Ershi Qi
Jiang Shen
Runliang Dou *Editors*

The 19th International Conference on Industrial Engineering and Engineering Management

Springer

The 19th International Conference on Industrial Engineering and Engineering Management

Ershi Qi • Jiang Shen • Runliang Dou
Editors

The 19th International Conference on Industrial Engineering and Engineering Management

Springer

Editors
Ershi Qi
Industrial Engineering
Institution of College of Management and Economics
Tianjin University
China, People's Republic

Jiang Shen
Industrial Engineering
Institution of College of Management and Economics
Tianjin University
China, People's Republic

Runliang Dou
Industrial Engineering
Institution of College of Management and Economics
Tianjin University
China, People's Republic

ISBN 978-3-642-37269-8 ISBN 978-3-642-37270-4 (eBook)
DOI 10.1007/978-3-642-37270-4
Springer Heidelberg New York Dordrecht London

Library of Congress Control Number: 2013937993

Printed on acid-free paper

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

Contents

1 **A New Waveform Design for Phase-Coded Quasi-CW Radar System** 1
Wei-gui Zeng, Ying-feng Sun, and Ming-gang Liu

2 **Research on Human Factors Engineering of Chinese Traditional Armchair with Four Protruding Ends** 13
De-hua Yu

3 **The Evaluation on Working Fatigue Based on Improved Fuzzy Borda Method** 23
Min-xia Liu, Jia-wei Ke, Jian Xie, Qing Xue, Huan Zhou, and Chen-xiao Guan

4 **A Study on Hydraulic Autofrettage Gun Tube Dynamic Stress by Finite Element Analysis** 33
Yan-mou Zhan, Guo-rui Huang, and Hao Cheng

5 **A Cooperative Game Model of Dynamic Alliance Under Grid Environment** 45
Xiang-bin Zhang and Min Wu

6 **Optimization of Injection Molding Shop Scheduling Based on the Two-Stage Genetic Algorithm** 59
Jin-ping Zhou and Hu Fu

7 **A Simulation-Based Analysis of Cluster Tools Scheduling with Plant Simulation** 71
Yan Min and Xiao-rui Lin

8 **A New Quant Model to Ensure Successful QFD Implementation in the Increasingly Uncertain Environment** 81
Wei Xiong, Kun Xiong, and Xiao-tun Wang

9 Multivariate Measurement System Analysis Based on Projection Pursuit Method 91
Xiaofang Wu, Liangxing Shi, and Zhen He

10 A Multivariate Synthetic Control Chart for Monitoring Covariance Matrix Based on Conditional Entropy 99
Li-ping Liu, Jian-lan Zhong, and Yi-zhong Ma

11 Multi-objective Robust Operation Model for a Supply Chain with Market Demands and Raw Material Price Uncertainty 109
Li-ping Yu, Li-jun Li, and Xiao-yuan Huang

12 The Application of Six Sigma to Enhance Product and Service Quality in Internet Marketing 121
Tai-Chang Hsia, Shui-Chuan Chen, and Kun-Sheng Chen

13 Production Process Improvement Using the Six Sigma DMAIC Methodology: A Case Study of a Laser Computer Mouse Production Process 133
S. Pimsakul, N. Somsuk, W. Junboon, and T. Laosirihongthong

14 Machine Vision-Aided Quality Decision System for Solder Joint Defect Evaluation 147
Chien-Chih Wang

15 Demand Analysis and Framework of Theory About Chinese Events Sports Logistics 159
Peng-hui Liu and Li Gao

16 Dual Selection of Competing Channels Based on Internet Channel and Power Structures 169
Jun Liu and De-qing Tan

17 Supply-Chain Ripple Effect Under Duopoly Retailers' Operating Uncertainty with Two Different Contracts 177
Qian Zhao, Hong Chen, and Zhong-he Wu

18 The Influence of Inter-enterprise Value Co-creation on Innovation Based on Resource Theories 187
Steven Ji-fan Ren, Qiong Bu, Ming-jian Zhou, and Cai-hong Hu

19 The Analysis of Fresh Food Safety Risks from the Cold Chain Logistics System 197
Fang Yang and Jian-ming Cai

20 Literate Review on Firm Network Capability: A Process Perspective 209
Yi-bo Lv, Qing-yang Mao, and Yu-wen Ding

21 **The Shortest Path Algorithm Based on Petri Net** 221
Yu-jie Zheng, Kai-hu Hou, Wei-zhen Liao, and Lin Yang

22 **Measuring Satisfaction of Two Level Employees in a Representative Chinese Steel Plant: A Structure Equation Modeling Approach** 231
Guo-wen Huang and Hong-wei Liu

23 **Performance Evaluation of CRM in Securities Investment Consulting Business Based on Fuzzy-AHP** 239
Qing-guo Zhao and Song-tao Jiang

24 **A Study on the Model of Selecting the Loyalty Card Issuance Method Based on Fuzzy Set Pair Analysis** 249
Yu Su and Jin Song

25 **Complex Network Community Detection Algorithm Based on Genetic Algorithm** 257
Yun Li, Gang Liu, and Song-yang Lao

26 **Complexity Analysis of Software Based on Function-Call Graph** . 269
Yang Guo, Zheng-xu Zhao, and Wei Wang

27 **Fault Diagnosis of Wet Flue Gas Desulphurization System Based on KPCA** 279
Yu-ping Zheng and Li-ping Zhang

28 **The Research on Negative Perfectionism Personality of Engineering Students from the Perspective of Game Theory** ... 289
Ying Sun, Ying-ying Yang, Chen-xu Feng, and Hui Zhang

29 **An Empirical Study on the Formation of Inter-firm eBusiness Capabilities: The Moderating Effect of Formal Governance** 299
Xiao-kuan Shen and Mao-mao Chi

30 **A Driven Model of IT Usage: Determinants and Moderating Effects of Situational Variables** 307
Rong-jie Lv, Li-jun He, Xiao-chun Chen, and Zhao Zhao

31 **A Semantic Ontology of Requirement – Product – Process – Resource for Modeling of Product Lifecycle Information** 319
Biao Yu and Han Zhao

32 **Research on Turning the Toll Mode of the Periphery Freeway into Congestion Charging** 331
Jian Tong, Yuan-hua Jia, and Zhong-ming Guo

33 **The Research of Information-Based Market Manipulation** 343
Xiao-long Liu and Wen-xiu Hu

34 **Research on Contractor's Claims Caused by Extension of Time** .. 355
Gong-jie Ao, Xiao-li Zhang, and Xiao-dong Zhang

35 **The Study on Sewage Treatment BOT Project Concessionaire Selection Criteria: Best Practice from China** 365
Wei-hua Yang, Ya-nan Yang, Feng-hai Zhang, and Xiu-shan Wang

36 **Study of Optimal Contract Under Asymmetric Information** 373
Ping Li and Hong-zhi He

37 **Intelligent Control via Power-Line Carrier for Illumination and Air Condition in Buildings** 383
Rong-Ching Wu, Ke Zhu, En-Chih Chang, and Jia-Chu Lee

38 **Design and Implementation of a Remote Equipment Monitoring and Management System Based on IOT** 395
Naren Gaowa, Ying Liu, and Mu-yan Li

39 **An Integrated Planning and Execution Model for Supporting Emergency Response** 405
Pan Tang and Jian Zhou

40 **Structural Effects on Consumer Purchase Intention Towards Slimming Underwear** 415
Wen Xiong, Hang Zhang, Shuang-jia Ren, and Jian-ping Wang

41 **Ant Colony Algorithm-Based Mixed-Model Assembly Line** 423
Yong-feng Xiong and Wen-sheng Yang

42 **Institutional Investors and Corporate Short-Term Debt Financing** 437
Fang-fei Ding, Chan Gu, and Jin-hua Chen

43 **Multi-product Supply Planning for Combat Units in Battlefield Environment** 445
Ji Ren, Xiao-lei Zheng, and Yue-jin Tan

44 **An Optimal Layout Design for Storage Yard of Container Terminal** 455
Jian-xun Tang and Li-xin Tang

45 **Analysis of Environmental Governance for Construction Project Based on a Dynamic Non-cooperative Game** 467
Jiao-jie Han and Rui Zhao

46 **Empirical Study on Stock Valuation Model Based on Multiple Linear Regression Analysis** 479
Qing Li and Ai-min Li

47 **The Research and Application of a Dynamic Dispatching Rule Selection Approach Based on BPSO-SVM for Semiconductor Production Line** 487
Kuo Tian, Yu-min Ma, and Fei Qiao

48 **Reallocation of Resources to Preserve Relative Efficiencies: Inverse CCR Model** 497
Saowanee Lertworasirikul and S.C. Fang

49 **Auxiliary Process for User-Interface Design Based on Genetic Algorithm: Using Multimedia Software as an Example** 509
Wei-Chang Yeh, Shang-Chia Wei, and Chin-Hao Yang

50 **A Mixed-Integer Linear Programming Model for an Integrated Planting, Harvesting, and Procuring in Pickled Ginger Industry** 521
A. Ritvirool and A. Pantong

51 **Reliability Analysis for Mechanical Components Subject to Degradation Process and Random Shock with Wiener Process** 531
Hui-bing Hao, Chun Su, and Zhong-zhou Qu

52 **A New Constructive Cost Model for Software Testing Project Management** 545
Yun-feng Lu and Yong-feng Yin

53 **Correction Coefficient of Catalytic Combustion Gas Detector** 557
Lu Pan, Yan-yan Wang, Zhi-hong Sun, Da-hu Zhang, and Lin Jiang

54 **Identification of Risk Factors for Casualty in Building Fires** 567
Chun-hua Yu, Li-ping Tian, and Zeng-chang Zhao

55 **Data-Driven Based Key Performance Index Residual Generation and Its Application on Complex Electrical Equipment** 577
Zhi-gang Yao, Li Cheng, and Yu-lei Wang

56 **Evolutionary Stability Analysis of Behavior Among Partners Under a Default Punishment Mechanism** 587
Rong Ding and Yan-ming Sun

57 **Stability and Hopf Bifurcation of a Four-Dimensional Dynamic Economic System** 599
Hong-liang Tu and Jun-hai Ma

58 **Spline Estimation for a Class of Time Series Variance Model** 613
Xin-qian Wu, Wan-cai Yang, and Shu-hong Zhang

59 Application of GM (1, N)-Markov Model in Shanghai Composite Index Prediction 625
Wan-cai Yang, Xin-qian Wu, Er-xin Zhang, and Guang Zhu

60 Research on Fixed-Point Maintenance Support Simulation of Mechanization Army Division Based on HLA 637
Chuang Gu, Xiao-ming Du, Ji-wei Cai, Bing Liu, and Gui-qi Wang

61 Evolution of Industrial Ecosystem with Government's Intervention: Integration of Evolutionary Game Model into Multi-agent Simulation 651
Kai-fang Zheng, Su-ling Jia, and Hong-sen Wang

62 The Impact of Major Event on Regime Switching of Financial Market Volatility Spillover: The Case of the 2011 Japanese Earthquake 663
Lu Wang, Jiong-lou Xu, and Mao Li

63 Comparative Study on the Performance of GEM Listed Companies Having or Not the Stock Right Incentive Mechanism 673
Gong-rong Chen, Cheng-yin Zheng, and Qi-shan Li

64 Benefit Embezzlement of Separation of Control Right and Cash Flow Right: Empirical Evidence from China Private Real Estate Listed Companies 683
Yan-hong Chen, Shi-jian Zhang, and Pei Zhu

65 Do Political Connections of Private Enterprises Impact Bank Loan Pricing 693
Gong-rong Chen, Qi-shan Li, and Cong-min Ding

66 How the Largest Shareholder Control Affects the Company's Cash Value 703
Zhi-xiong Ling and Chun-ling Li

67 Impact of the Corporate Income Tax Reform on Capital Structure Choices: Evidence from Data of Chinese Listed Firms 713
Qing-mei Tan

68 Optimization Model of Loan's Portfolio Based on Geometric Spectral Measure 725
Chong Duan and Xiu-min Jia

69 The Effect of Market Selection on Lagged Currency Risk Exposure 741
Xiao Wu and Long-wen Zhou

70 **Chairman Overconfidence and Dividend Payout Policy:** ***An Analysis Based on China's Listed Companies*** 749
Xiu-qun Hu, Rong-shen Lv, and Guo-liu Hu

71 **Shared Knowledge and Shared Understanding Between CIO and Top Management Team: A Literature Review** 759
Rong-jie Lv, Zhao Zhao, Xiao-chun Chen, and Li-jun He

72 **The Effect of Environmental Awareness on Consumers' Green Purchasing: Mediating Role of Green Perceived Value** 767
Jian-xin Li, Dan Liu, and Li Liu

73 **Developing an Analysis Model for Public Participation in Sustainable Construction** 777
Qiu-bo Yang, Xue-qing Wang, and Yuan Ren

74 **The Cooperative and Competitive Complexity Model of Emerging Strategic Marine Biological Medicine Virtual Industry Cluster Based on Value Net Theory** 787
Peng Du

75 **Research on Functional Demands of Urban Integrated Transportation Hub Project Based on the Stakeholder Theory:** ***A Case Study of Tianjin Station*** 797
Xiao-li Zhang, Yi-lin Yin, Gong-Jie Ao, and Xiao-Dong Zhang

76 **Environmental Policy and Enterprise Hiring Decisions** 807
Jin-ying Zhang

77 **Internal Mechanism Research on Synergy of Enterprise Network Organizations Based on Value Structure** 817
Jing-wen Han and Zheng-yin Peng

78 **The Application of Fuzzy Analytic Hierarchy Process and Grey Relational Analysis in the Taxi Passenger Satisfaction Evaluation** 827
Ting Liu

79 **The Impact of Consumption Emotions on Service Quality and Customer Satisfaction** 837
Guo-li Pu and Qin Su

80 **Exploratory Study on the Healthy Organization in Manufacturing Enterprise** 847
She-hong Liang, Lei Xing, Kan Shi, and Peng Gao

81 **Study on Customer Satisfaction of Express Enterprises Based on AHP-FCE** 855
Hai-yan Qi and Jiang-bo Zheng

82 **Service Enhancement Mode Study on Urban Rail Transportation Equipment Manufacturing Enterprises Cluster** .. 865
Hui-qun Gong, Qing-jing Xu, and Chao Huang

83 **Research on Comprehensive Evaluation of a Certain Type of Equipment's Service Life** 873
Qing Li, Xiao-nan Ye, Rong Wang, and Li Wang

84 **An Analysis and Empirical Study of the Impact of Customer Relationship Quality on Repurchase Intention in Online Shopping Context** 885
Mi-yuan Shan, Yue-hong Luo, and Ren-long Zhang

85 **An Empirical Research on Innovation Capability of Chinese Manufacturing Enterprises Based on Manufacturing Servitization** 893
Xue-gang Shi

86 **Research on Policy Adjustment of Employment Relationships of Private Enterprises in Heilongjiang Province** 903
Yong-hui Deng

87 **Optimization of Foreground Support Facilities for Commercial Banks Based on Queuing Theory** 913
Jianxin Shi and Jierong Wang

88 **Research of Sales Forecasting Model and Algorithm in Food and Beverage Enterprise** 925
Li Gao, Ke-lin Xu, and Dao-xiao Mao

89 **Waste Collection Vehicle Routing Problem: Case Study in Alexandria, Egypt** 935
A. Moustafa, A.A. Abdelhalim, A.B. Eltawil, and N. Fors

90 **Determining Lean Practices in Health Care Service Systems** 945
P. Punnakitikashem

91 **A Proposed Heuristic for Solving the Container Pre-marshalling Problem** 955
M.S. Gheith, A.B. El-Tawil, and N.A. Harraz

92 **Research on Optimization of Explosive Products Supply Support in Certain Base** 965
Jun Peng, Jian-jun Zhao, Xiao-yan Zhang, and Yang Sun

93 **An Exploration on the Organizational Structure Design of the Hong Kong-Zhuhai-Macao Bridge Authority** 975
Jia-qing Yu, Jie Huo, Qiang Shi, Gang Sheng, and Hao Zhang

94 **Co-innovation and Code of Value Appropriation** 987
Ping Xu and Yong-cheng Xu

95 **Study on Product Information Model for Collaborative Product Design Based on Ontology** 997
You-yuan Wang, Peng-biao Nie, Fa-lin Wang, Chao-yang Guo, and Xian-ming Dai

96 **An Analysis of Evaluating Enterprises' Ecological Management Information Systems** 1009
Jin-yan Sang, Zhen-fa Qi, and Ying-hua Zhang

97 **Knowledge Workers' Competence Evaluation Model Based on Extenics** 1017
Liang Tang

98 **Knowledge Know-How Management Model of High-Tech Enterprises – Dual Perspectives of Knowledge Management and Strategic Human Resource Management** 1027
Jing-yu Chen and Zu-guang Hu

99 **Ontology Technology-Supported English Teaching Scheme for TESOL** 1037
Xin-xing Gong and Ming-yang Zhong

100 **Strategy Planning for State-Owned New Energy Vehicle Industry of China** 1047
Li-ran Chen, Guo Yan, and Ying Zhang

101 **Research on Organization Tacit Knowledge Sharing: A Game Theoretic Analysis** 1059
Li Dan

Chapter 1
A New Waveform Design for Phase-Coded Quasi-CW Radar System

Wei-gui Zeng, Ying-feng Sun, and Ming-gang Liu

Abstract In this paper, we analyze the echo eclipse problem of phase-coded quasi-continuous wave (quasi-CW) radar using the echo eclipse ratio function and main lobe to side lobe ratio. And then we propose a waveform design method based on long short codes and synchronous alternation of carrier frequency. This method can not only solve the severe eclipse problem of short range target and simultaneously detect the whole range, but also increase the signal complexity and decrease the probability of interception. The simulation results show that the performance of phase coded quasi-CW radar can be effectively improved using the proposed waveform design method.

Keywords Echo eclipse • Long code • Phase-coded • Quasi-CW • Short code • Waveform design

1.1 Introduction

In the high-tech local war, battlefield situation awareness becomes more and more important, and radar is a key sensor which is a real time, fast and accurate access to the information of battlefield situation. However, along with the fast development of electronic technologies, the interception of radar signal, jamming, anti-radiation, and other techniques are increasingly diversified and becoming more and more effective. As a result, the countermeasure environment of radar is becoming more and more complicated. Low probability of interception (LPI) radar is a sensor of this kind that can adapt to complex countermeasure environment very well. LPI radar is a special radar system which owns special signal waveform, special antennas,

W. Zeng (✉) • Y. Sun • M. Liu
Department of Scientific Research, Naval Aeronautical and Astronautical University, Yantai, China
e-mail: zeng_yimel@hotmail.com

E. Qi et al. (eds.), *The 19th International Conference on Industrial Engineering and Engineering Management*, DOI 10.1007/978-3-642-37270-4_1,

controlled power, and low detected probability. At present, LPI radar system is continuous wave (CW) radar system or quasi-continuous wave (quasi-CW) radar system in most case (Zhang 2003; Liu et al. 2001; Sanmartin et al. 2000). CW radar has many excellent properties such as accurate velocity measurement, small size, and high mobility. However, the leakage of emission signal is very serious in CW radar. In order to avoid the leakage of transceiver, quasi-CW radar was introduced (Khan and Mitchel 1991; Gu et al. 1998; Yang et al. 2004). Quasi-CW radar is also called as interrupted continuous wave radar whose duty cycle is close to 0.5. Quasi-CW radar has low probability of interception and high transceiver isolation. Quasi-CW radar has a lower peak power and is more compatible with solid-state transmitter. Though quasi-CW radar has high transceiver isolation to avoid the leakage of emission signal, the high duty cycle of emission signal brings severe echo eclipse in some ranges. Because of echo eclipse problem, the main lobe level of echoes which originally has strongly self-correlation is decreased and the side lobe level is becoming relatively high, which has adverse effect on target detection. Therefore, solving the echo eclipse problem of quasi-CW radar system is becoming imperative (Yuan 2007; Wang et al. 2011). In this paper, the echo eclipse problem is studied from the perspective of waveform design, and a phase-coded quasi-CW waveform design method based on combination of long code with short code and synchronous CF alternation is proposed, which can solve the eclipse problem very well.

1.2 The Echo Eclipse Problem of Phase-Coded Quasi-CW Radar System

The signal of phase-coded quasi-CW has large time-bandwidth product, and the duty cycle is close to 50 % (Gao and Tian 2004). However, high duty cycle brings out the echo eclipse in some range (partly or completely). The echo eclipse in phase-coded quasi-CW radar system is shown in Fig. 1.1.

Let c denote the light speed, $\{c_m\}, m = 0, 1, \ldots, N-1$ denote the binary coding sequence, N denote the length of sequence, t_p denote the width of sub-code, T_r denote pulse repeat period, and R denote the target distance. According to the target range, there are three kinds of echo eclipse:

(1) When $R < \frac{cNt_p}{2}$, the anterior of echoes is eclipsed, just like echo 1 shown in Fig. 1.1; (2) when $\frac{cNt_p}{2} \leq R < \frac{c(T_r - Nt_p)}{2}$, the echo is not eclipsed, just like echo 2; (3) when $\frac{c(T_r - Nt_p)}{2} \leq R < \frac{cT_r}{2}$, the posterior of echo is eclipsed, just like echo 3.

1.2.1 Echo Eclipse Ratio Function

The echo eclipse of phase-coded quasi-CW radar can be evaluated by echo eclipse ratio function (Modarres-Hashemi et al. 1999). Let $x_T(t)$ denote the transmitting

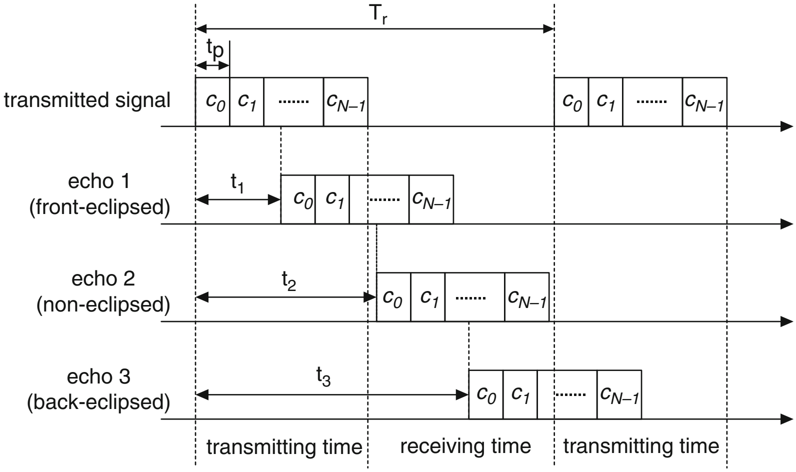


Fig. 1.1 The echo eclipse problem of phase-coded quasi-CW radar system

switch signal. When binary phase-coded signal is transmitted, $x_T(t) = 1$, or else $x_T(t) = 0$. The receiving switch signal is the opposite of transmitting switch signal, $x_R(t) = 1 - x_T(t)$. Echo eclipse ratio function $g(\tau)$ is defined as

$$g(\tau) = \frac{1}{T_r}\int_0^{T_r} x_T(t-\tau)(1-x_R(t))dt \tag{1.1}$$

where τ is the delay time of echo. When $\tau = 0$, $g(0) = 1$, the echo is completely eclipsed. When $\tau \neq 0$, the echo is partly eclipsed, and $0 \leq g(\tau) < 1$. When $g(\tau) = 0$, the echo is not eclipsed. As $g(\tau)$ grows larger, the echo eclipse becomes more serious and the energy of echo into radar receiver is less.

1.2.2 Main Lobe to Side Lobe Ratio of Eclipsed Echo

In the signal processing, the pulse compression process of eclipsed signal is a partial correlation process (Baden and Cohen 1990; Zeraster 1980). It therefore causes that the main lobe level of compressed signals is decreased and the side lobe level is increased relatively. Then we introduce peak side lobe level (PSL) and integrated side lobe level (ISL) to evaluate the main lobe to side lobe ratio of eclipsed signal. PSL and ISL are defined as follows:

$$PSL = 10\lg\frac{\max(\chi_m^2)}{\chi_0^2} \tag{1.2}$$

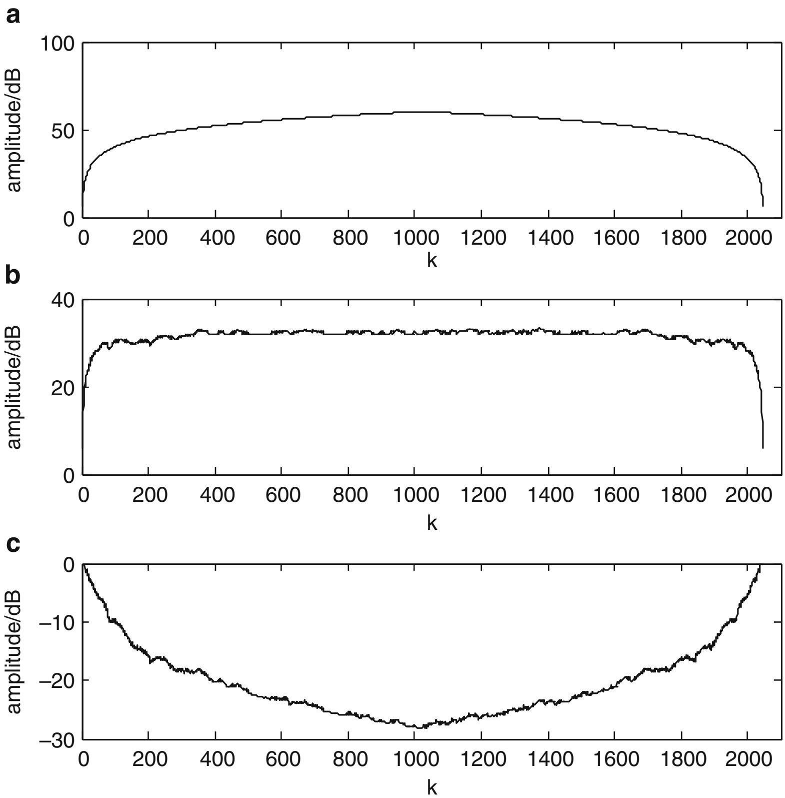


Fig. 1.2 The pulse compression performance simulation of eclipsed echoes for 1,023-elements *m* sequence code. (**a**) Main lobe amplitude, (**b**) Maximum sidelobe amplitude, (**c**) Peak sidelobe level

$$ISL = 10\lg\frac{\sum \chi_m^2}{\chi_0^2} \tag{1.3}$$

where χ_m $(-N+1 \leq m \leq N-1, m \neq 0)$ is m-th side lobe level and χ_0 is main lobe level.

Next, we make simulation to m sequence binary coded signal whose length is 1,023. Assume that the duty cycle of emission signal is 50 %, and the target distance is $\frac{cNt_p}{2}$ $(1 \leq k \leq 2N-1)$. We analyze the main lobe and maximum side lobe of the compressed signals. When $1 \leq k \leq 1,022$, the anterior of echoes is eclipsed. When $k = 1,023$, the echo is not eclipsed; when $1,024 \leq k \leq 2,045$, the posterior of echo is eclipsed. The simulation results are shown in Fig. 1.2. The amplitude of main lobe changing with the target location k is shown in Fig. 1.2a, the amplitude of maximum side lobe with the target location k is shown in Fig. 1.2b, and the PSL changing with k is shown in Fig. 1.2c.

1.3 The Waveform Design Method Based on Combination of Long Short Codes and Synchronous Alternation of CF

1.3.1 Wave Design

Normally, radar always has to detect targets at different range, but the binary phase-coded radar based on long coding sequence is not suitable for the detection of close targets as analyzed above. In this paper, by referring to the waveform design method of matching the target range, we propose a waveform design based on the combination of long short codes and CF alternation. The waveform design method is shown in Fig. 1.3.

In a repeat period T, the emission signals contains a binary phase-coded signal based on short coding sequence (the sequence length is M) and a binary phase-coded signal based on long coding sequence (the sequence length is N). The binary phase-coded signal based on short coding sequence which owns a very low side lobe is used to detect targets in short range. Because the number of sub-codes is relatively smaller and the pulse width is narrower, there is no severe echo eclipse in the detecting range. Though the energy of emission signal is lower, the energy of echoes reflected from close targets is usually much stronger. So, the relatively lower emission power does not influence detection performance. The binary phase-coded signal based on long coding sequence is used to detect targets located in far range, the pulse width of signal is much longer than that of short signal, and the total energy of emission signal is much stronger, which is useful to detect distant targets. While the code type of emission signal is being shifted, the CF of signals is synchronously being alternated. When the binary phase-coded signal based on short coding sequence is transmitted, the frequency of CF is f_1.

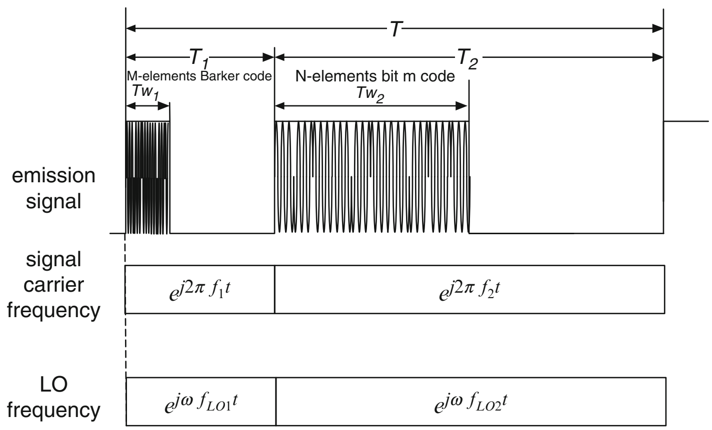


Fig. 1.3 The schematic diagram of proposed waveform design

When the binary phase-coded signal based on long coding sequence is transmitted, the frequency of CF is f_2. After that, we can expand the signal bandwidth and increase the signal complexity. And it is also useful to improve the signal's electronic counter-countermeasure (ECCM) ability and to reduce the probability of interception.

Normally, after the two kinds of binary phase-coded signals are transmitted, these two different kinds of signals will be reflected from target, and then enter radar receiver. As a result, it will probably cause the radar receiver saturated. Fortunately, that will not happen when the designed signals are transmitted, because the CF frequency of these signals is different. When the target appears near, the echoes mix with the LO (local oscillator) signal whose frequency is also synchronously alternated (at this time, the frequency of LO signal is f_{LO1}), and the long code signals is suppressed very well after passing through the low-pass filter (the CF frequency of long code signal is f_2, and $|f_{LO1} - f_2|$ exceeds the bandwidth of low-pass filter). Similarly, when the target is located in far distance, the short code signal in the echoes can also be effectively suppressed. Meanwhile, by transmitting the designed signals, radar can detect targets located in different range at the same time, the dynamic range and sensitivity of radar receiver can be improved, and the interference between signals is weaker.

Set the short code sequence as $\{c_m\}, m = 1, 2 \ldots, M - 1$, long code sequence as $\{\gamma_n\}, n = 1, 2, \ldots, N - 1$, and the sub-pulse width as τ. Then emission signal can be denoted by

$$x(t) = \begin{cases} \dfrac{1}{\sqrt{M}} \sum\limits_{m=0}^{M-1} c_m v(t - k\tau) e^{j2\pi f_1 t} & 0 < t < M\tau \\ \dfrac{1}{\sqrt{N}} \sum\limits_{n=0}^{N-1} \gamma_n v(t - k\tau - T_1) e^{j2\pi f_2 t} & T_1 < t < N\tau + T_1 \\ 0 & else \end{cases} \tag{1.4}$$

where T_1 is the pulse repetition interval (PRI) of short code signal, M is the number of short coding sequence, N is the number of long coding sequence, $v(t)$ is sub-pulse function, f_1 is the carrier frequency of the phase-coded binary signal based on short coding sequence, and f_2 is the carrier frequency of phase-coded binary signal based on long coding sequence.

1.3.2 Parameters Calculation

In waveform design, we need to calculate the waveform parameters such as PRI, pulse width, code length and so on according to the application. Assume that the detection range of radar is (R_1, R_2), and the delay of echo ranges from $\frac{2R_1}{c}$ to $\frac{2R_2}{c}$

accordingly. Make sure the short phase-coded echoes are received completely; the sub-pulse width is confined to the following condition:

$$\tau \leq \frac{2R_1}{Mc} \tag{1.5}$$

Then the pulse width of binary phase-coded signal based on short coding sequence Tw_1 can be given by

$$Tw_1 = M\tau \leq \frac{2R_1}{c} \tag{1.6}$$

Moreover, in order to ensure the correlation length and main lobe to peak side lobe ratio (MPSR) of long code sequence echoes which are reflected from a distant target, the eclipse ratio of echoes have to be less than $g_{\max_2}$. Assume that the target is located in R_2 and the eclipse ratio of echoes is $g_{\max_2}$, then

$$\frac{T_2 - t_2}{\eta T_2} = 1 - g_{\max_2} \tag{1.7}$$

$$T_2 = \frac{t_2}{1 - \eta + \eta\, g_{\max_2}} \tag{1.8}$$

where η is the duty cycle of long code sequence binary phase-coded signal. Let $t_2 = \frac{2R_2}{c}$, then

$$T_2 = \frac{2R_2}{(1 - \eta + \eta\, g_{\max_2})c} \tag{1.9}$$

So, the pulse width of long code sequence phase-coded binary signal Tw_2 is given by

$$Tw_2 = \eta T_2 = \frac{2\eta R_2}{(1 - \eta + \eta\, g_{\max_2})c} \tag{1.10}$$

From to (1.5) and (1.9), the sub-pulse number of long code sequence binary phase-coded signal $N^{'}$ can be given by

$$N' = \left\lfloor \frac{2\eta R_2}{(1 - \eta + \eta\, g_{\max_2})\tau c} \right\rfloor \tag{1.11}$$

Next, we select N which is close to N' to be the code length of long code sequence.

When close targets are detected, the anterior of echoes is eclipsed. Similarly, to ensure the correlation length and MPSR of long code sequence echoes, the eclipse ratio of echoes have to be less than $g_{\max_1}$. Let the target is located in R_3, at this time, the eclipse ratio is $g_{\max_1}$. Then

$$g_{\max_1} = \frac{Tw_2 - t_3}{Tw_2} \tag{1.12}$$

$$t_3 = (1 - g_{\max_1})Tw_2 \tag{1.13}$$

where the delay of echo is t_3.

Let (1.9) into (1.12), we can get

$$t_3 = \frac{2(1 - g_{\max_1})\eta R_2}{(1 - \eta + \eta\ g_{\max_2})c} \tag{1.14}$$

When the echo delay is less than t_3, the eclipse ratio is greater than $g_{\max 1}$, and the MPSR is becoming lower. At this situation, we need to use the binary phase-coded signal based on short coding sequence to detect target. In other words, R_3 is the maximum detection range of short code signal, and the eclipse ratio of short code sequence echoes is kept to 0. Therefore, the PRI of short code sequence signal T_1 is required to meet the following condition:

$$T_1 \geq t_3 + Tw_1 = \frac{2(1 - g_{\max_1})\eta R_2}{(1 - \eta + \eta\ g_{\max_2})c} + M\tau \tag{1.15}$$

When (1.5) and (1.14) both take mark of equality, we can get that

$$T_1 = \frac{2(1 - g_{\max_1})\eta R_2}{(1 - \eta + \eta g_{\max_2})c} + \frac{2R_1}{c} \tag{1.16}$$

1.4 Example and Simulation

Using the proposed waveform design method, we make simulation to phase-coded quasi-CW radar with a detection range of 0.2–50.0 km. The main parameters of radar are shown in Table 1.1.

