function image of A_7 and A_8

From the examples above, we can see that the priority index we introduce can rank the fuzzy numbers reasonably, so the method of ranking fuzzy numbers we introduce is effective.

5 Conclusion

This paper analyses the disadvantage of the methods of ranking fuzzy numbers based on deviation degree and centroid, introduces the concepts of left relative deviation degree, right relative deviation degree, relative centroid difference, and height weight

ratio, and gives a new method of ranking fuzzy numbers. Finally, some examples are given to prove the effectiveness of the methods we introduced.

Acknowledgments. This work is supported by Hebei Province Science Foundation (A2012502061).

References

1. Chou, S., Dat, L.Q., Yu, V.F.: A Revised Method for Ranking Fuzzy Numbers Using Maximizing Set and Minimization Set. Computer & Industrial Engineering 61, 1342–1348 (2011)
2. Wang, Y., Qian, W.: Possibility Degree Method for Ranking Triangular Fuzzy Numbers and Its Application. Journal of China Three Gorges University 32, 96–99 (2010)
3. Wang, Z., Gao, S., Li, J.: A New Method for Ranking Fuzzy Numbers. Guangxi Science 17, 97–99 (2010)
4. Nejad, A.M., Mashinchi, M.: Ranking the Fuzzy Numbers Based on the Areas on the Left and Right Sides of Fuzzy Number. Computer and Mathematics with Application 61, 431–442 (2011)
5. Gao, S., Li, J., Ruan, X.: A Method for Ranking Fuzzy Numbers Based on the Ideal Points. Journal of Shangdong University 44, 86–89 (2009)
6. Wang, Y., Yang, J., Xu, D., Chin, K.-S.: On the Centroids of Fuzzy Numbers. Fuzzy Sets and Systems 157, 919–926 (2006)
7. Wang, Y., Lee, H.-S.: The Revised Method of Ranking Fuzzy Numbers with an Area between the Centroid and Original Points. Computers and Mathematics with Applications 55, 2033–2042 (2008)
8. Zhao, J., Liu, Q.: Ranking Fuzzy Numbers Based on the Centroids of Fuzzy Numbers. Fuzzy Systems and Mathematics 22, 142–146 (2008)
9. Fan, Z., Li, J.: Ranking Method of Fuzzy Numbers Based on the Centroids and Deviation. Journal of Chongqing Institute of Technology 23, 152–157 (2009)
10. Xu, R., Xu, Z.: A New Method for Ranking Fuzzy Number. Mathematics in Practice and Theory 38, 111–120 (2008)
11. Wang, Z., Gao, S., Li, J.: A New Method for Ranking Fuzzy Numbers. Guangxi Science 17, 97–99 (2010)
12. Kim, K., Park, K.S.: Ranking fuzzy numbers with index of optimism. Fuzzy Sets Syst. 35, 143–150 (1990)
13. Chen, S.-H.: Ranking fuzzy numbers with maximizing set and minimizing set. Fuzzy Sets Syst. 17, 113–129 (1985)
14. Bortolan, G., Degani, R.: A review of some methods for ranking fuzzy subsets. Fuzzy Sets Syst. 15, 1–19 (1985)
15. Lee, K.M., Cho, C.H., Lee-Kwang, H.: Ranking fuzzy values with satisfaction function. Fuzzy Sets Syst. 64, 295–309 (1994)
16. Orlovsky, S.A.: Decision-making with a fuzzy preference relation. Fuzzy Sets Syst. 1, 155–167 (1978)
17. Liou, T.S., Wang, M.J.: Ranking fuzzy numbers with integral value. Fuzzy Sets Syst. 50, 247–255 (1992)
18. Fortemps, P., Roubens, M.: Ranking and defuzzification methods based on area compensation. Fuzzy Sets Syst. 82, 319–330 (1996)

19. Yoon, K.P.: A probabilistic approach to rank complex fuzzy numbers. Fuzzy Sets Syst. 80, 167–176 (1996)
20. Zimmermann, H.-J.: Fuzzy Set Theory and Its Applications (1985); Sengupta, D.-K.: Fuzzy preference and Orlovsky choice procedure. Fuzzy Sets Syst. 93, 231–234 (1998)
21. Lee, J.H., Lee-Kwang, H.: Comparison of fuzzy values on a continuous domain. Fuzzy Sets Syst.
22. Kaufmann, A., Gupta, M.M.: Introduction to Fuzzy Arithmetic: Theory and Applications. Van Nostrand Reinhold (1991)
23. Peneva, V., Popchev, I.: Comparison of clusters from fuzzy numbers. Fuzzy Sets Syst. 97, 75–81 (1998)
24. Cheng, C.H.: A new approach for ranking fuzzy numbers by distance method. Fuzzy Sets Syst. 95, 307–317 (1998)

Existence of Almost Periodic Solutions to a Class of Non-autonomous Functional Integro-differential Stochastic Equations

Lijie Li, Yu Feng, and Weiquan Pan

School of Mathematic and Information Sciences, Yulin Normal University, Yulin, 537000, China
239187472@qq.com
School of Statistics and Management, Shanghai University of Finance and Economics, Shanghai, 201900, China

Abstract. In this paper, a class of non-autonomous functional integro-differential stochastic equations in a real separable Hilbert space is studied. When the operators $A(t)$ satisfy Acquistapace-Terreni conditions, and with some suitable assumptions, the existence and uniqueness of a square-mean almost periodic mild solution to the equations are obtained.

Keywords: Stochastic differential equations, Square-mean almost periodic mild solution, Acquistapace-Terreni conditions.

1 Introduction

Recently, in [1], the concept of quadratic (or square) means almost periodicity was introduced and studied. And the paper is aimed at utilizing Schauder fixed point theorem to study the existence and uniqueness of a quadratic mean almost periodic solution to the class of stochastic differential equations

$$dX(t) = AX(t)dt + F(t, X(t))dt + G(t, X(t))dW(t).$$

Similarly, Bezandry [2], Fu [3] Chang [4, 5], Bezandry [6], and Chang [7] have considered the existence of quadratic (or square) mean almost periodic solutions to some autonomous stochastic differential equations.

Regarding to the research of almost periodic solutions to autonomous stochastic differential equation, the achievement is abundant, please consult the works mentioned above and the references therein .The research achievement of non-autonomous equations is relatively less to the autonomous. Bezandry [8] and Chang [9] have studied the existence and uniqueness of a quadratic mean almost periodic solution to non-autonomous semi-linear stochastic differential equations such as

$$dX(t) = A(t)X(t)dt + F(t, X(t))dt + G(t, X(t))dW(t),$$
$$dx(t) = A(t)x(t) + f(t, B_1x(t))dt + g(t, B_2x(t))dw(t),$$

Y. Yang, M. Ma, and B. Liu (Eds.): ICICA 2013, Part II, CCIS 392, pp. 605–614, 2013.

where $A(t)_{t\in R}$ is a family of densely closed linear operators satisfying Acquistapace-Terreni conditions.

Let $(H,\|\cdot\|)$ be a real (separable) Hilbert space. Motivated by the above mentioned works, the main purpose of this paper to deal with the existence and uniqueness of square-mean almost periodic solutions to the class of non-autonomous semi-linear functional stochastic integro-differential equation of the form:

$$\begin{aligned}\dot{X}(t) = A(t)X(t) + \int_{-\infty}^{t} C(t-u)G(u,X(u))dW(u) \\ + \int_{-\infty}^{t} B(t-u)F_2(u,X(u))du + F_1(t,X(t))\end{aligned} \tag{1}$$

where $A(t)$ for $t\in R$ is a family of densely defined closed linear operators satisfying the so-called Acquistapace-Terreni conditions [10], that is, there exist constants $\lambda_0 \geq 0,\ \theta\in\left(\frac{\pi}{2},\pi\right)$,

$L>0,\ K>0$ and $\alpha,\beta\in(0,1]$ with $\alpha+\beta>1$ such that

$$\Sigma_\theta \bigcup \{0\} \subset \rho(A(t)-\lambda_0); \| R(\lambda, A(t)-\lambda_0)\| \leq \frac{K}{1+|\lambda|} \tag{2}$$

and

$$\| (A(t)-\lambda_0)R(\lambda,A(t)-\lambda_0)\left[R(\lambda_0,A(t))-R(\lambda_0,A(s))\right]\| \leq L|t-s|^{\alpha}|\lambda|^{\beta}$$

for $t,s\in R, \lambda_0\in\Sigma_\theta := \{\lambda\in C-\{0\} : |\arg\lambda| \leq \theta\}$, where $R(\lambda,A) := (\lambda I - A)^{-1}$, $\rho(A) = \{\lambda\in C:$

$\lambda I - A$ is one to one ,and $(\lambda I - A)^{-1}$ is bounded$\}$; B and C are convolution-type kernels in $L^1(0,\infty)$ and $L^2(0,\infty)$, respectively, satisfying Assumptions 3.2 one page 361 in Kannan and Bharucha-Reid[11, 12]; $F_1, F_2 : R\times L^2(P,H)\to L^2(P,H)$ and $G: R\times L^2(P,H)\to L^2(P,L_2^0)$ are jointly continuous functions.