In the short distance range (200 m–9.0 km), the radar signal is 13-elements Barker code phase-coded signal, the echoes reflected from targets in this range is not eclipsed, and the compressed signals' MPSR is 22.3 dB. Meanwhile, 1,023-elements m sequence phase-coded signal is used to detect targets in the range of 9.0–50 km. The echoes reflected from targets in the range of 9.0–15.345 km are front eclipsed echoes, but the echo eclipse ratio of these echoes is less than 0.42,

Table 1.1 Example of X band phase-coded quasi-CW radar waveform parameters

Radar parameters	
Code type	Short code: 13-elements Barker sequence
	Long code: 1,023-elements *m* sequence
The width of sub-code (T_p)	0.1us
The frequency of CF (f)	$f = 9.170$ GHz, when $0 < t < T1$
	$f = 9.370$ GHz, when $T1 < t < T2$
Pulse repeat period (T)	Where, $T_{w1} = 0.1us, T_1 = 62.4us;$
	$T_{w2} = 102.3us, T_2 = 409.2us;$
Detection range (R)	Minimum detection range: 200 m
	Maximum detection range: 50 km
	Detection range section 1[a]:200 m–9.0 km
	Detection range section 2[b]:9.0–50.0 km
Echo eclipse ratio(g(t))	$g(t) = 0$, when $1.3us \le t \le 60.0us;$
	$0 \le g(t) < 0.42$, when $60.0us < t \le 102.3us;$
	$g(t) = 0$, when $102.3us < t \le 306.9us;$
	$0 \le g(t) < 0.26$, when $306.9us < t \le 333.3us;$
Emission power (P)	43 dBmW
MPSR	>22.0 dB

[a]Detection range of short code signal
[b]Detection range of long code signal

and the number of correlated sub-codes is greater than 594, and the MPSR is greater than 22.0 dB. The echoes reflected from targets in the range of 15.345–46.035 km are not eclipsed. The echoes reflected from targets in the range of 46.035–50.0 km are back eclipsed, but the echo eclipse ratio of these echoes is less than 0.26. The number of these correlated sub-codes is greater than 757; and MPSR is greater than 25.3 dB. When the target is located in 12.0, 27.0, and 50.0 km, the corresponding compressed signals are shown in Fig. 1.4.

1.5 Conclusion

In this paper, the echo eclipse problem of phase-coded quasi-CW radar system was analyzed, and a waveform design method which is based on combination of long code with short code and synchronous CF alternation was proposed. The echo eclipse problem was resolved very well, and the signal bandwidth was expanded and the ECCM ability was improved as well as the probability of interception was decreased using this method. The simulation results show that the radar signal designed by using the method proposed in this paper can effectively improve the performance of detection.

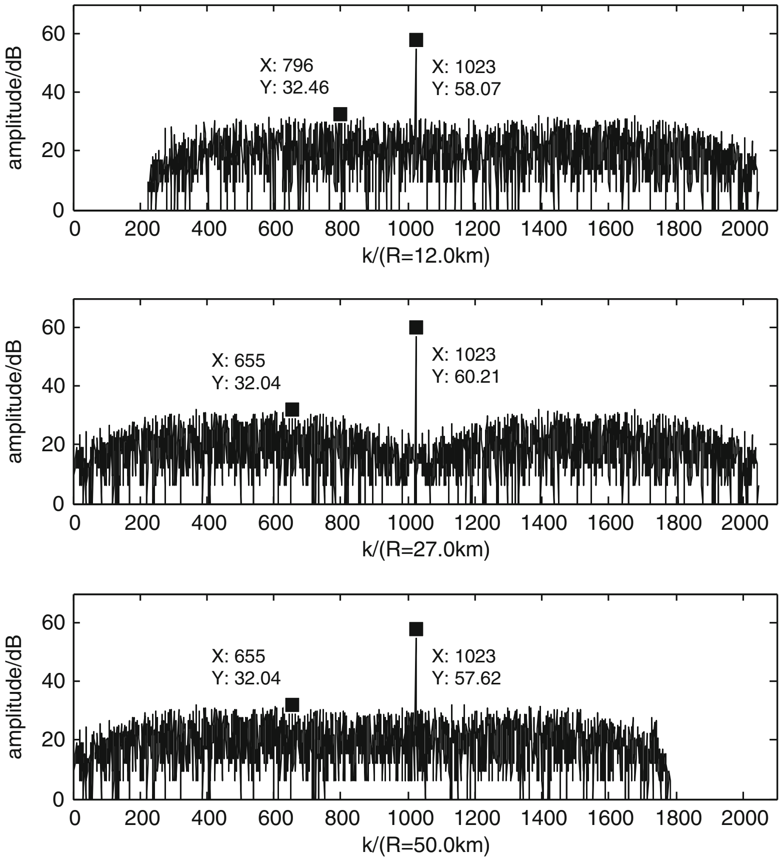


Fig. 1.4 The simulation of compressed signals reflected from targets in different range

References

Baden JM, Cohen LR (1990) Optimal peak side lobe filters for bi-phase pulse compression. In: Proceedings of 1990 IEEE international radar conference, New York, 1990, pp 249–252

Gao M, Tian L (2004) Radar signal processing for long phase-coded pseudo-continue waveforms (in Chinese). Syst Eng Electron 26(2):163–166 (in Chinese)

Gu H, Li X, Shang W et al (1998) A new approach of periodic square-wave interruption to solution of the leakage of CW radar. Acta Electron Sinica 26(12):7–10 (in Chinese)

Khan RH, Mitchel DK (1991) Waveform analysis for high frequency FMICW radar. IEE Process F 138(5):411–419

Liu G, Gu H, Su W et al (2001) The analysis and design of modern low probability of intercept radar. In: 2001 CIE international radar conference, Beijing, China, 2001, pp 120–124

Modarres-Hashemi M, Nayebi MM, Alavi H (1999) Performance evaluation of the phase-coded signal in LPD radars. In: IEEE military communications conference, Atlantic City, MILCOM-299, 1999, pp 796–800

Sanmartin J, Burgos M, Pérez F (2000) Radar sensor using low probability of interception SS-FH signals. IEEE Aerosp Electron Syst Mag 15(4):23–28

Wang H, Zhang J, Li Y, Zhu X (2011) Anti-range eclipse waveform optimization method for quasi-CW radar. J Data Acquis Process 26(5):531–535 (in Chinese)

Yang L, Wang R, Zhang D et al (2004) Research on quasi-continuous wave radar signal based on MAC sequence. Acta Electron Sinica 32(6):994–996 (in Chinese)

Yuan W (2007) Study on the design method of a novel quasi-CW radar waveform. Mod Radar 29(9):16–19 (in Chinese)

Zeraster S (1980) Minimum peak range sidelobe filters for binary phase-coded waveforms. IEEE Trans AES 6(1):214–217

Zhang X (2003) Development of LPI radar. Mod Radar 25(12):1–4 (in Chinese)

Chapter 2
Research on Human Factors Engineering of Chinese Traditional Armchair with Four Protruding Ends

De-hua Yu

Abstract Ming-style furniture is the elite of Chinese traditional furniture, which is doing well in human factors engineering. Chinese traditional armchair with four protruding ends has backrest, neckrest, armrest, footrest, and so on, which support the body properly. Based on the study of Chinese traditional armchair with four protruding ends, the practice of human factors engineering in Ming-style furniture is summarized, which is meaningful for the design of Chinese traditional furniture to develop ergonomics.

Keywords Chinese traditional armchair with four protruding ends • Design of Chinese traditional furniture • Human factors engineering • Ming-style furniture

2.1 Definition

Chinese traditional armchair with four protruding ends (Fig. 2.1) is the typical and common chair of Chinese traditional furniture, which is also called yoke back armchair because of the yoke-like headrest, and also the north official's armchair as the headrest looks like the official' hat (Shixiang Wang 2008). The four protruding ends of Chinese traditional armchair refer to the ends of two armrests and the ends of the headrest, compared with the Chinese traditional armchair whose armrest and headrest doesn't have protruding ends, which is also called southern official's armchair (Gustav Ecke 1991).

D. Yu (✉)
Art and Design School, Beijing Institute of Technology, Beijing, China
e-mail: yudehuabit@163.com

E. Qi et al. (eds.), *The 19th International Conference on Industrial Engineering and Engineering Management*, DOI 10.1007/978-3-642-37270-4_2,

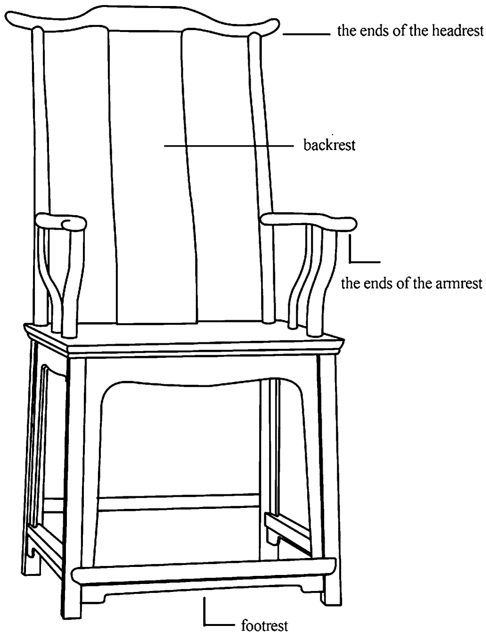


Fig. 2.1 Chinese traditional armchair with four protruding ends

2.2 Background

The ancient Chinese sat on woven mat more than 2,000 years ago, when there was only low furniture around them. It was not until the Han Dynasty (206 AD–220BC) that the high furniture spread into the Central Plains influenced by nomadic tribes and also the migration of Buddhism. According to the historical records, the Han emperor called Lingdi were much interested in the custom of the nomadic tribe called hu, including the folding chair (huchuang) (Fan Ye 2007). The folding chair was the early high seating, which became more and more popular among the nobilities of Han Dynasty. During the Northern and Southern Dynasties (386–586 AD) to the Tang dynasty (618–907 AD), high chairs were becoming more and more popular among the elite and then the common people, which was the transitional period from sitting on the mat to on the chair. By the Song Dynasty (960–1279 AD), high chairs and tables were already common among all the social classes (Yao Yang 2002a). Various types of high furniture like tables, chairs, beds, stools and so on were recorded in the paintings, murals and unearthed relic, which indicated that the transition from sitting on the mat to on the chair had finished, and high seating had already maintained the dominance. From then on, people were already accustomed to high furniture. In the late Ming and early Qing Dynasty, a unique style called Ming-style furniture came into being, which was brilliant in the whole world (Zongshan Li 2001).

Chinese traditional furniture in the ancient times indicated the masters' social status, wealth, and also the social etiquette, while the comfort was considered less, which was different from that of western countries. Shape, decoration, dimension and materials of the furniture varied in different classes. For example, the throne chair is the specific chair of the emperor, which is large enough to support the emperor, and indicates the authority and the status of the emperor (Jiaqing Tian 2003), although it is not comfortable and considerate. The only auxiliary thing for comfort is the cushion on the throne chair to support the emperor. In the pavilion, terrace, and the hall of the nobilities' houses, there were also chairs of different grades. In a standard hall of the house, the master usually sat in the chair of smart shape, well decoration and big dimension, like the armchair with round back, armchair with four protruding ends, etc., which highlighted the status of the master. Chairs of guests were not as particular as the master's chair, and usually simpler. And the last chair in the hall might be just a stool rather than a chair without the backrest and the armrest, which was prepared for the lowly person. The stools with four legs and seat surface were the common seating for the ordinary people.

Social status and etiquette was more important than the comfort when sitting in the chair in the ancient times. Upright sitting was the required posture for even the emperor and the nobles, and if the posture was sustained for a while, the lumbar muscle wound be under great strain, which was not comfortable.

However, the comfort was not the neglect factor when making a chair in the ancient times. Based on the consideration of the social status and etiquette, the comfort was also the important factor to make chairs comfortable and considerate. For example, the S-shaped backrest is compatible with the curved spine. Actually, carpenters in ancient China usually fit in with the basic principle of human factors engineering although without the guidance of modern theory of human factors engineering (Zengbi Chen 1981).

2.3 Analysis

Most of people nowadays spend most of the days sitting at a desk at work and also after work, and we call it sedentary behavior, which may make the irreparable damage on the body. Therefore, the correct sitting posture and chairs that are beneficial to the health is vital important for people.

Chinese traditional armchair with four protruding ends pays more attention to the comfort, and several details indicate the consideration of the human factors engineering.

Considered from the human factors engineering, the support from the seating includes three parts: the spine support, the buttock support, the foot support and the arm support (Yuanbo Sun 2010).

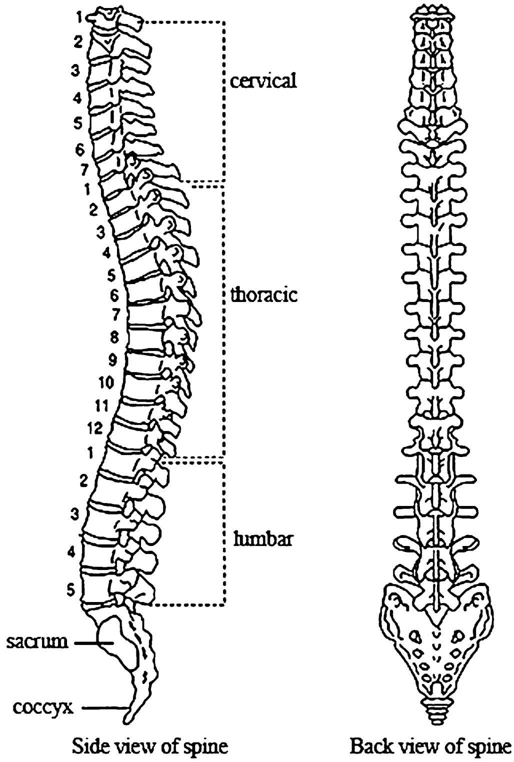


Fig. 2.2 Structure of the spine

2.3.1 The Spine Support

The spine provides structural support the body and protects the spine cord, which is made up of 7 cervical (neck) vertebrae, 12 thoracic (chest, dorsal) vertebrae, 5 lumbar (lower back) vertebrae, 5 sacral vertebrae (fused triangular bone) and 3 or 4 little vertebrae fused into a coccyx at the lower end of the spinal column (Fig. 2.2). The spine usually gets so many burdens due to bad posture while being seated, especially the long, motionless hours of sitting, which would lead to body damage, and become chronic and impossible to undo. The cervical (neck) regions, the 3–4 thoracic regions and the 4–5 lumbar regions are mostly under pressure, and easily wounded. The three regions should be paid more attention to when making a chair, especially the 3–4 lumbar regions which sustain the most pressure. If the backrest of the chair supports the 3–4 lumbar regions actively, people will be more comfortable (Leiqing Xu 2006).

Chinese traditional armchair with four protruding ends has curved backrest and headrest, which are prepared to support the curved spine properly. Chinese traditional armchairs with four protruding ends have different backrests of different curves influenced by the shape, the time and region that the furniture was made, including S-shaped, C-shaped and the straight backrests (Fig. 2.3). S-shaped

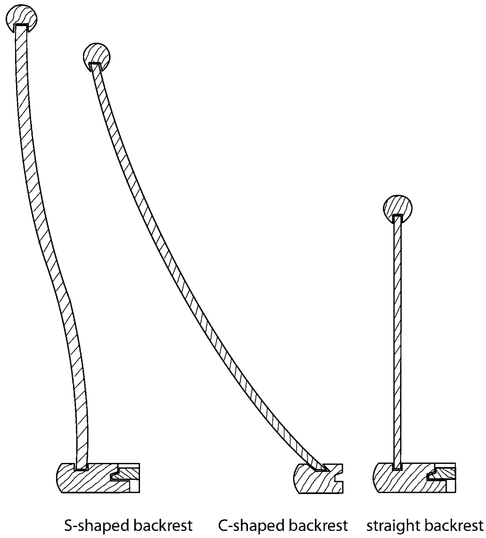


Fig. 2.3 Backrest of different shapes

backrests are more than the C-shaped, and the straight backrests are rare. The S-shaped backrests are the most compatible with the S-shaped curved spine. The 3–4 thoracic regions and the 4–5 lumbar regions of the spine get the proper support from the S-shaped backrest, and the cervical (neck) regions get the active support from the curved headrest (Fig. 2.4). The C-shaped backrest neglects the accurate support of the 4–5 lumbar regions, but its considerate C-shaped curve fit in with the whole spine properly, which wound also finish the function of comfort. The straight backrest is only fit for upright sitting, which is lack of the consideration of human factors engineering.

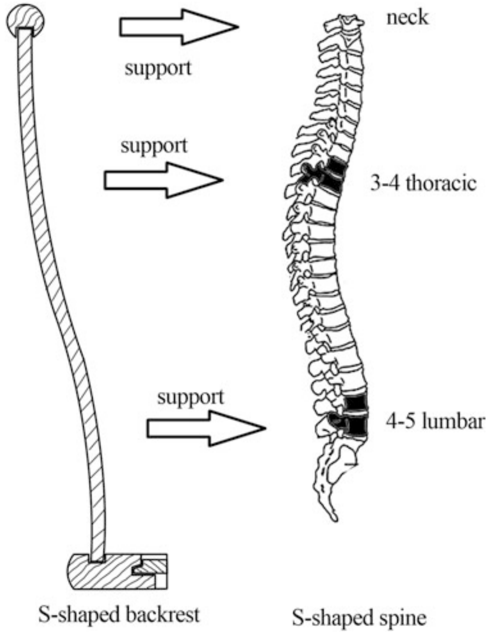
2.3.2 *The Arm Support*

Chinese traditional armchair with four protruding ends has the armrest to support the arm properly. The height of the armrest is so appropriate that the arm could be laid on it with ease. The armrest is also curve-shaped, and different chairs have armrests of different curves. The end of the armrest is round and smooth, which could help massage the palm of the hand. That would be useful for the health from the perspective of traditional Chinese medicine.

2.3.3 *The Buttock Support*

The maximum pressure of the buttock from the seat surface should be on the ischium, which could bear the burden most of the body. The pressure around the

Fig. 2.4 S-shaped backrest supports the S-shaped spine



ischium is gradually reduced. The back of knees should not bear any pressure from the front of the seat surface (Yulan Ding 2000). At the same time, the proper hardness of the seat surface is also important. If the seat surface is too loose, the pressure of the buttock and the thigh from the seat surface may increase, and it is not easy to change the seating posture, which may increase fatigue.

Chinese traditional armchair with four protruding ends has two different seat surfaces: the hard and the soft. The hard seat surface is made up of solid wood, and a soft and warm cushion would be put on it in winter, which is common in the northern area of China. The soft seat surface is woven by the hemp ropes, which is elastic and breathable. Therefore the soft seat surface develops well in the southern area of China.

2.3.4 The Foot Support

Chinese traditional armchair with four protruding ends has specific structure called the footrest stretcher to support the foot, which could help decrease the pressure of the back of the thigh. Therefore most of Chinese traditional chairs are a little taller than the ordinary chairs because of the footrest stretcher. There are two different ways to make use of the footrest stretcher. People can sit directly with their feet on the floor or on the footrest stretcher (Fig. 2.5). And the two ways can exchange with ease, which could make the body comfortable. If you are taller than the ordinary

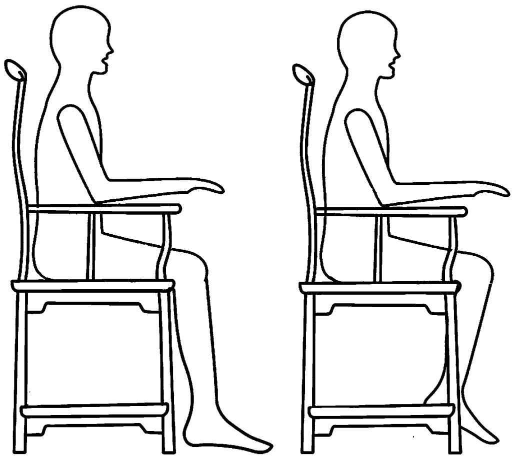


Fig. 2.5 Feet on the footrest and on the floor

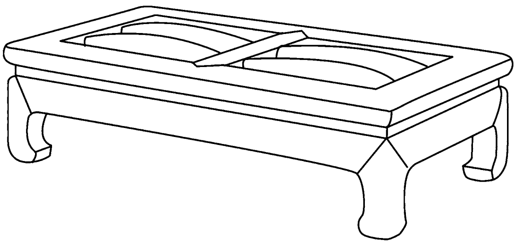


Fig. 2.6 Footstool

people, you can sit with your feet on the floor directly. And if you are shorter than ordinary people, you can sit with your feet on the footrest stretcher, which is considerate for people of different height.

There is also a particular category of Chinese traditional furniture called footstool, which is designed and made to support the feet (Desheng Hu 2010). When people sit in the chair or bed, in front of which the footstool is placed to provide the place for the feet. Some footstools even have the round sticks on the surface to message the feet when people trundle the stick with their feet continually, which is useful for their health (Fig. 2.6). Not all the people but high officials or noble lords could use these kinds of furniture like the footstool (Shixiang Wang 2005).

2.4 Sitting Posture

From the standpoint of the human factors engineering, we find that the protection of the muscle and the spine is contradictory when people sit on the chair (Sanders and McCormick 1985). Upright sitting is useful for the spine but presses the muscle. While bend sitting can make the muscle relax but useless for the spine. Therefore the

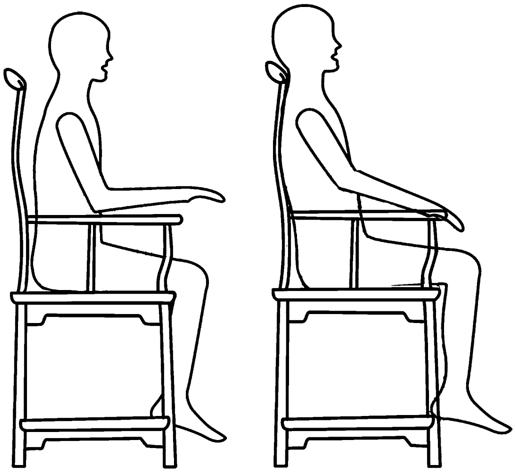


Fig. 2.7 Different postures to sit on the chair

exchange of upright sitting and bend sitting, the two different postures, in a certain period of time help both the muscle and the spine. Chinese traditional armchair with four protruding ends provide two corresponding postures to protect both the muscle and the spine (Fig. 2.7). We can sit in the chair with our back on the backrest and arms on the armrest for relaxation freely, and the spine could be supported by the backrest properly. Or just sit forward without the support of the backrest when we are busy at work. That means Chinese traditional armchair with four protruding ends is not only a chair for leisure, but also a chair for work.

2.5 Conclusion

Human factors engineering is employed to fulfill the two goals of health and productivity. Chinese traditional armchair with four protruding ends is one of the most common chairs of Ming-style furniture (Shixiang Wang 1985). It is not only a chair for leisure but also a chair for work, which take the health and productivity into consideration. Chinese traditional armchair with four protruding ends has the headrest, backrest, armrest, footrest and so on (Yao Yang 2002b), which take care of the body properly. We can sit forward on the chair and keep your mind on the work. After a while, we can sit with our back on the backrest and the arms on the armrest for a rest and relaxation, and the spine could be supported by the backrest properly, and the hand could be message gently.

In short, Ming-style furniture is doing well in human factors engineering. Based on the consideration of the social status and etiquette, the comfort was also the important factor to make furniture comfortable and considerate. Ming-style furniture usually fit in with the basic principle of human factors engineering although without the guidance of modern theory of human factors engineering.

Acknowledgment I would like to show my deepest gratitude to Mr. Yuanbo Sun, a respectable, responsible and resourceful scholar, who has provided me with valuable guidance. His keen and vigorous academic observation enlightens me not only in this thesis but also in my future study.

I shall extend my thanks to Mrs. Gong Xiaodong for all her kindness and help. I'd like to thank all my friends and my family for their encouragement and support.

References

Desheng Hu (2010) Twenty-four lectures of Ming and Qing palace furniture. The Imperial Palace Press, Beijing (in Chinese)

Fan Ye (2007) Hou Han Shu. Zhonghua Book Company, Beijing (in Chinese)

Gustav Ecke (1991) Chinese domestic furniture. Dizhen Press, Beijing (in Chinese)

Jiaqing Tian (2003) Appreciation and study of Ming and Qing dynasties. Cultural Relics Press, Beijing, pp 11–12 (in Chinese)

Leiqing Xu (2006) Human engineering and environment behavior studies. China Architecture and Building Press, Beijing (in Chinese)

Sanders MS, McCormick EJ (1985) Human factors in engineering and design. Wiley, New York

Shixiang Wang (1985) Appreciation of Ming-style furniture. Cultural Relics Press, Beijing, p 24 (in Chinese)

Shixiang Wang (2005) Collection of Ming-style furniture. Shanghai People's Publish House, Shanghai, p 117 (in Chinese)

Shixiang Wang (2008) Research on Ming-style furniture. Joint Publishing, Hong Kong, p 50 (in Chinese)

Yao Yang (2002a) Research on Ming-style furniture. China Architecture and Building Press, Beijing, pp 6–9 (in Chinese)

Yao Yang (2002b) Research on Ming-style furniture. China Architecture and Building Press, Beijing, pp 28–31 (in Chinese)

Yuanbo Sun (2010) Human factors engineering and design. Beijing Institute of Technology Press, Beijing, pp 130–136 (in Chinese)

YulanDing (2000) Human factors engineering. Beijing Institute of Technology Press, Beijing (in Chinese)

Zengbi Chen (1981) Function and styles of Ming-style furniture. Cult Relies 3:83–90 (in Chinese)

Zongshan Li (2001) History of Chinese traditional furniture. Hubei Fine Arts Publishing House, Hubei, pp 77–100 (in Chinese)

Chapter 3
The Evaluation on Working Fatigue Based on Improved Fuzzy Borda Method

Min-xia Liu, Jia-wei Ke, Jian Xie, Qing Xue, Huan Zhou, and Chen-xiao Guan

Abstract In order to promote the development of the social management mechanism on human working fatigue, and prevent the economic loss from working fatigue, it is necessary to make a quantitative evaluation on human working fatigue. First, based on the combination of entropy method, factor analysis method and comprehensive index method, this paper evaluates four fatigue indexes through questionnaires and experiments. Second, by using the improved fuzzy Borda method, this paper makes the combination evaluation of the results from the three comprehensive evaluation methods above. Finally, this paper draws the conclusion that, based on the result of the cluster analysis, the enterprise could improve its working tasks, incentive measures and working condition to reduce working fatigue.

Keywords Combination evaluation • Human factors engineering • Improved fuzzy Borda method • Working fatigue

3.1 Introduction

With the speeding up of economic development, people feel their life is becoming more and more stressful despite their material success. An increasing number of employees are affected by working fatigue which may jeopardize physical and mental health. A study points out that the working fatigue has been developed to a universal problem which would bring about social crisis. Besides, the study also puts forward the strategic planning to establish a perfect social crisis management system

M. Liu (✉) • J. Ke • J. Xie • Q. Xue • H. Zhou • C. Guan
School of Mechanical Engineering, Beijing Institute of Technology, Beijing, China
e-mail: liuminnxia@bit.edu.cn

E. Qi et al. (eds.), *The 19th International Conference on Industrial Engineering and Engineering Management*, DOI 10.1007/978-3-642-37270-4_3,

of fatigue (Guo Xianhong and Chen Xian 2008). Working fatigue is a systematic problem including complex factors. One of the most important task is to solve the problem that how to make it more reasonable to evaluate the working fatigue of the employees.

At present, the study of fatigue is mainly focus on the concept of fatigue, the theoretical model and influence factors (Zhan Fashang 2006). There is no method that can make a direct determination of fatigue (Qi Jiyan 2011). The evaluation of fatigue can be divided into three aspects, forecast beforehand, monitor during the process and evaluate after the event. Based on the decomposition of movement and the body's energy analysis, forecast beforehand focuses on estimation and simulation (Wang Xiangyin et al. 2010; Wang Wei et al. 2006). Besides, it evaluates fatigue model through the theoretical analysis (Gao Yongquan 2007); Monitor during the process mainly focuses on monitor fatigue through the physical and mental changes (Kang Weiyong et al. 2007) and how to provide real-time fatigue feedback and alert by using various fatigue monitoring device (Qi Jiyan 2011; Jin Jian 2002); Evaluate after the event mainly focuses on how to choose appropriate evaluation method to quantitatively analyze the fatigue (Jing Guoxun et al. 2006; Cai Na et al. 2009). This papers adopts the method of evaluate after the event to evaluate fatigue degree, and classify fatigue level, The result of evaluation will offer objective basis for enterprise to improve its working tasks, incentive measures and working condition in order to reduce working fatigue.

3.2 The Evaluation Method

Based on questionnaire survey method, CFF and grip strength test, this paper acquires four fatigue indexes including both subjective and objective aspects. After using entropy method, factor analysis and comprehensive index method to carry out a comprehensive evaluation, improved fuzzy Borda method is introduced to perform the combination analysis, and the cluster analysis method is used to divide fatigue level.

Considering that only one evaluation method lacks persuasion, this paper uses multimethod to make a comprehensive evaluation. By using the combination evaluation methods we could acquire a better fatigue evaluation of the employees. The methods which are commonly used include mean value method, Borda method, Copeland method and fuzzy Borda Method. Among which the fuzzy Borda method is more suitable to reflect the substance than other three methods for it takes account of the rank as well as the score (Gou Xianguang 1995). In addition, considering the fact that fuzzy Borda method cannot combine the evaluation value to the evaluation order, three improvement ideas is put forward, which made the evaluation more reasonable (Su Weihua and Chen Ji 2007). Moreover, In order to make a better combination evaluation, this paper evaluates working fatigue of employees who work on the assembly line by applying the improved fuzzy Borda method (Xu Ying and Hu Wei 2010; Sun Lijie 2011). Finally, cluster analysis is carried out using the result from the improved fuzzy Borda method, which could determine the level of fatigue degree.

3.3 The Experimental Method

Among the three experiments, Questionnaire investigation is subjective measurement while CFF experiment and grip strength experiment are objective measurements. Subjective measurement using questionnaire which lists questions about mental feeling and physical symptoms, and then investigation statistics method is applied to analyze the fatigue; Objective measurement measures the fatigue index by conducting experiment about human body. Four fatigue indexes are obtained from these fatigue measurements, they are Score of physical fatigue, Score of mental fatigue, Daytime change rate of Critical frequency and Daytime change rate of grip strength (Table 3.1).

3.3.1 Subjective Measurement

Subjective measurement reflects the subjective feeling of employees by using the questionnaire survey. Questionnaire survey was divided into two stages. In the first stage, an anonymous questionnaire survey was conducted. In the second stage, a registered questionnaire survey related to the grip strength experiment and CFF experiment is carried out. Thirty-seven people were asked to answer the anonymous questionnaire. We eliminated questionnaires incomplete answer and that most

Table 3.1 Data tables: four evaluation indexes

	Questionnaire		Experiment	
Employee	Score of physical fatigue	Score of mental fatigue	Daytime change rate of Critical frequency (%)	Daytime change rate of grip strength (%)
1	4.43	3.00	2.77	16.72
2	4.43	2.00	4.87	21.62
3	4.43	3.50	6.23	29.03
4	3.14	3.00	0.33	6.89
5	3.00	3.50	4.56	17.05
6	4.29	2.67	2.69	13.47
7	3.71	3.00	2.87	12.85
8	4.29	3.00	−0.31	13.07
9	4.00	2.67	2.99	13.17
10	4.14	2.50	5.16	13.45
11	4.43	3.00	4.3	17.17
12	3.86	2.33	3.93	11.98

Table 3.2 The reliability and validity test tables

Test method	Reliability test	Validity test		
		Content validity	Project validity	Structure validity
Result	Crowns Bach Alpha reliability coefficient = 0.777	Single topic options and overall options are related	Most of the questions are difficult and question 14 and 16 have low degree of identification	The accumulative contribution rate of 13 variables above 60 %
Evaluation	>0.75, credible	Effective contents	Modify the difficulty and identification degree of questions	Slightly higher than the medium level, the structure is effective

questions are checked the same. Thirty-three valid questionnaires were remained and the effective rate was 89.2 %. The second stage involved 12 people, and all of the 12 questionnaires were effective.

In order to ensure the accuracy of the questionnaire survey, it is necessary to examine the rationality and reliability of this questionnaire before the survey. Thus, a reliability and validity check is needed (Ke Huixin and Shen Hao 2005) (Table 3.2).

Overall, this table proves that the questionnaire is reliable and valid, and it is also suitable for the registered questionnaire survey.

3.3.2 *Objective Measurement*

In the second stage of the questionnaire survey, every five working days are grouped for a test unit. Twelve employees participated in the CFF experiment and grip strength experiment before and after work. The details of the experiment procedures are as follows.

- The procedure of grip strength experiment

First, the experimenter opened the lock of the hand dynamometer and reset the pointer. Then, the subject stood erectly, spreading legs and dropping arms naturally. Meanwhile, the subjects held the hand dynamometer with his palm inward and dial outward. After that, the subject gripped the dynamometer using his maximum strength in one time, and the experimenter filled in the data table with grip value.

- The procedure of CFF experiment

The experimenter set the red light for the test, 1/4 for the light intensity of back ground, 1:3 for the proportion of black and white. The data of each subject would be measured when the intensity of light in the highlights was 1/2, 1/8 and 1/32. The subject was asked to observe the highlight in the middle of the vision with his eyes clinging to the observation tube. During the test, the subject could regulate the frequency by twisting the frequency regulator. At beginning of the observation, if the subject found the highlight was flickering, he should increase the frequency to make the highlight seems just not flickering, and the experimenter recorded the frequency value; If the subject found the highlight was stable, he should decrease the frequency to make the highlight seems just flickering, and the experimenter records the frequency value.

3.4 The Collection of Date and Its Initialization

3.4.1 The Data of Questionnaire

Since fatigue including physical fatigue and mental fatigue, the data is divided into two parts. One includes seven questions relevant to physical fatigue. The other includes six questions relevant to mental fatigue. This paper assumes that the proportion of fatigue reflected by each question is equal and that options A$\sim$E separately correspond to $1 \sim 5$ points. Thus the weighted average value of each part would serve as the index of the subjective assessment. The two results separately represent the employee's scores of physical fatigue and mental fatigue, which are presented in Table 3.1.

The table shows that employee 1 only gets 3.00 scores in mental fatigue which is lower than others. Since the questionnaire survey shows that employee 1 has a shorter working experience (between 3 month and half a year) and has a good habit of eating breakfast, it can be regard that working experience and breakfast eating habit have certain influence on mental fatigue.

3.4.2 The Data of Experiment

Through measuring the maximum value of grip strength of each employee before and after work in 5 days and getting the average value of these data, we got the results as the experimental data before and after work in this stage. In the experiment of CFF, the data of each employee when the intensity of light in the highlights is 1/2, 1/8 and 1/32 was measured. Then, the average critical flicker fusion frequency of each employee before and after class was calculated. Next, we respectively calculated the daytime change rate of the maximum value of grip strength and that of the critical flicker fusion frequency. The daytime change rate is used as the evaluation index of fatigue. Formula of the daytime change rate is represented as follows:

$$\text{Daytime change rate} = \frac{\text{The average before work} - \text{The average after work}}{\text{The average before work}} \qquad (3.1)$$

The daytime change rate of the maximum value of grip strength and that of the critical flicker fusion frequency of each employee are presented in Table 3.1.

3.5 Combination Evaluation and Analysis

3.5.1 Comprehensive Evaluation

Using the entropy method, factor analysis method and comprehensive index method, 12 employees in the production line were taken into account by the comprehensive evaluation. The result of the comprehensive evaluation is presented in Table 3.3. From the table, it shows that there is a little difference between three methods.

3.5.2 The Consistency Check

The sort results acquired from the three methods above do not belong to the normal distribution. So the method of non-parametric spearman rank correlation analysis related inspection was used to test its consistency and the result is presented in Table 3.4. The spearman correlation coefficient shown on the table are all above 0.800 which explain that the sort results acquired from the three comprehensive evaluation method above has the consistency. Therefore, the improved fuzzy Borda method is feasible for combination evaluation.

Table 3.3 Evaluation result of the 12 employees' fatigue degree and ranking

	Entropy method		Factor analysis		Comprehensive index method		Fuzzy Borda method	
Employee	Score	Order	Score	Order	Score	Order	Score	Order
1	0.08	7	0.09	4	0.98	6	7.18	6
2	0.11	2	−0.89	5	1.27	2	10.69	2
3	0.14	1	3.29	1	1.65	1	12.00	1
4	0.03	12	−0.94	12	0.45	12	1.00	12
5	0.10	3	2.27	2	1.17	3	10.38	3
6	0.07	10	−0.89	9	0.88	10	3.17	10
7	0.07	9	0.04	7	0.89	9	4.58	9
8	0.04	11	−0.97	11	0.53	11	2.00	11
9	0.07	8	−0.71	8	0.90	8	5.00	8
10	0.10	4	−0.57	6	1.16	5	8.28	5
11	0.10	5	0.51	3	1.17	4	8.88	4
12	0.08	6	−1.22	10	0.96	7	6.19	7

Table 3.4 Spearman rank correlation coefficient

	Entropy method	Factor analysis	Comprehensive index method
Entropy method	1.000	0.832	0.986
Factor analysis	0.832	1.000	0.895
Comprehensive index method	0.986	0.895	1.000

3.5.3 The Combination Evaluation by Using the Improved Fuzzy Borda Method

The improved Fuzzy Borda method has a change in rank with the normal one. The original linear function is replaced by the linear function. (See function 3.2)

$$Q_h = n - h + 1 \tag{3.2}$$

Where, Q_k is the score after the change, n is the sample size, h is the rank.

The procedures of improved fuzzy Borda method are described below:

1. By using the range transform method, the membership degree of comprehensive evaluation for every member is worked out.
2. Calculating the fuzzy frequency of the evaluation member which ranks h $(1 \leq h \leq n)$.
3. Calculating the score of the member which rank h $(1 \leq h \leq n)$.
4. Calculating the score of improved fuzzy Borda for every member and giving them the order based on this score.

The result is presented in Table 3.4. It is approximate to the order from the comprehensive evaluation.

3.5.4 Systematic Cluster Analysis

By using the average distance between classes of systematic cluster method, we applied the 12 employees' score acquired from combination evaluation to carry out the cluster analysis. The result is presented in Fig. 3.1. From the tree diagram, when the cluster distance is less than 5, the top three employees are 3, 2 and 5 which is the first kind belonging to the kind of the most tired. Ranking 4–7 of all employees are 10, 11, 1 and 12, which is the second kind. Ranking 8–10 of all employees are 9, 7 and 6, which is the third kind. Ranking 11–12 of all employees are 8 and 4, which is the fourth kind.

The result from the comprehensive evaluation of the employees' working fatigue is influenced by various factors, such as the choice of the evaluation index system,

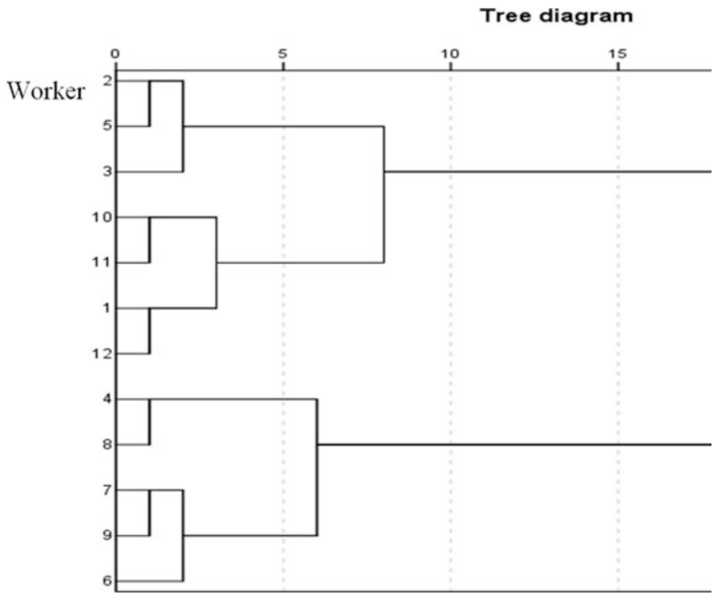


Fig. 3.1 Tree diagram of cluster analysis

Table 3.5 The stratification diagram of the degree of working fatigue of employees

The degree of working fatigue	Employee
The first kind	3, 2, 5
The second kind	11, 10, 1, 12
The third kind	9, 7, 6
The fourth kind	8, 4

the selection of evaluation methods and the accuracy of the data etc. Based on the survey of questionnaire and objective experiments, combining the scores of physical fatigue and mental fatigue in the questionnaires with the daytime rata of change of critical frequency and grip strength in the experiments, this paper applied the comprehensive evaluation method including entropy method, factor analysis method and comprehensive index method to make evaluation of the employees' fatigue. Then, we made combination evaluation by using the improved fuzzy Borda method and got the final score and ranking. In order to stratify the degree of fatigue of the employees, the method of systematic cluster analysis was adopted to classify the degree of fatigue of the employees into four kinds. The results of the stratification are presented in Table 3.5.

Since 3, 2 and 5 are featured with that all of the three indexes are ranking before, the result of the combination evaluation is considered in accordance with the actual situation. Thus, the model of the evaluation of the improved fuzzy Borda method is reasonable.

3.6 Conclusion

Applying the idea of the combination evaluation and cluster analysis, quantify the degree of the employees' working fatigue in the production line and classify the kinds of the fatigue have a great application prospect in the management of production. Based on the questionnaire survey and objective experiment, the results acquired from combination evaluation are relatively objective and accurate. According to the kinds classified by systematic cluster analysis, the enterprise can ensure the relative degree of fatigue among employees. And on the basis of this, enterprise can reduce the degree of fatigue through rational allocation of tasks, taking the measure of incentives and improving the working condition. Moreover, by doing this, not only the workers can improve their enthusiasm, but also the enterprises can improve production efficiency, optimize the allocation of resources and promote the virtuous cycle of the production.

The method of combination evaluation can sum up the influences of the overall effect on all relevant factors and make the evaluation of the employees' fatigue more objectively and accurately. However due to the limit of the experimental equipments and condition, the experiment just choose the daytime change rate of grip strength and that of the critical frequency as the evaluation index, which could cause the result lack of accuracy. For this, improvement will be carried out in the future research.

References

Cai Na, Hu Jiangbi, Liu Xiaoming (2009) The analysis on fatigue of traffic police during Beijing Olympic games. J Chin People's Public Secur Univ (Sci Technol) 3:63–67 (in Chinese)

Gao Yongquan (2007) A tiredness model of monotone operation based on ergonomics. In: Mechanical management and development academic annual conference, 2007, pp 116–117 (in Chinese)

Guo Xianguang (1995) A new integrated evaluation method – combinatorial evaluation method. Stat Methods Med Res 5:56–59 (in Chinese)

Guo Xianhong, Chen Xian (2008) Crisis management and prevention of employees' work-fatigue. China Occup Saf Health Assoc 3:46–52 (in Chinese)

Jin Jian (2002) Research on driving fatigue mechanism and feedback and selection mode. Southwest Jiaotong University, Chengdu (in Chinese)

Jing Guoxun, Zhang Yongquan, Zhang Junbo (2006) Comprehensive evaluation on human fatigue based on multi-stage fuzzy synthetic judgment. China Saf Sci J 16(6):55–59 (in Chinese)

Kang Weiyong, Yuan Xiugan, Liu Zhongqi, Dong Dayong (2007) Analysis of relations between changes of pupil and mental workloads. Space Med Med Eng 20(5):364–366 (in Chinese)

Ke Huixin, Shen Hao (2005) Statistic in science research. Communication University of China Press, Beijing (in Chinese)

Qi Jiyan (2011) Analysis and monitoring of driving fatigue. Automob Sci Technol 1:34–38 (in Chinese)

Su Weihua, Chen Ji (2007) The defects of fuzzy Borda method and its improvement. Stat Methods Med Res 7:58–64 (in Chinese)

Sun Lijie (2011) The evaluation of the construction degree of Dalian' technology innovation system based on fuzzy Borda method. In: IEEE proceedings on international conference on management and service science, 978-1-4244-6581-1/11, 2011

Wang Wei, Tang Xiaofeng, Li Yan (2006) Human fatigue analysis based on maintenance simulation. J Syst Simul 18:245–248 (in Chinese)

Wang Xiangyin, Liu Jian, Yu Dejie, Pu Hairong (2010) Research on operator fatigue improvement in assembly line based on metabolic energy expenditure. J Hunan Univ (Nat Sci) 5:31–34 (in Chinese)

Xu Ying, Hu Wei (2010) The application of hospital quality evaluation based on improved fuzzy Borda method. Chin J Health Stat 27(2):187–188 (in Chinese)

Zhan Fashang (2006) Present study of fatigue. Chin J Behav Med Sci 15(2):188–189 (in Chinese)

Chapter 4
A Study on Hydraulic Autofrettage Gun Tube Dynamic Stress by Finite Element Analysis

Yan-mou Zhan, Guo-rui Huang, and Hao Cheng

Abstract Gun tube dynamic strength design, first of all, is to enable the identification of dynamic stress on each section of tube, and the dynamic stress calculation in analytical method or numerical method is relatively complex. For most non mechanical professional design staff, the use of finite element soft- ware interactively in the general structure of the model has certain difficulty. Therefore, the realization of gun tube dynamic strength design is an urgent need to develop a general engineering program, so that the designer can easily use this program to achieve the dynamic stress on every section of tube, and truly realize the dynamic strength design of tube. With the APDL language of ANSYS finite element analysis software platform, a suitable program for tube dynamic stress calculation procedures is developed. This program provides an engineering foundation to realize the tube dynamic strength design.

Keywords Dynamic stress • Finite element analysis • Hydraulic autofrettage

4.1 Introduction

In the past, the tube design and strength calculation are mostly performed by static analytic estimation of tube stress and deformation. As the artillery weapon lightweight requirements increase, static strength design has not been able to precisely meet the modern gun design requirements (Shim et al. 2010), the application of modern analytical technology for gun structure to carry out accurate analysis is imperative, in which the finite element technique is the most widely-used method (Huang and Cui 2006). However, due to the finite element technology for users

Y. Zhan (✉) • G. Huang • H. Cheng
New Star Research Institute of Applied Technology, Hefei, China
e-mail: zym_196310@163.com; huangdava@sina.com; cho25086@163.com

E. Qi et al. (eds.), *The 19th International Conference on Industrial Engineering and Engineering Management*, DOI 10.1007/978-3-642-37270-4_4,

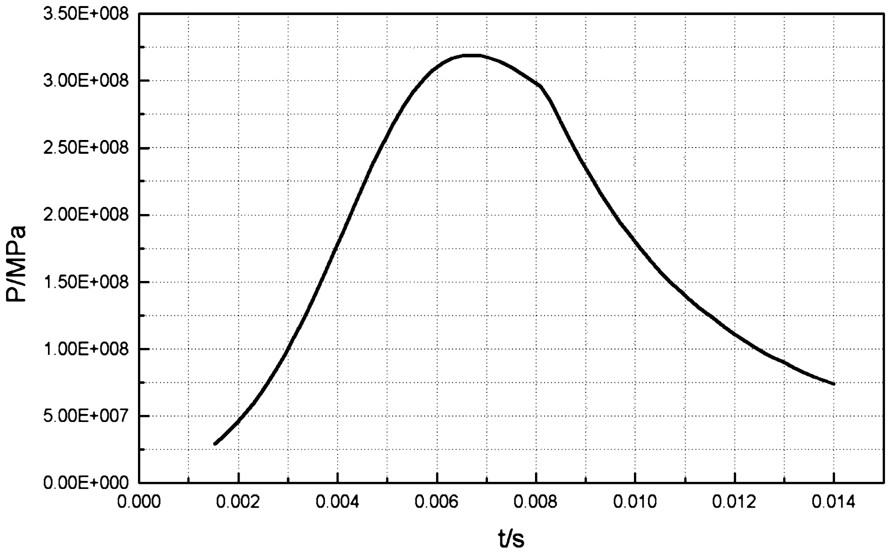


Fig. 4.1 Radial load of tube beforetime

with high demands, one must have a solid mechanical expertise, be familiar with the basic functions of the application software and application of skills and have a wealth of engineering experience. These are difficult for a designer in terms of both. Therefore, there is an urgent need to develop to facilitate the dynamic strength of Universal Barrel finite element program used by the designers (Nelson and Petillo 2004; Cao et al. 2011).

Studies have shown that different impulse load types generate different dynamic effects for the same structure (Swardt and Rolf 2006). Therefore, the dynamic strength design of tube is based on the correct definition of radial load types on different tube positions. In the past, the strength design of tube was simplified statically because the load of tube was taken merely as a single pulse type (Coco et al. 2007; Zhang et al. 2009). The tube radial loads on different parts along the axial direction were simplified as a pulse type shown in Fig. 4.1 (Shi et al. 2008). Theoretical studies have shown that this approach does not correctly reflect the load types on the different parts of the tube, as a matter of fact; the load type is different in the parts of the tube. For the chamber section, the load types of all sections can be simplified as shown in Fig. 4.1. For the straight tube section, due to the projectile forward movement, the load type can be simplified as shown in Fig. 4.2 or Fig. 4.3 (Shi et al. 2008). Before the maximum tube pressure point, it can be shown in Fig. 4.2, and after the point, it can be shown in Fig. 4.3. Therefore, the radial load types on the inner surface of tube can be basically simplified as the three types shown in Figs. 4.1 and 4.2 and Fig. 4.3. From all three graphics seen, they are all or part of the pressure curve and can generate different dynamic responses. Load type shown in Fig. 4.1 can show pulse dynamic effects, while load types shown

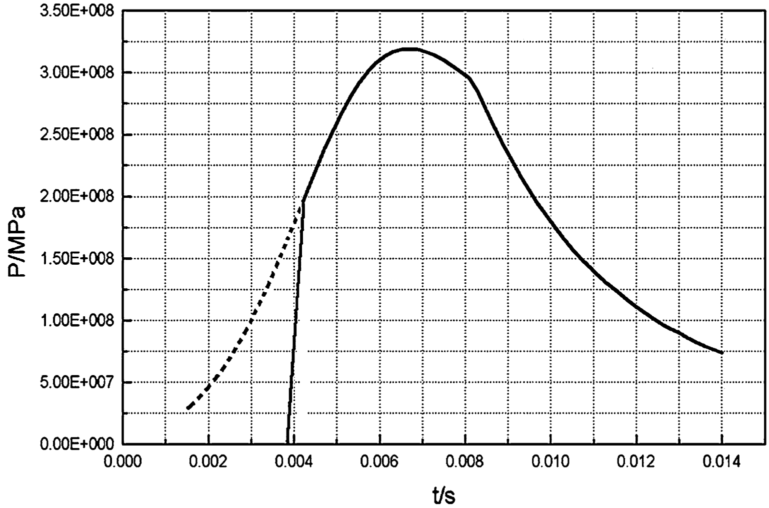


Fig. 4.2 Radial load before the maximum pressure point in the straight tube section

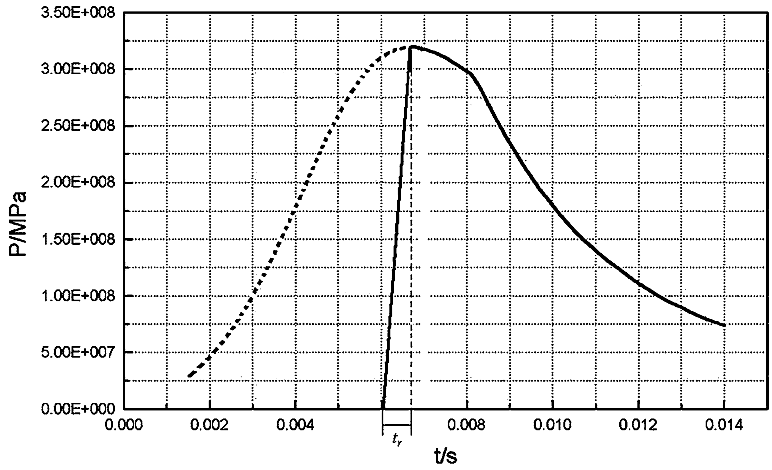


Fig. 4.3 Radial load after the maximum pressure point in the straight tube section

in Figs. 4.2 and 4.3 can show step pulse dynamic effects. Studies have shown that dynamic loads for a majority of tubes can be simplified as static loads if we deal with the cross-section load as shown in Fig.4.1. However, if we deal with the load as shown in Fig. 4.2 or Fig. 4.3, the tube cross-section loads must be handled as the dynamic problem. In fact, the tube radial load consists of three types as shown in Figs. 4.1, 4.2 and 4.3. Therefore, tube strength design according to the dynamic problem is in line with the actual design (Sunwall et al. 2012; Liu et al. 2010).

4.2 General Finite Element Program for Tube Dynamic Stress

Gun tube dynamic strength design, first of all, is to enable the identification of dynamic stress on each section of tube, while dynamic stress calculation for both analytical methods or numerical methods are relatively complex. For most non-mechanical professional design staff, the use of finite element software interactively in the general structure of the model has certain difficulty (Li et al. 2005; Tan et al. 2010). Therefore, the realization of gun tube dynamic strength design is an urgent need to develop a common engineering program (Maleki et al. 2010), enable the designer to facilitate the use of the program obtained tube every section of the dynamic stress, and truly realize the dynamic strength design of tube. It is based on this thought and spirit of reliable, convenient, practical principles, with the help of the software ANSYS platform, using ANSYS software for the APDL language, suitable for the development of the rifled musket, and hydraulic self tightening gun barrel dynamic stress calculation program. This program provides an engineering foundation to realize the tube dynamic strength design.

According to the design of the tube, the use of this procedure only need the material parameters (elastic modulus, Poisson's ratio, density, yield limit and tangential modulus), geometric parameters (inner and outer radius calculation section, line and line width, depth, rifling chamfer radius) and data load in the interface and deposit in the designated path, start the ANSYS program, from ANSYS software interface in the file menu to read into the specified path under the APDL program files (File → Read input from . . . To file name > OK). The ANSYS software can automatically set the barrel section of dynamic stress, while the output stress varies with time and the corresponding data file. Designers need not be finite element modeling, meshing and applying dynamic load and complex before the process can get the tube dynamic stress calculation result. The program design block diagram was shown in Fig. 4.4.

Program design of the overall process is as follows:

1. software and display settings
2. building the tube cross-section desired parameter setting
3. the material parameter setting
4. unit selection and material specification
5. if a rifled gun, rifle geometric data set

```
*if,nnum,gt,0,then
/prep7
hland =        !Width of land line (m)
hgroove =      !Width of groove line (m)
hdep =         !Depth of tube (m)
drad =     !Junction chamfer of land line and groove line (m)
```

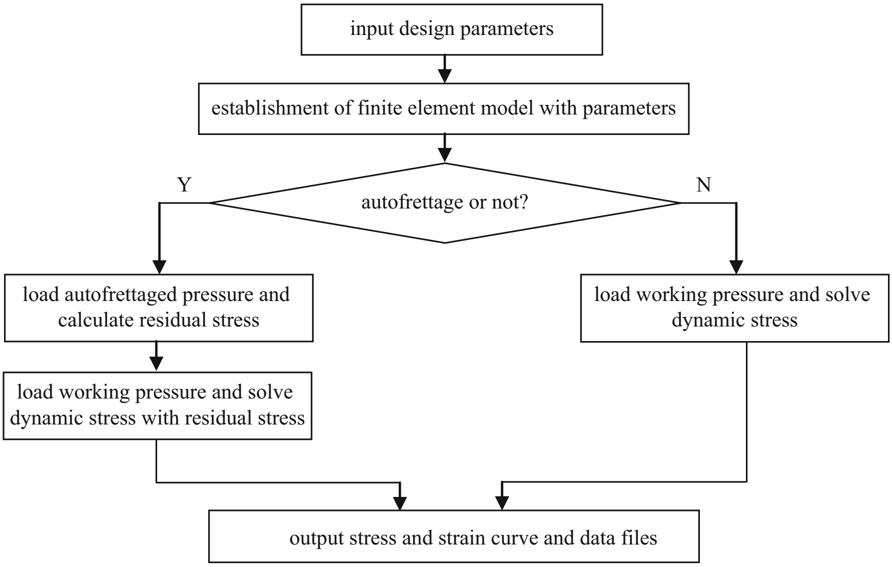


Fig. 4.4 The tube dynamic stress analysis diagram

6. if a rifled gun, need to establish the tube section model
7. if land line and groove line junction chamfer, need to establish the chamfer model

```
*if,drad,gt,0,then
LFILLT,13,9,drad,,
LFILLT,10,14,drad,,
lsel,s,,,15,18,3,1
cm,fuzhu2,line
asbl,2,fuzhu2
NUMCMP,AREA
adele,2,3,1,1
NUMCMP,AREA
alls
……
```

8. if it is a smooth tube gun, need to establish the tube section model
9. if the autofrettaged pressure of existence, undertake the autofrettaged residual stress calculation
10. input the chamber pressure and the projectile travel data
11. according to the tube pressure curve interpolation to solve the tube section of the actual load curve, load and transient calculation
12. the results of processing
13. output of the key points of the stress

4.3 The Tube Dynamic Stress Calculation

4.3.1 Tube Material Parameters

Modulus of elasticity $E = 2.1 \times 10^5$MPa, Poisson's ratio $\mu = 0.3$, Density $\rho = 7,800\ \mathrm{kg/m^3}$.

The design parameters of each section of tube are shown in Table 4.1. Static stress calculation formula is as follows (Zeng et al. 2007):
the tangential stress on inner surface

$$\sigma_{\theta a} = \frac{b^2 + a^2}{b^2 - a^2} p \tag{4.1}$$

the tangential stress on outer surface

$$\sigma_{\theta b} = \frac{2a^2}{b^2 - a^2} p \tag{4.2}$$

if axial stress is zero, equivalent stress on inner surface

$$\sigma_{d4} = \frac{\sqrt{3b^4 + a^4}}{b^2 - a^2} p \tag{4.3}$$

a is the inner radius of tube, b is the outer radius of tube, p is the inner pressure of the tube.

Note: load rise time is the duration through which the load rises from zero to the original tube pressure value, expressed as t_r. $\sigma_{\theta a}$ is the tangential stress on inner surface, σ_{ra} is the radial stress on inner surface, σ_{d4} is the equivalent stress on inner surface, $\sigma_{\theta b}$ is the tangential stress on outer surface.

4.3.2 Calculation Results

Taking a gun tube as an example and applying the program and formulas (4.1), (4.2) and (4.3), the dynamic and static stresses on different sections along the tube are given in Table 4.1. Dynamic stress curves are shown in Figs. 4.5, 4.6, 4.7, 4.8, 4.9 and 4.10. In these figures, the unit of stress is Pa; the unit of time is sec.

4.3.3 Analysis of Calculation Results

Found by computation, the results of dynamic stress and static stress are almost the same (the error is less than 1 %). That is to say, the chamber section, whether by

Table 4.1 The calculation results of tube dynamic and static stress

Section location and tube size			Calculation parameters/MPa	Dynamic stress/MPa	Static stress/MPa	Ratio of dynamic and static stresses
Chamber section	$a = 87.4$ mm	Load rise time $t_r = 6.5$ ms	$\sigma_{\theta a}$	548.6	547.0	1.003
	$b = 170$ mm		σ_{ra}	−318.3	−318.3	1
	$l = 44$ mm		σ_{d4}	752.1	758.1	0.992
	$p = 318.3$ MPa					
Maximum pressure section	$a = 78.6$ mm	Load rise time $t_r = 0.1$ ms	$\sigma_{\theta a}$	687.3	497.0	1.383
	$b = 168$ mm		σ_{ra}	−318.5	−318.5	1
	$l = 1915$ mm		σ_{d4}	874.0	711.9	1.228
	$p = 318.5$ MPa	Load rise time $t_r = 0.5$ ms	$\sigma_{\theta a}$	540.0	497.0	1.087
			σ_{ra}	318.5	−318.5	1
			σ_{d4}	744.6	711.9	1.046
Section at test point	$a = 87.4$ mm	Load rise time $t_r = 0.1$ms	$\sigma_{\theta a}$	602.4	461.2	1.212
	$b = 170$ mm		σ_{ra}	−227.8	−143	1
	$l = 44$ mm		σ_{d4}	721.9	538.7	1.123
	$p = 227.7$ MPa		$\sigma_{\theta b}$	301.0	233.5	1.289
		Load rise time $t_r = 0.5$ ms	$\sigma_{\theta a}$	496.7	452.8	1.064
			σ_{ra}	−227.8	−227.8	1
			σ_{d4}	628.5	608.0	1.033
			$\sigma_{\theta b}$	250.3	233.5	1.072
			measured value of $\sigma_{\theta b}$	(238.8)	233.5	1.023

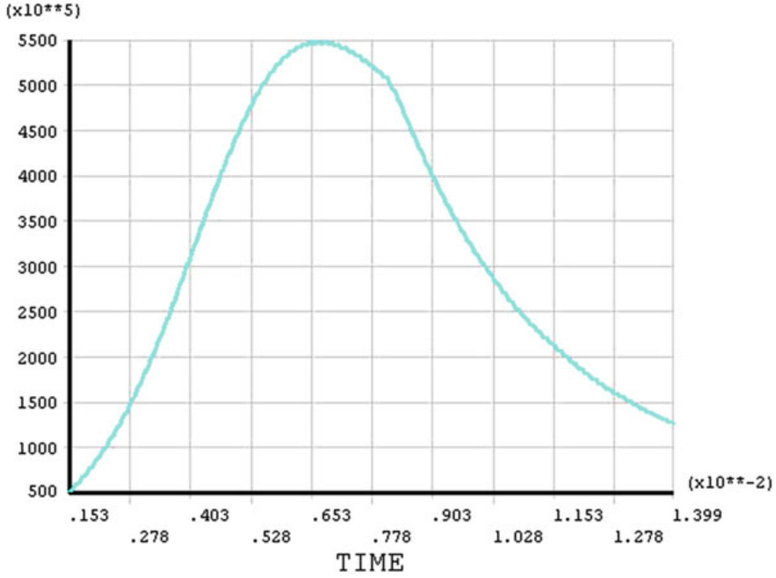


Fig. 4.5 The tangential stress on inner surface of the tube chamber section

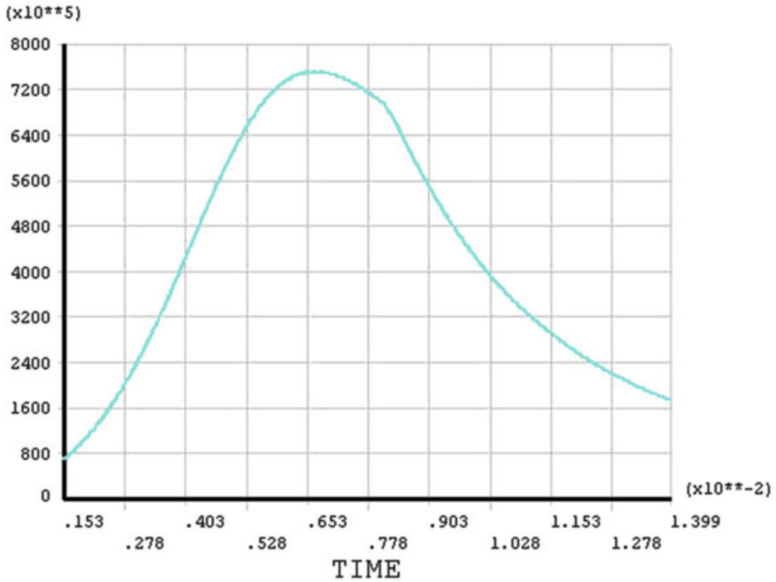


Fig. 4.6 The equivalent stress on inner surface of the tube chamber section

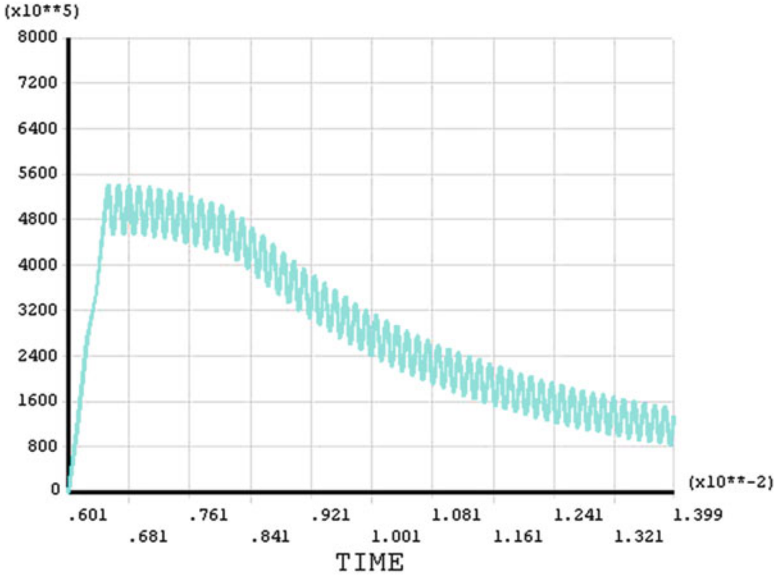


Fig. 4.7 The tangential stress on inner surface of the tube maximum pressure section ($t_r = 0.5$ ms)

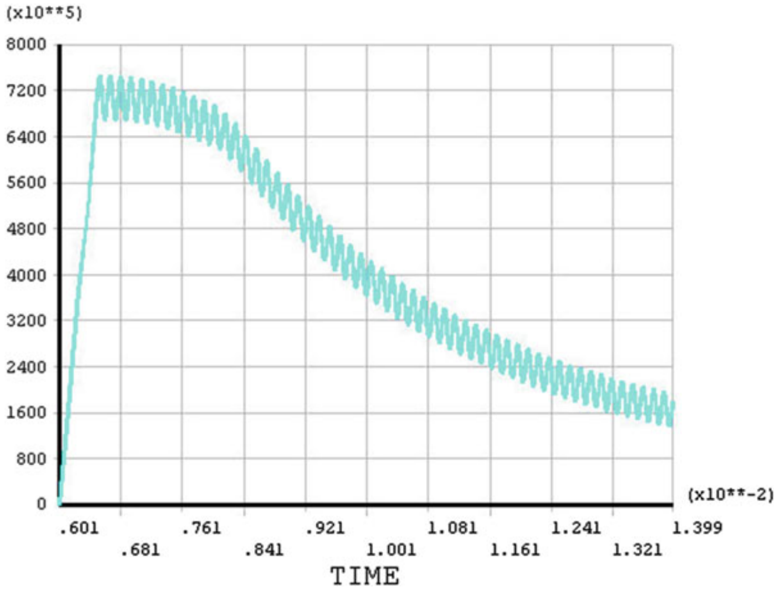


Fig. 4.8 The equivalent stress on inner surface of the tube maximum pressure section ($t_r = 0.5$ ms)

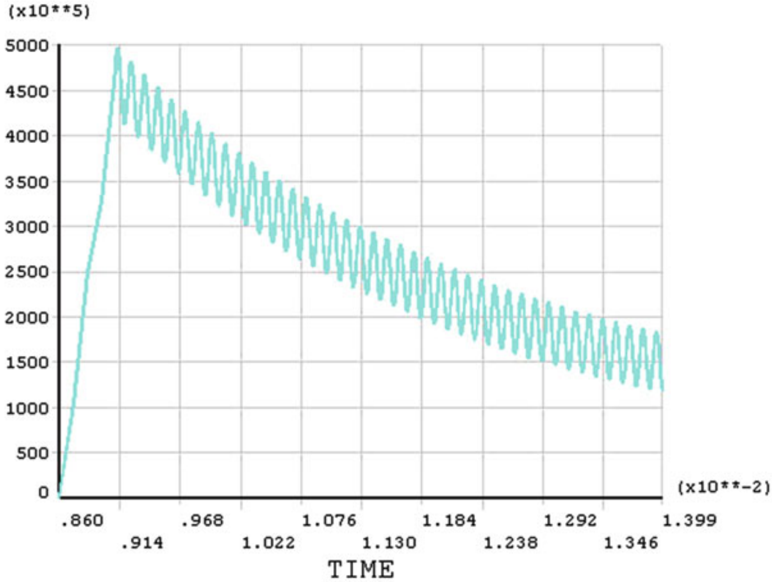


Fig. 4.9 The tangential stress on inner surface of the tube section at test point ($t_r = 0.5$ ms)

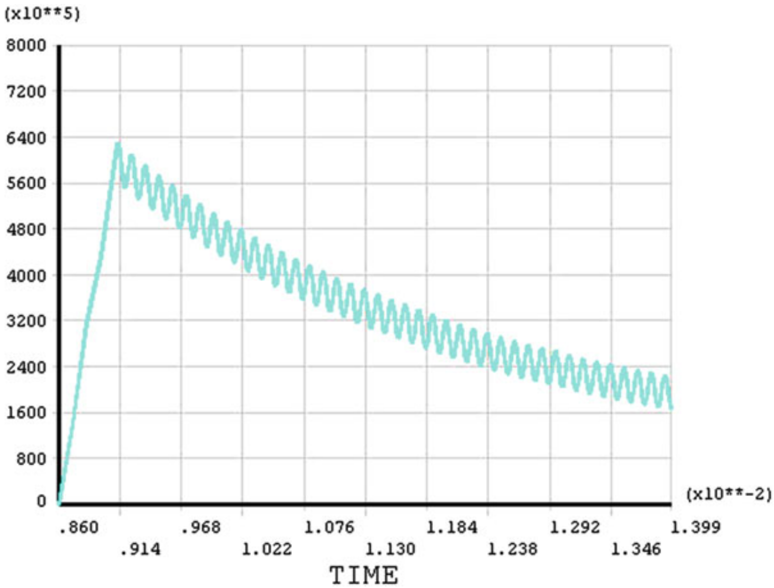


Fig. 4.10 The equivalent stress on inner surface of the tube section at test point ($t_r = 0.5$ ms)

dynamic method or by the static method, the results are basically the same. The chamber section is always loaded with a non-step pulse pressure shown in Fig. 4.1. At the same time, the maximum value (T) of the inherently radial vibration cycle of the section is about 0.14 ms, the load duration (τ) is about 14 ms, and load rise time (t_r) is about 6.5 ms. According to the general engineering judgment criteria (Bahre and Brunnet 2011) of the dynamic and static system response, the response of a non step pulse load can be simplified as a static problem if $\tau/T \geq 4$. For this example, $\tau/T = 100$, is much higher than 4. Therefore, the calculation results by static method are almost identical with the results by dynamic method, and proved by the facts.

Because the straight tube section is always loaded with a step pulse pressure shown in Fig. 4.2. or Fig. 4.3, the stress response is closely related to the rise time. For the tube dynamic stress calculation, the key problem is to determine the load rise time (t_r) according to the tube firing conditions. According to the general engineering judgment criteria (Bahre and Brunnet 2011) of the dynamic and static system response, the response of a step pulse load can be simplified as a static problem if $t_r/T \geq 3$ and as a dynamic problem if $t_r/T < 3$. For the straight tube section, the calculation results shown in Table 4.1 have been certificated the effect of load rise time on the system response and the preceding criteria.

4.4 Conclusion

For general dynamics problem, its response is judged by the ratio of load duration and maximum system inherent cycle, but to the tube under the step load pressure, its effect should be judged by the ratio of load rise time and maximum system inherent cycle. Because the straight tube section is always loaded with a step pulse pressure, the key problem is to determine the load rise time (t_r) according to the tube firing conditions. Step load is different from non step load, it should be considered as a dynamic problem, theoretical and experimental studies have also proved this point.

References

Bahre D, Brunnet H (2011) Simulation of removing autofrettage-induced residual stress loaded layers by finite element analysis. Procedia Eng 19:9–15

Cao LN, Han XQ, Zhang YG, Cao YX (2011) Finite element analysis of detonation wave interference among perforating bullets. Baozha Yu Chongji/Explos Shock Waves 31:220–224

Coco S, Laudani A, Pollicino G, Dionisio R, Martorana R (2007) 3-D finite-element analysis of TWT grid electron guns. IEEE Trans Magn 43:1233–1236

Huang XP, Cui WC (2006) Effect of Bauschinger effect and yield criterion on residual stress distribution of autofrettaged tube. J Press Vessel Technol Trans ASME 128:212–216

Li ZY, Cheng SK, Zheng P, Feng M, Cao YJ (2005) Finite-element analysis for the active electromagnetic armor projectile interceptor. IEEE Trans Magn 41:453–455

Liu SB, Ruan JJ, Du ZY, Huang DC (2010) Analysis of inductive coil gun performance based on field coupling circuit method. Diangong Jishu Xuebao/Trans China Electro Tech Soc 25:1–7
Maleki M, Farrahi GH, Haghpanah JB, Hosseinian E (2010) Residual stress analysis of autofrettaged thick-walled spherical pressure vessel. Int J Press Vessels Piping 87:396–401
Nelson EM, Petillo JJ (2004) An analysis of the basic space-charge-limited emission algorithm in a finite-element electrostatic gun code. IEEE Trans Plasma Sci 32:1223–1235
Shi LP, Zhong YS, Yu J, He XD (2008) Finite element modeling of residual stress of Ti/TiAl multilayer composites prepared by electron beam physical vapor deposition technology. J Comput Theor Nanosci 5:1704–1707
Shim WS, Kim JH, Lee YS, Cha KU, Hong SK (2010) A study on hydraulic autofrettage of thick-walled cylinders incorporating bauschinger effect. Exp Mech 50:621–626
Sunwall DA, Speece MA, Pekar SF (2012) Advances in on-sea-ice seismic reflection methods using an air gun: McMurdo sound, Antarctica. Geophysics 77:19–30
Swardt D, Rolf R (2006) Material models for the finite element analysis of materials exhibiting a pronounced Bauschinger effect. J Press Vessel Technol Trans ASME 128:201–207
Tan LB, Cai L, Yuan RS (2010) Finite element analysis and test research for dynamic strength of thick-wall cylinder. Dandao Xuebao/J Ballist 22:86–88
Zeng ZY, Ning BF, Wang ZS (2007) A finite element solution for barrel dynamic stress. J China Ordnance 3:11–14
Zhang QM, Wang XY, Huang FL, Chen L, Guo XH (2009) An experimental and numerical study of the dynamic response of a free-free aluminum beam under high velocity transverse impact. Int J Impact Eng 36:1385–1393

Chapter 5
A Cooperative Game Model of Dynamic Alliance Under Grid Environment

Xiang-bin Zhang and Min Wu

Abstract This paper mainly researches on how members of dynamic alliance make a cooperative game under manufacturing grid environment. Determined the contribution level which is made by resources suppliers to demanders by adjusting the proportion of resources provided by suppliers, so that to maximize the total revenue of this alliance. First, this paper establishes a differential dynamic equation for members of this alliance. Then it determines the performance index function of alliance's total revenue. Finally, by using the pontryagin's maximum principle to solve this problem, it gets the Nash equilibrium and through an example to verify the feasibility of this method.