2 Preliminaries

For a symmetric nonnegative operator $Q\in L_2(K,H)$ with finite trace we assume that $\{W(t), t\in R\}$ is a Q-Wiener process defined on $(\Omega,\mathcal{F},P)$ and with values in K. Recall that W can obtained as follows: let $W_i(t),\ t\in R,\ i=1,2$, be independent K-valued Q-Wiener processes, then

$$W(t) = \begin{cases} W_1(t), & if \quad t\geq 0, \\ W_2(-t), & if \quad t\leq 0 \end{cases}$$

is Q-wiener process with R as time parameter. We let $\mathcal{F} = \sigma\{W(s), s\leq t\}$.

The collection of all strongly measurable, square-integrable H -valued random variables, denoted by $L^2(P,H)$, is a Banach space equipped with norm $\| X \|_{L^2(P,H)} = (E \| X \|^2)^{\frac{1}{2}}$, where the expectation E is defined by $E[g] = \int_\Omega g(\omega) dP(\omega)$.

Let $K_0 = Q^{\frac{1}{2}} K$ and $L_2^0 = L_2(K_0, H)$ with respect to the norm $\| \Phi \|_{L_2^0}^2 = \| \Phi Q^{\frac{1}{2}} \|_2^2 = \mathrm{Tr}(\Phi Q \Phi^*)$.

We suppose that the system

$$\begin{cases} u'(t) = A(t)u(t), \ t \geq s, \\ u(s) = x \in L^2(P;H), \end{cases}$$

Has an associated evolution family of operators $\{U(t,s) : t \geq s$ with $t, s \in R\}$, which is uniformly asymptotically stable.

If B_1, B_2 are Banach spaces, then the notation $\mathcal{L}(B_1, B_2)$ stands for the Banach space of bounded linear operators from B_1 into B_2 . When $B_1 = B_2$, this is simply denoted $\mathcal{L}(B_1)$.

Definition 1. A family of bounded linear operators $\{U(t,s) : t \geq s$ with $t, s \in R\}$ on $L^2(P;H)$ is called an evolution family of operators for (3) whenever the following conditions hold:

(a) $U(t,s)U(s,r) = U(t,r)$ for every $r \leq s \leq t$;

(b) For each $x \in X$ the function $(t,s) \to U(t,s)x$ is continuous and $U(t,s) \in \mathcal{L}(L^2(P;H), D)$ for every $t > s$;

(c) the function $(s,t] \to \mathcal{L}(L^2(P;H)), \ t \to U(t,s)$ is differentiable with

$$\frac{\partial}{\partial t} U(t,s) = A(t)U(t,s) .$$

Definition 2. [1] A continuous stochastic process $X : R \to L^2(P;B)$ is said to be square-mean almost periodic if for each $\varepsilon > 0$ there exists $l(\varepsilon) > 0$ such that any interval of length $l(\varepsilon)$ contains at least a number τ for which $\sup_{t \in R} E \| X(t+\tau) - X(t) \|^2 < \varepsilon.$

The collection of all stochastic processes $X : R \to L^2(P;B)$ which are square-mean almost periodic is then denoted by $AP(R; L^2(P;B))$.

The next lemma provides some properties of square-mean almost periodic processes.

Lemma 3. [1] If X belongs to $AP(R;L^2(P;B))$, then

(i) the mapping $t \to E\|X(t)\|^2$ is uniformly continuous;

(ii) there exists a constant $M > 0$ such that $E\|X(t)\|^2 \le M$, for all $t \in R$.

Let $CUB(R;L^2(P;B))$ denote the collection of all stochastic processes $X : R \to L^2(P;B)$, which are continuous and uniformly bounded. It is then easy to check that $CUB(R;L^2(P;B))$ is a Banach space when it is equipped with the norm: $\| X \|_\infty = \sup_{t\in R}\left(E \| X(t) \|^2\right)^{\frac{1}{2}}$.

Lemma 4. [1] $AP(R;L^2(P;B)) \subset CUB(R;L^2(P;B))$ is a closed subspace.

In view of the above, the space $AP(R;L^2(P;B))$ of square-mean almost periodic processes equipped with the norm $\|\cdot\|_\infty$ is a Banach space.

Let $(B_1,\|\cdot\|_{B_1})$ and $(B_2,\|\cdot\|_{B_2})$ be Banach spaces and let $L^2(P;B_1)$ and $L^2(P;B_2)$ be their corresponding L^2-spaces, respectively.

Definition 5. [1] A function $F : R\times L^2(P;B_1) \to L^2(P;B_2), (t,Y) \mapsto F(t,Y)$, which is jointly continuous, is said to be square-mean almost periodic in $t \in R$ uniformly in $Y \in K$, where $K \subset L^2(P;B_1)$ is a compact, if for any $\varepsilon > 0$, there exists $l(\varepsilon,K) > 0$ such that any interval of length $l(\varepsilon,K) > 0$ contains at least a number τ for which

$$\sup_{t\in R} E\|F(t+\tau,Y) - F(t,Y)\|^2_{B_2} < \varepsilon$$

for each stochastic process $Y : R \to K$.

Theorem 6. [1] Let $F : R\times L^2(P;B_1) \to L^2(P;B_2), (t,Y) \mapsto F(t,Y)$ be a square-mean almost periodic process in $t \in R$ uniformly in $Y \in K$, where $K \subset L^2(P;B_1)$ is compact. Suppose that F is Lipschitz in the following sense:

$$E \| F(t,Y) - F(t,Z) \|^2_{B_2} \le ME \| Y - Z \|^2_{B_1}$$

for all $Y,Z \in L^2(P;B_1)$ and for each $t \in R$, where $M > 0$. Then for any square-mean almost periodic process $\Phi : R \to L^2(P;B_1)$, the stochastic process $t \mapsto F(t,\Phi(t))$ is square-mean almost periodic.

3 Application to Stochastic Integro-Differential Equations

In this section we make extensive use of the results recalled in the previous section to study the existence and uniqueness of a square-mean almost periodic solution to (1).

Throughout the rest of this section, we impose the following conditions:

(H0) The operators $A(t), U(r,s)$ commute and that the evolution family $U(t,s)$ is asymptotically stable. Namely, there exist some constants $M,\delta > 0$ such that

$\|U(t,s)\| \le Me^{-\delta(t-s)}$

for every $t \ge s$. In addition, $R(\lambda_0, A(\cdot)) \in AP(R; \mathcal{L}(L^2(P,H)))$ for λ_0 in (1.2).

(H1) The function $F_i : R \times L^2(P;H) \to L^2(P;H), (t,X) \mapsto F_i(t,X)(i=1,2)$ be a square-mean almost periodic in $t \in R$ uniformly in $X \in \vartheta_i$ ($\vartheta_i \subset L^2(P;H)(i=1,2)$ being a compact subspace). Moreover, F is Lipschitz in the following sense: there exists $K_i > 0$ for which

$$E\|F_i(t,X) - F_i(t,Y)\|^2 \le K_i E\|X - Y\|^2$$

for all stochastic processes $X, Y \in L^2(P;H)$ and $t \in R$.

(H2) The function $G : R \times L^2(P;H) \to L^2(P;L_2^0), (t,X) \mapsto G(t,X)$ be a square-mean almost periodic in $t \in R$ uniformly in $X \in \vartheta_3$ ($\vartheta_3 \subset L^2(P;H)$ being a compact subspace). Moreover, G is Lipschitz in the following sense: there exists $K_3 > 0$ for which

$$E\|G(t,X) - G(t,Y)\|_{L_2^0}^2 \le K_3 E\|X - Y\|^2$$

for all stochastic processes $X, Y \in L^2(P;H)$ and $t \in R$.

Lemma 7. [11] Suppose $A(t)$ satisfies the Acquistapace-Terreni conditions, $U(t,s)$ is exponentially stable and $R(\lambda_0, A(\cdot)) \in AP(R; \mathcal{L}(L^2(P,H)))$. Let $h > 0$. Then, for any $\varepsilon > 0$, there exists $l(\varepsilon) > 0$ such that every interval of length $l(\varepsilon)$ contains at least a number τ with the property that $\| U(t+\tau, s+\tau) - U(t,s) \| \le \varepsilon e^{-\frac{\delta}{2}(t-s)}$ for all $t - s \ge h$.