Keywords Cooperative game • Dynamic alliance • Manufacturing grid • Maximum principle

5.1 Introduction

Manufacturing grid is a specific application of Grid technology in manufacturing industry, aiming at providing the industry with a scalable resource sharing and collaborative working environment by taking advantages of the openness and flexibility of the grid, so that distributed and heterogeneous manufacturing resources can be shared and work collaboratively to achieve the goal of responding quickly

---

Funds Project: National Nature Science Foundation of China (70792083)

X. Zhang (✉) • M. Wu
Department of Economics and Management, Nanjing University of Posts and Telecommunications, Nanjing, China
e-mail: shadowkimi@sina.com

E. Qi et al. (eds.), *The 19th International Conference on Industrial Engineering and Engineering Management*, DOI 10.1007/978-3-642-37270-4_5,

to market, reducing processing costs and improving competitive of resources (Zhou and Chen 2007). In the paper of "The Anatomy of the Grid: Enabling Scalable Virtual Organizations", Ian Foster pointed out that "The real and specific problem that underlies the Grid concept is coordinated resource sharing and problem solving in dynamic, multi-institutional virtual organizations (Ian et al. 2001)". As a dynamic alliance –the main form of manufacturing virtual organizations under gird environment (Liu et al. 2009), manufacturing gird need to try more to solve the problem it encounters. On one side, enterprises under a dynamic alliance not only have manufacturing resources with respectively characteristics but also have different resource allocation methods and management mechanism (Piao and Zhu 2005), one the other side, they also have the features of coordinated resource sharing and problem solving (Ian and Carl 2004), and by the sharing and allocation of resource they can maximize the profit of the alliance. And the reason why enterprises choose to be an alliance member is that they realize if they have relevant targets, self-serving need partners while not for altruism (Dean and David 1993). Therefore it becomes an important issue for dynamic alliance under grid environment, which is how to optimize resource allocation strategies to achieve a win-win goal of maximization of individual and alliance benefit via this kind of cooperative competition.

For the issues of resource allocation under grid environment: the Hesam Izakian et al. (2010) proposed a continuous double auction method (CDA), which allows demanders and suppliers make decision independently (Izakian et al. 2009, 2010), the (He et al. 2006) pointed out that resource allocation is the bottleneck point and core problem of manufacturing grid resource management nodes and proposed a TQCS-based multi-objective integer programming algorithm (Liu et al. 2003, 2004; He et al. 2006; Shi et al. 2004), Buyya et al. (2002) proposed the basic architecture of the grid resource allocation based on market mechanisms (Buyya et al. 2002; Buyya and Murshed 2002). However, these models or algorithms haven't taken this situation into consideration, which is manufacturing grid as a virtual organization; its members will be driven by self-interest, and consequently a conflict which betweens individual rationality and collective rationality during the pursuit of profit maximization will happen. So the cooperative game theory (Zhang et al. 2009) is an effective method to solve this problem as it seeks to satisfy both the individual rationality and the collective rationality. Since now this theory has been applied to many fields: the Krus and Bronisz (2000) discussed a procedure supporting multi-criteria analysis of the cost allocation problem via using the concept it (Krus and Bronisz 2000; Krus 2004, 2009), the Jiang and Wang (2008) analyzed the cooperative game situation of three-stage supply chain and come to an conclusion that the profit of each participant is increased significantly and the proportion of income distribution is more reasonable (Jiang and Wang 2008). In summary, although there are many researches on manufacturing grid resource allocation home and abroad, still little literature on solving this problem via cooperative game theory.

In this paper, we consider taking the proportion of resources provided by suppliers as bound variables to guide decision-making of the alliance members. Firstly, we build a cooperative game model of this dynamic alliance on the basis of the problem description and hypothesis. Then we solve this model with pontryagin's maximum principle and draw the relationship between the parameters and decision variables.

5.2 Construction of Competitive Model of Enterprises Dynamic Alliance Under Grid Environment

5.2.1 Description and Hypothesis of Problem

Suppose a dynamic alliance under gird environment is composed of N resource nodes, which includes n resource demanders and m resource suppliers. Besides internal cooperation, there is no external cooperation for all the alliance members. The parameters of cooperative game model are as follows:

x_i – represent enterprise i's contribution level (such as total assets, production capacity), $x_i > 0\,(i = 1, 2, \ldots, n)$,
r_i – represent the attenuation coefficient of enterprise i's contribution level (such as obsolescence of the new technology, physical loss of the materials),
R_j – represent the amount of resources that supplier j can provide, $R_j \geq 0\,(j = n + 1, n + 2, \ldots, N)$,
α_{ij} – represent the proportion of resource that a supplier willing to provide to a demander, which forms a m*n order matrix, $0 \leq \alpha_{ij} \leq 1\ (i = 1, 2, \ldots n;\ j = n + 1, n + 2, \ldots, N)$,
s_i – represent the growth factor of enterprise i's contribution level which is caused by resource transactions $(i = 1, 2, \cdots, N)$,
c_i – represent the cost coefficient of supplier i $(i = 1, 2, \ldots, n)$,
p_j – represent the cost coefficient of demander j $(j = n + 1,\ n + 2, \ldots, N)$,
ρ – represent the discount rate, $\rho > 0$,
θ_i – represent enterprise i's self-propagation ability of contribution level (such as capital investment, own resources investment),

Taking the characteristics of resource allocation and resource trading process into consideration, this paper makes the following hypotheses:

Hypothesis 5.1 the amount of resources can be provided by a supplier is the number of remaining resources after deducting the amount used by itself and the contribution level's increasement caused by own resource's consumption is included in θ_i,

Hypothesis 5.2 in resource trading process, there's a square relation between the total cost and the amount of trading resources for both suppliers and demanders, which can be set as $C_D = \frac{1}{2}\sum\limits_{i=1}^{n}\sum\limits_{j=n+1}^{N} c_i(\alpha_{ij}R_j)^2$ and $C_s = \frac{1}{2}\sum\limits_{i=1}^{n}\sum\limits_{j=n+1}^{N} p_j\left(\alpha_{ij}R_j\right)^2$,

Hypothesis 5.3 the resource trading brings increasement to both suppliers' and demanders' contribution level, and the increase rate is in direct ratio to the amount of treading resources, while the proportional coefficient is s_i $(i = 1, 2, \cdots, N)$,

Hypothesis 5.4 the profit of a member not only in proportion to its own interest, but also in proportion to the total interest of the alliance. For simplicity, the total revenue function after dimensional processing can be set as $W_i = x_i \sum_{i=1}^{N} x_i$.

5.2.2 Establishment of Model

We set the rate of firm's interest change is determined by the increase and attenuation of its own interest and the contribution level of other members to it. So the state equation of this alliance is:

$$x_i = -r_i x_i + \sum_{j=m+1}^{N} s_i \alpha_{ij} R_j + \theta_i,\ i = 1, 2, \cdots n$$

$$x_j = -r_j x_j + \sum_{i=1}^{m} s_j \alpha_{ij} R_j + \theta_j,\ j = n+1, \cdots N$$

$$x_i(0) = x_{i0},\ i = 1, 2, \cdots N \tag{5.1}$$

The total revenue function of alliance is:

$$\sum_{i=1}^{N} W_i = \sum_{i=1}^{N} x_i \sum_{i=1}^{N} x_i = \left(\sum_{i=1}^{N} x_i\right)^2,\ i = 1, 2, \cdots, N$$

So the alliance's total profit is:

$$\max \sum_{i=1}^{N} J_i = \int_0^T e^{-\rho s} \left(\sum_{i=1}^{N} W_i - C_D - C_S\right) d_s$$

That is:

$$max \sum_{i=1}^{N} J_i = \int_0^T e^{-\rho s} \begin{bmatrix} \left(\sum_{i=1}^{N} x_i\right)^2 - \frac{1}{2} \sum_{i=1}^{n} \sum_{j=n+1}^{N} c_i \left(\alpha_{ij} R_j\right)^2 - \\ \frac{1}{2} \sum_{i=1}^{n} \sum_{j=n+1}^{N} p_j \left(\alpha_{ij} R_j\right)^2 \end{bmatrix} d_s \tag{5.2}$$

$0 \leq \alpha_{ij} \leq 1;$

The alliance members maximize profit (5.2) by adjusting α_{ij} in (5.1), but every member also cares about the action of others that is because the contribution level of others can be affected by their strategy, which have an effect on the interest of this member. So the strategic interaction between members appears. (5.1) and (5.2) constitute a differential game between the N members and eventually come to an equilibrium status, which is Nash equilibrium.

5.3 Model Solution

This paper takes the proportion of resource that a supplier provides α_{ij} as control variable and system (5.1) as the basic state equation, and determines the admission control function in the control domain. Then it makes the solution of system (5.1) meet the condition and also maximize the performance indicator (5.2). That is to mean maximizing the target functional by seek the optimal contribution proportion. This is a linear optimal control problem. So we take the Pontryagin's maximum principle to solve this problem.

Construct the Hanmilton function:

$$\begin{aligned} H(\lambda_i, k_j, x_i, \alpha_{ij}) = e^{-\rho t} & \left[\left(\sum_{i=1}^{N} x_i \right)^2 - \frac{1}{2} \sum_{i=1}^{n} \sum_{j=n+1}^{N} \left(c_i + p_j\right) \left(\alpha_{ij} R_j\right)^2 \right] \\ & + \sum_{i=1}^{n} \lambda_i \left(-r_i x_i + \sum_{j=m+1}^{N} s_i \alpha_{ij} R_j + \theta_i \right) \\ & + \sum_{j=n+1}^{N} k_j \left(-r_j x_j + \sum_{i=1}^{m} s_j \alpha_{ij} R_j + \theta_j \right) \end{aligned} \tag{5.3}$$

Here λ_i, k_j are both covariant.

According to Pontryagin's maximum principle:

$$\frac{\partial H_i \left[t, \lambda^*(t), x_i^*(t), \alpha_{ij}^*(t) \right]}{\partial \alpha_{ij}} = 0$$

The solution is:

$$\alpha_{ij}^* = \frac{e^{\rho t} \left(\lambda_i s_i + k_j s_j \right)}{\left(c_i + p_j \right) R_j} \tag{5.4}$$

Set $u_i = e^{\rho t} \lambda_i, v_j = e^{\rho t} k_j$.

Then (5.4) can be reduced to:

$$\alpha_{ij} = \frac{u_i s_i + v_j s_j}{\left(c_i + p_j\right) R_j} \tag{5.5}$$

In the mean while, according to Pontryagin's maximum principle:

$$\lambda_i = -\frac{\partial H\left(\lambda_i, k_j, x_i, \alpha_{ij}\right)}{\partial x_i} = -\left[e^{-\rho t}\left(2\sum_{i=1}^{N} x_i\right) - \lambda_i x_i\right] \tag{5.6}$$

$$K_j = -\frac{\partial H\left(\lambda_i, k_j, x_i, \alpha_{ij}\right)}{\partial x_j} = -\left[e^{-\rho t}\left(2\sum_{i=1}^{N} x_i\right) - k_j x_j\right] \tag{5.7}$$

According to $u_i = e^{\rho t}\lambda_i$, $v_j = e^{\rho t}k_j$ we can get the first derivative of λ_i, k_j,

$$\lambda_i = \frac{u_i - \rho u_i}{e^{\rho t}} \quad k_j = \frac{v_j - \rho v_j}{e^{\rho t}} \tag{5.8}$$

Substitutes (5.8) into (5.6) and (5.7), then we can get

$$u_i = (\rho + r_i)\, u_i - 2\sum_{i=1}^{N} x_i, i = 1, 2, \ldots n$$

$$v_j = \left(\rho + r_j\right) v_j - 2\sum_{i=1}^{N} x_i, j = n+1, n+2, \ldots N \tag{5.9}$$

Based on the above calculations, (5.1) and $\dot{x_i}, \dot{u_i}, \dot{x_j}, \dot{v_j}$ all equals to 0 when they in a stable state, we can get the following differential equation:

$$0 = -r_i x_i + \sum_{j=n+1}^{N} \frac{s_i\left[u_i s_i + v_j s_j\right]}{c_i + p_j} + \theta_i, i = 1, 2, \ldots n \tag{5.10a}$$

$$0 = -r_i x_j + \sum_{i=1}^{n} \frac{s_j\left[u_i s_i + v_j s_j\right]}{c_i + p_j} + \theta_j, j = n+1, n+2, \ldots N \tag{5.10b}$$

$$(\rho + r_i)\, u_i = 2\sum_{i=1}^{N} x_i, i = 1, 2, \ldots n \tag{5.10c}$$

$$(\rho + r_i)\, v_j = 2\sum_{i=1}^{N} x_i, j = n+1, n+2, \ldots N \tag{5.10d}$$

By (5.10c) and (5.10d) we can get:

$$u_i = \frac{2P}{\rho + r_i};\ i = 1, 2, \ldots n;$$
$$v_j = \frac{2P}{\rho + r_j};\ j = n+1, n+2 \ldots N; \tag{5.11}$$

Substituting (5.11) into (5.5), and after finishing can get a Nash equilibrium solution of the system:

$$\alpha_{ij} = \frac{\frac{2P}{\rho+r_i} s_i + \frac{2P}{\rho+r_j} s_j}{\left(c_i + p_j\right) R_j},$$

$$\text{and}\quad P = \frac{\sum_{i=1}^{N} \frac{\theta_i}{r_i}}{1 - \left(\sum_{i=n}^{n} \frac{a_i}{r_i} + \sum_{j=n+1}^{N} \frac{f_j}{r_j}\right)};\ a_i = 2\sum_{j=n+1}^{N} \frac{\frac{s_i^2}{\rho + r_i} + \frac{s_i s_j}{\rho + r_j}}{c_i + p_j}$$

$$f_j = 2\sum_{i=1}^{n} \frac{\frac{s_i s_j}{\rho + r_i} + \frac{s_j^2}{\rho + r_j}}{c_i + p_j} \tag{5.12}$$

By (5.12), we learn that the proportion of resources provided by supplier is a decreasing function of the cost coefficient of suppliers p_j and the amount of resources can be provided by suppliers R_j. Because the more resources to be sold, the less interest the alliance can get, which reducing the competitiveness of the alliance. At the same time, the alliance's productive capacity is limited within a certain period of time, so the resource they need to maximum the profit is limited, which means the bigger the R_j is, the smaller the proportion is.

By (5.12), we learn that the proportion of resources provided by supplier is a decreasing function of the cost coefficient of demander i. Because that the higher the cost of demander is, the more reluctant suppliers to sell their resources and the lower the proportion of providing resources is. So they will get less interest in the transaction process and the total revenue for this alliance will be reduced accordingly. So the suppliers are willing to sell resources to demander which has a lower cost coefficient.

By (5.12), we learn that the proportion of resources provided by supplier is an increasing function of the suppliers and demanders' self-propagation ability θ_i. Because the higher the enterprise' self-propagation ability is, the more inward investment is, which means the enterprise can handle more resource and more resource is allocated to the enterprise so that the alliance can maximum the profit.

By (5.12), we learn that the proportion of resources provided by supplier is a decreasing function of the attenuation coefficient of enterprise i's contribution level. Because the higher the attenuation coefficient, the weaker the enterprise' competitiveness, which means the enterprise is hard to survive long and the suppliers are less willing to provide resource to it so that to increase the alliance' total profit.

By (5.12), we learn that the proportion of resources provided by supplier is an increasing function of s_i – the growth factor of enterprise i's contribution level which is caused by resource transactions. Because the bigger the growth factor is, the more the increasement of the enterprise's contribution level is, which also increases the alliance's profit.

5.4 Example

A dynamic alliance of the manufacturing gird is composed of five resources nodes, which includes two resources demanders and three resources suppliers. Set discount rate ρ as 8 % and the remaining parameters are shown in Table 5.1:

Through calculation based on the above parameters, we can get:

$$a_1 = 0.0217,\ a_2 = 0.0142,\ f_3 = 0.0134,\ f_4 = 0.0074,\ f_5 = 0.0137, \mathrm{P} = 39.9544$$

$$\alpha_{13} = \frac{\dfrac{2P}{\rho + r_1} s_1 + \dfrac{2P}{\rho + r_3} s_3}{(c_1 + p_3)\, R_3} = \frac{\dfrac{2 * 39.9544 * 0.05}{0.08 + 0.08} + \dfrac{2 * 39.9544 * 0.05}{0.08 + 0.08}}{(4 + 4.5) * 15} \approx 39.17\ \%$$

Similarly: $\alpha_{14} \approx 27.30\%, \alpha_{15} \approx 19.90\%, \alpha_{23} \approx 32.48\ \%,\ \alpha_{24} \approx 21.94\ \%,\ \alpha_{25} \approx 16.52\ \%$.

Figure 5.1 shows α_{ij} (α_{13} for an example in the figure) – the proportion of resources provided by supplier decrease with increasing cost coefficient of supplier p_j and amount of resources can be provided by supplier.

Figure 5.2 shows the proportion of resources provided by supplier α_{ij} decrease with increasing c_i cost coefficient of demander i.

Figure 5.3 shows the proportion of resources provided by supplier α_{ij} constant increase with increasing suppliers and demanders' self-propagation ability θ_i.

Table 5.1 The values of c_i, r_i, R_j, p_j, θ_i, s_i units: Piece, Yuan

Parameters	1	2	3	4	5
c_i	4	5			
r_i (%)	8	7	8	7	7.5
θ_i	0.05	0.04	0.05	0.04	0.03
R_j			15	20	30
s_i	0.05	0.04	0.05	0.03	0.05
p_j			4.5	3.5	4.5

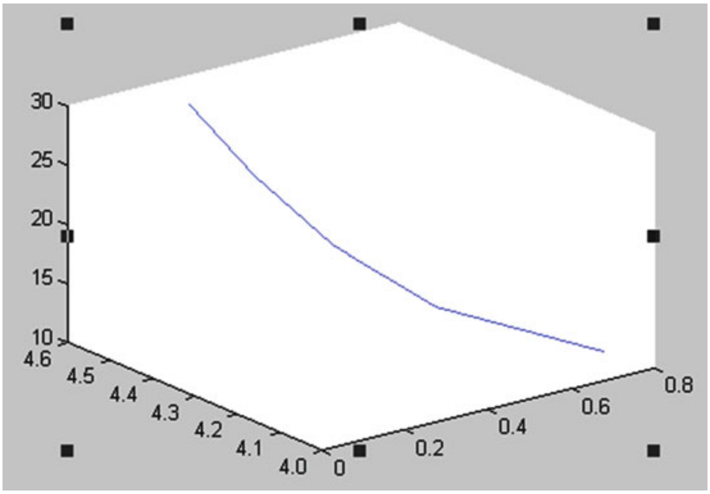


Fig. 5.1 The functional relationship between α_{ij} and R_j, p_j

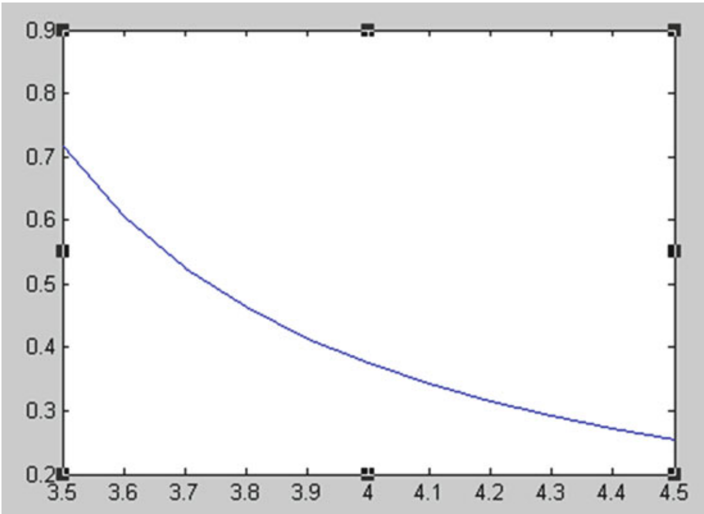


Fig. 5.2 The functional relationship between α_{ij} and c_i

Figures 5.4 and 5.5 shows the proportion of resources provided by supplier α_{ij} constant decrease with increasing attenuation coefficient of enterprise i's contribution level r_i.

Figures 5.6 and 5.7 shows the proportion of resources provided by supplier α_{ij} constant increase with increasing growth factor of enterprise i's contribution level which is caused by resource transactions s_i.

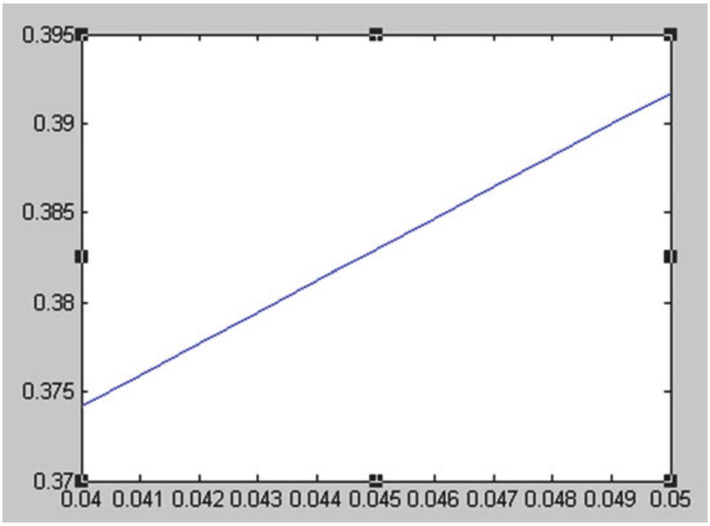


Fig. 5.3 The function relationship between α_{ij} and θ_i

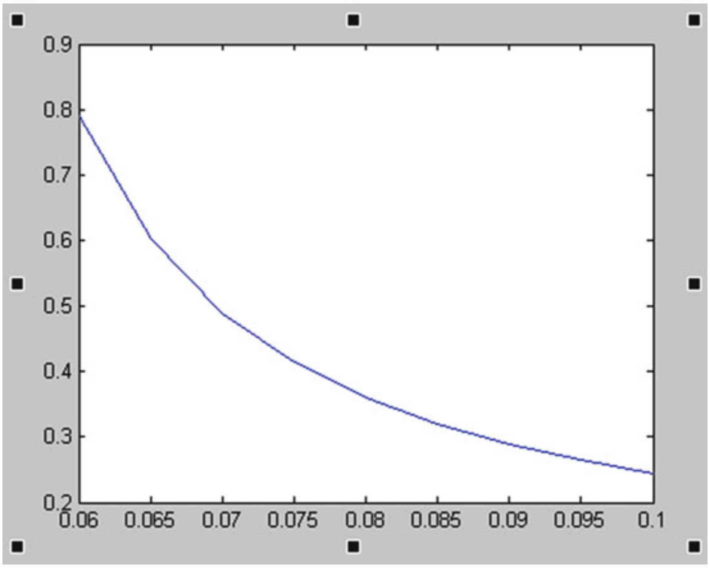


Fig. 5.4 The function relationship between α_{ij} and r_i $(i = 1, 2)$

5.5 Conclusion

To sum up, this paper comes to the following conclusions: (1) it establishes a cooperative game model about alliance members' contribution level. The members can achieve a win-win goal of maximization of individual and alliance benefit due

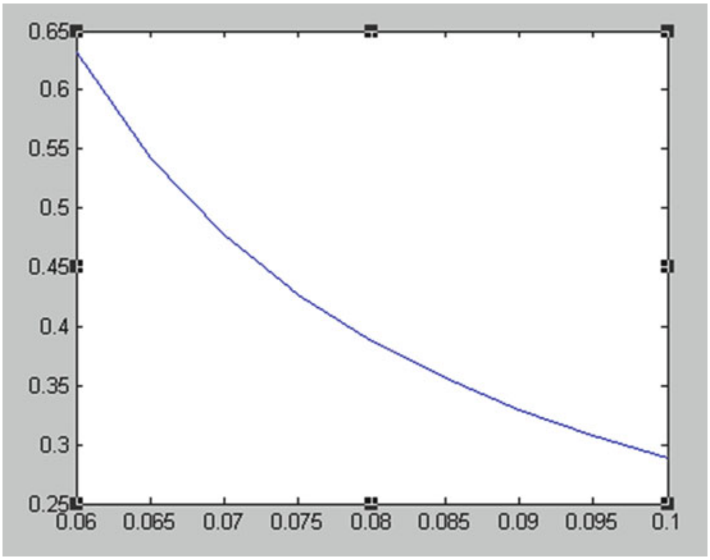


Fig. 5.5 The function relationship between α_{ij} and r_i $(i = 3, 4, 5)$

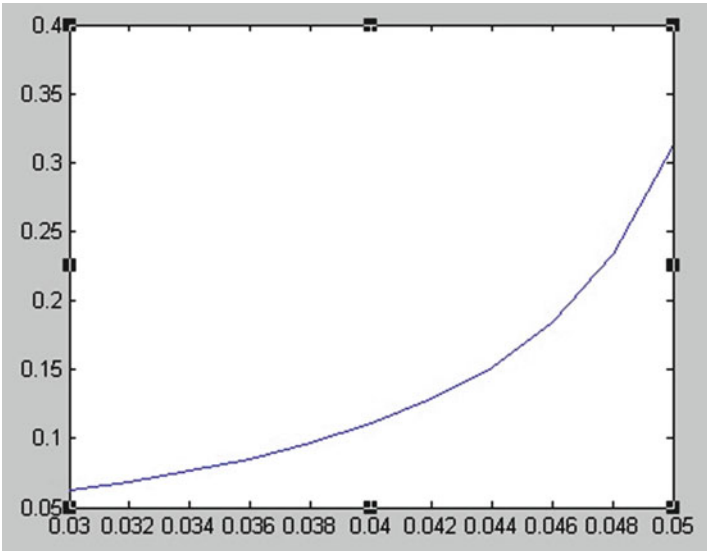


Fig. 5.6 The function relationship between α_{ij} and s_i $(i = 1, 2)$

to strategic interaction via adjusting its own resources proportion, (2) according to the model solution, on one hand the proportion of resources provided by supplier is positive correlated with enterprise i's self-propagation ability of contribution level and the growth factor of enterprise i's contribution level which is caused by resource

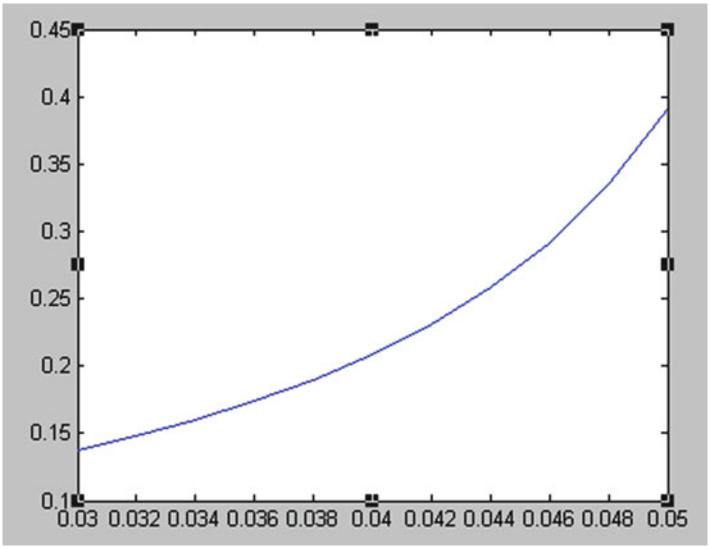


Fig. 5.7 The function relationship between α_{ij} and s_i $(i = 3, 4, 5)$

transactions, on the other hand it's negative correlated with the amount of resources can be provided by suppliers, the cost coefficient of suppliers and the cost coefficient of demanders and the attenuation coefficient of enterprise i's contribution level. And then it have an effect on the total profit of the alliance, (3) the group competitiveness of dynamic alliance can be improved via strengthening management, improving the matching of internal resources, reducing the operation cost and raising one's own competition which can make the alliance in an invincible position in the increasingly fierce market competition.

References

Buyya R, Murshed M (2002) GridSim: a toolkit for the modelling and simulation of distributed resource management and scheduling for grid computing [J]. Concurr Comput Pract Exp 14:1175–1220

Buyya R, Abramson D, Giddy J, Stockinger H (2002) Economic models for resource management and scheduling in grid computing [J]. Concurr Comput Pract Exp 14:1507–1542

Dean T, David W (1993) Cooperative and competitive networking by entrepreneurs: a critical incident study [J]. J Small Bu Manage 31(1):11–22

He Y, Yu T, Liu L, Sun H (2006) Research on manufacturing resource discovery and scheduling in manufacturing grid. Knowledge enterprise strategies in product design, manufacturing and management [C], 2006, pp 902–907

Ian F, Carl K (2004) The grid: blueprint for a new computing infrastructure [M]. Morgan and Kaufmann, San Francisco

Ian F, Carl K, Steven T (2001) The anatomy of the grid: enabling scalable virtual organizations [J]. Int J High Perform Comput Appl 15(3):200–222

Izakian H, Ladani BT, Zamanifar K, Abraham A, Snasel V Ramsar branch, Islamic Azad University, Ramsar (2009) A continuous double auction method for resource allocation in computational grids. Computational intelligence in scheduling, 2009. In: IEEE symposium on [C], Nashville, 2009, pp 29–35

Izakian H, Abraham A, Ladani BT (2010) An auction method for resource allocation in computational grids. Future Generat Comput Syst 26:228–235

Jiang PF, Wang Z (2008) Analysis of cooperative games on a three-stage supply chain [J]. J Shandong Univ (Nat Sci) 43(1):103–106

Krus L (2004) A computer based system supporting analysis of cooperative strategies. In: ICAISC international conference on artificial intelligence and soft computing [C], vol 3070, Springer, Berlin/Heidelberg, pp 891–896

Krus L (2009) On a partition function from game motivated by a cost allocation. In: CSM'2009 22nd workshop on methodologies and tools for complex system modeling [C], Springer, Berlin/Heidelberg, 2009, pp 28–30

Krus L, Bronisz P (2000) Cooperative game solution concepts to a cost allocation problem [J]. Eur J Oper Res 122(2):258–271

Liu LL, Yu T, Shi ZB, Fang ML (2003) Self-organization manufacturing grid and its task scheduling algorithm [J]. Comput Integr Manuf Syst 9(6):449–455

Liu L-L, Yu T, Cao H-W, Shi Z-B (2004) A TQCS-based scheduling approach for manufacturing grid [J]. J Donghua Univ 21(6):43–48

Liu SJ, Wu L, Meng XX (2009) Manufacturing grid [M]. Publishing House of Electronics Industry, Beijing, pp 39–40

Piao J, Zhu YL (2005) Research on manufacturing resources scheduling of virtual enterprise based on manufacturing grid [J]. Appl Res Comput 11(7):1019–1024

Shi ZB, Yu T, Liu LL (2004) Manufacturing grid and its resource configuration algorithm [J]. Comput Eng 30(5):117–119

Zhang PZ, Ye HX, Xue YW, Liu RJ, Hou GS (2009) Non-complete common interest group cooperation & game theory and application [M]. Shanghai Jiao Tong University Press, Beijing, pp 22–28

Zhou YL, Chen YW (2007) Research of allocation and management of manufacturing grid resource based on beneficial drive [J]. Mod Manuf Eng 1(11):20–23

Chapter 6
Optimization of Injection Molding Shop Scheduling Based on the Two-Stage Genetic Algorithm

Jin-ping Zhou and Hu Fu

Abstract Injection molding shop Scheduling is a large-scale parallel machine scheduling with process constraints, time constraints, earliness/tardiness penalties and due window constraints. Given the complexity of injection molding shop scheduling, a two-stage genetic algorithm is presented: the first stage is to partition jobs to machines, and the second stage is to sequence jobs for each machine. A simulation model for solving injection molding shop scheduling problem is proposed. For determining the optimal starting time of a single machine, a rule-based heuristic algorithm is also proposed. The application demonstrates the reliability and validity of the algorithm and simulation model.

Keywords Due window • Earliness/tardiness penalties • Injection molding • Parallel machine scheduling • Procedure constraint

6.1 Introduction

The general processes of the plastic injection company which makes to order are: injection, painting, printing, assembly. According to the differences of orders and product requirements, we can choose one of the general processes or all. But the injection process must be the first one. The different processing abilities of machines between the injection and follow-up crafts necessities the temporary storage for products after plastic injection. For direct molding products, the time of temporary storage determines the inventory cost and other related costs. The indirect molding products which have some rest processes need temporary packaging (dust requirements) and storage, transportation. If the injection process is delayed,

J. Zhou (✉) • H. Fu
School of Mechatronics Engineering, Guangdong University of Technology, Guangzhou, China
e-mail: iezhou@gdut.edu.cn

E. Qi et al. (eds.), *The 19th International Conference on Industrial Engineering and Engineering Management*, DOI 10.1007/978-3-642-37270-4_6,

the follow-up processing plan is often disrupted, and even causes the tardiness. Therefore the scheduling of injection molding shop is the most important aspects of reducing the overall production costs.

Injection molding shop scheduling (IMSS) has two critical parameters in consideration: the injection pressure and the product's color. The injection pressure determines the type of injection molding machine (process constraints), the product's color determines the setup time (processing time constraints). The aim of IMSS is how to partition the existing jobs (orders) to the injection molding machine reasonably and find out the best process sequence of these jobs to ensure follow-up process work normally and minimize the earliness/tardiness (E/T) penalties.

6.2 Problem Description

IMSS can be stated as follows:

1. There are n independent jobs (defined by set J) working on m independent machines (defined by set M);
2. Due to the injection pressure, J_{i} ($I = 1, 2, \ldots, n$) are only allowed to work on some machines like M_{j} ($j = 1, 2, \ldots, m$), called special process constraint;
3. There are some identical parallel machines in set M; others are non identical parallel machines. So the processing time on the different machines is not the same. But when the jobs have partition to machine, the processing time of the jobs is identified, which is called processing time constraint;
4. Although the fluctuation of the processing time is existed, we still assume the processing time is constant because the jobs are finished as a batch.
5. The setup time is associated with jobs and colors of products. Because the jobs are finished as a batch and the total time of finishing a batch is long, setup time can be merged into processing time.
6. Delivery is determined by customers and earliness/tardiness penalty exist. The earliness penalty is determined by the injection molding company itself, and it's related with holding costs for finished goods, deterioration of perishable goods and opportunity cost. Unlike the earliness penalty, the tardiness penalty is relatively complex and it's related to the breach of contract and loss of future sales and rush shipping cost and customer credit losses.
7. An interval $[e_{\mathrm{i}}, l_{\mathrm{i}}]$ is defined as the due window. The delivery ahead of e_{i} will cause the earliness penalty. And the delivery behind of l_{i} will get the tardiness penalty. It is assumed the interval is shorter than processing time of an job, and if p_{i} represents the processing time of the job i. so $p_{\mathrm{i}} > l_{\mathrm{i}} - e_{\mathrm{i}}$.
8. When the set of jobs is processing on the machine, continuous processing or intermittent processing is depend on the earliness/tardiness and the needs of Just-In-Time (JIT). That means the idle time of machine is allowed in processing.
9. The start time of a job processing is generally nonzero due to the effect by the existing tasks on the machine. So the machine's start time is a constraint of IMSS.

Therefore the IMSS can be considered as a parallel machine scheduling (PMS) problem with a special process constraint, processing time constraint, start time constraint and the due window constraint with minimum earliness/tardiness penalties. Because the numbers of the jobs and machines are huge, it also can be seen as a large-scale scheduling problem.

6.3 Solutions of PMS Problems

There is less specific literature for IMSS problem. A production schedule algorithm was presented based on heuristics and tabu search algorithms, which can not only in priority satisfy the job's delivery, but also can make setup time minimum in the injection production line (Zhong-yao Liu et al. 2008). A hybrid genetic algorithm (GA) is investigated to solve PMS problem (Cheng and Gen 1997), and presented an encoding method which consists of a job symbol list and a partitioning symbol list (it is called extended sequence encoding method). A heuristic algorithm is presented to solve non identical multi-machine scheduling problems with a common due window and earliness/tardiness (De-cai Huang et al. 2001), its goal is to minimize system makespan. It shows a hybrid genetic algorithm based on three genetic coding methods, which solve the earliness/tardiness scheduling problem of finding optimal common due date and optimal scheduling on parallel machines (Min Liu and Cheng Wu 2002).

Parallel machine earliness/tardiness scheduling problem can be divided into two categories: one category is common due window scheduling problem, and the other is independent due window scheduling problem. The study of the first problems has proved that there is no idle time inserted and the optimal schedule is V-shaped structure, as well as at least one workpiece completed within its due window properties (Xing-chu Liu et al. 2000).

On study of the second category, the most majority of the research is assumptions: the start processing time is zero, and there is no idle time between the work piece processing on the machine (Hua Zhao et al. 2007; Li Wang et al. 2002). To solve the different deliveries on non-identical PMS problem, the segment encoding of the genetic algorithm is presented in Jia-quan Gao et al. (2007).

Although the major study shows that the parallel machine earliness/tardiness scheduling problem is fit to JIT production (Min Liu et al. 2001; Feldmann and Biskup 2003; Sourd 2009; Bean 1994; Balin 2011; Si-min Huang et al. 2010; Kocamaz et al. 2009), but the assumption of the start of processing time is zero and there's no idle time inserted between two processes on one machine. This is not fit to JIT, because the significant feature of the E/T scheduling problems is to allow the idle time inserted between two processes on one machine (Xing-chu Liu et al. 2000). Therefore, to solve the PMS problem with earliness/tardiness, it is necessary to partition jobs to machines and sequence jobs for each machine, but also need to figure out the best start time. Reference Xing-chu Liu et al. (2000) shows a genetic algorithm for optimal process sequence and best start time, and reference Feldmann and Biskup (2003) shows a heuristic algorithm for minimizing earliness/tardiness penalties. The branch and bound algorithms to solve the PMS problem is presented (Sourd 2009), and the algorithm can be able to expand to multi-machine situations, but there are not expansion methods and case studies.

So, if the existing PMS model and algorithm is directly applied to the IMMS problem, we must simplify the constraints which cannot be simplified in reality. Especially, the surmise of E/T scheduling is contrary to JIT idea.

This paper proposes the IMMS problem as a PMS with special process constraint, constraint of processing time and due window constraint. The aim is to minimize earliness/tardiness penalties. The corresponding mathematical model and genetic algorithm based on two -stage chromosomes encoding method read as follows: the first stage is to partition jobs to machines, and the second stage is to sequence jobs for each machine, and at the same time to calculate the optimal start time.

6.4 Mathematical Model

Nonnegative constants α_{i} and β_{i} is defined as unit earliness/tardiness penalties cost of job i. The due window is $[e_{\mathrm{i}}, l_{\mathrm{i}}]$. For the random machine M_{j}, the set J_{j} shows the jobs which partition to the machine and satisfies the constraints, obviously $J_j \subseteq J$. $J_{\mathrm{jk}}\left(k = 1, 2, \ldots K_{\mathrm{j}}, j = 1, 2, \ldots, m, n = \sum_{j=1}^{m} K_j\right)$ can shows the jobs in J_{j} set. p_{jk} is defined as processing time and s_{jk} as start time ($s_{\mathrm{j}1} = s_{\mathrm{j}}$, s_{j} is the start time of M_{j}). The due window of job i will translate to $[e_{jk}, l_{jk}]$. The unit earliness/tardiness penalty cost is α_{jk} and β_{jk}. So the finish time of J_{jk} is:

$$C_{jk} = \sum_{q=1}^{k} (s_{jq} + p_{jq})\ (q = 1, 2, \ldots, k, k = 1, 2, \ldots K_j) \tag{6.1}$$

We assume s_j is the start time of machine M_j and there's no idle time between two jobs. That means $s_{jk} = 0$, so the complete time of order J_{jk} is:

$$C_{jk} = s_j + \sum_{q=1}^{k} p_{jq}(q = 1, 2, \ldots, k, k = 1, 2, \ldots K_j) \tag{6.2}$$

Set E_{jk} and T_{jk} as earliness/tardiness time:

If $C_{jk} < e_{jk}$, that means earliness happens, so earliness time is $E_{jk} = e_{jk} - C_{jk}$, $T_{jk} = 0$;
If $C_{jk} > l_{jk}$, that means tardiness happens, so tardiness time is $T_{jk} = C_{jk} - l_{jk}$, $E_{jk} = 0$;
If $C_{jk} >= e_{jk}$ and $C_{jk} <= l_{jk}$, $E_{jk} = 0$ and $T_{jk} = 0$.

For random machine M_j, total earliness/tardiness cost is:

$$f(M_j) = \sum_{k=1}^{K_j} (\alpha_{jk} \times E_{jk} + \beta_{jk} \times T_{jk}) \tag{6.3}$$

For a certain scheduling (σ), total earliness/tardiness cost is:

$$f(\sigma) = \sum_{j=1}^{m} \sum_{k=1}^{K_j} (\alpha_{jk} \times E_{jk} + \beta_{jk} \times T_{jk}) \tag{6.4}$$

$f(\sigma)$ is called as penalty function and is also the fitness function of the genetic algorithm employed in the paper.

6.5 Algorithm

In summary, the IMSS has three fundamental problems:

1. How to partition the orders to the appropriate injection molding machines, which means how to form the orders set J_{j} in machine M_{j}.
2. How to sequence jobs for each machine. The sequence is determined by set J_{jk}.
3. How to get the optimal start time for each machine. The result is determined by set s_{jk}.

Two ways to find the sequence of jobs currently are first to assign tasks, and then to sequence the jobs, or first to sequence jobs, and then to assign tasks.

6.5.1 Encode Design

Cheng and Gen (1997) proposed an extended sequence encoding method to solve the n-job m-machine problem. For a n-job m-machine problem, a legal chromosome contains n job symbols and $m - 1$ partitioning symbols ("*"), resulting in a total size of $(n + m - 1)$. Take a simple case with nine jobs and three machines as an example. Suppose there is a schedule, and the chromosome can be represented as [9 1 * 3 5 4 2 * 8 6 7]. "9/1" means job 9 and job 1 partition to M_1 (machine No.1). "3/5/4/2" means job 3/5/4/2 are partition to M_2. "8/6/7" means the job 8/6/7 are partition to M_3 and the processing sequence is 8-6-7.

The benefit of applying the extended sequence encoding is the chromosome which not only indicates the job processing sequence, but also represents the assigning of the jobs to machines. This is a typical first-partition- and-then-sequence problem. However, the encoding method may lead to the change in sequence and partition through crossover or mutation. The change in the partition jobs to machine can also change the processing sequence in the machine, and it is not conducive to the neighborhood search of combinatorial optimization problems. At the same time, the encoding method is not very suitable for the scheduling problem with process constraint.

Bean (1994) proposed a random key representation encoding method, its code length of n-genes within the chromosome from 1 to $m-1$ the real number (not including the border), the coding method effectively reflect the partition jobs to the machine sequence jobs of each machine. The encoding method is that partition and sequence are completed at the same time, but it's also not suitable for scheduling problems with process constraint.

The genetic algorithm chromosome coding mentioned in the above method is all using single encoding. This method mainly solves the problem of partition jobs to machine. Obviously, the sequence may not be the best sequence. As in the above 9-job 3-machine example, the sequence order 3/5/4/2 may be not the best sequence.

Thus, in order to solve large-scale injection molding E/T scheduling problem, a two-stage (level) chromosome encoding method is proposed. First, the chromosome A encoding the machine number to complete partition jobs to machine, a set $\{J_{\mathrm{j}}\}$ is a job set of the machine M_{j} which some jobs have been partition. Then, a set $\{J_{\mathrm{j}}\}$ of jobs is regarded as chromosome B which sequence jobs for each machine. It is summarized as follows:

The job set is $J=\{1,2, \ldots n\}$, and the machine set is $M=\{1,2, \ldots m\}$. Chromosome A encoding is machine-based representation, such as $[M^{(1)}, M^{(2)}, \ldots, M^{(n)}]$. $M^{(i)}$ means the job i can be partitioned to the machine (process constraints). For the 15-job and 4-machine encoding example, chromosome gene no. is 1–15, corresponding chromosome A (machine no.) as (1 3 3 2 1 4 2 2 1 1 4 3 1 4 4).

This encoding can be applied to partition all jobs to the machine. Described in the above example, the job sets that are assigned to four machines are: $\{J_1\}=\{1,5, 9,10,13\}$, $\{J_2\}=\{4,7,8\}$, $\{J_3\}=\{2,3, 12\}$, $\{J_4\}=\{6,11, 14,15\}$. The genetic algorithm which partitions of all jobs to the machine is the first stage genetic algorithm.

The chromosome B encoding is job-based representation and the encoding length is variable. For example, the current machine is M_1, and the jobs set is $\{J_1\}$, a feasible chromosomes B encoding is (10 1 9 5 13).

Genetic algorithm based on job-based representation is the second stage genetic algorithm. When completing the first stage genetic algorithm, the second stage genetic algorithm is applied to sequence jobs for each machine.

6.5.2 Fitness Function

The two-stage genetic algorithm defines a job set of machine and a sequence of jobs in each machine. Finally, a heuristic method used to figure out the job's optimal start time. There is a basic principle of "priority of earliness to tardiness". The fitness calculation steps reads as follows:

1. To employ the first genetic algorithm to partition jobs to the machine by chromosome A. For machine set $M_{\mathrm{j},}$ the job set J_{jk} ($k=1, 2, \ldots K_{\mathrm{j}}, j=1, 2, \ldots,$ $m, n = \sum_{j=1}^{m} K_j$) constitutes the chromosome B.
2. To sequence the chromosome B through the second stage genetic algorithm.
3. For a scheduling plan of machine M_{j}, let s_j as the initial processing time. There is no idle time inserted between the two jobs processed on a machine. The finish time of job J_{jk} process on machine M_j is calculated by (6.2).
4. For job J_{jk} ($k=1,2, \ldots K_j$) on machine M_j, when on the condition of $k=K_j$, if $c_{jk} >= l_{jk}$, that means no machine idle time exists and $f(M_j)$ may include the cost of tardiness penalties, then go to 6).
5. If $c_{jk} < l_{jk}$, we can get $c_{jk} = l_{jk}$, $a_{jk} = c_{jk} - t_{jk}$. Because of the delay start time of J_{jk}, there appears a "neutral" between the J_{jk} and J_{jk-1}. That means c_{jk-1} may be latter. There are two circumstances:

   (a) If $c_{jk} - 1 >= l_{jk} - 1$, that means J_{jk-1} is already delayed (or just the latest delivery time). Assigning the c_{jk-1} just will increase penalties cost, so we retain the "neutral"
   (b) If $c_{jk-1} < l_{jk-1}$ and $l_{jk-1} < a_{jk}$, it results in $c_{jk-1} = l_{jk-1}$ and moving forward part of "neutral", then let $k=k-1$ and return 5). Otherwise let $c_{jk-1} = a_{jk}$ and all of the "neutral" are forward.

6. Calculating the jobs total earliness/tardiness penalties cost of M_j by (6.3).
7. Evaluating whether the result meet the end condition of the second stage genetic algorithm. If it does, let it go to 8), otherwise go back to 3).
8. Evaluating whether the two-stage genetic algorithm is completed. If it does, let it turn to 9), otherwise go back to 2).
9. Calculating the orders' total earliness/tardiness penalties cost as the total one scheduling penalties cost by (6.4). Then evaluating whether the result meet the end condition of the first stage genetic algorithm. If it does, we already get the minimum E/T indicators scheduling, otherwise go back to 1) to continue.

Therefore, the calculating of the optimal start time and the fitness is fit to JIT production. The essence of this algorithm is to nest sequence genetic algorithm into partition genetic algorithm. The algorithm is also applied to sequence jobs for each machine.

6.5.3 Other Genetic Operators

As the first stage and the second stage genetic algorithm both are general partition and sequence algorithm, the crossover operator of the chromosome A/B can select the partially mapped crossover (PMX). But the Mutation operation and selecting the operation use the standard genetic operators.

Table 6.1 Start time of injection molding machine

Machine no.	Start process time	Machine no.	Start process time
1	12:00:00	5	1:10:00
2	0	6	0
3	5:00:00	7	2:00:00
4	7:40:00	8	3:30:00

6.6 Case Study

Take a medium scale injection molding company as an example. There are over 100 injection molding machines with processing capacity from 100 to 1,300 t. When the company considers the production schedule, they generally separate jobs with direct molding without subsequent processing and jobs with follow-up process. And the injection molding machines are assigned to different operating groups: the direct molding without subsequent processing and follow-up painting process. To the group of follow-up painting process, the delivery time is the job schedule of the follow-up painting process. So this type of job scheduling is figure out how to satisfy the painting process.

In this case, there are 40 orders and 8 injection molding machines now and it is to decide the schedule of the next 3 days. The eight machines have different processing capacities and some of them are working for current jobs. So the 40 orders can only process on the part of machines. Table 6.1 shows the start time of these 40 orders. Table 6.2 is the constraints of the orders.

The corresponding genetic algorithm and simulation model can be developed by the application of the Tecnomatix series software which produced by Siemens company, shows as Fig. 6.1.

The population size of chromosome A is set as 30/60 (the first generation of the parent is 30). The maximum generation is 60. The crossover probability P_c is 0.5. The mutation probability P_m is 0.15. The population size of chromosome B is set as 10/20 (the first generation of the parent is 10). The maximum generation is 10 (which can automatically adjust according to the number of orders). The crossover probability P_c is 0.8. The mutation probability P_m is 0.1. The result of simulation is shown in Fig. 6.2 and the Gantt chart shown in Fig. 6.3.

6.7 Conclusion

This paper proposes a two-stage genetic algorithm based on the recent study of genetic algorithm for parallel machine to solve the large scale PMS problem with special process constraint, processing time constraint and the due window constraint: the first stage is partition jobs to machines, and the second stage is sequence

Table 6.2 Earliness/tardiness penalty factor, due window, process and processing time constraint

Job	α_i	β_i	e_i	I_i	Machine & proc time constraint					
J1	0.68	1.09	2:09:12:00	2:10:40:00	M1	7:06:00	M3	16:18:00	M6	20:06:00
J2	0.92	1.47	17:58:00	19:48:00	M4	20:24:00	M5	13:54:00	M8	5:06:00
J3	0.68	1.09	18:35:00	19:10:00	M4	15:42:00	M5	19:48:00	M7	5:54:00
J4	0.96	1.54	1:17:25:00	1:19:15:00	M5	20:06:00	M6	4:36:00	M7	12:36:00
J5	0.84	1.34	1:23:04:00	2:01:13:00	M1	16:30:00	M2	7:18:00	M5	11:36:00
J6	0.96	1.54	18:10:00	19:27:00	M4	17:48:00	M5	4:06:00	M6	15:54:00
J7	1.04	1.66	1:09:16:00	1:11:04:00	M1	17:24:00	M2	13:48:00	M8	19:12:00
J8	0.92	1.47	16:15:00	17:41:00	M2	6:18:00	M4	13:12:00	M8	22:42:00
J9	0.68	1.09	2:18:06:00	2:19:30:00	M2	19:06:00	M4	11:30:00	M7	11:36:00
J10	0.72	1.15	1:01:25:00	1:02:05:00	M3	21:18:00	M6	6:24:00	M8	23:48:00
J11	0.68	1.09	2:08:02:00	2:08:21:00	M2	22:54:00	M4	16:36:00	M6	22:30:00
J12	1.12	1.79	2:05:21:00	2:05:45:00	M1	20:12:00	M2	3:12:00	M8	13:24:00
J13	1.04	1.66	1:08:09:00	1:08:45:00	M1	20:30:00	M2	10:06:00	M7	19:42:00
J14	0.60	0.96	1:08:40:00	1:10:43:00	M5	18:18:00	M6	6:30:00	M8	19:12:00
J15	0.84	1.34	15:55:00	17:12:00	M1	21:54:00	M2	19:18:00	M7	9:48:00
J16	0.72	1.15	2:09:55:00	2:11:43:00	M2	12:00:00	M4	23:18:00	M5	5:48:00
J17	0.88	1.41	11:17:00	11:32:00	M1	9:36:00	M6	3:48:00	M8	10:18:00
J18	1.08	1.73	1:18:21:00	1:20:22:00	M5	11:54:00	M7	6:00:00	M8	3:18:00
J19	1.04	1.66	1:20:28:00	1:21:15:00	M1	14:18:00	M2	7:18:00	M4	10:42:00
J20	1.12	1.79	2:17:03:00	2:18:22:00	M3	9:18:00	M4	20:00:00	M8	9:30:00
J21	0.80	1.28	1:06:53:00	1:08:09:00	M4	5:18:00	M7	23:24:00	M8	6:54:00
J22	0.60	0.96	2:18:05:00	2:19:40:00	M5	3:06:00	M7	10:36:00	M8	7:54:00
J23	1.12	1.79	1:00:15:00	1:02:31:00	M1	5:42:00	M2	10:30:00	M7	14:06:00
J24	0.88	1.41	1:02:10:00	1:04:14:00	M3	15:06:00	M5	10:00:00	M7	20:30:00
J25	0.64	1.02	16:51:00	19:02:00	M1	7:00:00	M3	20:30:00	M6	3:00:00
J26	0.80	1.28	1:20:19:00	1:22:58:00	M4	13:48:00	M5	6:42:00	M6	7:00:00
J27	1.16	1.86	1:19:38:00	1:21:44:00	M3	12:30:00	M4	17:42:00	M8	14:18:00
J28	1.16	1.86	1:23:03:00	2:00:44:00	M2	7:36:00	M4	22:06:00	M7	3:00:00
J29	0.80	1.28	1:15:27:00	1:15:57:00	M2	14:00:00	M3	13:30:00	M7	14:00:00
J30	1.12	1.79	1:18:00:00	1:19:14:00	M3	19:06:00	M5	18:00:00	M6	13:36:00
J31	1.08	1.73	2:12:46:00	2:15:11:00	M1	11:24:00	M3	9:18:00	M7	14:42:00
J32	1.00	1.60	11:26:00	11:27:00	M1	11:30:00	M3	14:12:00	M6	11:18:00
J33	1.12	1.79	1:01:00:00	1:01:21:00	M1	17:06:00	M3	19:12:00	M8	20:00:00
J34	1.04	1.66	2:13:07:00	2:14:45:00	M1	10:00:00	M5	7:48:00	M7	16:00:00
J35	0.60	0.96	2:06:44:00	2:09:07:00	M2	5:18:00	M3	9:42:00	M6	16:48:00
J36	0.96	1.54	1:15:56:00	1:16:33:00	M2	9:42:00	M4	19:30:00	M6	23:30:00
J37	0.84	1.34	1:06:34:00	1:08:31:00	M1	6:06:00	M3	12:12:00	M5	7:30:00
J38	0.76	1.22	2:11:32:00	2:14:25:00	M1	7:24:00	M3	9:06:00	M4	5:48:00
J39	0.88	1.41	16:46:00	18:22:00	M3	7:00:00	M4	7:48:00	M8	9:12:00
J40	0.80	1.28	2:11:25:00	2:12:39:00	M3	16:48:00	M7	7:06:00	M8	12:42:00

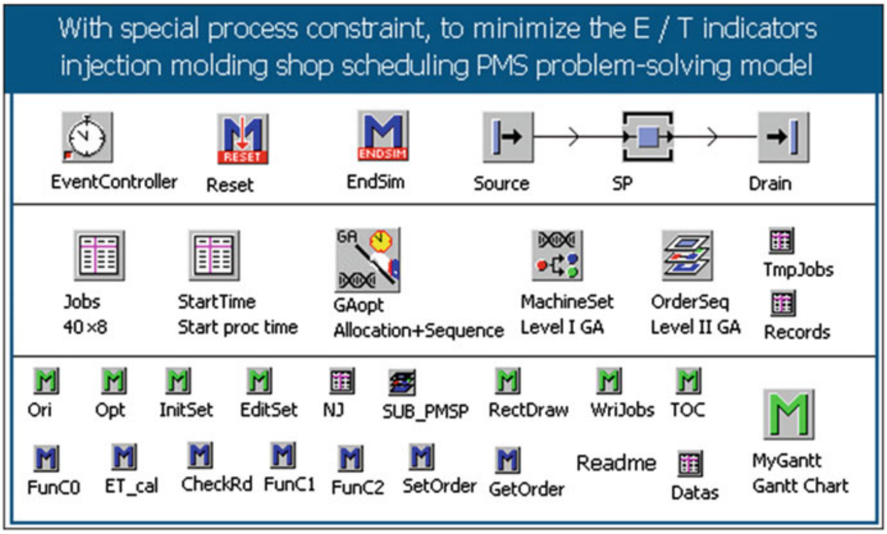
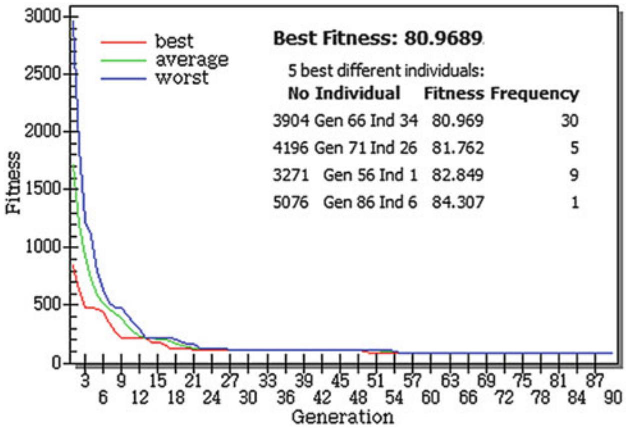


Fig. 6.1 Simulation model for IMSS



M	*Jobs Set*						
1	23	37	19	38	31		
2	29	8	13	36	28	12	35
3	33	30	20				
4	39	21	11	9			
5	2	6	24	5	26	16	34
6	32	17	25	10	14	4	1
7	15	3	18	40			
8	7	27	22				

Fig. 6.2 Solution results of simulation example

jobs for each machine. And this method is totally fit to the JIT production which requires minimum earliness/tardiness penalties. Practice shows that the application of this method can get more practical operability result than the application of the *first-partition-then-sequence* method or the *sequence-plus-partition* method. The follow-up study will focus on how to achieve the collaborative work plan and the optimization of production planning of entire processes: injection—painting—printing—packaging.

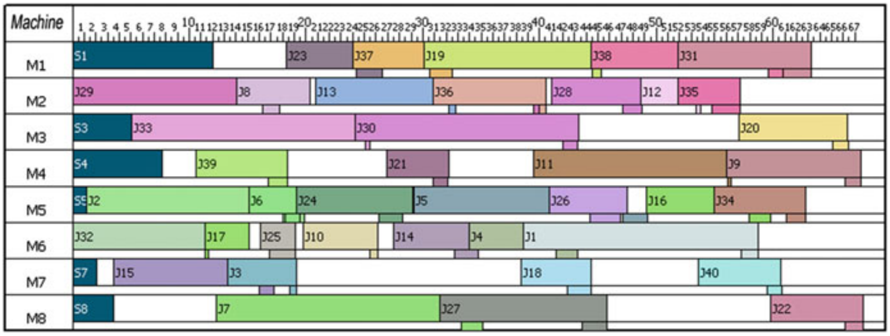


Fig. 6.3 Gantt chart of scheduling example

References

Balin S (2011) Non-identical parallel machine scheduling using genetic algorithm. Expert Syst Appl 38:6814–6821

Bean JC (1994) Genetic algorithms and random keys for sequencing and optimization. INFORMS J Comput 6(2):154–160

Cheng RW, Gen M (1997) Parallel machine scheduling problems using memetic algorithms. Comput Ind Eng 33:761–764

De-cai Huang, Yi-hua Zhu, Wan-liang Wang (2001) Earliness-tardiness job scheduling problem with a common due window on non-uniform machines. Syst Eng Theory Pract 21(4):64–69 (in Chinese)

Feldmann M, Biskup D (2003) Single-machine scheduling for minimizing earliness and tardiness penalties by meta-heuristic approaches. Comput Ind Eng 44:307–323

Hua Zhao, Ren-jun Wen, Xiao-hui Chen (2007) A study on the parallel machine scheduling problem based on individuality production mode. China Mech Eng 18(17):2058–2061 (in Chinese)

Jia-quan Gao, Yu-shun Wang, Gui-xia He (2007) Earliness-tardiness job scheduling problem with a common due window on non-uniform machines. Comput Eng Appl 42(2):15–17 (in Chinese)

Kocamaz M, Cicekli UG, Soyuer H (2009) A developed encoding method for parallel machine scheduling with permutation genetic algorithm. In: European and Mediterranean conference on information systems, EMCIS’2009, Izmir, 2009, pp 1–8

Li Wang, Da-wei Li, Meng-guang Wang (2002) Genetic algorithm for parallel machine scheduling with due windows. J Syst Eng 17(1):45–49 (in Chinese)

Min Liu, Cheng Wu (2002) Parallel genetic algorithm for solving parallel machine scheduling problem with different due window. Acta Automat Sin 26(2):258–262 (in Chinese)

Min Liu, Cheng Wu, Wen-jun Yin (2001) Solving identical parallel machine production line scheduling problem with special procedure constraint by genetic algorithm. Acta Automat Sin 27(3):381–386 (in Chinese)

Si-min Huang, Lin-ning Cai, Xiao-yue Zhang (2010) Parallel dedicated machine scheduling problem with sequence-dependent setups and a single server. Comput Ind Eng 58:165–174

Sourd F (2009) New exact algorithms for one-machine earliness-tardiness scheduling. INFORMS J Comput 21(1):167–175

Xing-chu Liu, Qian-chuan Zhao, Da-zhong Zheng (2000) Single machine earliness and tardiness scheduling problem with distinct due window using genetic algorithms. J Tsinghua Univ (Sci Technol) 40(7):59–62 (in Chinese)

Zhong-yao Liu, Chong-jia Peng, Nai-qi Wu (2008) Production schedule based on tabu search algorithms. Contr Automat 24(6):56–58 (in Chinese)

Chapter 7
A Simulation-Based Analysis of Cluster Tools Scheduling with Plant Simulation

Yan Min and Xiao-rui Lin

Abstract Cluster tools are configurable and integrated equipment. They are widely used in semiconductor fabrication in recent years. It is difficult to effectively operate a cluster tools and schedule it due to wafer residency time constraints and activity time variation of cluster tools. The schedule effects on the time performance of cluster tools operation. In this paper, a simulation model based on Plant simulation is presented to discuss the schedule of cluster tools in steady state. This simulation model could describe and simulate the activities of robot, wafer processing and their corresponding timing of cluster tools operation. Hence, it is a useful method to analyze the time performance of cluster tools such as scheduling. Two examples are used to test the simulation model and show the proposed application of the simulation model.

Keywords Cluster tools • Modeling • Plant simulation • Production scheduling • Semiconductor fabrication • Simulation

7.1 Introduction

Cluster tools are widely used in semiconductor fabrication in recent years. With single wafer processing technology, cluster tools make semiconductor manufacturing flexible, reconfigurable and efficient (Bader et al. 1990; Burggraaf 1995), result in higher yield (Newboe 1990), shorter cycle time (Newboe 1990; Singer 1995; McNab 1990), better utilization of space (Burggraaf 1995; Singer 1995), and

Y. Min (✉) • X. Lin
School of Electro-Mechanical Engineering, Guangdong University of Technology, Guangzhou, China
e-mail: miny@gdut.edu.cns

E. Qi et al. (eds.), *The 19th International Conference on Industrial Engineering and Engineering Management*, DOI 10.1007/978-3-642-37270-4_7,

lower capital cost (Singer 1995). Because of wafer residency time constraints and activities time variation of cluster tools, how to effectively operate a cluster tools is a challenge.

7.2 Architecture and Operations of Cluster Tools

7.2.1 Architecture of Cluster Tools

A cluster tools (Fig. 7.1) contains the following major components: an aligner module, several processing modules (PMs), a wafer handling robot (single-arm or dual-arm) that loads wafer into or unloads wafer from LL/PM and moves wafer between PMs, and one or two loadlocks (LL) for wafer cassette loading (Wu et al. 2008).

7.2.2 Operations of Cluster Tools

A raw wafer is unloaded from LL by robot initially and then loaded into a PM. After finishing process, it is unloaded from that PM and moved to other PMs in specific sequence as needed. It is returned to LL as the completed wafer at finial stage eventually. In general, 25 wafers are in a cassette of LL.

The particular processing route for different type wafers is denoted by the wafer processing flow. The form of wafer processing flow without wafer revisiting is (m_1, m_2, …, m_n) (Kim et al. 2003), where n is the number of processing steps, m_i is the number of parallel PMs at process step i.

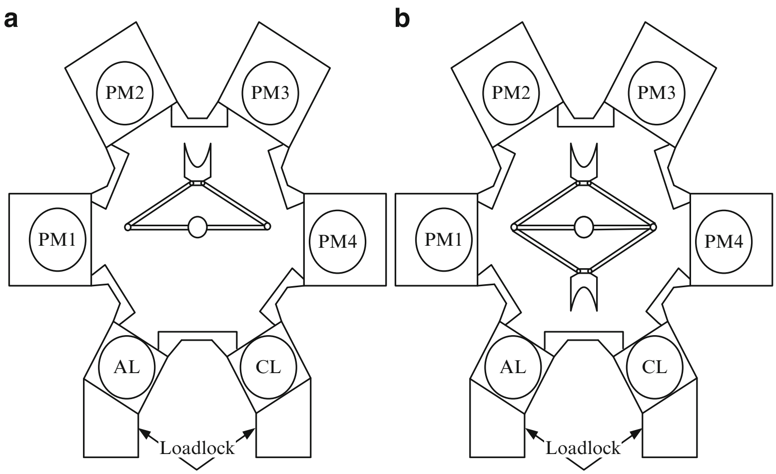


Fig. 7.1 Cluster tools: (**a**) single-arm robot; (**b**) dual-arm robot

7.2.3 Operations Requirement

Generally, wafers in a cassette have the same processing route (Kim et al. 2003; Lee and Park 2005). Thus, a cluster tools just operates in steady state at most of the time due to its continuous operation with two loadlocks.

A type of the various constraints for wafer fabrication is wafer residency time constraints (Lee 2008). These constraints require that a wafer just stays in a PM within a given time range after completion of process. Otherwise a wafer would be damaged. The wafer residency time constraints are very restrict for some process. They affect the schedule of cluster tools operation.

In order to maximize the throughput of a cluster tools, it needs scheduling a cluster tools to satisfy the wafer residency time constraints.

7.2.4 Cluster Tools Modeling

Time performance of cluster tools operation is the most concerned quantitative property. Important efforts have been made to analyze modeling and time performance of cluster tools. These efforts have indicated that the configuration of a cluster tools has effect on the performance (Lopez and Wood 1998, 2003; Koehler and Seppanen 1999; Perkinson et al. 1994).

Numerous researches have done investigation about modeling of cluster tools schedule. These analyses are based on timing diagram (Perkinson et al. 1996; Srinivasn 1998), marked graph (Zuberek 2001, 2004; Wu and Zhou 2010a, b, c, d), colored resource-oriented Petri net (CROPN) (Wu and Zhou 2001, 2010a; Wu 1999; Wu et al. 2011a, b) and simulation analysis (Koehler and Seppanen 1999; Artherton et al. 1990; Li and Wu 2009; Pan and Wu 2009). Being able to describe the dynamic behavior of cluster tools, CROPN and simulation method are powerful in modeling.

Another difficulty of scheduling and modeling cluster tools is to find the feasible and optimal solution effectively and precisely. Some efforts have done on it. The typical algorithms include heuristic, mathematical programming, analytical method and simulation method. Basing on the consideration of computationally efficient and solving accuracy, analytical and simulation method have more advantages (Kim et al. 2003; Lee and Park 2005; Zuberek 2001, 2004; Wu and Zhou 2001, 2010b, c, d; Wu 1999; Li and Wu 2009; Pan and Wu 2009).

Simulation analysis is an effective method for modeling of cluster tools operation because it can describe the dynamic behavior of cluster tools with wafer residency time constraints. It is also a simple and feasible way to find the solution of cluster tools scheduling.

7.3 Simulation Model for Cluster Tools Schedule with Plant Simulation

Plant Simulation is a system developed by Siemens PLM Software. It is a computer application for simulation modeling, optimization of production system and process, etc. It is an object and event oriented graphical software.

The simulation model discussed in this paper is based on Plant Simulation. The system of simulation model focuses on the activities of robot and the timing of wafer processing. It consists of loadlock (LL), processing modules (PMs), robot and raw wafers. The methods provided by the system of this simulation model have their corresponding SimTalk programs that simulate the activities of robot and wafer processing route. The Table files provided by system record the absolute time when a wafer is loaded into LL/PMs or unloaded from LL/PMs. The residency time of a wafer in a PM is calculated based on these records.

The correspondence between the components of a cluster tools and the objects of simulation model with Plant Simulation is shown in Table 7.1.

There are several system parameters of simulation model. They correspond to actual processing parameters of cluster tools. These parameters are defined as global variables of the Plant Simulation system. The definitions of these parameters are described in Table 7.2.

Cluster tools are reconfigurable. For simulation model with Plant Simulation, they are configured by choosing different objects and modifying the SimTalk programs of control methods. To do so, simulation model could simulate all possible activities of robot and wafer processing, describe the timing of robot activities and wafer processing of cluster tools with various processing flow. Thus, it can implement the simulation of any schedule of cluster tools. The simulating process of cluster tools with single-arm robot in steady state is shown in Fig. 7.2.

Table 7.1 Correspondence between components and objects

		Cluster tools
System part	Plant simulation object	Component/activity/parameter
Platform	Frame	Cluster tools
Product entering	Source	Loading wafer into system
Product	Entity	Wafer
Equipment	Single Proc	PMs/LL/Robot
Table	Table File	Technical parameter Processing time Residency time
Control	Method	Processing flow control Data statistics
Product leaving	Drain	Unloading wafer from system
Event controller	Event controller	Generate processing time

Table 7.2 Definition of variable object

Variable object	Definition
LoadingTime	Wafer loading time into LL/PM
UnloadingTime	Wafer unloading time from LL/PM
RotatingTime	Robot moving time between LL/PM
BatchSize	Wafer lot size
BatchNo	Batch amount of processed wafers
BatchCount	Sequence number of processing wafer in current batch
RobotOccupied	Status of robot: holding with/without wafer

7.4 Key Assumption of Simulation

The schedule of cluster tools discussed in this paper is in steady state. The simulation model is also running in steady state specifically to match it. However, before the system reaching steady state, there is an initial transient process, similarly, there is a final transient process when the system needs to stop. Same as it in the steady state, the key of scheduling a cluster tools in transient process is to determine when a wafer should be loaded into or unloaded from LL/PM. A special type of dummy wafer called W_0 is introduced into the system when the first wafer is loaded into the system, similarly, when the last wafer in LL is unloaded from it. To do that, the behavior of system in initial transient and finial transient are same as it in steady state.

7.5 Experimentation of Simulation for Cluster Tools

In order to test the simulation model discussed in this paper, two examples of cluster tools with single-arm robot are used as the application for this simulation model.

Here are some details of examples.