Definition 8. A $\mathcal{F}_t$-progressively process $\{X(t)\}_{t \in R}$ is called a mild solution of (1.1) on R if

$$X(t) = U(t,s)X(s) + \int_s^t U(t,\sigma)\int_s^\sigma C(\sigma - u)G(u, X(u))dW(u)d\sigma$$
$$+ \int_s^t U(t,\sigma)\int_s^\sigma B(\sigma - u)F_2(u, X(u))dud\sigma + \int_s^t U(t,u)F_1(u, X(u))du$$

for each $t \ge s$ for each $s \in R$.

Theorem 9. Under assumptions (H0)-(H2), then (1.1) has a unique square-mean almost period mild solution, which can be explicitly expressed as follows:

$$X(t) = \int_{-\infty}^t U(t,\sigma)\int_{-\infty}^\sigma C(\sigma - u)G(u, X(u))dW(u)d\sigma$$
$$+ \int_{-\infty}^t U(t,\sigma)\int_{-\infty}^\sigma B(\sigma - u)F_2(u, X(u))dud\sigma + \int_{-\infty}^t U(t,u)F_1(u, X(u))du$$

Whenever $\Theta := \frac{3M^2}{\delta^2}\left(K_1 + K_2 \| B \|_{L^1_{(0,\infty)}}^2 + K_3 \| C \|_{L^2_{(0,\infty)}}^2 \mathbf{Tr}Q\right) < 1.$ (3.4)

Proof. First of all, note that

$$X(t)=\int_{-\infty}^{t}U(t,\sigma)\int_{-\infty}^{\sigma}C(\sigma-u)G(u,X(u))dW(u)d\sigma$$

$$+\int_{-\infty}^{t}U(t,\sigma)\int_{-\infty}^{\sigma}B(\sigma-u)F_2(u,X(u))dud\sigma+\int_{-\infty}^{t}U(t,u)F_1(u,X(u))du \quad (3.5)$$

is well defined at satisfies

$$X(t)=U(t,s)X(s)+\int_{s}^{t}U(t,\sigma)\int_{s}^{\sigma}C(\sigma-u)G(u,X(u))dW(u)d\sigma$$

$$+\int_{s}^{t}U(t,\sigma)\int_{s}^{\sigma}B(\sigma-u)F_2(u,X(u))dud\sigma+\int_{s}^{t}U(t,u)F_1(u,X(u))du$$

for all $t\geq s$ for each $s\in R$, and hence X given by (3.5) is a mild solution to (1.1).

Define

$$\Phi(t)=\int_{-\infty}^{t}U(t,u)F_1(u,X(u))du,$$

$$\Psi(t)=\int_{-\infty}^{t}U(t,\sigma)\int_{-\infty}^{\sigma}B(\sigma-u)F_2(u,X(u))dud\sigma,$$

$$\Lambda(t)=\int_{-\infty}^{t}U(t,\sigma)\int_{-\infty}^{\sigma}C(\sigma-u)G(u,X(u))dW(u)d\sigma.$$

Let us first show that $\Phi(\cdot)$ is square-mean almost periodic where X is. Indeed, assuming that X is square-mean almost periodic and using assumption (H1), Theorem 2.6, and Lemma 3.1, one can easily see that $u\mapsto F_1(u,X(u))$ is square-mean almost periodic. Therefore, given by $\varepsilon>0$, one can find $l_1(\varepsilon)>0$ such that any interval of length $l_1(\varepsilon)$ contains at least τ_1 for which

$$\|U(t+\tau_1,s+\tau_1)-U(t,s)\|\leq \varepsilon e^{-\frac{\delta}{2}(t-s)}$$

for all $t-s\geq\varepsilon$, and

$$E\|F_1(u+\tau_1,X(u+\tau_1))-F_1(u,X(u))\|^2<\eta_1(\varepsilon)$$

for each $u\in R$, where $\eta_1(\varepsilon)\to 0$ as $\varepsilon\to 0$.

Moreover, it follows from Lemma2.3 (ii) that there exists a positive constant N_1 such that

$$\sup_{u\in R}E\|F_1(u,X(u))\|^2\leq N_1.$$

Now, using assumption (H0) and *Hölder*'s inequality, we obtain

$$E\|\Phi X(t+\tau_1)-\Phi X(t)\|^2$$

$$\leq E\left(\int_{-\infty}^{t}\left\|U(t+\tau_1,u+\tau_1)F_1(u+\tau_1,X(u+\tau_1))-U(t,u)F_1(u,X(u))\right\|du\right)^2$$

$$\leq 3E\left(\int_{-\infty}^{t}\left\|U(t+\tau_1,u+\tau_1)\right\|\left\|F_1(u+\tau_1,X(u+\tau_1))-F_1(u,X(u))\right\|du\right)^2$$

$$+3E\left(\int_{t-\varepsilon}^{t}\left\|U(t+\tau_1,u+\tau_1)-U(t,u)\right\|\left\|F_1(u,X(u))\right\|du\right)^2$$

$$+3E\left(\int_{-\infty}^{t-\varepsilon}\left\|U(t+\tau_1,u+\tau_1)-U(t,u)\right\|\left\|F_1(u,X(u))\right\|du\right)^2$$

$$\leq 3\left(\int_{-\infty}^{t} Me^{-\delta(t-u)}du\right)^2 \sup_{u\in R} E\left\|F_1(u+\tau_1, X(u+\tau_1)-F_1(u,X(u))\right\|^2$$

$$+3\left(\int_{-\infty}^{t-\varepsilon} 2Me^{-\delta(t-u)}du\right)^2 \sup_{u\in R} E\left\|F_1(u,X(u))\right\|^2$$

$$+3\left(\int_{-\infty}^{t-\varepsilon} \varepsilon e^{-\frac{\delta}{2}(t-s)}du\right)^2 \sup_{u\in R} E\left\|F_1(u,X(u))\right\|^2$$

$$\leq 3M^2\frac{1}{\delta^2}\eta_1 + 12M^2\varepsilon^2 N_1 + 3\varepsilon^2\frac{1}{\delta}N_1,$$

where $\eta_1(\varepsilon)\to 0$ as $\varepsilon\to 0$.

In view of the above, $\Phi(\cdot)$ is square-mean almost periodic.

Similarly, assuming that X is square-mean almost periodic and using assumption (H1), Theorem 2.6, and Lemma 3.1, one can easily see that $u\mapsto F_2(u,X(u))$ is square-mean almost periodic. Therefore, given by $\varepsilon>0$, one can find $l_2(\varepsilon)>0$ such that any interval of length $l_2(\varepsilon)$ contains at least τ_2 for which

$$\|U(t+\tau_2,s+\tau_2)-U(t,s)\|\leq \varepsilon e^{-\frac{\delta}{2}(t-s)}$$

For all $t-s\geq\varepsilon$, and

$$E\|F_2(u+\tau_2,X(u+\tau_2))-F_2(u,X(u))\|^2<\eta_2(\varepsilon)$$

For each $u\in R$, where $\eta_2(\varepsilon)\to 0$ as $\varepsilon\to 0$.

Moreover, it follows from Lemma2.3 (ii) that there exists a positive constant N_2 such that

$$\sup_{u\in R} E\|F_2(u,X(u))\|^2\leq N_2.$$

Now, using assumption (H0) and *Hölder* 's inequality, we obtain

$$E\|\Psi X(t+\tau_2)-\Psi X(t)\|^2$$

$$\leq E\left(\int_{-\infty}^{t}\left\|U(t+\tau_2,\sigma+\tau_2)\int_{-\infty}^{\sigma}B(\sigma-u)F_2(u+\tau_2,X(u+\tau_2))du\right.\right.$$

$$\left.\left.-U(t,\sigma)\int_{-\infty}^{\sigma}B(\sigma-u)F_2(u,X(u))du\right\|\sigma\right)^2.$$

$$\leq 3E\left(\int_{-\infty}^{t}\left\|U(t+\tau_2,\sigma+\tau_2)\right\|\int_{-\infty}^{\sigma}\left\|B(\sigma-u)\right\|\left\|F_2(u+\tau_2,X(u+\tau_2))-F_2(u,X(u))\right\|du d\sigma\right)^2$$

$$+3E\left(\int_{t-\varepsilon}^{t}\left\|U(t+\tau_2,\sigma+\tau_2)-U(t,\sigma)\right\|\int_{-\infty}^{\sigma}\left\|B(\sigma-u)\right\|\left\|F_2(u,X(u))\right\|du d\sigma\right)^2$$

$$+3E\left(\int_{-\infty}^{t-\varepsilon}\left\|U(t+\tau_2,\sigma+\tau_2)-U(t,\sigma)\right\|\int_{-\infty}^{\sigma}\left\|B(\sigma-u)\right\|\left\|F_2(u,X(u))\right\|du d\sigma\right)^2$$

$$\leq 3\left(\int_{-\infty}^{t}Me^{-\delta(t-\sigma)}d\sigma\right)^2\left(\int_{-\infty}^{+\infty}\left\|B(u)\right\|du\right)^2\sup_{u\in R}\left\|F_2(u+\tau_2,X(u+\tau_2))-F_2(u,X(u))\right\|$$

$$+3\left(\int_{t-\varepsilon}^{t}2Me^{-\delta(t-\sigma)}d\sigma\right)^2\left(\int_{-\infty}^{+\infty}\left\|B(u)\right\|du\right)^2\sup_{u\in R}\left\|F_2(u,X(u))\right\|$$

$$+3\left(\int_{-\infty}^{t-\varepsilon}\varepsilon e^{-\frac{\delta}{2}(t-\sigma)}d\sigma\right)^2\left(\int_{-\infty}^{+\infty}\|B(u)\|du\right)^2\sup_{u\in R}\|F_2(u,X(u))\|$$

$$\le 3M^2\frac{1}{\delta^2}\|B\|^2_{L^1(0,\infty)}\eta_2(\varepsilon)+12M^2\varepsilon^2\|B\|^2_{L^1(0,\infty)}N_2+3\varepsilon^2\frac{1}{\delta}\|B\|^2_{L^1(0,\infty)}N_2,$$

where $\eta_2(\varepsilon)\to 0$ as $\varepsilon\to 0$.