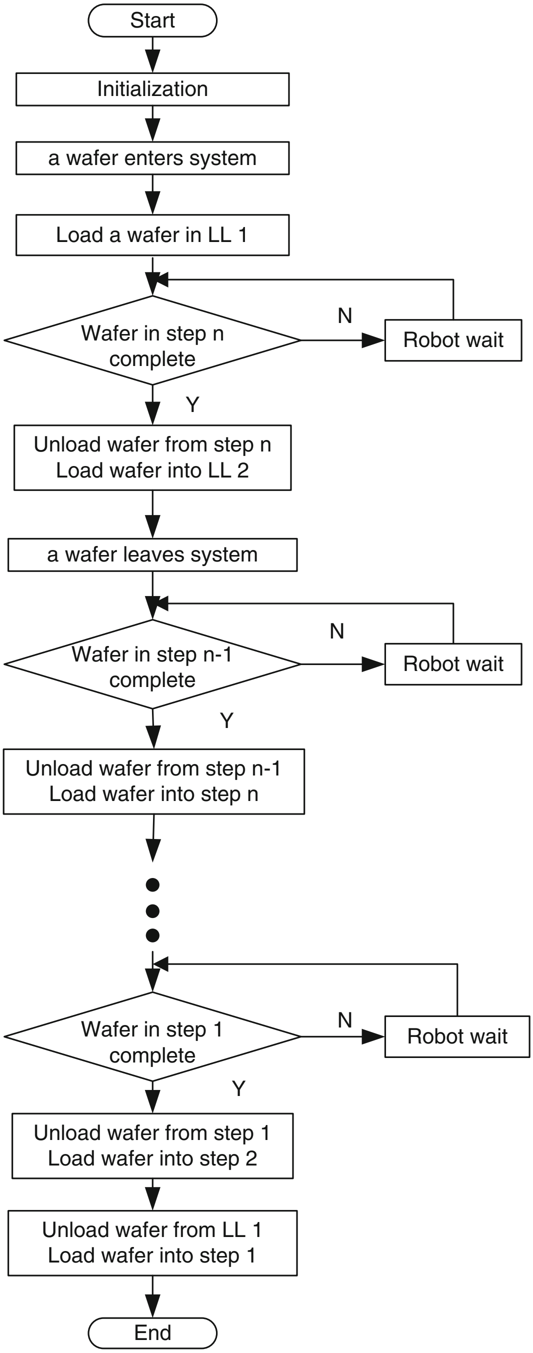
Example 7.1 The flow pattern is (1,1,1,1), the batch lot size is 25, the processing time for step 1, 2, 3 and 4 are 169, 139, 145, 129 s, after completion of processing, a wafer can stay in step 1, 2, 3 and 4 for 15, 20, 15 and 25 s, loading a wafer into LL/PM takes 10 s and unloading a wafer from LL/PM takes 15 s, respectively, it takes 2 s for robot moving a wafer from one LL/PM to another.

The simulation model is initialized by inputting processing parameters of a cluster tools. The interface of simulation model after completion of initialization is illustrated in Fig 7.3.

To start the simulating process just needs to double click icon Eventcontroller. Once the program is started, it can be stopped anytime later when it has been running at least one cycle. The result of simulating process is shown in Fig. 7.4.

The timing of wafer processing is recorded in Table file Statistic. The detail information is shown in Fig. 7.5.

Fig. 7.2 Simulation process in steady state



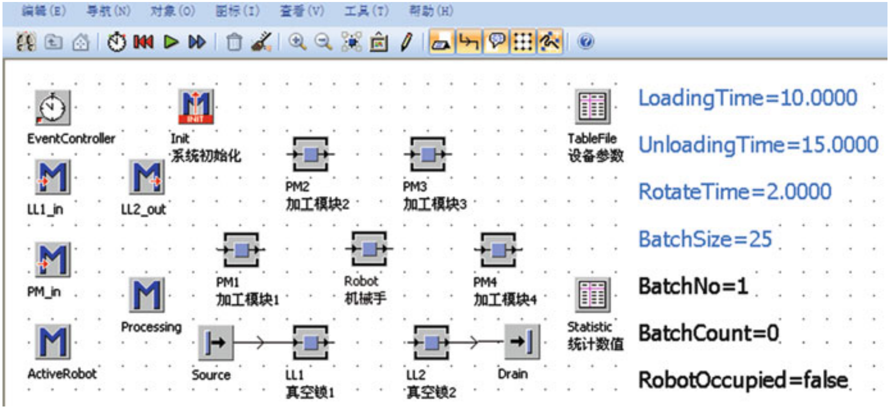


Fig. 7.3 Interface of simulation model

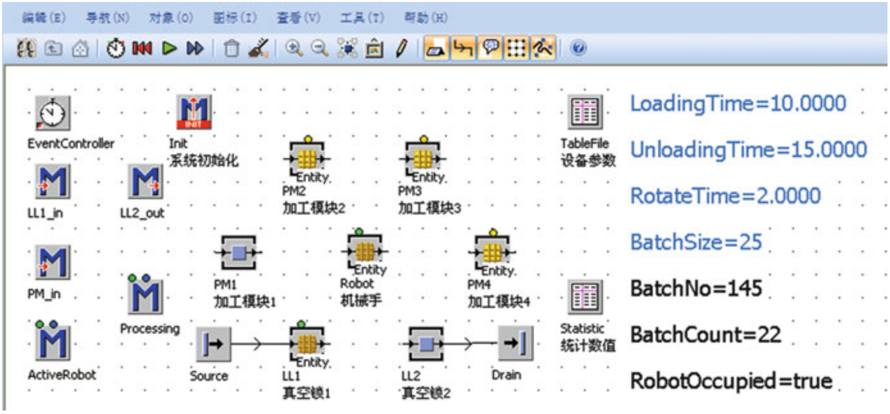


Fig. 7.4 Application result of simulation model

	object 0	integer 1	time 2	time 3	time 4	time 5	time 6	time 7	time 8	time 9	time 10	time 11	time 12	time 13
string	Object	BatchNo	PM1in	PM1out	PM1span	PM2in	PM2out	PM2span	PM3in	PM3out	PM3span	PM4in	PM4out	PM4span
1	(?)	1	29.0000	3:30.0000	3:01.0000	3:47.0000	6:18.0000	2:31.0000	6:35.0000	9:12.0000	2:37.0000	9:29.0000	12:02.0000	2:33.0000
2	(?)	1	4:16.0000	7:17.0000	3:01.0000	7:34.0000	10:05.0000	2:31.0000	10:22.0000	12:59.0000	2:37.0000	13:16.0000	15:49.0000	2:33.0000
3	(?)	1	8:03.0000	11:04.0000	3:01.0000	11:21.0000	13:52.0000	2:31.0000	14:09.0000	16:46.0000	2:37.0000	17:03.0000	19:36.0000	2:33.0000
4	(?)	1	11:50.0000	14:51.0000	3:01.0000	15:08.0000	17:39.0000	2:31.0000	17:56.0000	20:33.0000	2:37.0000	20:50.0000	23:23.0000	2:33.0000
5	(?)	1	15:37.0000	18:38.0000	3:01.0000	18:55.0000	21:26.0000	2:31.0000	21:43.0000	24:20.0000	2:37.0000	24:37.0000	27:10.0000	2:33.0000
6	(?)	1	19:24.0000	22:25.0000	3:01.0000	22:42.0000	25:13.0000	2:31.0000	25:30.0000	28:07.0000	2:37.0000	28:24.0000	30:57.0000	2:33.0000
7	(?)	1	23:11.0000	26:12.0000	3:01.0000	26:29.0000	29:00.0000	2:31.0000	29:17.0000	31:54.0000	2:37.0000	32:11.0000	34:44.0000	2:33.0000
8	(?)	1	26:58.0000	29:59.0000	3:01.0000	30:16.0000	32:47.0000	2:31.0000	33:04.0000	35:41.0000	2:37.0000	35:58.0000	38:31.0000	2:33.0000
9	(?)	1	30:45.0000	33:46.0000	3:01.0000	34:03.0000	36:34.0000	2:31.0000	36:51.0000	39:28.0000	2:37.0000	39:45.0000	42:18.0000	2:33.0000
10	(?)	1	34:32.0000	37:33.0000	3:01.0000	37:50.0000	40:21.0000	2:31.0000	40:38.0000	43:15.0000	2:37.0000	43:32.0000	46:05.0000	2:33.0000

Fig. 7.5 Processing time of PMs

The residency time of a wafer in a PM would fluctuate with activity time variation such that the wafer residency time constraints may be violated when a feasible schedule obtained by simulation model discussed above is applied. The simulation model could analyze the scheduling feasibility and point out the feasible region of processing parameters, respectively.

	object 0	integer 1	time 2	time 3	time 4	time 5	time 6	time 7	time 8	time 9	time 10	time 11	time 12	time 13
string	Object	BatchNo	PM1in	PM1out	PM1span	PM2in	PM2out	PM2span	PM3in	PM3out	PM3span	PM4in	PM4out	PM4span
1	*.MUs.Entity:1	1	29.0000	3:40.0000	3:11.0000	3:57.0000	6:28.0000	2:31.0000	6:45.0000	9:22.0000	2:37.0000	9:39.0000	12:32.0000	2:53.0000
2	*.MUs.Entity:2	1	4:26.0000	7:37.0000	3:11.0000	7:54.0000	10:25.0000	2:31.0000	10:42.0000					
3	*.MUs.Entity:3	1	8:23.0000	11:34.0000	3:11.0000	11:51.0000								
4	*.MUs.Entity:4	1	12:20.0000											
5	*.MUs.Entity:5	1												

Fig. 7.6 Processing time of PMs with wafer extended residency

Fig. 7.7 Alert information

Example 7.2 Modify the processing time of step 1 from 169 to 179 s and then re-run the program. The simulating result of cluster tools operation collected during the test is shown in Fig. 7.6.

According to the simulating result of example 7.2, its schedule is obviously infeasible. When a schedule is infeasible, the simulation model would issue alert information and calculate exceeding wafer sojourn time. They are illustrated in Fig. 7.7.

7.6 Conclusion

A simulation model of steady state scheduling for cluster tools has been presented above. The simulation model could simulate the activities of cluster tools operation and record their timing sequence.

Two examples are applied to test the simulation model. The results of simulating process indicate that when a cluster tools operates under normal operating condition that subjects to wafer residency time constraints, the feasible schedule of cluster tools operation could be found. If the schedule violated the wafer residency time constraints, simulation model would issue the alert information and calculate the wafer sojourn time exceeding allowed wafer residency time.

According to above discussion, simulation model discussed in this paper could describe the activities of robot and wafer processing of cluster tools such that it can analyze the quantitative property of cluster tools such as scheduling. It is a useful way to analyze the quantitative performance of cluster tools operation. It is noted that this type of simulation model is suitable for modeling and scheduling cluster tools with single-arm or dual-arm robot just by adapting objects, the control methods and corresponding programs of SimTalk.

References

Artherton RF, Turner FT, Pool MA (1990) Performance analysis of multi-process semiconductor manufacturing equipment. In: Proceedings of IEEE/SEMI advanced semiconductor conference, Danvers, MA, 1990, pp 131–136

Bader M, Hall R, Strasser G (1990) Integrated processing equipment. Solid State Technol 33(5):149–154

Burggraaf P (1995) Coping with the high cost of wafer fabs. IEEE Trans Semiconduct Int 18(38):45–50

Kim J-H, Lee T-E, Lee H-Y, Park D-B (2003) Scheduling analysis of timed-constrained dual-armed cluster tools. IEEE Trans Semiconduct Manuf 16(3):521–534

Koehler EJ, Seppanen MS (1999) Evaluation of cluster tool throughput for thin film head production. In: Proceedings of 1999 winter simulation conference, Phoenix, 1999, pp 714–719

Lee T-E (2008) A review of scheduling theory and methods for semiconductor manufacturing cluster tools. In: Proceedings of 2008 winter simulation conference, Miami, 2008, pp 2127–2135

Lee T-E, Park S-H (2005) An extended event graph with negative places and tokens for timed window constraints. IEEE Trans Automat Sci Eng 2(4):319–332

Li F, Wu N-q (2009) Virtual cluster tool in wafer fabrication based on eM-Plant. Comput Eng 35(10):232–234 (in Chinese)

Lopez MJ, Wood SC (1998) Systems of multiple cluster tools: configuration and performance under perfect reliability. IEEE Trans Semiconduct Manuf 11(3):465–474

Lopez MJ, Wood SC (2003) Systems of multiple cluster tools: configuration, reliability, and performance. IEEE Trans Semiconduct Manuf 16(2):170–178

McNab TK (1990) Cluster tools, part 1: emerging processes. IEEE Trans Semiconduct Int 13(9):58–63

Newboe B (1990) Cluster tools: a process solution. IEEE Trans Semiconduct Int 13(8):82–88

Pan C-r, Wu N-q (2009) Modeling and simulation of single-arm cluster tool in wafer fabrication. J Syst Simul 21(2):554–562 (in Chinese)

Perkinson TL, MacLarty PK, Gyurcsik RS, Cavin RK III (1994) Single-wafer cluster tool performance: an analysis of throughput. IEEE Trans Semiconduct Manuf 7(3):369–373

Perkinson TL, Gyurcsik RS, MacLarty PK (1996) Single-wafer cluster tool performance: an analysis of the effects of redundant chambers and revisitations sequences on throughput. IEEE Trans Semiconduct Manuf 9:384–400

Singer P (1995) The driving forces in cluster tool development. Semiconduct Int 18(8):113–118

Srinivasn RS (1998) Modeling and performance analysis of cluster tools using Petri nets. IEEE Trans Semiconduct Manuf 11(3):394–403

Wu NQ (1999) Necessary and sufficient conditions for deadlock-free operation in flexible manufacturing systems using a colored Petri net model. IEEE Trans Syst Man Cybern Part C 29(2):192–204

Wu NQ, Zhou MC (2001) Avoiding deadlock and reducing starvation and blocking in automated manufacturing systems based on a Petri net model. IEEE Trans Robot Automat 17(5):658–669

Wu NQ, Zhou MC (2010a) Colored timed petri nets for modeling and analysis of cluster tools. Asian J Contr 12(3):253–266

Wu NQ, Zhou MC (2010b) A closed-form solution for schedulability and optimal scheduling of dual-arm cluster tools with wafer residency time constraint based on steady schedule analysis. IEEE Trans Automat Sci Eng 7(2):303–315

Wu NQ, Zhou MC (2010c) Analysis of wafer sojourn time in dual-arm cluster tools with residency time constraint and activity time variation. IEEE Trans Semiconduct Manuf 23(1):53–64

Wu NQ, Zhou MC (2010d) Petri net-based scheduling of time-constrained dual-arm cluster tools with bounded activity time variation. In: Proceedings of 6th IEEE conference on automation science and engineering, Toronto, 21–24 Aug 2010, pp 465–470

Wu N, Chu C, Chu F, Zhou M (2008) A Petri net method for schedulability and scheduling problem in single-arm cluster tools with wafer residency time constraint. IEEE Trans Semiconduct Manuf 21(2):224–237
Wu NQ, Chu F, Chu CB, Zhou MC (2011a) Petri net-based scheduling of single-arm cluster tools with reentrant atomic layer deposition processes. IEEE Trans Automat Sci Eng 8(1):42–55
Wu NQ, Chu F, Chu CB, Zhou MC (2011b) Petri net-based cycle time analysis of dual-arm cluster tools with wafer revisiting and swapping strategy. In: Proceedings of 2011 IEEE international conference on robotics and automation, Shanghai, 2011, pp 5499–5504
Zuberek WM (2001) Timed Petri nets in modeling and analysis of cluster tools. IEEE Trans Robot Automat 17(5):562–575
Zuberek WM (2004) Cluster tools with chamber revisiting – modeling and analysis using timed Petri nets. IEEE Trans Semiconduct Manuf 17(3):333–344

Chapter 8
A New Quant Model to Ensure Successful QFD Implementation in the Increasingly Uncertain Environment

Wei Xiong, Kun Xiong, and Xiao-tun Wang

Abstract QFD (Quality Function Deployment) can improve the quality of product and increase customer satisfaction, leading to superior performance of companies. But as global economy develops, the conditions wherein QFD is applied are becoming increasingly uncertain, which causes a lot of problems. Particularly, traditional techniques of QFD cannot deal with the vagueness, subjectivity and uncertainty of the linguistic information provided by customers as well as the R&D teams. This paper proposed a new model to solve such problems. We first integrated rough set theory into AHP (Analytic Hierarchy Process) to build one part of the model, which mainly deal with the uncertain information provided by customers. Then we applied linguistic information based GDM (Group Decision Making) theory to the expert assessment process to build another part of the model, which deal with the uncertain information within the R&D team.

Keywords Linguistic information based group decision making theory • Quality function deployment • Rough set theory • Uncertain environment

8.1 Introduction

As global economy develops, customers are gaining clouts in the market. The powerful customers force companies pay more attention to the VOC (voice of customer), which is critical to customer satisfaction (Akao and Mazur 2003). In order to gain competitive advantage, companies must rely on the sustainable product innovations that cater to the CNs (customer needs) (Li and Atuahene-Gima 2001).

W. Xiong (✉) • K. Xiong • X. Wang
School of Management, Zhejiang University, Hangzhou, China
e-mail: xiongk1987@hotmail.com

E. Qi et al. (eds.), *The 19th International Conference on Industrial Engineering and Engineering Management*, DOI 10.1007/978-3-642-37270-4_8,

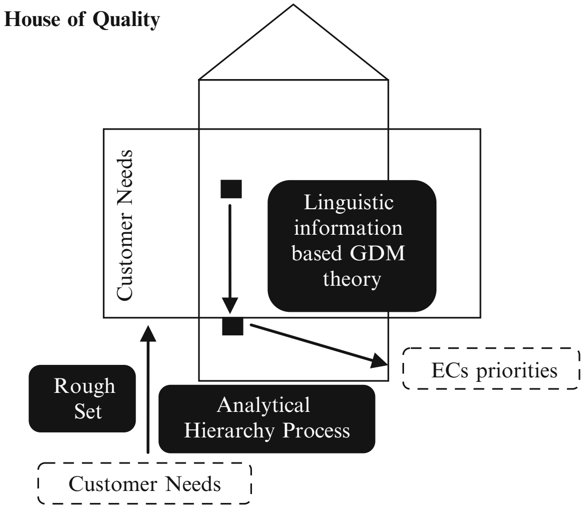


Fig. 8.1 Quant model to tackle uncertainty

That is the reason why QFD, a method to develop product and control the quality based on customer needs, is gaining attention in the literature (Akao and Mazur 2003; Carnevalli and Miguel 2008).

QFD is a "method to transform CNs into ECs (Engineering Characteristics), which can be easily monitored by the structured process (e.g. house of quality)" (Akao and Mazur 2003). Previous researchers have demonstrated that QFD can shorten product development cycle (Franceschini and Rossetto 1998), facilitate intern team communication (Sanford 2005), enhance quality of product and increase customer satisfaction, leading to superior performance of firms (Herrmann et al. 2006).

However, in the increasingly uncertain environment the traditional techniques used in QFD have showed a lot of drawbacks, especially its inability to deal with the vagueness, subjectivity and uncertainty of the linguistic information provided by customers as well as R&D teams (Van De Poel 2007).

Although scholars have tried hard to solve such problems (Kim et al. 2000; Raharjo et al. 2008; Fung et al. 2006), the methods they use might not be suitable in some conditions. For example, Kim et al. (2000) and Fung et al. (2006) applied fuzzy set theory and fuzzy regression in QFD, but the results might be unreliable to some extent because of the subjective selection of membership function used in their models.

To solve this problem we build a new model Fig. 8.1, integrating the rough set theory, which is a powerful knowledge discovery tool in uncertain conditions, the AHP and the linguistic information based GDM theory into QFD application.

The rest of this paper is organized as follows: Sect. 8.2 describes how we integrate rough set theory and AHP to get the RAHP method, which is used to analyze customer needs particularly to determine the comparative importance of different customer needs. In Sect. 8.3, linguistic information based GDM theory is applied in the correlation determination process to determine the comparative importance of ECs. We conclude this paper in Sect. 8.4.

8.2 The R-AHP Method for CNs Analysis

8.2.1 The Vague, Subjective and Uncertain Information of Customers

Customers are the critical factor in the product development process, especially in the early stage. This makes the subjective judgments perceptions and assessments unavoidable at the early stage of product development process. Therefore, the QFD analysis of customers' needs is inherently subjective and uncertain (Van De Poel 2007).

Traditionally, analyses of customers' needs include two steps (Kwong and Bai 2003): first, acquire initial customer needs by questionnaires; second, apply hierarchical method such as AHP (analysis of hierarchical process) to structuralize customers' needs and to determine the comparative importance of different customer needs. Currently the first step is relatively mature and certain. But the determination of the comparative importance of customer needs is much more difficult due to the large amount of ambiguous and vague information. To solve this problem we propose a new model that combines the rough set theory and AHP.

8.2.2 Rough Set Theory and Rough Number

AHP is easy to implement in the second stage of customer needs analysis, but it cannot deal with the vague and subjective linguistic-information. On the other hand the rough set theory can solve the vagueness and subjectivity without any preliminary information (e.g. the membership function in fuzzy theory) (Pawlak 1997). This provides with the solid foundation to the combination.

Before we propose our model, we give some basic definition used in the model.

Definition 8.2.1 U is a set of real numbers, according to rough set theory (Pawlak 1997) $B_*(c_i)$ and $B^*(c_i)$ are the B-lower and B-upper approximation of number C_i, wherein B is the attributes of the objects in U (here B has one object: 'value'). $BN_B(c_i) = B^*(c_i) - B_*(c_i)$ is the boundary of $B(c_i)$. Define:

$$\underline{\mathrm{Lim}}(c_i) = \frac{1}{|B_*(c_i)|} \sum x | x \in B_*(c_i)$$

$$\overline{\mathrm{Lim}}(c_i) = \frac{1}{|B^*(c_i)|} \sum x | x \in B^*(c_i)$$

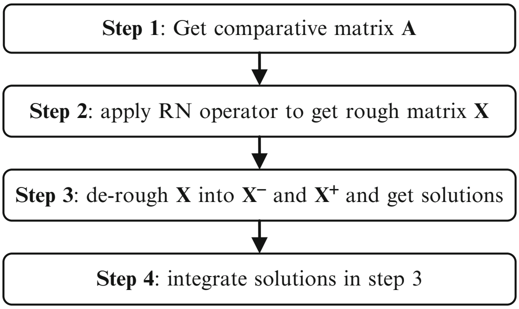


Fig. 8.2 R-AHP model for CNs analysis

as the maximum approximation and minimum approximation of set $B(c_i)$, wherein c_i is an object of U and satisfies $c_i < c_j(i < j)$. We further define the rough number as follows:

$$RN\left(c_i\right) = \left[\underline{Lim}\left(c_i\right), \overline{Lim}\left(c_i\right)\right]$$

$\underline{Lim}\left(c_i\right)$ and $\overline{Lim}\left(c_i\right)$ are denoted as c_i^- and c_i^+ for short.

8.2.3 Integrate Rough Number and AHP to Assess CNs

Now we can present the model proposed to assess CNs in Fig. 8.2

Step 1: determine and structuralize CNs, and use AHP questionnaire to get comparative matrix A. Suppose we have s customers to assess n CNs. We get comparative matrix $\mathbf{A^i}$ $(i = 1, 2 \ldots s)$ as follows:

$$\mathbf{A^i} = \begin{bmatrix} \mathbf{1} & \mathbf{x^i_{12}} & \ldots & \mathbf{x^i_{1n}} \\ \mathbf{x^i_{21}} & 1 & \ldots & \mathbf{x^i_{2n}} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{x^i_{n1}} & \mathbf{x^i_{n2}} & \ldots & \mathbf{1} \end{bmatrix}$$

Step 2: in order to construct the rough number matrix, we first integrate $\mathbf{A^i}$ into $\mathbf{A^*}$

$$\mathbf{A^*} = \begin{bmatrix} \mathbf{1} & \mathbf{X^*_{12}} & \ldots & \mathbf{X^*_{1n}} \\ \mathbf{X^*_{21}} & 1 & \ldots & \mathbf{X^*_{2n}} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{X^*_{n1}} & \mathbf{X^*_{n2}} & \ldots & \mathbf{1} \end{bmatrix}$$

Wherein $\mathbf{X^*_{ij}} = (\mathbf{x^1_{ij}}, \mathbf{x^2_{ij}}, \ldots, \mathbf{x^s_{ij}})$.

Then use $\mathbf{RN}\left(\mathbf{C_i}\right)$ operator on $\mathbf{X^*_{ij}}$, denote as $\mathbf{RN}\left(\mathbf{X^*_{ij}}\right) = [\mathbf{x^-_{ij}}, \mathbf{x^+_{ij}}]$. And therefore

$$\mathbf{RN}\left(\mathbf{A^*}\right) = \mathbf{X} = \begin{bmatrix} [\mathbf{1},\mathbf{1}] & [\mathbf{x^-_{12}}, \mathbf{x^+_{12}}] & \cdots & [\mathbf{x^-_{1n}}, \mathbf{x^+_{1n}}] \\ [\mathbf{x^-_{21}}, \mathbf{x^+_{21}}] & [\mathbf{1},\mathbf{1}] & \cdots & [\mathbf{x^-_{2n}}, \mathbf{x^+_{2n}}] \\ \vdots & \vdots & \ddots & \vdots \\ [\mathbf{x^-_{n1}}, \mathbf{x^+_{n1}}] & [\mathbf{x^-_{n2}}, \mathbf{x^+_{n2}}] & \cdots & [\mathbf{1},\mathbf{1}] \end{bmatrix}$$

Step 3: to get the comparative importance of CNs, we can break the rough AHP matrix X into $\mathbf{X^-}$ and $\mathbf{X^+}$, and calculate respectively.

$$\mathbf{X^-} = \begin{bmatrix} \mathbf{1} & \mathbf{x^-_{12}} & \cdots & \mathbf{x^-_{1n}} \\ \mathbf{x^-_{21}} & \mathbf{1} & \cdots & \mathbf{x^-_{2n}} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{x^-_{n1}} & \mathbf{x^-_{n2}} & \cdots & \mathbf{1} \end{bmatrix}$$

$$\mathbf{X^+} = \begin{bmatrix} \mathbf{1} & \mathbf{x^+_{12}} & \cdots & \mathbf{x^+_{1n}} \\ \mathbf{x^+_{21}} & \mathbf{1} & \cdots & \mathbf{x^+_{2n}} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{x^+_{n1}} & \mathbf{x^+_{n2}} & \cdots & \mathbf{1} \end{bmatrix}$$

We can obtain the comparative importance of CNs from $\mathbf{X^-}$ and $\mathbf{X^+}$ (Kwong and Bai 2003; Saaty 2003)

$$\mathbf{W^-} = \left(\mathbf{w^-_1}, \mathbf{w^-_2}, \ldots, \mathbf{w^-_n}\right)$$
$$\mathbf{W^+} = \left(\mathbf{w^+_1}, \mathbf{w^+_2}, \ldots, \mathbf{w^+_n}\right)$$

Step 4: standardize and integrate the results in step 3 to get the comparative importance of CNs.

$$w^*_i = \frac{1}{2}\left(\frac{w^-_i}{\sum_{k=1}^{n} w^-_k} + \frac{w^+_i}{\sum_{k=1}^{n} w^+_k}\right)$$

8.3 Linguistic Information Based Method For ECs Assessment

8.3.1 Linguistic Information Theory

One of the most important steps to determine ECs is to figure out the preliminary priority of them by analyzing the relationships between customer requirements and ECs. In previous literatures the essential methods to determine such relationships are expert interviews and decisions (Chan and Wu 2002). But due to the complex and ambiguous decision making situation and uncertain judgment process of human, the experts often cannot give accurate quantitative assessment when they evaluate the situation. In order to solve this problem, we apply linguistic information group decision making theory into the experts' decision.

Definition 8.3.1 define linguistic information measurement

$$S^{(k)} = \left\{ S_\alpha^{(k)} \,\middle|\, \begin{array}{c} \alpha = 1-k, \frac{2}{3}(2-k), \ldots, 0, \ldots, \frac{2}{3}(k-2) \\ , \ldots, k-1 \end{array} \right\} \tag{8.1}$$

Wherein k is positive integer, and $S_\alpha^{(k)}$ satisfies:

1. if $\alpha > \beta$, *then* $S_\alpha^{(k)} > S_\beta^{(k)}$;
2. It has negative operator neg $\left(S_\alpha^{(k)}\right) = S_{-\alpha}^{(k)}$

For calculation purpose, we further extend discrete set $S^{(k)}$ to continuous set $\bar{S}^{(k)} = \left\{S_\alpha^{(k)} | \alpha \in [-t, t]\right\}, t\,(t \geq k)$ is a nature number.

Definition 8.3.2 for any two linguistic terms, $S_{\alpha_1}^{(k)}, S_{\alpha_2}^{(k)} \in \bar{S}^{(k)}, \lambda \in [0, 1]$ we have following algorithm

$$S_{\alpha_1}^{(k)} \oplus S_{\alpha_2}^{(k)} = S_{\alpha_2}^{(k)} \oplus S_{\alpha_1}^{(k)} = S_{\alpha_1+\alpha_2}^{(k)}$$
$$\lambda S_{\alpha_1}^{(k)} = S_{\lambda\alpha_1}^{(k)}$$

Definition 8.3.3 (Xu 2006a) a projection from n dimensions vector to single dimension vector, $\bar{S}^n \to \bar{S}$, we define LWA (Linguistic Weighted Algorithm) as:

$$\begin{aligned} &\mathrm{LWA}\left(S_{\alpha_1}, S_{\alpha_2}, \ldots, S_{\alpha_n}\right) \\ &\quad = w_1 S_{\alpha_1} \oplus w_2 S_{\alpha_2} \oplus \ldots \oplus w_n S_{\alpha_n} \\ &\quad = S_{\bar{\alpha}} \end{aligned} \tag{8.2}$$

Wherein $\bar{a} = \sum_{j=1}^{n} w_j \alpha_j$ and $w_j \in [0, 1], \sum_{j=1}^{n} w_j = 1$

Definition 8.3.4 (Xu 2006b) a projection from n dimensions vector to single dimension vector $\bar{S}^n \rightarrow \bar{S}$; and $\boldsymbol{\nu} = (\nu_1, \nu_2, \ldots, \nu_n)$ is the weighted location vector, $\nu_j \in [0, 1], \sum_{j=1}^{n} \nu_j = 1$. We define LHA (Linguistic Hybrid Algorithm) as:

$$LHA(S_{\alpha_1}, S_{\alpha_2}, \ldots, S_{\alpha_n}) = \nu_1 S_{\beta_1} \oplus \nu_2 S_{\beta_2} \oplus \ldots \oplus \nu_1 S_{\beta_n} \tag{8.3}$$

Wherein S_{β_j} is the j^{th} factor in weighted average set $\bar{S}' = \{S'_{\alpha_i} | S'_{\alpha_i} = nw_i S_{\alpha_i}, i = 1, 2, \ldots, n\}$, here n is the balance coefficient; $\mathbf{w} = (w_1, w_2, \ldots, w_n)$ is the weight vector of S_{α_i} $(i = 1, 2, \ldots, n)$, $\sum_{j=1}^{n} w_i = 1$.

Often in time, the experts give linguistic information assessment with different granular, due to their different experience and knowledge base. To make the calculation of linguistic information consistent we should standardize the linguistic information assessments that have different granular. Assume we have continuous linguistic information assessment $\bar{S}^{(k_1)} = \left\{S_\alpha^{(k_1)} | \alpha \in [1 - k_1, k_1 - 1]\right\}$ and $\bar{S}^{(k_1)} = \{S_\beta^{(k_2)} | \beta \in [1 - k_2, k_2 - 1]\}$, define the transfer function (Xu 2009) F as:

$$F : S^{(k_1)} \rightarrow S^{(k_2)} \tag{8.4}$$

$$\boldsymbol{\beta} = F(\boldsymbol{\alpha}) = \boldsymbol{\alpha} \frac{k_2 - 1}{k_1 - 1} \tag{8.5}$$

$$F^{-1} : S^{(k_2)} \rightarrow S^{(k_1)} \tag{8.6}$$

$$\alpha = \mathbf{F}^{-1}(\boldsymbol{\beta}) = \boldsymbol{\beta} \frac{\mathbf{k_1 - 1}}{\mathbf{k_2 - 1}} \tag{8.7}$$

8.3.2 Apply Linguistic Information Theory into ECs Analysis

Now that we have basic tool for linguistic information assessment, we can build up our linguistic information based model in Fig. 8.3.

Step 1: Assume we have m ECs, $EC_i (i = 1, 2, \ldots, m)$, which can be determined by current product standards or cause-effect analysis such as DOE (Design of Experiment).

Step 2: Assume that we have t experts to determine the relationships $Ex_\xi (\xi = 1, 2, \ldots, t)$. The weight of Ex_ξ is $\mathbf{w} = (w_1, w_2, \ldots, w_t)^T$, which satisfies $w_\xi \geq 0$, and $\sum_{\xi=1}^{t} w_\xi = 1$. They consider n CNs of relationships for each EC $CN_\xi (\xi = 1, 2, \ldots, n)$ and its weights vector $= (d_1, d_2, \ldots, d_n)^T d_j \geq 0, \sum_{j=1}^{n} d_j = 1$. This generates relationships values $r_{i,j}^{(k^\xi)} (\xi = 1, 2, \ldots, t; i = 1, 2, \ldots, m; j = 1, 2, \ldots, n)$.

We define expert matrix $\mathbf{R} = \left(R_{i,j}\right)$ wherein $R_{i,j} = \{r_{i,j}^{k^1}, r_{i,j}^{k^2}, \ldots, r_{i,j}^{k^t}\}$.

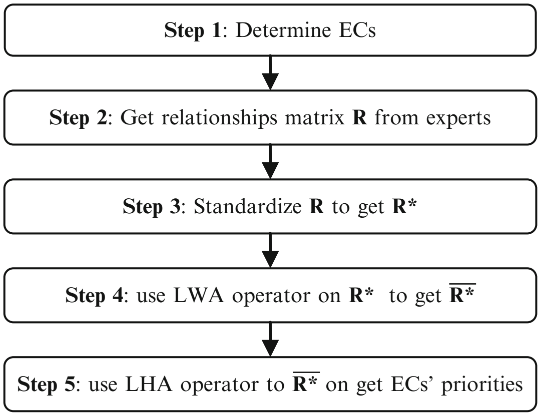


Fig. 8.3 Linguistic information based model for ECs correlation analysis

Note that the matrix **R** is a three-dimension matrix and $\mathbf{k}^{\boldsymbol{\xi}}\,(\boldsymbol{\xi} = \mathbf{1}, \mathbf{2}, .., \mathbf{t})$ indicate the different granular the decision maker uses.

Step 3: To get the ultimate priorities of ECs we should standardize k^{ξ}, using algorithm (8.4), (8.5), (8.6) and (8.7) mentioned above to get a unified-granular decision making matrix $\mathbf{R}^* = \left(\mathrm{R}_{\mathrm{i,j}}{}^*\right)$.

Step 4: Then we can apply LWA operator (8.2) to $\mathbf{R}^*$ to get $\overline{\mathbf{R}^*} = (\mathrm{r}_{\mathrm{i,j}})_{\mathrm{m}\times\mathrm{n}}$.

Note, here LWA operator makes the three-dimensional matrix $\mathbf{R}^*$ a two-dimension matrix $\overline{\mathbf{R}^*}$, which indicate the relationships between CNs and ECs.

Step 5: We finally apply LHA operator (8.3) to $\overline{\mathbf{R}^*}$ to calculate the ultimate priorities value $(\mathrm{P_i}(\mathrm{i} = 1, 2, \ldots, \mathrm{m}))$ of each EC.

8.4 Discussion and Conclusion

Firstly, compared with traditional methods to analyze CNs, our R-AHP method does not need any preliminary information such as subjective pre-judgment, hypothesis or membership function. All the calculations are based on the data acquired through AHP questionnaire. This ensures the objectivity of original data. On the other hand, the rough number, which is based on rough set theory, considers accordingly the vague subjective and uncertain information provided by customer. Therefore, the model has much more validity and effectiveness than traditional methods.

Secondly, the method proposed for correlation analysis between ECs and CNs is integrated with linguistic information based GDM theory, which ensures the better understanding and description of the experts' survey than traditional methods. Firstly, it can properly deal with the different linguistic information granular used by experts. Secondly, it also appropriately leverages the knowledge base of different experts, by assigning importance weights and location weights when conducting calculation.

References

Akao Y, Mazur GH (2003) The leading edge in QFD: past, present and future. Int J Qual Reliab Manage 20(1):20–35
Carnevalli JA, Miguel PC (2008) Review, analysis and classification of the literature on QFD—types of research, difficulties and benefits. Int J Product Econ 114(2):737–754
Chan LK, Wu ML (2002) Quality function deployment: a literature review. Eur J Oper Res 143(3):463–497
Franceschini F, Rossetto S (1998) Quality function deployment: how to improve its use. Total Qual Manage 9(6):491–500
Fung RYK, Chen Y, Tang J (2006) Estimating the functional relationships for quality function deployment under uncertainties. Fuzzy Set Syst 157(1):98–120
Herrmann A et al (2006) An empirical study of quality function deployment on company performance. Int J Qual Reliab Manage 23(4):345–366
Kim KJ et al (2000) Fuzzy multicriteria models for quality function deployment. Eur J Oper Res 121(3):504–518
Kwong C, Bai H (2003) Determining the importance weights for the customer requirements in QFD using a fuzzy AHP with an extent analysis approach. Iie Trans 35(7):619–626
Li H, Atuahene-Gima K (2001) Product innovation strategy and the performance of new technology ventures in china. Acad Manage J 1:1123–1134
Pawlak Z (1997) Rough set approach to knowledge-based decision support. Eur J Oper Res 99(1):48–57
Raharjo H, Brombacher AC, Xie M (2008) Dealing with subjectivity in early product design phase: a systematic approach to exploit quality function deployment potentials. Comput Ind Eng 55(1):253–278
Saaty TL (2003) Decision-making with the AHP: why is the principal eigenvector necessary. Eur J Oper Res 145(1):85–91
Sanford J (2005) How useful is QFD. Qual Prog 38(1):51–59
Van De Poel I (2007) Methodological problems in QFD and directions for future development. Res Eng Des 18(1):21–36
Xu Z (2006a) On generalized induced linguistic aggregation operators. Int J Gen Syst 35(1):17–28
Xu Z (2006b) A note on linguistic hybrid arithmetic averaging operator in multiple attribute group decision making with linguistic information. Group Decis Negot 15(6):593–604
Xu Z (2009) An interactive approach to multiple attribute group decision making with multigranular uncertain linguistic information. Group Decis Negot 18(2):119–145

Chapter 9
Multivariate Measurement System Analysis Based on Projection Pursuit Method

Xiaofang Wu, Liangxing Shi, and Zhen He

Abstract With the improvement of the automation of the measurement processes and the complexity of products, measurement system analysis is becoming increasingly important (Supported by National Natural Science Foundation of China (No.71102140, 70931004)). However, there exists some difficulty in directly application of univariate measurement system analysis for multiple measured quality characteristics with correlation and the univariate measurement system capability index cannot be used in multivariate measurement system. Therefore, in this paper projection pursuit is used to analyze the multivariate measurement system. The best projection direction is obtained by optimizing the projection direction with Genetic Algorithm, the relationship between multivariate data and there projection is analyzed. Then three common measurement system capability indices are extended to the multivariate measurement system with the projection of the raw data in order to evaluate multivariate measurement system capability, at last the method proposed was proved by an example.

Keywords Analysis of variance (ANOVA) • Measurement system analysis • Multivariate capability • Projection pursuit

9.1 Introduction

The manufactures make decisions for quality of products and process by process capability indices and control charts. However before process capability analysis, the measurement system capability should firstly be analyzed to make sure the

---

Supported by National Natural Science Foundation of China (No.71102140, 70931004).

X. Wu (✉) • L. Shi • Z. He
College of Management and Economics, Tianjin University, Tianjin, China
e-mail: shuanger@tju.edu.cn

E. Qi et al. (eds.), *The 19th International Conference on Industrial Engineering and Engineering Management*, DOI 10.1007/978-3-642-37270-4_9,

effective of the measurement data. The measurement system for univariate quality characteristic has been studied and applied in depth, but there still exist that the gauge measures some quality characteristics at once in many measurement process, such as the two-dimensional measured values of pump rotation device (Joseph 2003; Shannon 2007), the measurement results of three coordinate measurement system for the location (Rajamohan et al. 2011), the measurement system for the radio frequency function in mobile phone production process which measures more than one index at once (Su et al. 2006).

To assess measurement system independently for the multivariate measurement system cannot be accurately reflected the capability of the multivariate measurement system, if there exists correlate between quality characteristics. Based on thorough analysis of evaluation indices for the measurement system capability, a new method based on projection pursuit is proposed for multivariate measurement system capability evaluation.

9.2 Review

The main concern of the measurement system analysis is whether the measurement process can accurately and reliably reflect the variation of the measured object, so we need to assess the measurement system capability. The most common evaluation indices for measurement system capability are Precise-to-Tolerance ratio (PTR), %R&R, Signal-Noise Ratio (SNR) (Automotive Industry Action Group 2002). PTR is the ratio of variation of measurement system to quality characteristic tolerance, proposed by AIAG. The definition of %R&R and SNR indicate that the accuracy of the gauge error is not relied on the specified limits (Woodall and Borror 2008).

$$
\begin{aligned}
PTR &= \frac{5.15\sigma_G}{TOL} \times 100\ \% \\
\%R\&R &= \left(\frac{\sigma_G}{\sigma_M}\right) \times 100\ \% \\
SNR &= \sqrt{2}\left(\frac{\sigma_Q}{\sigma_G}\right)
\end{aligned}
\tag{9.1}
$$

Where σ_G, σ_M and σ_Q are standard deviation of the gauge error, measurement value and unknown true value of quality characteristics, respectively. TOL are the specified limits of the quality characteristics.

According the guidelines proposed by AIAG, the measurement system capability is acceptable when PTR value is less than 0.10. Whether the capability is acceptable depends on the importance of the measurement object, the cost of measurement device and repair cost when PTR value is between 0.10 and 0.30. When PTR value is more than 0.30, the measurement system capability cannot be acceptable and should be improved. When SNR value is not less than 5, the measurement system can be accepted (Burdick et al. 2003).

These evaluation indices of measurement system capability can be only used in univariate measurement system. They cannot be used when there are multi-dimensional quality characteristics in multivariate measurement system. Burdick etc. suggested the necessity of research in depth on multivariate system analysis after review on measurement system analysis (Burdick et al. 2005). Multivariate statistics methods, such as Projection Pursuit (PP), are applied to the quality engineering practice to resolve the multivariate capability problems (Johnson and Wichern 2007). Ch'ng etc. proposed multivariate process capability indices based on weighting average of univariate index (Ch'ng et al. 2005). Shinde etc. considered analyzing multivariate process capability by using Principal Component Analysis method (Shinde and Khadse 2009). However these multivariate statistic methods have been seldom applied to multivariate measurement system. The PTR, %R&R and SNP, as measurement system capability indices, can be extended to multivariate measurement system using multivariate statistic method. Majeske etc. presented a new method to establish the evaluation indices of multivariate measurement system capability based on the volume ratio of ellipsoid and obtain the estimators of multivariate evaluation indices to assess the multivariate measurement system (Majeske 2008). However, when the measured quality characteristics are high dimensional, the efficiency and effectiveness of multivariate measurement system analysis would decrease using this method.

In this paper, the Genetic Algorithm is used to optimize the projection direction in the projection pursuit method to obtain the best projection direction and the projection relationship between multivariate sample data and projection data is analyzed, and then multivariate measurement system capability is evaluated.

9.3 Multivariate Measurement System Analysis

Projection pursuit method, which can be applied to high dimensional data analysis, is widely used in multiple regressions, multiple process control, etc. The main characteristics of projection pursuit method are, high dimensional data can be turned into low dimensional, interference of variables which have no relationship with data structure and characteristics can be excluded, nonlinear problems can be solved (Miyoshi et al. 1999). Projection pursuit method projects the high dimensional data through some combination into low dimensional subspace by computer technology, and find out the best projection which can reflect the original data structure and characteristics of multivariable through minimization or maximization of a projection index in order to achieve the purposes of high dimensional data analysis and research (Montanari and Lizzani 2001).

Assume $M = (M_1, M_2, \ldots, M_n)^T_{n\times p}$ represent the n-dimensional measured data of p quality characteristics which follow the normal distribution $M \sim N_p\left(\mu_{1\times p}, \sigma^2_{p\times p}\right)$ and $M_k (k = 1, 2, \ldots, n)$ are the k-th sample data, $Q = (Q_1, Q_2, \cdots, Q_n)^T_{n\times p}$ represent the true values of measured quality characteristics which are unknown, $G = (G_1, G_2, \ldots, G_n)^T_{n\times p}$ are gauge errors, $a = \left(a_1, a_2, \ldots, a_p\right)^T$ are an p-dimensional unit vector where $\|a\| = 1$.

For the evaluation indices calculation of multivariate system capability, the measured values of quality characteristics need to be projected linearly, that is, from different angle to observe data and find out the most optimal projection direction which is fully mining data features. The projection values z which M_i projected on the projection direction a are $z_i = a^T M_i^T = \sum\limits_{j=1}^{p} a_j m_{ij}$, and z follow the normal distribution $z \sim N\left(a^T \mu^T, a^T \sigma^2 a\right)$, where

$$\hat{\mu}_z = \frac{1}{n}\sum_{i=1}^{n} z_i = \frac{1}{n} a^T M_i^T = a^T \hat{\mu}^T$$

$$\hat{\sigma}_z^2 = \frac{1}{n-1}\sum_{i=1}^{n} (z_i - \hat{\mu}_z)^2 = \frac{1}{n-1} a^T (M - \hat{\mu})^T (M - \hat{\mu})\, a \tag{9.2}$$

According the features of multivariate measurement system capability, the measured data are projected to the direction of the largest extent discrete, that is, the projection indices function is constructed as $Q(a) = \sigma_z$, where σ_z is the standard error of projection values z. The best projection direction can reflect the data variation to the maximum extent, so the optimal projection direction can be estimated through solving the maximal problem of projection indices function, that is,

$$\begin{aligned} &\max\ Q(a) = \sigma_z \\ &s.t.\ \sum_{j=1}^{n} a_j^2 = 1 \end{aligned} \tag{9.3}$$

The projection direction can be determined by the genetic algorithm. In p-dimensional space random numbers a_i $(i = 1, 2, \ldots, p)$ of N groups in $0 \sim 1$ interval are randomly selected according to population scale as optimization codes, and normalize a, that is $a = \frac{a}{\|a\|}$, then calculate the projection indices $Q(a) = \sigma_z$. According to the principal in favor of projection indices to increase, selection, hybridization and variation is operated, and then calculate the values of evaluation function. If the values do not meet the conditions, continue to the next turn of operation, or else a obtained is the best projection direction.

Finally the multivariate measurement system capability indices *MPTR*, *MR&R* and *MSNR* can be obtained, which are latent univariate indices of the projection values z.

$$\begin{aligned} MPTR &= PTR_z \\ MR\&R &= \%R\&R_z \\ MSNR &= SNR_z \end{aligned} \tag{9.4}$$

Where for the calculation of PTR_Z, the specified limits TOL_Z for projection values are,

$$USL_Z = a^T USL_m$$
$$LSL_Z = a^T LSL_m \tag{9.5}$$

In order to estimate $MPTR$, $MR\&R$ and $MSNR$, manufacturer should do the gauge study. In multivariate measurement system analysis, the test design is similar to the single variable test design. Therefore, the one factor, two factors or three factors study is chosen according to the specific application (Montgomery and Runger 1993). However, when do the gauge study, each operator should measure n quality characteristics values for each product each time.

$$\widehat{\sigma}^2{}_{QZ}7 = \widehat{\sigma}^2{}_{OZ} + \widehat{\sigma}^2{}_{POZ} + \widehat{\sigma}^2{}_{EZ} = \frac{MS_{P_Z} - MS_{PO_Z}}{or}$$
$$\widehat{\sigma}^2{}_{GZ} = \widehat{\sigma}^2{}_{OZ} + \widehat{\sigma}^2{}_{POZ} + \widehat{\sigma}^2{}_{EZ}$$
$$= \frac{MS_{O_Z} - MS_{PO_Z}}{pr} + \frac{MS_{PO_Z} - MS_{E_Z}}{r} + MS_{E_Z}$$
$$\widehat{\sigma}^2{}_{MZ} = \widehat{\sigma}^2{}_{QZ} + \widehat{\sigma}^2{}_{GZ} \tag{9.6}$$

In the paper, we choose a cross test design model of random effect two factors. The estimators of $\widehat{\sigma}_{Q_z}, \widehat{\sigma}_{G_z}, \widehat{\sigma}_{M_z}$ can be obtained by analysis of variation for z. Then the evaluation indices estimators of multivariate measurement system can be acquired by using Eq. (9.4) to assess the multivariate measurement system.

9.4 Example

An automobile manufacturer did a gauge study for four critical characteristics M1, M2, M3 and M4 on a sheet-metal body panel. This example is from the reference (Majeske 2008), which chose five parts, two operators and measured four quality characteristics each part three times. Table 9.1 contains the raw data from the reference Majeske (2008) and the projection values. In this paper we apply the new method based on projection pursuit proposed before to analyze the multivariate measurement system capability.

According to the raw data, the variance-covariance vector of the four quality characteristics are obtained, see Table 9.2. And then use genetic algorithm method to get the projection direction $a = (0.498, 0.427, -0.645, -0.393)^T$. The projection values can be calculated by $z_i = a^T M_i^T = \sum_{j=1}^{p} a_j m_{ij}$, see Table 9.1.

Table 9.1 Raw data of measurement systems and projection values

P	O	M1	M2	M3	M4	z
1	1	−0.86	−0.40	0.44	−1.29	−0.3759
1	1	−0.81	−0.30	0.25	−1.10	−0.2604
1	1	−0.78	−0.32	0.24	−1.09	−0.2515
1	2	−0.78	−0.23	0.16	−1.01	−0.1929
1	2	−0.77	−0.33	0.21	−1.14	−0.2119
1	2	−0.81	−0.33	0.21	−1.16	−0.2239
2	1	−0.92	−0.73	−0.17	−0.01	−0.6563
2	1	−0.96	−0.75	−0.22	−0.03	−0.6446
2	1	−0.95	−0.74	−0.19	−0.07	−0.6390
2	2	−0.99	−0.80	−0.17	−0.08	−0.6935
2	2	−0.95	−0.82	−0.17	−0.10	−0.6743
2	2	−0.96	−0.77	−0.18	−0.08	−0.6593
3	1	−0.70	0.35	−0.83	1.38	−0.2061
3	1	−0.79	0.14	−0.69	1.19	−0.3563
3	1	−0.73	0.23	−0.76	1.30	−0.2860
3	2	−0.76	0.09	−0.63	1.09	−0.3621
3	2	−0.78	0.01	−0.59	1.03	−0.4084
3	2	−0.79	−0.03	−0.58	0.99	−0.4210
4	1	−0.66	−0.98	−0.04	−1.12	−0.2812
4	1	−0.63	−1.02	0.03	−1.19	−0.3010
4	1	−0.62	−1.02	0.02	−1.21	−0.2817
4	2	−0.69	−1.07	0.09	−1.24	−0.3712
4	2	−0.64	−1.07	0.10	−1.26	−0.3449
4	2	−0.71	−1.13	0.10	−1.32	−0.3818
5	1	−0.99	−1.01	0.49	−0.82	−0.9181
5	1	−0.96	−1.00	0.52	−0.86	−0.9025
5	1	−0.98	−0.97	0.52	−0.81	−0.9193
5	2	−0.96	−1.01	0.55	−0.81	−0.9458
5	2	−1.01	−1.01	0.52	−0.79	−0.9592
5	2	−0.97	−1.04	0.55	−0.85	−0.9478

Table 9.2 The variance-covariance of quality characteristics

	M1	M2	M3	M4
M1	0.01590	0.01471	−0.01935	−0.00463
M2	0.01471	0.21359	−0.13445	0.29958
M3	−0.01935	−0.13445	0.17727	−0.33019
M4	−0.00463	0.29958	−0.33019	0.82392

Through the analysis of variation for z, $\widehat{\sigma}^2_{QZ} = 0.0809$, $\widehat{\sigma}^2_{GZ} = 0.0037$ and $\widehat{\sigma}^2_{MZ} = 0.0846$ using Eq. (9.6). Finally MSNR and MR&R can be calculated by Eqs. (9.4) and (9.5) (Table 9.3). It gave no the specific limits for quality characteristics, so the index MPTR cannot be calculated (Majeske 2008).

In comparison with the indices of reference (Majeske 2008), both of the MSNR and MR&R values are less than the indices values of reference (Majeske 2008).

Table 9.3 The values of MSNR and MR&R

Index	MSNR	MR&R
Method based on PP	6.62	6.61 %
Method of Majeske's	11.30	12.30 %

In the multivariate measurement system analysis, the new method proposed in the paper, which is based on the projection pursuit method, is better than the method based on the volume ratio, especially when the data are high dimensional. Also the proposed method is applied to engineering practice more easily.

9.5 Conclusion

A new method of the multivariate measurement system analysis based on the projection pursuit method is put forward in the paper. This method extends the measurement system capability indices to the multivariate measurement system through turning the high dimensional quality characteristics with correlation into low dimensional data with no correlation by projection pursuit method, and then using the ANOVA model analyze the projection values. The high dimensional data with correlation can be projected to low dimension by the method. This method is proved feasible through the case analysis. Further research will include considering the importance of the quality characteristics in the multivariate measurement system analysis.

Acknowledgment The authors thank the editor and the anonymous referees for their comments and suggestions.

References

Automotive Industry Action Group (2002) Measurement systems analysis, 3rd edn. Automotive Industry Action Group, Detroit

Burdick RK, Borror CM, Montgomery DC (2003) A review of methods for measurement systems capability analysis. J Qual Technol 35(4):342–354

Burdick RK, Borror CM, Montgomery DC (2005) Design and analysis of gauge R&R studies. ASA-SIAM, Philadelphia

Ch'ng CK, Quah SH, Low HC (2005) Index in multiple response optimization. Qual Eng 17:165–171

Johnson RA, Wichern DW (2007) Applied multivariate statistical analysis, vol 6. Prentice Hall, Upper Saddle River

Joseph GV (2003) Gauge R&R analysis for two-dimensional data with circular tolerances. J Qual Technol 35(2):153

Majeske KD (2008) Approval criteria for multivariate measurement systems. J Qual Technol 40(2):140–153

Miyoshi T, Ichihashi H, Nagasaka K (1999) Fuzzy projection pursuit density estimation by eigenvalue method. Int J Approx Reason 20(3):237–248

Montanari A, Lizzani L (2001) A projection pursuit approach to variable selection. Comput Stat Data Anal 35(4):463–473

Montgomery DC, Runger GC (1993) Gauge capability and designed experiments, part I: basic methods. Qual Eng 6:115–135

Rajamohan G, Shunmugam MS, Samuel GL (2011) Effect of probe size and measurement strategies on assessment of freeform profile deviations using coordinate measuring machine. Measurement 44(5):832–841

Shannon S (2007) Analysis of two-dimensional gage repeatability and reproducibility. Qual Eng 19:29–37

Shinde RL, Khadse KL (2009) Multivariate process capability using principal component analysis. Qual Reliab Eng Int 25:69–77

Su CT, Chen LS, Chiang TL (2006) A neural network based information granulation approach to shorten the cellular phone test process. Comput Ind 57:412–423

Woodall WH, Borror CM (2008) Some relationships between gage R&R criteria. Qual Reliab Eng Int 24:99–106

Chapter 10
A Multivariate Synthetic Control Chart for Monitoring Covariance Matrix Based on Conditional Entropy

Li-ping Liu, Jian-lan Zhong, and Yi-zhong Ma

Abstract In multivariate statistical process control field, besides monitoring the changes in the mean vector of a multivariate process, it is important to detect the changes in the covariance matrix of a multivariate process. This paper proposes a multivariate synthetic control chart for monitoring the changes in the covariance matrix of a multivariate process under multivariate normal distribution. The proposed control chart is a combination of the traditional control chart based on conditional entropy and the conforming run length chart. The operation and design of this control chart are described.

Keywords Entropy • Multivariate control charts • Quality control • Statistical process control • Synthetic control chart

10.1 Introduction

Nowadays, multivariate statistical quality control is widely concerned, i.e. the simultaneously monitoring several correlated quality characteristics of a process (Zou and Tsung 2011). In fact, in many industrial applications more than a single variable is involved in the quality control problems.

The majority of the multivariate statistical quality control focuses on monitoring the shifts in process mean vector. Some control charts including Hotelling T^2 control chart, the MCUSUM control chart and the MEWMA control chart have been introduced. Except for monitoring the changes in the mean vector of a multivariate process, it's vital to detect the changes in the multivariate process covariance matrix, which reflects the changes in multivariate process variation. Actually, an increase

L. Liu (✉) • J. Zhong • Y. Ma
School of Economics and Management, Nanjing University of Science of Technology, Nanjing, China
e-mail: lpliu1016@163.com

E. Qi et al. (eds.), *The 19th International Conference on Industrial Engineering and Engineering Management*, DOI 10.1007/978-3-642-37270-4_10,

in process variation, which causes the quality of the process output to be more scattered, should be detected; while a decrease in process variation should also be detected since it causes the quality of the process output to be more concentrated and may be a clue to improve the quality of the process output.

The $|s|$ control chart is the basic tool for monitoring the shifts in multivariate process covariance matrix (Jackson 1985). Guerrero-Cusumano proposes a control chart based on the conditional entropy (E-chart) (Guerrero-Cusumano 1995), this type of control chart is much simpler than the $|s|$ control chart because the E-chart just requires simple algebraic operations for a given correlation matrix. In addition, a variety of the strategies used to improved the detection performance to monitor the mean vector have been propose to monitor the multivariate process covariance matrix, such as MEWMA control chart (Yeh et al. 2005).

Recently, the synthetic chart (Wu and Spedding 2000), combining the traditional Shewhart control chart or some other control chart with the conforming run length (*CRL*) control chart (Bourke 1991), has drawn the attention of researchers in literature (Wu et al. 2001; Huang and Chen 2005; Scariano and Calzada 2003; Gadre and Rattihalli 2005; Khoo et al. 2008; Bourke 2008; Castagliola and Khoo 2009). It turned out that the synthetic control chart is superior to the Shewhart $\bar{X}$ (or Hotelling T^2) control chart over the whole range of changes in the process mean (or multivariate process mean vector). For monitoring the multivariate process covariance matrix, Ghute and Shirke propose a synthetic chart combining the $|s|$ control chart with the *CRL* control chart for monitoring covariance matrix (Ghute and Shirke 2008). Machado et al. propose a synthetic chart to monitor the bivariate process covariance matrix with two sample variances (Machado et al. 2009).

The purpose of this paper is to put forwards a strategy to monitor the covariance matrix by the way of using the concept of synthetic control chart that is recently developed. In this paper, a synthetic control chart is developed which combines the control chart based on conditional entropy with the *CRL* control chart. The rest of this paper is organized as follows. In Sect. 10.2, the proposed multivariate synthetic control chart for monitoring the multivariate process covariance matrix is brief introduced. In Sect. 10.3, the operation of the multivariate synthetic control chart is described. In Sect. 10.4, the design of the multivariate synthetic control chart is presented. Finally, conclusions are drawn in the last section.

10.2 The Multivariate Synthetic Control Chart for Monitoring Multivariate Process Covariance Matrix

In this section, a multivariate synthetic control chart is developed to monitor the covariance matrix of a multivariate process that follows multivariate normally distribution. The synthetic control chart is an integration of the control chart based on conditional entropy and the *CRL* chart. A brief review of the two kinds of control chart is introduced and then the integrated control chart is developed.

10.2.1 E-Chart Based on Conditional Entropy

The entropy of a vector X may be viewed as a descriptive quantity which measures the extent to which the multivariate probability density function (PDF) is concentrated on a few points or dispersed over many points. Thus the entropy is a measurement of multivariate process variation, similar as the standard deviation in the single variate process case. For a continuous p-variate random vector X, the entropy is defined as:

$$H(X) = -\int f(X) \ln f(X) dX = E_f[-\ln f(X)] \tag{10.1}$$

where $f(X)$ is the PDF of random vector X. If $X \sim N_p(\mu_0, \Sigma_0)$, then the entropy is given by:

$$H(X) = \frac{1}{2}[p \ln(2\pi e) + \ln |\Sigma_0|] \tag{10.2}$$

where $|\cdot|$ denotes the determinant of matrix. Guerrero-Cusumano (1995) suggests that Eq. (10.2) can be written as:

$$H(X) = \frac{1}{2}\left[p \ln(2\pi e) + \sum_{i=1}^{p} \ln\left(\sigma_0^{(ii)}\right)\right] - T(X) \tag{10.3}$$

Let Σ_0 denote the desired covariance matrix, and $\sigma_0^{(ii)}$, $i = 1, 2, \ldots, p$ denote the ith diagonal element of Σ_0. $T(X)$ is called the mutual information of the vector X.

Let the vectors, $X_1, X_2, \ldots, X_n$ be a random sample which represent measurements of p process characteristics and $n > p$ to ensure that sample covariance matrix has full matrix. We are focus on detecting shifts in Σ. The conditional entropy which measure the difference between the sample and theoretical entropy can be obtained. Then, the charting statistic is given by:

$$E' = \sqrt{\frac{n-1}{2p} \sum_{i=1}^{p} \ln\left(\frac{s_i}{\sigma_0^{ii}}\right)} \tag{10.4}$$

where s_i denotes variance for the ith component of sample covariance matrix. The statistic E' equals to a normalizing constant multiple the conditional entropy (Guerrero-Cusumano 1995) and the mean and variance of E' are given by

$$E[E'] = gp\left[G'\left(\frac{n-1}{2}\right) - \ln\left(\frac{n-1}{2}\right)\right] \tag{10.5}$$

$$Var[E'] = g^2\left[pG''\left(\frac{n-1}{2}\right) + \frac{2}{n-1} trac(Q_0 - I)^2\right] \tag{10.6}$$

where $g = [2(n-1)/p]^{1/2}$ and $G''(\bullet)$ and $G''(\bullet)$ is the first and second derivative of the natural logarithm of the gamma function respectively, $Q_0 = \Sigma_{d_0}^{-1} \Sigma_0 \Sigma_{d_0}^{-1}$ is the correlation matrix, $\Sigma_{d_0} = diag(\sigma_{i0})$, $trac(\bullet)$ denotes the trace of a matrix, I is the $p \times p$ identify matrix.

Let

$$E = \frac{E' - E[E']}{\sqrt{Var[E']}} \tag{10.7}$$

The sample statistics E, which is calculated using independent groups of observations from the process output, are plotted on the control chart against the upper control limit (UCL) and the lower control limit (LCL):

$$UCL = Z_{\alpha}/2 \tag{10.8}$$

$$LCL = -Z_{\alpha}/2 \tag{10.9}$$

z_{α} is the $1-\alpha$ percentage point of $N(0, 1)$ and α is the probability of type I error of this kind of control chart. The ARL values of this kind of chart can be given by:

$$ARL = \frac{1}{p} \tag{10.10}$$

where $p = 1 - \Pr(LCL \leq E \leq UCL|\Sigma)$. The on-target value of p can be obtained by:

$$\begin{aligned} p_0 &= 1 - \Pr(LCL \leq E \leq UCL|\Sigma = \Sigma_0) \\ &= 2\Phi_0(UCL) - 1 \end{aligned} \tag{10.11}$$

where $\Phi_0(Y) = \Pr(x \leq Y)$ denotes the cumulative distribution function of a univariate standard normal distribution.

The off-target value of p is determined as:

$$\begin{aligned} p_1 &= 1 - \Pr(LCL \leq E \leq UCL|\Sigma \neq \Sigma_0) \\ &= 2\Phi_1(UCL) - 1 \end{aligned} \tag{10.12}$$

where $\Phi_1(Y) = \Pr(x \leq Y)$ denotes the cumulative distribution function of a univariate normal distribution with mean e and standard variation 1, where e reflects the mean shift size.

10.2.2 Conforming Run Length Chart

The conforming run length control chart, which was firstly proposed by Bourke (1991), was originally developed for monitoring attribute processes. When all the product of the process is inspected, the conforming run length (CRL) is the total number of the conforming items between two consecutive nonconforming items which include the end nonconforming item. Since the random variable *CRL* follows a geometric distribution with parameter θ, its mean and the cumulative probability function can be obtained (Bourke 1991).

$$\mu_{CRL} = \frac{1}{\theta} \tag{10.13}$$

$$F(CRL) = 1 - (1-\theta)^{CRL}, CRL = 1, 2, \ldots \tag{10.14}$$

If an increase in θ wanted to be detected, only a lower control limit L is required for the *CRL* control chart. The expression of L is given as follows:

$$L = \frac{\ln(1-\alpha_{CRL})}{\ln(1-\theta_0)} \tag{10.15}$$

Where $\alpha_{CRL} = F(L\,|\theta = \theta_0)$ is the type I error of the *CRL* control chart and θ_0 is the fraction nonconforming when the process is in-control, i.e., the *CRL* control chart gives a signal if $CRL < L$.

The average number of *CRL* items, which is needed to detect an increase in θ, of the *CRL* control chart is derived as follows:

$$ARL_{CRL} = \frac{1}{F(L)} = \frac{1}{1-(1-\theta)^L} \tag{10.16}$$

The average number of inspected items, AII, which is required to detect a change in θ of the *CRL* control chart, equals the product of Eqs. (10.13) and (10.16):

$$AII = \mu_{CRL} \times ARL_{CRL} = \frac{1}{\theta} \times \frac{1}{1-(1-\theta)^L} \tag{10.17}$$

10.3 Operation of the Multivariate Synthetic Control Chart for Monitoring Multivariate Process Covariance Matrix

The proposed multivariate synthetic control chart is composed of E-chart and the *CRL* control chart. The proposed control chart attempts to monitor the covariance matrix of multivariate normally distributed processes. The operating procedures of the synthetic control chart are as follows:

(1) Determining the control limits *UCL*(*LCL*) and the sample size n of the E-chart, and the control limit L of the *CRL* control chart. The design procedure of those parameters will be described in detail in Sect. 10.4.
(2) Randomly drawing a group of sample with size n and calculating the E-chart statistics, E, at each inspection point.

(3) When $LCL \leq E \leq UCL$, we can obtain the sample is a conforming item in the *CRL* control chart, and the procedure return to step (2). When $E > UCL$ or $E < LCL$, the sample is regarded as a nonconforming item, and the procedure goes to step (4).
(4) Recording the total number of E-chart items between the current and previous nonconforming sample. And this number is regarded as a *CRL* sample of the *CRL* control chart.
(5) If $L \leq CRL$, we can conclude that the multivariate process is in-control, and the procedure goes back to step (2). Otherwise, we can draw a conclusion that the multivariate process is out-of-control; the operating procedure goes to the next step.
(6) Giving the out-of-control signal. And searching for and eliminating the assignable cause, then, the operating procedure goes back to step (2).

The in-control *ARL*, ARL_0 of the proposed multivariate synthetic chart is given as follows:

$$ARL_0 = \frac{1}{P_0} \times \frac{1}{1 - (1 - P_0)^L} \tag{10.18}$$

Let ARL_1 denote the average number of E samples required for the proposed multivariate synthetic chart to give an out-of-control signal when a change in the multivariate process covariance matrix Σ. Then, ARL_1 can be calculated by replacing θ in Eq. (10.17) with Eq. (10.12):

$$ARL_1 = \frac{1}{P_1} \times \frac{1}{1 - (1 - P_1)^L} \tag{10.19}$$

10.4 Design of the Multivariate Synthetic Control Chart for Monitoring Multivariate Process Covariance Matrix

The design of the proposed multivariate synthetic control chart is based on *ARL*. Similar to other kinds of control chart, the in-control *ARL* should be large enough to keep the false alarm rate at a acceptable level, while the out-of-control *ARL* should be small to detect changes in the multivariate process covariance matrix promptly. To design this multivariate synthetic control chart requires well design the control limits for both E-chart and *CRL* chart such that the out-of-control *ARL* is minimized given that the in-control *ARL* equals a specified value.

The optimal design procedure for the proposed multivariate synthetic control chart is given as follows:

(1) Set up specification n, p, e and in-control *ARL*, ARL_0.
(2) Set up L with initial value 1.
(3) Obtain p_0 by making Eq. (10.18) equal to ARL_0 based on current value of L.

Table 10.1 Different sets of parameters of control charts and corresponding ARL_1 for $e = 2$

L	CL	ARL_1
1	1.9432	3.6605
2	2.0846	2.9985
3	2.1638	2.8052
4	2.2186	2.7428
5	2.2602	2.7338
6	2.2937	2.7502
7	2.3216	2.7795
8	2.3456	2.8155
9	2.3665	2.8546
10	2.3850	2.8950

(4) We can derive the parameter of control limit UCL from this value of p_0, and calculate ARL_1 for Σ from current L and CL by solving Eq. (10.19).
(5) If $L \neq 1$, the design procedure goes to the step (6); If $L = 1$, L is increased from one in step of one and the design procedure goes back to step (3).
(6) If the calculated value of ARL_1 for Σ is bigger than the preceding one, go to the next step; otherwise, L is increased in step of one and go back to step (3).
(7) Choose the current L, CL as the optimal design parameters of the proposed multivariate synthetic control chart.

To illustrate the design of the multivariate synthetic control chart, consider the example that $n = 5$, $p = 3$, $e = 2$ and $ARL_0 = 370$. In fact, as can be seen from Eq. (10.7) that statistic E is the normalized form of statistic E' and the control limits for monitoring statistic E is unrelated with n and p. The sets of parameters of control chart and corresponding ARL_1 for $e = 2$ are given in Table 10.1.

From Table 10.1, the ARL_1 first decreases and then increases as L increases. The ARL_1 reaches the minimum at 2.7338 when the parameters L and CL equal to 5 and 2.2602, respectively. So $L = 5, CL = 2.7338$, are the optimal parameters of this example.

Furthermore, the optimal design parameters and corresponding ARL_1 of the proposed multivariate synthetic control chart for ARL_0 equals 200, 370, 500 and 1,000 with e equals 1, 1.5, 2, 2.5, 3 are presented in Table 10.2.

10.5 Conclusion

In this paper, a new multivariate synthetic control chart for monitoring the covariance matrix of multivariate normal processes based on conditional entropy is suggested. The proposed multivariate synthetic control chart is a combination of the E-chart based on conditional entropy and the *CRL* chart. The proposed multivariate synthetic control chart can be a good alternative to the standard control chart based on conditional entropy for monitoring the covariance matrix of a multivariate process from a multivariate normal distribution.

Table 10.2 Optimal design parameters of the proposed multivariate synthetic control chart

	$ARL_0 = 200$			$ARL_0 = 370$			$ARL_0 = 500$			$ARL_0 = 1,000$		
e	L	CL	ARL_1	L	CL	ARL_1	L	CL	ARL_1	L	CL	ARL_1
1.0	14	2.3234	14.3648	19	2.4945	20.0435	21	2.5667	23.6438	29	2.7409	33.7738
1.5	7	2.1992	4.8176	8	2.3456	6.0521	9	2.4235	6.7846	9	2.5501	8.9524
2.0	4	2.0932	2.3563	5	2.2602	2.7338	5	2.3187	2.9479	6	2.4804	3.5334
2.5	3	2.0366	1.5246	3	2.1638	1.6665	3	2.2238	1.7472	4	2.4089	1.9550
3.0	2	1.9545	1.1998	2	2.0846	1.2603	3	2.2238	1.2937	3	2.3569	1.3757

Acknowledgment This work is supported by the National Natural Science Foundation of China (NSFC) under grant No. 70931002 and China Postdoctoral Science Foundation under grant No. 2011 M500928.

References

Bourke P (1991) Detecting a shift in fraction nonconforming using run-length control charts with 100 % inspection. J Qual Technol 23(3):225–238

Bourke P (2008) Performance comparisons for the synthetic control chart for detecting increases in fraction nonconforming. J Qual Technol 40(4):461–475

Castagliola P, Khoo M (2009) A synthetic scaled weighted variance control chart for monitoring the process mean of skewed populations. Commun Stat Simul Comput 38(8):1659–1674

Gadre M, Rattihalli R (2005) Unit and group-runs chart to identify increases in fraction nonconforming. J Qual Technol 37(3):199–209

Ghute V, Shirke D (2008) A multivariate synthetic control chart for process dispersion. Qual Technol Quant Manag 5(3):271–288

Guerrero-Cusumano J (1995) Testing variability in multivariate quality control: a conditional entropy measure approach. Inform Sci 86(1–3):179–202

Huang H, Chen F (2005) A synthetic control chart for monitoring process dispersion with sample standard deviation. Comput Ind Eng 49(2):221–240

Jackson J (1985) Multivariate quality control. Commun Stat Theory Methods 14(11):2657–2688

Khoo M, Wu Z, Atta A (2008) A synthetic control chart for monitoring the process mean of skewed populations based on the weighted variance method. Int J Reliab Qual Saf Eng 15(3):217–245

Machado M, Costa A, Rahim M (2009) The synthetic control chart based on two sample variances for monitoring the covariance matrix. Qual Reliab Eng Int 25(5):595–606

Scariano S, Calzada M (2003) A note on the lower-sided synthetic chart for exponentials. Qual Eng 15(4):677–680

Wu Z, Spedding T (2000) A synthetic control chart for detecting small shifts in the process mean. J Qual Technol 32(1):32–38

Wu Z, Yeo S, Spedding T (2001) A synthetic control chart for detecting fraction noncon- forming increases. J Qual Technol 33(1):104–111

Yeh A, Huwang L, Wu C (2005) A multivariate EWMA control chart for monitoring process variability with individual observations. IIE Trans 37(11):1023

Zou C, Tsung F (2011) A multivariate sign EWMA control chart. Technometrics 53(1):84–97

Chapter 11
Multi-objective Robust Operation Model for a Supply Chain with Market Demands and Raw Material Price Uncertainty

Li-ping Yu, Li-jun Li, and Xiao-yuan Huang

Abstract In this paper, the operation of a multi-product and multi-period supply chain involving one producer and one supplier with uncertain market demands and raw material price is considered. With variations in the market demands and raw material price described by using an interval uncertainty method, a multi-objective robust optimization model is established using a robust linear programming approach. A numerical example is used to verify the proposed model, and the optimum robust operating strategy is determined for worst case supply chain conditions during an uncertain market demands and raw material price. When robust measures are adopted in the objectives of supply chain coordination and profit was maximized for all participants, the effect of market demands and raw material price uncertainty on objective values decreased significantly.