In view of the above, $\Psi X(\cdot)$ is square-mean almost periodic.

Assuming that X is square-mean almost periodic and using assumption (H1), Theorem 2.6, and Lemma 3.1, one can easily see that $u\mapsto G(u,X(u))$ is square-mean almost periodic. Therefore, given by $\varepsilon>0$, one can find $l_3(\varepsilon)>0$ such that any interval of length $l_3(\varepsilon)$ contains at least τ_3 for which

$$\|U(t+\tau_3,s+\tau_3)-U(t,s)\|\le\varepsilon e^{-\frac{\delta}{2}(t-s)}$$

for all $t-s\ge\varepsilon$, and

$$E\|G(u+\tau_3,X(u+\tau_3))-G(u,X(u))\|^2<\eta_3(\varepsilon)$$

for each $u\in R$, where $\eta_3(\varepsilon)\to 0$ as $\varepsilon\to 0$.

Moreover, it follows from Lemma2.3 (ii) that there exists a positive constant N_2 such that

$$\sup_{u\in R}E\|G(u,X(u))\|^2\le N_3.$$

The next step consists in proving the square-mean almost periodicity of $\Lambda X(\cdot)$. This more complicated that the previous case because of the involvement of the Brownian motion W. To overcome such a difficulty, we make extensive use of the properties of $\tilde{W}$ defined by

$$\tilde{W}(u):=W(u+\tau)-W(\tau)$$

for each u. Note that $\tilde{W}$ is also a Brownian motion and has the same distribution as W.

Now, make a change of variable to get

$$E\|\Lambda X(t+\tau_3)-\Lambda X(t)\|^2$$

$$\le E\left(\int_{-\infty}^{t}\left\|U(t+\tau_3,\sigma+\tau_3)\int_{-\infty}^{\sigma}C(\sigma-u)G(u+\tau_3,X(u+\tau_3))dW(u+\tau_3)\right.\right.$$

$$\left.\left.-U(t,\sigma)\int_{-\infty}^{\sigma}C(\sigma-u)G(u,X(u))dW(u)\right\|d\sigma\right)^2$$

$$\le 3E\left(\begin{array}{l}\int_{-\infty}^{t}\|U(t+\tau_3,\sigma+\tau_3)\|\int_{-\infty}^{\sigma}\|C(\sigma-u)\|\\ \|G(u+\tau_3,X(u+\tau_3))-G(u,X(u))\|d\tilde{W}(r)d\sigma\end{array}\right)^2$$

$$+3E\left(\int_{t-\varepsilon}^{t}\|U(t+\tau_3,\sigma+\tau_3)-U(t,\sigma)\|\int_{-\infty}^{\sigma}\|C(\sigma-u)\|\|G(u,X(u))\|dW(u)d\sigma\right)^2$$

$$+3E\left(\int_{-\infty}^{t-\varepsilon}\|U(t+\tau_3,\sigma+\tau_3)-U(t,\sigma)\|\int_{-\infty}^{\sigma}\|C(\sigma-u)\|\|G(u,X(u))\|dW(u)d\sigma\right)^2.$$

Then using an estimate on the Ito integral established to obtain

$$
\begin{aligned}
&E\,\|\Lambda X(t+\tau_3)-\Lambda X(t)\|^2\\
&\le 3\mathrm{Tr}QE\left(\int_{-\infty}^{t} Me^{-\delta(t-\sigma)}\int_{-\infty}^{\sigma}\|C(\sigma-u)\|\|G(u+\tau_2,X(u+\tau_2))-G(u,X(u))\|dud\sigma\right)^2\\
&+3\mathrm{Tr}QE\left(\int_{t-\varepsilon}^{t} 2Me^{-\delta(t-\sigma)}\int_{-\infty}^{\sigma}\|C(\sigma-u)\|\|G(u,X(u))\|dud\sigma\right)^2\\
&+3\mathrm{Tr}QE\left(\int_{-\infty}^{t-\varepsilon}\varepsilon e^{-\frac{\delta}{2}(t-\sigma)}\int_{-\infty}^{\sigma}\|C(\sigma-u)\|\|G(u,X(u))\|dud\sigma\right)^2\\
&\le 3\mathrm{Tr}QM^2\left(\int_{-\infty}^{t}e^{-\delta(t-\sigma)}d\sigma\right)^2\left(\int_{-\infty}^{+\infty}\|C(u)\|du\right)^2\sup_{u\in R}\|G(u+\tau_2,X(u+\tau_2))-G(u,X(u))\|\\
&+12\mathrm{Tr}QM^2\left(\int_{t-\varepsilon}^{t}e^{-\delta(t-\sigma)}d\sigma\right)^2\left(\int_{-\infty}^{+\infty}\|C(u)\|du\right)^2\sup_{u\in R}\|G(u,X(u))\|\\
&+3\mathrm{Tr}Q\varepsilon^2\left(\int_{-\infty}^{t-\varepsilon}e^{-\frac{\delta}{2}(t-\sigma)}d\sigma\right)^2\left(\int_{-\infty}^{+\infty}\|C(u)\|du\right)^2\sup_{u\in R}\|G(u,X(u))\|\\
&\le 3\mathrm{Tr}QM^2\frac{1}{\delta^2}\|C\|^2_{L^2(0,\infty)}\eta_3(\varepsilon)+12\mathrm{Tr}QM^2\varepsilon^2\|C\|^2_{L^2(0,\infty)}N_3+3\mathrm{Tr}Q\varepsilon^2\frac{1}{\delta}\|C\|^2_{L^2(0,\infty)}N_3,
\end{aligned}
$$

where $\eta_3(\varepsilon)\to 0$ as $\varepsilon\to 0$.

In view of the above, $\Lambda X(\cdot)$ is square-mean almost periodic

Define

$$
\begin{aligned}
&LX(t)=\int_{-\infty}^{t}U(t,\sigma)\int_{-\infty}^{\sigma}C(\sigma-u)G(u,X(u))dW(u)d\sigma\\
&+\int_{-\infty}^{t}U(t,\sigma)\int_{-\infty}^{\sigma}B(\sigma-u)F_2(u,X(u))dud\sigma+\int_{-\infty}^{t}U(t,u)F_1(u,X(u))du.
\end{aligned}
$$

Finally, we will show that $LX(\cdot)$ is a contraction. Since $(a+b+c)^2\le 3a^2+3b^2+3c^2$, we obtain

$$
\begin{aligned}
&E\,\|LX(t)-LY(t)\|^2\\
&\le 3E\left(\int_{-\infty}^{t}\|U(t,u)\|\|F_1(u,X(u))-F_2(u,X(u))\|du\right)^2\\
&+3E\left(\int_{-\infty}^{t}\|U(t,\sigma)\|\int_{-\infty}^{\sigma}\|B(\sigma-u)\|\|F_2(u,X(u))-F_2(u,Y(u))\|dud\sigma\right)^2\\
&+3E\left(\int_{-\infty}^{t}\|U(t,\sigma)\|\left\|\int_{-\infty}^{\sigma}C(\sigma-u)[G(u,X(u))-G(u,Y(u))]dW(u)\right\|d\sigma\right)^2\\
&\le 3M^2\left(\int_{-\infty}^{t}e^{-\delta(t-\sigma)}d\sigma\right)\sup_{u\in R}\|F_1(u,X(u))-F_1(u,Y(u))\|^2\\
&+3M^2\left(\int_{-\infty}^{t}e^{-\delta(t-\sigma)}d\sigma\right)^2\left(\int_{-\infty}^{+\infty}\|B(u)\|du\right)^2\sup_{u\in R}\|F_2(u+\tau_2,X(u+\tau_2))-F_2(u,X(u))\|\\
&+3M^2\mathrm{Tr}Q\left(\int_{-\infty}^{t}e^{-\delta(t-\sigma)}d\sigma\right)^2\left(\int_{-\infty}^{+\infty}\|C(u)\|du\right)^2\sup_{u\in R}\|G(u,X(u))-G(u,Y(u))\|^2_{L_2^0}
\end{aligned}
$$

$$\leq \left(\frac{3K_1}{\delta^2} M^2 + \frac{3K_2}{\delta^2} M^2 \|B\|^2_{L^1_{(0,\infty)}} + \frac{3K_3}{\delta^2} M^2 \|C\|^2_{L^2_{(0,\infty)}} \operatorname{Tr} Q \right) \|X - Y\|_\infty ,$$

and therefore

$$\|LX(t) - LY(t)\|_\infty \leq \left[\frac{3M^2}{\delta^2} \left(K_1 + K_2 \|B\|^2_{L^1_{(0,\infty)}} + K_3 \|C\|^2_{L^2_{(0,\infty)}} \operatorname{Tr} Q \right) \right] \|X - Y\|_\infty$$

$$= \Theta \|X - Y\|_\infty .$$

Consequently, if $\Theta < 1$, then (1.1) has unique fixed point, which obviously is the unique square-mean almost periodic solution to (1.1).

Acknowledgement. This article is funded by the National Natural Science Foundation of China (11161052), Guangxi Natural Science Foundation of China (2011jjA10044) and Guangxi Education Hall Project (201012MS183)

References

1. Bezandry, P., Diagana, T.: Existence of almost periodic solutions to some stochastic differential equations. Appl. Anal. 86(7), 819–827 (2007)
2. Bezandry, P.: Existence of almost periodic solutions to some functional integro-differential stochastic evolution equations. Stat. Prob. Lett. 78, 2844–2849 (2008)
3. Fu, M.M., Liu, Z.X.: Square-mean almost automorphic solutions for some stochastic differential equations. Proc. Amer. Math. Soc. 138, 3689–3701 (2010)
4. Chang, Y.K., Zhao, Z.H., N'Guérékata, G.M.: A new composition theorem for square-mean almost automorphic functions and applications to stochastic differential equations. Nonlinear Anal. 74, 2210–2219 (2011)
5. Chang, Y.K., Ma, R.Y., Zhao, Z.H.: Almost periodic solutions to a stochastic differential equations in Hilbert space. Results in Mathematics (2011) doi: 10.1007/s00025-011-0207-9
6. Bezandry, P., Diagana, T.: Existence of quadratic-mean almost periodic solutions to some stochastic hyperbolic differential equations. Electron. J. Differential Equations 2009(111), 1–14 (2009)
7. Chang, Y.K., Zhao, Z.H., N'Guérékata, G.M.: Square-mean almost automorphic mild solutions to some stochastic differential equations in a Hilbert space. Advances in Difference Equations 9, 1–12 (2011)
8. Bezandry, P., Diagana, T.: Square-mean almost periodic solutions nonautonomous stochastic differential equations. Electron. J. Differ. Equations 117, 1–10 (2007)
9. Chang, Y.K., Zhao, Z.H., N'Guérékata, G.M.: Square-mean almost automorphic mild solutions to non-autonomous stochastic differential equations in Hilbert spaces. Computers and Mathematics with Applications 61, 384–391 (2011)
10. Acquistapace, P.: Evolution operators and strong solution of abstract linear parabolic equations. Differential Integral Equations 1, 433–457 (1988)
11. Maniar, L., Schnaubelt, R.: Almost Periodicity of Inhomogeneous Parabolic Evolution Equations. Lecture Notes in Pure and Appl. Math. 234, 299–318 (2003)
12. Kannan, D.: On a stochastic integro-differential evolution equation of Volterra type. J. Integral Equations 10, 351–379 (1985)

A Commuting Pair in Quantum Qroupoids

Bing-liang Shen

Department of public education, Shanghai University of Finance and Economics Zhejiang College, Jinhua 321013, China
1712004584@qq.com

Abstract. Let H be a weak Hopf algebra (quantum groupoid) in the sense of [1]. Then in this note we first introduce weak Drinfeld double $D(H)$ and character algebra $C(H)$ of H. Next we define a weak $D(H)$-module algebra and a $C(H)$-module on $C_H(H_S)$, the centralizer of H_S. Finally we prove that if H is a semisimple weak Hopf algebra over an algebraically closed field k of characteristic 0, then the action of $D(H)$ on $C_H(H_S)$ and the action of $C(H)$ on $C_H(H_S)$ form a commuting pair, which generalizes the main result in [11].

Keywords: Weak Hopf algebra (quantum groupoid), Weak Drinfeld double, Character algebra, Commuting pair.

1 Introduction

Weak Hopf algebras (quantum groupoids) which are generalizations of ordinary Hopf algebras, were introduced in Bohm et al. [1, 2]. A weak Hopf algebra is a vector space that has both algebra and coalgebra structures related to each other in a certain self-dual way and that possesses an antipode. The main difference between ordinary and weak Hopf algebras comes from the fact that the comultiplication of the latter is no longer required to preserve the unit (equivalently, the counit is not required to be an algebra homomorphism) and results in the existence of two canonical subalgebras playing the role of ‘‘non-commutative bases’’ in a ‘‘quantum groupoid’’. The axioms of a weak Hopf algebra are self-dual, which ensures that when H is finite dimensional, the dual vector space H* has a natural structure of a weak Hopf algebra.

The initial motivation to study weak Hopf algebras was their connection with the theory of algebra extensions. It was explained that weak Hopf C*-algebras naturally arise as symmetries of finite depth von Neumann subfactors. The other main motivation for studying weak Hopf algebras comes from quantum field theory, operator algebras and representation theory. We refer the reader to [5-8] for the further study.

Another important application of weak Hopf algebras is that they provide a natural framework for the study of dynamical twists (cf. P. Etingof and D. Nikshych, Dynamical quantum groups at roots of 1, Duke Math. J. 108(2001), 135–168.) in Hopf algebras. It was proved that every dynamical twist in a Hopf algebra gives rise

Y. Yang, M. Ma, and B. Liu (Eds.): ICICA 2013, Part II, CCIS 392, pp. 615–624, 2013.

to a weak Hopf algebra. Also a family of dynamical quantum groups (weak Hop f algebras corresponding to dynamical twists in quantum groups at roots of unity) was constructed. These weak Hopf algebras were shown to be quasi-triangular with non-degenerate R-matrices and, therefore, self-dual.

Perhaps the easiest example of a weak Hopf algebra is a groupoid algebra; other examples are face algebras, quantum groupoids, generalized Kac algebras and so on. It has turned out that many results of classical Hopf algebra theory can be generalized to weak Hopf algebras. Despite it, the structure of a weak Hopf algebra is much more complicated than that of a Hopf algebra.

Drinfeld double and character algebra are very important tools for studying Hopf algebras. Zhu [11] studied the Drinfeld double and the character algebra, and found a famous result that the action of the Drinfeld double and the action of the character algebra form a commuting pair, which is powerful to classify finite dimensional Hopf algebras.

D. Nikshych developed the theory of semisimple weak Hopf algebra and got analogues of a number of classical results for ordinary semisimple Hopf algebras (see [4], [8]). The definition of character algebra of a weak Hopf algebra is formed and some basic properties are obtained.

The motivation of this paper is to generalize the main results of [11] into the case H is a quantum groupoid (weak Hopf algebra).

This paper is organized as follows.

In Section 2, we recall basic definitions and results related to weak Hopf algebras, weak Drinfeld double and character algebra.

In Section 3, we first review some identities of $C_H(H_S)$. Next we give a weak $D(H)$ -module algebra structure and the action of character algebra $C(H)$ on $C_H(H_S)$. Finally we obtain the main result of this paper: When H is a semisimple weak Hopf algebra over an algebraically closed field k of characteristic 0, then the action of weak Drinfeld double $D(H)$ and the action of $C(H)$ on $C_H(H_S)$ form a commuting pair. As corollary, we obtain that the dimension of every simple $D(H)$ -submodule in H is a divisor of $\dim(H)$.

2 Preliminaries

In this section we recall some basic definitions and results related to the paper.

Throughout the paper, all algebraic systems are supposed to be over a field k. k denotes a fixed field, and we will always work over k. The tensor product $\otimes = \otimes_k$ and Hom is always assumed to be over k. The reader is referred to Bohm et al. [1, 2] and D. Nikshych et al. [5]-[7] as general references about weak Hopf algebras, to Sweedler [9] about Hopf algebras. If C is a coalgebra, we use the Sweedler-type notation for the comultiplication:

$\Delta(c) = c_1 \otimes c_2$, for all $c \in C$

2.1 Weak Hopf Algebras

Recall from [1] that a weak bialgebra consists of an algebra structure $(m, 1)$ and a coalgebra structure (Δ, ε) such that $\Delta(hk) = \Delta(h)\Delta(k)$ and

$$\Delta^2(1) = 1_1 \otimes 1_2 1'_1 \otimes 1'_2 = 1_1 \otimes 1'_1 1_2 \otimes 1'_2 = 1_1 \otimes 1_2 \otimes 1_3,$$
$$\varepsilon(hkl) = \varepsilon(hk_1)\varepsilon(k_2 l) = \varepsilon(hk_2)\varepsilon(k_1 l),$$

for all $h, k, l \in H$, where $1'$ stands for another copy of 1.

We now summarize the elementary properties of weak bialgebras as follows. We have the target map ε_t and the source map ε_s defined by

$$\varepsilon_t(h) = \varepsilon(1_1 h)1_2; \quad \varepsilon_s(h) = 1_1 \varepsilon(h1_2).$$

And their images H_t and H_s are called the target space and the source space. They are separable subalgebras of H which are playing the role of non-commutative numbers in many respects.

Recall that for an algebra A and a coalgebra C, we have the convolution algebra Conv (C, A)= Hom (C, A) as space, with the multiplication given by $(f * g)(c) = f(c_1)g(c_2)$, for all $f, g \in Hom(c, a), c \in C$.

A weak Hopf algebra H is a weak bialgebra together with a k-linear map $S : H \to H$ (called the antipode) satisfying

$$S * id_H = \varepsilon_s, \; id_H * S = \varepsilon_t, \; S * id_H * S = S,$$

Where $*$ is the usual convolution product. It follows immediately that $S = \varepsilon_s * S = S * \varepsilon_t$. If an antipode S exists, then it is unique and S is an anti-algebra and anti-coalgebra morphism. If H is a finite dimensional weak Hopf algebra, then the antipode S is automatically bijective and the dual $H^* = Hom(H, k)$ has a natural structure of a weak Hopf algebra with the structure operations dual to those of H. We will always assume that H is finite dimensional. Now we list some properties of S :

$$h_1 \otimes \varepsilon_t(h_2) = 1_1 h \otimes 1_2, \quad h_1 \otimes \varepsilon_s(h_2) = h1_1 \otimes S(1_2),$$
$$\varepsilon_t(h_1) \otimes h_2 = S(1_1) \otimes 1_2 h, \quad \varepsilon_s(h_1) \otimes h_2 = 1_1 \otimes h1_2,$$
$$h_1 y \otimes h_2 = h_1 \otimes h_2 S(y), \quad \forall\, y \in H_s,$$
$$h_1 \otimes z h_2 = S(z)h_1 \otimes h_2, \quad \forall\, z \in H_t.$$

Every weak Hopf algebra H contains a canonical minimal weak Hopf subalgebra $H_{\min}$ generated, as an algebra, by H_s and H_t [5, Section 3]. Obviously, H is an ordinary Hopf algebra if and only if $H_{\min} = k1$. The restriction of S^2 to $H_{\min}$ is always an inner automorphism of $H_{\min}$, see [5].

Remark 1. In what follows we will consider only weak hopf algebras satisfying the following natural property:.

$$S^2 \mid_{H_{\min}} = id.$$

We will call a weak Hopf algebra satisfying above regular. It was shown in [6, 6.1] that every weak Hopf algebra can be obtained by twisting a regular Hopf algebra with the same algebra structure.

2.2 Weak Drinfeld Double

Let H be a finite dimensional weak Hopf algebra over a field k. We define a map

$$J : H^* \to H, J(f \otimes h) = (\varepsilon \otimes 1_H)(f \otimes h) = (f \otimes h)(\varepsilon \otimes 1_H),$$

where the multiplication in $H^* \to H$ is the twisted smash product, i.e.,

$$(f \otimes a)(g \otimes b) = f(a_1 \to g \leftarrow S^{-1}(a_3)) \otimes a_2 b.$$

Let $[f \odot h]$ denotes the class of $f \otimes h$ in $D(H) = H^* \otimes H / \ker J$. $D(H)$ is a weak Hopf algebra with the following structures:

$$[f \odot a][g \odot b] = [f(a_1 \to g \leftarrow S^{-1}(a_3)) \odot a_2 b],$$
$$\Delta([f \odot h]) = [f_2 \odot h_1] \otimes [f_1 \odot h_2],$$
$$\varepsilon([f \odot h]) = \langle f, 1_1 \rangle \varepsilon(1_2 h),$$
$$S([f \odot h]) = [\varepsilon \odot S(a)][S^{*-1}(f) \odot 1_H].$$

Here $\to$ and $\leftarrow$ mean the left and right H-module action on H^* defined as:

$$a \to g = g_1 \langle g_2, a \rangle, \quad a \leftarrow g = g_2 \langle g_1, a \rangle,$$

for all $g \in H^*, a \in H$..

2.3 Character

We refer the definition of the character from [3].

If ρ is a representation of a group G acting on a vector space V / k, then the k-valued function on G defined by

$$\Upsilon_\rho : g \mapsto tr\rho(g)$$

Where $tr\rho(g)$ is the trace of the linear transformation $\rho(g)$, is called the character of G afforded by ρ.

If H is a semisimple weak Hopf algebra and the ground field is algebraically closed and of characteristic 0, then the character algebra $C(H)$ of H may be regarded as a subalgebra of H^* with unit $\varepsilon_1 \varepsilon_2$ (here we denote $\Delta(\varepsilon) = \varepsilon_1 \otimes \varepsilon_2$, and $C(H)$ is semisimple consisting of the elements $v \in H^*$ that are cocommutative, i.e., $v_1 \otimes v_2 = v_2 \otimes v_1$ (cf. [8])

3 A Commuting Pair in Weak Hopf Algebras

We first recall the definition about the weak module algebras

3.1 Definition

Let H be a weak Hopf algebra and A an algebra. A left H -module A is called a left weak H -module algebra if A satisfies

$$h\cdot(ab)=(h_1\cdot a)(h_2\cdot b),\quad h\cdot 1_A=\varepsilon_t(h)\cdot 1_A,$$

for all $h\in H$ and $a,b\in A$.

3.2 Example

Let H be a finite dimensional weak Hopf algebra. Then H is a left weak H^* - module algebra with action $g\to a=\langle g,a_2\rangle a_1$, for all $a\in H$, $g\in H^*$. Also H^* is a left weak H -module algebra with action $a\to g=\langle g_2,a\rangle g_1$, for all $a\in H$, $g\in H^*$.

Recall from [7], the adjoint action of $1\in H$ gives rise to a non-trivial map $H\to H$:

$h\mapsto 1_1 hS(1_2)=Ad_1^l(h),\quad h\in H.$

3.3 Lemma

([7]) The map Ad_1^l is a linear projection from H onto $C_H(H_S)$, the centralizer of H_S, i.e., $(Ad_1^l)^2=Ad_1^l$ and $Ad_1^l(H)=C_H(H_S)$

3.4 Lemma

Let H be a finite dimensional weak Hopf algebra. Then $C_H(H_S)$ is a left weak module algebra of $D(H)$ under the following action:

$[f\odot a]\cdot b=(a_1bS(a_2))\leftarrow S^{*-1}(f)$, for all $f\in H^*, a\in H$ and $b\in C_H(H_S)$. Here $a\leftarrow f=\langle f_1,a\rangle f_2,\forall a\in H, f\in H^*$

Proof. (1) We first prove the left $D(H)$ -module action is well-defined.

1) For all $y\in H_s$, we check

$$\begin{aligned}
&y([f\odot a]\cdot b)=y[(a_1bS(a_2))\leftarrow S^{*-1}(f)]\\
&=\langle S^{*-1}(f),a_1b_1S(a_4)\rangle ya_2b_2S(a_3)\\
&=\langle S^{*-1}(f),a_1b_1S(a_4)\rangle y1_1a_2b_2S(a_3)S(1_2)\\
&=\langle S^{*-1}(f),a_1b_1S(a_4)\rangle 1_1a_2b_2S(a_3)S(1_2)y\\
&=\langle S^{*-1}(f),a_1b_1S(a_4)\rangle a_2b_2S(a_3)y\\
&=([f\odot a]\cdot b)y
\end{aligned}$$

We get that $[f \odot a] \cdot b$ belongs to $C_H(H_S)$.

2) We verify

$$[f \odot \varepsilon_t(x)h] \cdot b = (1_1 \varepsilon_t(x) h_1 b S(1_2 h_2)) \leftarrow S^{*-1}(f)$$
$$= (\varepsilon_t(x) 1_1 h_1 b S(1_2 h_2)) \leftarrow S^{*-1}(f)$$
$$= (\varepsilon_t(x) h_1 b S(h_2)) \leftarrow S^{*-1}(f),$$

and

$$[f(x \to \varepsilon) \odot h] \cdot b = [f \varepsilon_1 \langle \varepsilon_2, x \rangle \odot h] \cdot b$$
$$= [(h_1 b S(h_2)) \leftarrow S^{*-1}(f \varepsilon_1)] \langle \varepsilon_2, x \rangle$$
$$= \langle S^{*-1}(\varepsilon_1), h_1 b_1 S(h_4) \rangle \langle \varepsilon_2, x \rangle [(h_2 b_2 S(h_3)) \leftarrow S^{*-1}(f)]$$
$$= \varepsilon(S^{-1}(h_1 b_1 S(h_4)) x) [(h_2 b_2 S(h_3)) \leftarrow S^{*-1}(f)]$$
$$= \varepsilon(S(x) 1'_1) \varepsilon(1'_2 1_1 h_1 b_1 S(h_4)) [(1_2 h_2 b_2 S(h_3)) \leftarrow S^{*-1}(f)]$$
$$= \varepsilon(S(x) 1_1) [(1_2 h_1 b S(h_2)) \leftarrow S^{*-1}(f)]$$
$$= (\varepsilon_t(x) h_1 b S(h_2)) \leftarrow S^{*-1}(f).$$

Similarly we can prove $[f \odot \varepsilon_s(y)h] \cdot b = [f(\varepsilon \leftarrow y) \odot h] \cdot b$.

(2) Next we show $C_H(H_S)$ is a left $D(H)$-module. We verify

$[\varepsilon \odot 1_H] \cdot b = (1_1 b S(1_2)) \leftarrow \varepsilon = 1_1 b S(1_2) = b 1_1 S(1_2) = b$ using the fact $b \in C_H(H_S)$, and

$$([f \odot a][g \odot b]) \cdot c = [f(a_1 \to g \leftarrow S^{-1}(a_3)) \odot a_2 b] \cdot c$$
$$= (a_2 b_1 c S(a_3 b_2)) \leftarrow S^{*-1}(f(a_1 \to g \leftarrow S^{-1}(a_4)))$$
$$= \langle S^{*-1}(f(a_1 \to g \leftarrow S^{-1}(a_6))), a_2 b_1 c_1 S(a_5 b_4) \rangle a_3 b_2 c_2 S(a_4 b_3)$$
$$= \langle f, a_6 b_5 S^{-1}(c_2) S^{-1}(a_3 b_2) \rangle \langle g, S^{-1}(a_8) a_7 b_6 S^{-1}(c_1) S^{-1}(b_1) S^{-1}(a_2) a_1 \rangle a_4 b_3 c_3 S(a_5 b_4)$$
$$= \langle f, a_5 b_5 S^{-1}(c_2) S^{-1}(a_2 b_2) \rangle \langle g, S^{-1}(\varepsilon_s(a_6)) b_6 S^{-1}(c_1) S^{-1}(b_1) S^{-1}(\varepsilon_s(a_1)) \rangle a_3 b_3 c_3 S(a_4 b_4)$$
$$= \langle f, a_4 b_5 S^{-1}(c_2) S^{-1}(a_1 1_2 b_2) \rangle \langle g, S^{-1}(\varepsilon_s(a_5)) b_6 S^{-1}(c_1) S^{-1}(b_1) S^{-1}(1_1) \rangle a_2 b_3 c_3 S(a_3 b_4)$$
$$= \langle f, a_4 1_1 b_5 S^{-1}(c_2) S^{-1}(a_1 b_2) \rangle \langle g, S^{-1}(S(1_2)) b_6 S^{-1}(c_1) S^{-1}(b_1) \rangle a_2 b_3 c_3 S(a_3 b_4)$$
$$= \langle f, a_4 b_5 S^{-1}(c_2) S^{-1}(a_1 b_2) \rangle \langle g, b_6 S^{-1}(c_1) S^{-1}(b_1) \rangle a_2 b_3 c_3 S(a_3 b_4)$$
$$= \langle S^{*-1}(f), a_1 b_2 c_2 S(a_4 b_5) \rangle \langle g, b_6 S^{-1}(b_1 c_1) \rangle a_2 b_3 c_3 S(a_3 b_4)$$
$$= \langle g, b_4 S^{-1}(b_1 c_1) \rangle a_1 b_2 c_2 S(a_2 b_3) \leftarrow S^{*-1}(f)$$
$$= [a_1 \langle S^{*-1}(g), b_1 c_1 S(b_4) \rangle a_1 b_2 c_2 S(b_3) S(a_2)] \leftarrow S^{*-1}(f)$$
$$= [a_1((b_1 c S(b_2)) \leftarrow S^{*-1}(g)) S(a_2)] \leftarrow S^{*-1}(f)$$
$$= [f \odot a] \cdot ((b_1 c S(b_2)) \leftarrow S^{*-1}(g))$$
$$= [f \odot a] \cdot ([g \odot b] \cdot c).$$

(3) Finally it is not hard to check that $C_H(H_S)$ is a left weak $D(H)$-module algebra. We finish the proof.

3.5 Lemma

Let H be a semisimple weak Hopf algebra over an algebraically closed field k of characteristic 0. Then $C_H(H_S)$ is a left $C(H)$ -module via the action

$$g \to a = \langle g, a_2 \rangle a_1 ,$$

for all $g \in C(H)$ and $a \in C_H(H_S)$.

Proof. (1) We first prove the left module action is well-defined. For all $y \in H_s$, we must show $y(g \to a) = (g \to a)y$, i.e., $\langle g, a_2 \rangle ya_1 = \langle g, a_2 \rangle a_1 y$. Now taking any $f \in H^*$, we compute

$$\begin{aligned}
&\langle g, a_2 \rangle \langle f, ya_1 \rangle = \langle g, a_2 \rangle \langle f_1, y \rangle \langle f_2, a_1 \rangle \\
&= \langle f_2 g, a \rangle \langle f_1, y \rangle = \langle f_2 g, a \rangle \langle f_1, \varepsilon_s(y) \rangle \\
&= \langle f_2 g, a \rangle \langle \varepsilon_s^*(f_1), y \rangle = \langle f \varepsilon_2 g, a \rangle \langle \varepsilon_1, y \rangle \\
&= \langle fg_2, a \rangle \langle S^{*-1} \varepsilon_t^*(g_1), y \rangle = \langle fg_2, a \rangle \langle S^* \varepsilon_t^*(g_1), y \rangle \\
&= \langle fg_1, a \rangle \langle S^* \varepsilon_t^*(g_2), y \rangle = \langle f \varepsilon_1 g, a \rangle \langle S^*(\varepsilon_2), y \rangle \\
&= \langle f_1 g, a \rangle \langle \varepsilon_s^*(f_2), y \rangle = \langle f_1 g, a \rangle \langle f_2, \varepsilon_s(y) \rangle \\
&= \langle f, a_1 y \rangle \langle g, a_2 \rangle .
\end{aligned}$$

So we get $\langle g, a_2 \rangle ya_1 = \langle g, a_2 \rangle a_1 y$ and this means $g \to a \in C_H(H_S)$.

(2) It is easy to get $f \to (g \to a) = (fg) \to a$, for all $f, g \in C(H)$. Next we verify $\varepsilon_1 \varepsilon_2 \to a = a$ for all $a \in C_H(H_S)$. We use the fact $\Delta(ya) = a_1 \otimes ya_2 = a_1 \otimes a_2 y$, here $a \in C_H(H_S), y \in H_s$.

$$\begin{aligned}
&\varepsilon_1 \varepsilon_2 \to a = \langle \varepsilon_1 \varepsilon_2, a_2 \rangle a_1 = \varepsilon(a_2 a_3) a_1 \\
&= \varepsilon(\varepsilon_s(a_2) a_3) a_1 = \varepsilon(S(1_2) a_2) a_1 1_1 \\
&= \varepsilon(S(1_2) a_2 S(1_1)) a_1 = \varepsilon(S(1_2) a_2 1_1) a_1 \\
&= \varepsilon(a_2 S(1_2) 1_1) a_1 = \varepsilon(a_2 S^{-1}(1_2) 1_1) a_1 \\
&= \varepsilon(a_2) a_1 = a.
\end{aligned}$$

We finish the proof.

The proof of the following Lemma is Similar to the proof of the Lemma 1 in [11]. For the integrality of the paper, we give the prove

3.6 Theorem

Let H be a weak Hopf algebra. If $T \in End_k(H)$ commutes with the action $\leftarrow$ of H^* on $C_H(H_S)$, then there exists $v \in H^*$ such that $T(a) = v \to a$ for all $a \in C_H(H_S)$.

Proof. This is a version of the following well-known fact: if A is an associative algebra, $T \in End_k(A)$ commutes with the left multiplication r_a for all $a \in A$, then T is a right multiplication for some $b \in A$. To apply this fact, we notice that the

transpose action of $\rightarrow$ is the left multiplication of H* on H*, while the transpose action of $\leftarrow$ is the right multiplication of H* on H*. T commutes with the action $\leftarrow$ of H* on $C_H(H_S)$, implies that T* commutes with the left multiplications on H*. Therefore T* is given by a right multiplication, and therefore there exists $v \in H^*$ such that $T(a) = v \rightarrow a$ for all $a \in C_H(H_S)$. The proof is completed.

Now we are to state our main theorem in this paper

3.7 Theorem

If H is a semisimple weak Hopf algebra over an algebraically closed field k of characteristic 0, then the action of $D(H)$ given by Lemma 3.4 and the action $\rightarrow$ of $C(H)$ on $C_H(H_S)$ given by Lemma 3.5 form a commuting pair.

Proof. We note the semisimplicity of H implies that $D(H)$ is semisimple ([5-8]). The semi- simplicity of H also implies that $C(H)$ is a semisimple algebra (see Lemma 4.1.3 of [8]). In particular the images of $D(H)$ and $C(H)$ in $End_k(H)$ are semisimple algebras. Therefore it suffices to prove that $T \in End_k(H)$ commutes with the action of $D(H)$ if and only if T is in the image of $C(H)$.

Assume that T commutes with the action of $D(H)$, we need to prove T is in the image of $C(H)$. We note that $H^* \subset D(H)$ acts on H by restriction: this action is just the action $\leftarrow$ of H^* on H twisted by S^{-1}. By Lemma 3.6, there exists a unique $v \in H^*$ such that

$$T(b) = v \rightarrow b = \langle v, b_2 \rangle b_1,$$

for every $b \in H$.

For T as in above, T commutes with the action of $H \subset D(H)$ implies that

$$\langle v, a_2 b_2 S(a_3) \rangle a_1 b_1 S(a_4) = \langle v, b_2 \rangle a_1 b_1 S(a_2),$$

for every $a, b \in H$. Apply the counit map to both sides of above equation, we obtain

$$\langle v, a_1 b S(a_2) \rangle = \langle v, b_2 \rangle \varepsilon(a_1 b_1 S(a_2)).$$

This further implies that

$$\begin{aligned}
&\langle v, ab \rangle = \langle v, ab1_1 S(1_2) \rangle = \langle v, 1_1 abS(1_2) \rangle \\
&= \langle v, a_1 bS(\varepsilon_t(a_2)) \rangle = \langle v, a_1 ba_3 S(a_2) \rangle \\
&= \langle v, b_2 a_4 \rangle \varepsilon(a_1 b_1 a_3 S(a_2)) \\
&= \langle v, b_2 a_3 \rangle \varepsilon(a_1 b_1 S(\varepsilon_t(a_2))) \\
&= \langle v, b_2 a_2 \rangle \varepsilon(a_1 b_1) = \langle v, b_2 a_2 \rangle \varepsilon(\varepsilon_s(a_1)\varepsilon_t(b_1)) \\
&= \langle v, 1'_2 b_2 a_2 1_2 \rangle \varepsilon(1_1 S(1'_1)) = \langle v, 1'_2 b_2 a_2 1_2 \rangle \varepsilon(1_1 1'_1) \\
&= \langle v, 1_2 baS(1_1) \rangle = \langle v, S(1_1) ba1_2 \rangle = \langle v, ba \rangle.
\end{aligned}$$

This proves that v is cocommutative or $v \in C(H)$.

Conversely, if $v \in C(H)$, we need to prove that the action "$v \rightarrow$" commutes with the action of $D(H)$. For all $[f \odot h] \in D(H)$, we have

$$T([f \odot a] \cdot b) = v \rightarrow [(a_1 bS(a_2)) \leftarrow S^{*-1}(f)]$$
$$= [v \rightarrow (a_1 bS(a_2))] \leftarrow S^{*-1}(f).$$

Since v is cocommutative,

$$v \rightarrow (a_1 bS(a_2)) = \langle v, a_2 b_2 S(a_3)\rangle a_1 b_1 S(a_4)$$
$$= \langle v, b_2 a_2 S(a_3)\rangle a_1 b_1 S(a_4)$$
$$= \langle v, b_2 \varepsilon_t(a_2)\rangle a_1 b_1 S(a_3)$$
$$= \langle v, b_2 1_2\rangle 1_1 a_1 b_1 S(a_2)$$
$$= \langle v, b_2 1_2\rangle a_1 b_1 1_1 S(a_2)$$
$$= \langle v, b_2\rangle a_1 b_1 S(a_2)$$
$$= a_1 (v \rightarrow b_1) S(a_2)$$

and

$$T([f \odot a] \cdot b) = (a_1 (v \rightarrow b_1) S(a_2)) \leftarrow S^{*-1}(f)$$
$$= [f \odot a] \cdot (v \rightarrow b) = [f \odot a] \cdot T(b).$$

We get "$v \rightarrow$" commutes with the action of $D(H)$. The proof is completed

3.8 Corollary

If H is a semisimple weak Hopf algebra over an algebraically closed field k of characteristic 0 with $\varepsilon_t = S^{-1} \circ \varepsilon_s$, then the action of $D(H)$ given by Lemma 3.4 and the action $\rightarrow$ of $C(H)$ on H form a commuting pair.

Proof. We only need to show that $C_H(H_S) = H$ and $S^2|_{H_{\min}} = id$. For all $h \in H$ and $y \in H_s$, we compute

$$yh = \varepsilon_s(y)h = 1_1 h\varepsilon(y1_2) = h_1 \varepsilon(y\varepsilon_t(h_2))$$
$$= h_1 \varepsilon(yS^{-1} \circ \varepsilon_s(h_2)) = h1_1 \varepsilon(y1_2) = h\varepsilon_s(y) = hy,$$

and

$$S^2(yz) = S^2(y)S^2(z) = [SS\varepsilon_s(y)][SS\varepsilon_t(z)]$$
$$= [S\varepsilon_t S(y)][S\varepsilon_s S(z)] = [\varepsilon_s S(y)][\varepsilon_t S(z)]$$
$$= [S\varepsilon_t(y)][S^{-1}\varepsilon_s S(z)] = \varepsilon_s(y)[S^{-1}S\varepsilon_t(z)] = yz,$$

for all $y \in H_s$ and $z \in H_t$. The proof is completed.

If H is an ordinary Hopf algebra, we can get the main result in [11, Theorem1]

3.9 Corollary

If H is a semisimple Hopf algebra over algebraically closed field k of characteristic 0, then the action of $D(H)$ given by Lemma 3.4 and the action $\rightarrow$ of $C(H)$ on H form a commuting pair

3.10 Corollary

Let H be a semisimple weak Hopf algebra over an algebraically closed field k of characteristic 0 with $\varepsilon_t = S^{-1} \circ \varepsilon_s$, and let H be the $D(H)$ -module defined above. Then the dimension of every simple $D(H)$ -submodule in H is a divisor of $\dim(H)$.

Proof. We use the conditions in Corollary 3.8 and the proof is similar to the proof of Corollary in [11], we omit it here

Acknowledgement. This work was supported by Foundation of Zhejiang Educational Committee (Y201121955).

References

1. Bohm, G., Nill, F., Szlachanyi, K.: Weak Hopf algebras I. Integral theory and C*-structure. J. Algebra 221, 385–438 (1999)
2. Bohm, G., Szlachanyi, K.: A coassociative C*-quantum group with nonintegral dimensions. Lett. Math. Phys. 35, 437–456 (1996)
3. Jacobson, N.: Basic algebra II, San Francisco (1980)
4. Larson, R.G., Radford, D.E.: Semisimple cosemisimple Hopf algebras. Amer. J. Math. 110, 381–385 (1988)
5. Nikshych, D.: On the structure of weak Hopf algebras. Adv. Math. 170, 257–286 (2002)
6. Nikshych, D., Vainerman, L.: Finite quantum groupoids and their applications. In: New Directions in Hopf Algebras, vol. 43, pp. 211–262. MSRI Publications (2002)
7. Nikshych, D., Turaev, V., Vainerman, L.: Invariants of knots and 3-manifolds from quantum groupoids. Topology and its Application 127, 91–123 (2003)
8. Nikshych, D.: Semisimple weak Hopf algebras. J. Algebra 275, 639–667 (2004)
9. Sweedler, M.: Hopf Algebras, Benjamin, New York (1969)
10. Zhu, Y.C.: Hopf algebras of prime dimension. International Math. Res. Notices 1, 53–59 (1994)
11. Zhu, Y.C.: A commuting pair in Hopf algebras. Proceedings of American Mathematical Society 125, 2847–2851 (1997)

Author Index