Keywords Objective programming • Raw material price • Robust optimization • Supply chain operation • Uncertainty

11.1 Introduction

In supply chain, logistics and information flow can form a complex network for many suppliers, manufacturers and distributors to interconnect each other (Lee and Billington 1993). Supply chain itself with uncertainty of the dynamic properties, such as customer needs, the raw material supply, production capacity, transportation time, manufacturing time, cost, quality, payment date (limit payment time), priority, lost information, and fuzzy information and the bullwhip effect, etc. (Lee and Billington 1993; Davis 1993; Arns et al. 2002; Geary et al. 2002; Kouvelis and

L. Yu (✉) • L. Li • X. Huang
School of Business Administration, Northeastern University, Shenyang, China
e-mail: Lpyu@mail.neu.edu.cn

E. Qi et al. (eds.), *The 19th International Conference on Industrial Engineering and Engineering Management*, DOI 10.1007/978-3-642-37270-4_11,

Milner 2002). Changes in the economy will increase the uncertainty of the supply chain operation process and uncertain parameters can through the supply chain network transmission (Van der Vorst and Beulens 2002), increase the difficulty to establish the model for the supply chain, which is much more challenging.

At present, there are numerous researches on uncertain environment of supply chain operation. It stress future supply chain management must respond to the demand uncertainty (Christopher and Towill 2002). It established multi-objective optimization model for the many members supply chain with uncertain product demand and prices (Cheng-Liang and Wen-Cheng 2004). It proposes short life cycle of the products coordinated order decisions with delivery time and demand uncertainty (Kevin Weng and McClurg 2003). It researches the optimal control problem for the supply chain with uncertain requirements (Dimitris and Aurelie 2006).

In this paper, a multi-objective robust optimization model will be established in an uncertain market demands and raw material price for a multi-product and multi-period supply chain involving one producer and one supplier. Using an interval uncertainty method to describe variations in the market demands and raw material price, and using robust linear programming method and the goal programming model to describe the operation of the supply chain, the optimal robust operation strategy will be obtained for worst case supply chain conditions during an uncertain market demands and raw material price.

11.2 Problem Description

In the supply chain involving a producer and a supplier, supplier's products are producer's raw material. In pursuit of the goal the profit maximization conditions, the producer decided to its each phase to the supplier ordering the number of raw materials. The same, to pursue maximum profit as the objective, the supplier decide the phase to the producer delivered quantity. The supplier has plenty of production capacity, but because of the price level to change, the producer's raw material price is uncertain, and consumption market demand is uncertain. In the supply chain operation process, both parties need to face the uncertain raw material price and the market demand to determine the optimal operation of the strategy. In order to guarantee the supply chain coordination, at each specific stage, the both sides of the order and the number of delivery should be equal, and this is the supply chain operation process must first priority, it is the highest aim of supply chain operation. The producer pursuit the maximum profits, and the supplier realize the maximum income based on the supply chain coordination. This three goals priority level is in turn reduce.

Model's symbols and its meaning are as follows:

Indices h denotes product index ($h = 1, \ldots, H$); i denotes producer's raw materials index ($i = 1, \ldots, I$); t denotes time period ($t = 1, \ldots, T$); k denotes supplier's raw materials index ($k = 1, \ldots, K$).

Decision variables s_{ht} denotes the number of sold product h at time t; z_{ht} denotes the number of production product h at time t; z_{ht}^{L} denotes the inventory level of product h at time t; e_{it}^{L} denotes the inventory level of raw materials i at time t; b_{it} denotes the number of purchased raw material i at time t; q_{it} denotes the number of supplier delivery product i at time t; g_{it} denotes the number of supplier production product i at time t; g_{it}^{L} denotes the inventory level of supplier's product i at time t.

Parameters p_{ht} denotes the price of product h at time t; w_{it} denotes the price of raw materials i at time t; c_h^z denotes the variable unit production cost of product h; u_h^z denotes the unit inventory holding cost of product h; u_i^y denotes the unit inventory holding cost of raw materials i; α_h^j denotes the production capacity rate of product h; $J^{\max}$ denotes the maximum capacity of manufacturers; $z_{h0}^{L'}$ denotes the initial inventory level of product h; o_h^z denotes the unit inventory level of product h; $z^{L^{\max}}$ denotes the total inventory ability of product; v_{ih}^e denotes "bill-of-material (BOM)" of the final product h for raw materials i; $e_{i0}^{L\prime}$ denotes the initial inventory of raw materials i; o_i^e denotes the unit raw materials i occupies the inventory; $e^{L^{\max}}$ denotes the total inventory ability of raw materials; r_{kt} denotes the price of raw materials k at time t; c_i^x denotes the unit variable costs for product i; v_{ki}^r denotes "bill-of-material (BOM)" of the product i for raw material k; u_i^x denotes the unit inventory holding cost of product i; α_i^g denotes the production capacity rate of product i; $N^{\max}$ denotes the maximum capacity of supplier available; $g_{i0}^{L\prime}$ denotes the initial inventory level of product i; o_i^g denotes the unit product i occupies the supplier's inventory; $g^{L^{\max}}$ denotes product total inventory ability.

11.3 Robust Operating Model

Based on the basic model in the reference Dimitris and Aurelie (2006), this paper use Dimitris and Melvyn proposed robust linear programming method in another (Dimitris and Melvyn 2004) to establish the robust operating model for a supply chain with uncertain market demands and raw material price.

Assume consumption market demand d_{ht} is uncertain, let $\left[d_{ht} - \widehat{d}_{ht}, d_{ht} + \widehat{d}_{ht}\right]$ be a symmetrical bounded interval for uncertainty market demand, and d_{ht} is consumption market benchmark needs, $\widehat{d}_{ht}$ is the largest amount of market demand changes. At the same time, assume the producers' raw materials' price w_{it} also is uncertain, and let $\left[w_{it} - \widehat{w}_{it}, w_{it} + \widehat{w}_{it}\right]$ be a symmetrical bounded interval for uncertainty raw materials' price, and w_{it} is the manufacturers' benchmark raw material prices, $\widehat{w}_{it}$ is the largest amount of the raw materials' price changes.

Raw material price volatility will directly influence the increase in producer's earnings, and inevitably lead to variation in the producer order quantity. Assume variation between the producer quantity order and raw material prices is the most simple linear corresponding relation, that can be expressed as $b_{it} = b_{it}^0 - \phi_{it} \times \Delta_{it}$. Among them, ϕ_{it} is fixed constants, it denotes variation in the producer order quantity for raw material prices every a unit change, Δ_{it} denotes variation in the raw material prices, b_{it}^{0} denotes producer order quantity for the benchmark raw material prices.

Supply chain operation consider the following three operation objectives

1. In each specific stage, the supply chain operating pursue participate in both sides the operation coordination, that is the supplier delivery is equal to the quantity of manufacturers. By model can be expressed as follows.

$$\begin{aligned} \min \quad & P_T \times \sum\nolimits_{t=1}^{T} \sum\nolimits_{i=1}^{I} \left(d_{it}^{-} + d_{it}^{+}\right) \\ s.t. \quad & -b_{it} + q_{it} + d_{it}^{-} - d_{it}^{+} = 0 \quad \forall i, t \end{aligned} \tag{11.1}$$

P_T denotes priority factor, it is a relatively big constant, d_{it}^{-} and d_{it}^{+} respectively denotes shortage of delivery and excess amount delivered for supplier to producer the raw materials i in t stage.

2. The second objective is to maximize the producer goal profit, the model can be expressed as follows.

$$\begin{aligned} \pi^{P} = \sum\nolimits_{t=1}^{T} \Bigg(& \sum\nolimits_{h=1}^{H} (p_{ht} s_{ht} - c_h^z z_{ht} - u_h^z z_{ht}^L) \\ & - \sum\nolimits_{i=1}^{I} (u_i^y e_{it}^L + w_{it} b_{it}) \Bigg) \rightarrow \max \end{aligned}$$

And the goal programming form can be expressed as follows.

$$\begin{aligned} \min \quad & P_P \times d_P^{-} \\ s.t. \quad & \pi^{P} + d_P^{-} - d_P^{+} = M_P \end{aligned} \tag{11.2}$$

$$\begin{aligned} \pi^{P} + \sum\nolimits_{t=1}^{T} \Bigg(& \sum\nolimits_{h=1}^{H} (-p_{ht} s_{ht} + c_h^z z_{ht} + u_h^z z_{ht}^L) \\ & + \sum\nolimits_{i=1}^{I} (u_i^y e_{it}^L + w_{it} b_{it}) \Bigg) \leq 0 \end{aligned} \tag{11.3}$$

P_P denotes Priority factor, it is a relatively big constant; M_P denotes producer's expected profits, it is a given constant; d_P^{-} and d_P^{+} respectively denotes the insufficient value and the more than value of the target profit.

For producer, its operation is influenced by raw material price uncertainty, thus, there is several of the coefficient uncertainty perturbation in Constraints (11.3), and the number of the coefficient is $T^{*}I$. The coefficient uncertainty perturbation can be expressed as follows.

$$\max \quad \sum\nolimits_{t=1}^{T} \sum\nolimits_{i=1}^{I} \widehat{w}_{it} b'_{it}$$
$$(s.t. \quad -b'_{it} \leq b_{it} \leq b'_{it}, \quad b'_{it} \geq 0 \quad \forall i, t)$$

After join perturbation terms, (11.3) type equivalent to the following (11.4, 11.5, and 11.6) type, b'_{it}, ε_{it} and z all are negative variables.

$$\pi^P + \sum\nolimits_{t=1}^{T} \left(\sum\nolimits_{h=1}^{H} (-p_{ht} s_{ht} + c_h^z z_{ht} + u_h^z z_{ht}^L) + \sum\nolimits_{i=1}^{I} (w_{it} b_{it} + u_i^y y_{it}^L) \right) + (T^* I) x + \sum\nolimits_{t=1}^{T} \sum\nolimits_{i=1}^{I} \varepsilon_{it} \leq 0 \tag{11.4}$$

$$x + \varepsilon_{it} \geq \widehat{w}_{it} b'_{it} \quad \forall i, t \tag{11.5}$$

$$-b'_{it} \leq b_{it} \leq b'_{it} \quad \forall i, t \tag{11.6}$$

And the goal programming form can be expressed as follows.

$$\min \quad P_S \times d_S^-$$
$$s.t. \quad \pi^S + d_S^- - d_S^+ = M_S \tag{11.7}$$

$$\pi^S + \sum\nolimits_t \sum\nolimits_i (-w_{it} q_{it} + c_i^x g_{it} + u_i^x g_{it}^L) + \sum\nolimits_t \sum\nolimits_i \left(\sum\nolimits_k r_{kt} v_{ki}^r \right) g_{it} \leq 0 \tag{11.8}$$

P_S denotes priority factor, it is a relatively big constant; M_S denotes supplier's expected profits, it is a given constant; d_S^- and d_S^+ respectively denotes the insufficient value and the more than value of the target profit.

For supplier its operation is influenced by it provides to the producer's raw material price uncertainty, thus, there are several of the coefficient uncertainty perturbation in constraints (11.8), and the number of the coefficient is $T^* I$. The coefficient uncertainty perturbation can be expressed as follows.

$$\max \quad \sum\nolimits_{t=1}^{T} \sum\nolimits_{i=1}^{I} \widehat{w}_{it} q_{it}$$
$$(s.t. \quad -q'_{it} \leq q_{it} \leq q'_{it}, \quad q'_{it} \geq 0 \quad \forall i, t)$$

After join perturbation terms, (11.8) type equivalent to the following (11.9), (11.10) and (11.11) type, q'_{it}, ξ_{it} and y all are negative variables.

$$\pi^S + \sum_t \sum_i (-w_{it}q_{it} + c_i^x g_{it} + u_i^x g_{it}^L)$$
$$+ \sum_t \sum_i \left(\sum_k r_{kt} v_{ki}^r\right) g_{it}$$
$$+(T^* I)y + \sum_{t=1}^{T} \sum_{i=1}^{I} \xi_{it} \leq 0 \tag{11.9}$$

$$y + \xi_{it} \geq \widehat{w}_{it} q'_{it} \quad \forall i, t \tag{11.10}$$

$$-q'_{it} \leq q_{it} \leq q'_{it} \quad \forall i, t \tag{11.11}$$

The supply chain operating first has to reach the supply chain channels coordination, ensure that the producer can meet the supplier order quantity. At the same time, we have to consider the supply chain is pulled by consumer market demand, supplier operational targets depends on producer order quantity, and supplier operation objectives in turn affect producer order quantity. We can think in supply chain operation, producer operation objectives prior to supplier operation objectives. The supply chain operating can be described as the following three goal programming model, the three priority factor P_T P_P and P_S meet $P_T >> P_P >> P_S$. Thus, the objective function of multi-objective robust operation model for a supply chain can be expressed as follows.

$$\min \quad P_T \times \sum_t \sum_i (d_{it}^- + d_{it}^+) + P_P \times d_P^- + P_S \times d_S^-$$

The constraint conditions of model except (11.1, 11.2, 11.3, 11.4, 11.5, 11.6, 11.7, 11.8, 11.9, 11.10, and 11.11) besides, still bound by the following:

Producer production capacity constraint at each stage:

$$\sum_h \alpha_h^j z_{ht} \leq J^{\max} \quad \forall j, t \tag{11.12}$$

The producer end product inventory at each stage:

$$z_{ht}^L = z_{h,t-1}^L + z_{ht} - s_{ht} \quad \forall h, t \tag{11.13}$$

$$z_{h0}^L = z_{h0}^{L'} \quad \forall h\, z \tag{11.14}$$

$$\sum_h o_h^z z_{ht}^L \leq z^{L^{\max}} \quad \forall h, t \tag{11.15}$$

Producer inventory of raw materials:

$$e_{it}^L = e_{i,t-1}^L + b_{it} - \sum_h v_{ih}^e z_{ht} \quad \forall i, t \tag{11.16}$$

$$e_{i0}^L = e_{i0}^{L'} \quad \forall i \tag{11.17}$$

$$\sum_i o_i^e e_{it}^L \leq e^{L^{\max}} \quad \forall i, t \tag{11.18}$$

Producer actual order restriction:

Under the raw material price uncertainty disturbance, consider the most conservative of the order, we assume variations of raw material price meet $\Delta_{it} = \widehat{w}_{it}$,so in the worst case, the producer actual quantity meet:

$$b_{it} \leq b_{it}^{0} + \phi_{it} \times \widehat{w}_{it} \quad \forall i, t \tag{11.19}$$

Producer actual sales constraint:

$$d_{ht} - \widehat{d}_{ht} \leq s_{ht} \leq d_{ht} + \widehat{d}_{ht} \tag{11.20}$$

Supplier production capacity constraint:

$$\sum_{i} \alpha_i^g g_{it} \leq N^{\max} \quad \forall i, t \tag{11.21}$$

The supplier inventory:

$$g_{it}^{L} = g_{i,t-1}^{L} + g_{it} - q_{it} \quad \forall i, t \tag{11.22}$$

$$g_{i0}^{L} = g_{i0}^{L'} \quad \forall i \tag{11.23}$$

$$\sum_{i} o_i^g g_{it}^{L} \leq g^{L^{\max}} \quad \forall i, t \tag{11.24}$$

Above all of the variables are bound take not a negative value.

Because the model of target function and all the constraint conditions are linear, so at least we can use linear programming method to get the optimal solution of the model.

11.4 Numerical Example

Using a numerical example, we will illustrate the robustness of the proposed model. We assume that there are two types of products ($H = 2$), one type of producer production materials ($I = 1$), two types of raw materials ($K = 2$), and four stages ($T = 4$). Parameter values are as follows:

$$J^{\max} = 3450, N^{\max} = 7500, z^{L^{\max}} = 600$$

$$e^{L^{\max}} = 600, g^{L^{\max}} = 600$$

$$z_{h0}^{L'} = 0, e_{i0}^{L'} = 0, g_{i0}^{L'} = 0 \quad \forall h, i$$

$$c_i^x = 20, c_h^z = 25, w_t = 50, \quad \forall h, i, t$$

Table 11.1 Standard prices of final products and supplier's raw materials

Price / Stage	Raw materials		Products	
	Raw materials 1	Raw materials 2	Product 1	Product 2
1	10	20	260	255
2	20	30	270	275
3	20	30	270	270
4	10	20	260	265

Table 11.2 Standard demand final products

	1	2	3	4
Product 1	2,400	2,400	2,400	2,400
Product 2	2,550	2,700	2,700	2,550

$$u_i^x = 2, u_i^y = 4, u_h^z = 5, \alpha_i^g = 2, \alpha_h^j = 1 \quad \forall h, i$$
$$v_{11}^r = 0.6, v_{21}^r = 0.4, v_{11}^y = 1, v_{12}^y = 1$$
$$o_i^g = 1, o_i^e = 1, o_h^z = 1 \quad \forall h, i$$

Priority factor P_T, P_P and P_S Respectively is 10^6, 1,000, 1,000, producer and supplier's target profit M_P and M_S all take 4×10^7. The final products and suppliers' raw materials unit price at each stage is shown in Table 11.1.

Each phase of the final product standard demand as shown in Table 11.2, and the demand variations for range is 15 and 80 % respectively.

According to the proposed model, certainty purchase price (fluctuation of price is 0) and fluctuations of price 1 unit two kinds of situations are calculated. The operating strategy as raw materials prices certain and as raw materials prices uncertain are shown in Tables 11.3 and 11.4.

It is easily noticed that as the fluctuation of raw materials prices are 1 %, suppliers and producer profits than the certain purchase price state changes 1.39 and 2.56 % respectively, and both sides of the supply chain operation strategy does not produce change. The price of a product are the most big wave momentum for 2, 3, ..., 19 total 18 kinds of situations of the calculation results show that, when the fluctuation of product price no more than a range, the supply chain operating strategy remains the same, these cases the model optimal solutions is close to the raw materials prices certainty optimal solution, both order and delivery quantity reduce does not exceed 22 %, and keep equal; suppliers profits changes no more than 24 %, producer profits reduces no more than 21 %. The raw materials price uncertainty only affect supply chain and its members the operation performance. As the results shown, the model is robust, it can effectively ensure the supply chain operation robustness.

Table 11.3 The operating strategy as raw materials prices certainty

	Producer (profit: 2,651,475)								Supplier (profits:141,000)		
	Final products production		Final products sales		Final products inventory		Raw material inventory	Raw material order	Delivery of products	Products production	Products inventory
Stage	Product 1	Product 2	Product 1	Product 2	Product 1	Product 2					
1	2,685	765	2,085	765	600	0	0	3,450	3,450	3,450	0
2	120	3,330	720	3,330	0	0	0	3,450	3,450	3,450	0
3	720	2,730	720	2,730	0	0	0	3,450	3,450	3,450	0
4	720	2,730	720	2,730	0	0	0	3,450	3,450	3,750	300

Table 11.4 The operating strategy as the fluctuation of raw materials prices are 1 %

	Producer (profit: 2,614,615)								Supplier (profits:144,620)		
	Final products production		Final products sales		Final products inventory		Raw material inventory	Raw material order	Delivery of products	Products production	Products inventory
Stage	Product 1	Product 2	Product 1	Product 2	Product 1	Product 2					
1	1,965	1,365	1,965	765	0	600	80	3,410	3,410	3,410	0
2	720	2,730	720	3,330	0	0	40	3,410	3,410	3,410	0
3	720	2,730	720	2,730	0	0	0	3,410	3,410	3,410	0
4	720	2,690	720	2,690	0	0	0	3,410	3,410	3,750	340

11.5 Conclusion

Uncertainty throughout the whole process of supply chain operation, and how to study this uncertainty influence on the supply chain operation is the supply chain management inevitable problems. In this paper, we developed a robust optimization model to optimize the supply planning function. Using interval uncertainty describe raw material prices and market demand fluctuations and a robust optimization method, the uncertain raw materials price to producer and supplier target profit nonlinear disturbances change into a group of linear constraints, and using multiple goals to describe the supply chain and its members operation performance, we get the supply chain operation multi-objective robust operating model. The result of a numerical example shows that, a model of optimal solution can be found in raw material price changes to the worst situation, an effective method is provided to solve the supply chain operation uncertainty, the supply chain operation robustness is ensured.

References

Arns M, Fischer M, Kemper P, Tepper C (2002) Supply chain modeling and its analytical evaluation. J Oper Res Soc 53(4):885–894

Cheng-Liang C, Wen-Cheng L (2004) Multi-objective optimization of multi-echelon supply chain networks with uncertain product demands and prices. Comput Chem Eng 28(10):1131–1144

Christopher M, Towill DR (2002) Developing market specific supply chain strategies. Int J Phys Distrib Logist Manage 13(1):1–14

Davis T (1993) Effective supply chain management. Sloan Manage Rev 34(1):35–46

Dimitris B, Aurelie T (2006) A robust optimization approach to inventory theory. Oper Res 54(1):150–168

Dimitris B, Melvyn S (2004) The price of robustness. Oper Res 52(1):35–53

Geary S, Childerhouse P, Towill D (2002) Uncertainty and the seamless supply chain. Supply Chain Manage Rev 6(1):52–60

Kevin Weng Z, McClurg T (2003) Coordinated ordering decisions for short life cycle products with uncertainty in delivery time and demand. Eur J Oper Res 151(1):12–24

Kouvelis P, Milner J (2002) Supply chain capacity and outsourcing decision: the dynamic interplay of demand and supply uncertainty. IIE Trans 34(3):717–728

Lee H, Billington C (1993) Material management in decentralized supply chains. Oper Res 41(4):835–847

Van der Vorst J, Beulens A (2002) Identifying sources of uncertainty to generate supply chain redesign and strategies. Int J Phys Distrib Logist Manage 32(2):409–430

Chapter 12
The Application of Six Sigma to Enhance Product and Service Quality in Internet Marketing

Tai-Chang Hsia, Shui-Chuan Chen, and Kun-Sheng Chen

Abstract The Six Sigma quality control systems may be used to enhance consumer satisfaction with products and services on the Internet. This study applies the methodology in five steps: design, measure, analyze, improve, and control. First, we conducted a consumer acceptance survey, then we constructed a product quality matrix to determine whether any product or service is considered defective. Using a cause-and-effect diagram, we identified the major causal categories of consumer dissatisfaction, prioritized product improvement targets, implemented a holistic analysis of the defective service, and generated specific strategies. We also initiated concrete improvements to the defective products and services, established standard operating procedures and a knowledge management system to standardize operations and maintain long-term control.

Keywords Internet marketing • Product quality • Service quality • Standard operating procedure

12.1 Introduction

The rapid development of computer technology and the internet, which involves no-limit on time or distance and low-cost home delivery, has made the internet marketing very popular among many customers. However, customers tend to be more concerned about the product and service quality associated with internet

T.-C. Hsia (✉)
Department of Industrial Engineering and Management, Chienkuo Technology University, Changhua, Taiwan
e-mail: tchsia@ctu.edu.tw

S.-C. Chen • K.-S. Chen
Department of Industrial Engineering and Management, National Chin-Yi Institute of Technology, Taichung, Taiwan

E. Qi et al. (eds.), *The 19th International Conference on Industrial Engineering and Engineering Management*, DOI 10.1007/978-3-642-37270-4_12,

retailers than those obtained from traditional marketers. Therefore, product and service quality represent two fundamental aspects that may help enhance the effectiveness of internet marketing.

Previous studies of product and service quality have developed performance measures for both product (Chen et al. 2002; Pearn and Chen 1997) and service (Hung et al. 2003; Parasuraman et al. 1985, 1991; Rosen et al. 2003) quality. However, these studies each focus on a single aspect—namely, the quality of either the product or the service—to measure customer satisfaction with the market offering. Furthermore, as they failed to integrate manufacturing and post-sale operations together, these studies cannot identify the key elements that affect customers' acceptance of overall product/service qualities. In response to these gaps in previous literatures, we integrate product and service quality using Pyzdek, Michael and Thomas' (Pyzdek 2001; Michael 2002; Thomas 2003) Six Sigma DMAIC (Define, measurement, analysis, improvement, and control) quality control system to assess customer satisfaction associated with internet marketing and provide suggestions to enhance consumer satisfaction so as to improve business performance.

12.2 Defining Importance and Satisfaction for Quality

Consumers usually consider their potential satisfaction with the product and its importance to them when they go shopping. From the perspective of product quality, the product will be acceptable if the consumers are satisfied with the product. With regard to service quality, the service is considered good enough if the customer believes that his or her satisfaction with the service will be higher than its importance (Chen et al. 2005). Therefore, when the customer thinks the merchandise has acceptable product quality and service quality, he or she may likely make a purchase.

To gather data for our study, we conduct a survey of consumers who purchased products from an Internet marketing company and had used those products for a specified period of time. Following Lambert and Sharma (1990) and Parasuraman et al. (1985, 1991), we adapt a questionnaire to investigate four aspects of internet marketing: product quality importance, satisfaction, service quality importance and satisfaction. The following four indices thus constitute our methodological approach:

$$\text{Importance of product quality } (\mathrm{I_{PI}}) = \frac{\mu_{PI} - \min}{R}, \tag{12.1}$$

$$\text{Satisfaction index of product quality } (\mathrm{I_{PS}}) = \frac{\mu_{PS} - \min}{R}, \tag{12.2}$$

$$\text{Importance of service quality } (\mathrm{I_{SI}}) = \frac{\mu_{SI} - \min}{R}, \text{ and} \tag{12.3}$$

$$\text{Satisfaction index of service quality } (\mathrm{I_{SS}}) = \frac{\mu_{SS} - \min}{R} \tag{12.4}$$

To understand product and service quality evaluations from both consumers' and manufacturers' viewpoints, a performance evaluation matrix offers a good method. By adapting the performance evaluation matrix provided by Hung et al. (2003) and Lambert and Sharma (1990), we construct a revised performance evaluation matrix for this study (Hsia et al. 2009).

12.3 Measurement Methods

Using the database of our anonymous internet marketing company, we gather information about computer, communication and consumer electrics (3C) items sold during the 12 months prior to the study and then choose the first two items from seven randomly selected categories for 14 market offerings in total (Table 12.1).

To investigate the product and service quality of these items, we conduct a survey that relies on five-point Likert-type scales. The online questionnaire, sent via e-mail to consumers who had bought products from the company, consists of two stages. First, consumers indicated the product importance and their satisfaction with the products they had bought. Second, they noted the service importance and their satisfaction on the basis of 15 items pertaining to presales, during sales, post-sales, and overall service associated with the products they bought. We mailed 50 questionnaires for each of the 14 selected products, for a total of 700, and received 302 returned questionnaires, of which 286 were valid, for an effective return rate of 41 %.

According to Gay and Airasian (1992), a Cronbach's α coefficient of .80 or greater is acceptable, and a coefficient of .90 or greater indicates very high reliability. However, researchers such as De Vellis (1991) and Nunnally and Ira (1978) consider a Cronbach's α coefficient of .70 the minimum reliability threshold. On the basis of SPSS software reliability analyses, our results provide an overall Cronbach's α of .92.

Table 12.1 Product quality importance and satisfaction indices

Item	Product item	μ_{PI}	μ_{PS}	I_{PI} (Y_i)	I_{PS} (X_i)
1	Kodak CX7430 4.0 megapixel digital camera	4.01	4.18	0.75	0.80
2	CASIO EX257 4.0 megapixel digital camera	3.92	3.78	0.73	0.70
3	Panasonic min cassette recorder RQ-L31	3.17	3.55	0.54	0.64
4	Panasonic CD\MP 3player SL-CT520	3.19	3.71	0.55	0.68
5	Panasonic KX-FT901 telephone/Facsimile	4.76	4.88	0.94	0.97
6	SANYO B93-F012-T telephone/Facsimile	4.57	4.38	0.89	0.85
7	SONY DCR-PC108 digital video	3.89	3.37	0.72	0.59
8	SONY DCR-HC30 digital video	3.47	3.04	0.62	0.51
9	TECD TD 2024VK DVD player	3.34	3.21	0.69	0.35
10	Panasonic DVDS30 DVD player	2.99	2.78	0.50	0.45
11	View sonic VA91 2LCD monitor	3.24	3.77	0.56	0.69
12	CMV CT-720D LCD monitor	3.34	3.52	0.59	0.63
13	Tyan Tiger MPX S2466-4M motherboard	4.41	2.08	0.85	0.27
14	MSI K8 NSLI motherboard	3.57	3.12	0.64	0.53

Table 12.2 Service quality importance and satisfaction indices

Item	Status	Product item	μ_{SI}	μ_{SS}	I_{SI} (Y_i)	I_{SS} (X_i)
1	Presales	Access to log onto this company through the Internet	3.07	3.02	0.52	0.51
2		Design of the Web site is user friendly	3.14	3.10	0.54	0.50
3		Level of variety of the products on the Internet	3.18	3.14	0.55	0.54
4		Design of hyperlinks	3.01	3.08	0.50	0.52
5	During sales	Ease of choosing different products	3.20	3.09	0.55	0.52
6		Providing customized service	2.08	4.41	0.27	0.85
7		Security of electronic transaction	3.08	3.22	0.52	0.56
8		Cooperation of logistic circulation	3.24	3.12	0.56	0.53
9	Postsales	Delivery on schedule and safe	3.17	3.02	0.54	0.51
10		Product delivery and customer confirmation	2.02	2.98	0.51	0.50
11		Warranty, service, and repair of product	3.91	1.78	0.73	0.20
12		Reasonableness of transaction cost for purchase and repair	3.11	3.28	0.52	0.57
13	Overall sales	Acceptability of product repair time	3.08	3.03	0.52	0.51
14		Overall quality and flow of electronic commerce	4.11	2.32	0.78	0.33
15		Overall service performance	2.90	2.97	0.48	0.49

Next, we calculated the means of product quality importance, product quality satisfaction, service quality importance, and service quality satisfaction on the basis of the survey results, which we then use to determine the values of the four indices, as we demonstrate in Eqs. 12.1, 12.2, 12.3 and 12.4 (Tables 12.1 and 12.2). That is, we constructed a product quality performance matrix on the basis of the values of I_{PI} and I_{PS} in Table 12.1 (Fig. 12.1), as well as a service quality performance matrix on the basis of I_{SI} and I_{SS} in Table 12.2 (Fig. 12.2). The performance matrices in Figs. 12.1 and 12.2 thus offer a means to investigate the quality of different product and service items.

In the product quality performance matrix in Fig. 12.1, the plots of the importance and satisfaction indices show that of the 14 items, Items 13 and 9 are located outside the target zone (Zone A) in the insufficient resources zone (Zone I). These two items thus represent defective products. In Fig 12.3, we also demonstrate that the product quality of most products sold by the focal Internet marketing company is acceptable. Similarly, in the service quality performance matrix in Fig. 12.2, the plots of the importance and satisfaction indices show that three of the 15 service items are not located in target zone; Items 11 and 14 appear in the insufficient resources zone, and Item six is in the overloaded resources zone (Zone D).

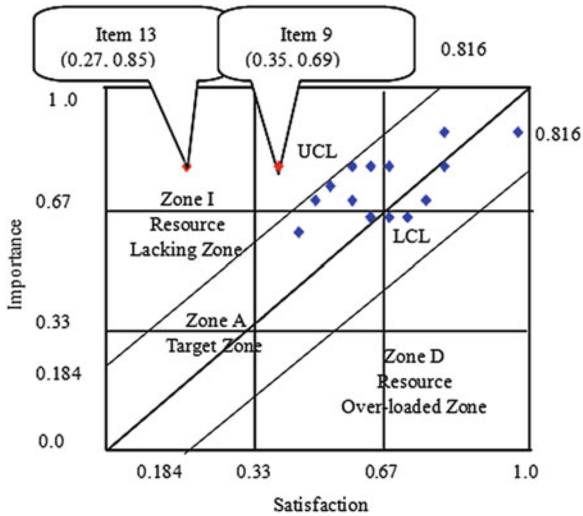


Fig. 12.1 Product quality performance matrix

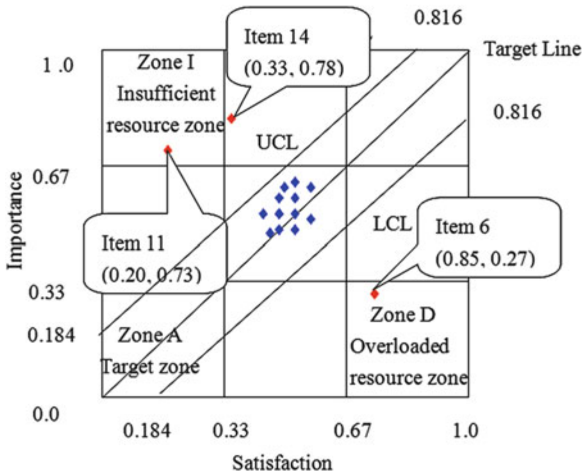


Fig. 12.2 Service quality performance matrix

12.4 Analysis and Improvement of Defective Products and Services

12.4.1 Defective Product

According to our previous prioritization, the first defective product item to be examined is Item 13, the Tyan Tiger MPX S2466-4M motherboard, followed by Item 9, the TECD TD2024VK DVD player. Participants indicated that the importance of these two products was higher than their satisfaction with them, so their producers must enhance product quality satisfaction. We use the Tyan Tiger motherboard as an example in our following discussion.

Fig. 12.3 Performance matrix of four major causal categories before improvement

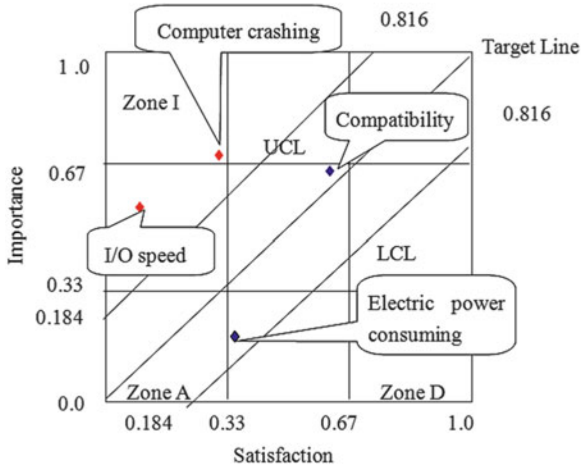


Table 12.3 Importance and satisfaction indices of four major causal categories before improvement

	Major causal category	Importance index (I_{PI})	Satisfaction index (I_{PS})	Remarks
Tyan tiger motherboard	I/O speed	0.61	0.12	In Zone I, critical causal category
	Computer crashing	0.76	0.31	In Zone I, critical causal category
	Electric power consuming	0.30	0.35	In Zone A
	Compatibility	0.67	0.60	In Zone A

When consumers perceive a product as reliable, efficient, and user friendly, they evince higher satisfaction. To analyze the Tyan Tiger motherboard, we first invited engineers from the motherboard manufacturer to provide ideas to enhance its reliability, efficiency, and ease of use in response to the four major causal categories of dissatisfaction: frequent computer crashing, unacceptable electric power consumption, low I/O speed, and insufficient compatibility. Then, within these four major causal factors, the engineers identified minor causes and arranged them in a cause-and-effect diagram. Finally, maintaining the focus on these four major causal categories, we asked designers to take a consumer point of view to evaluate the quality importance and satisfaction of the four major causal categories at the level of the production line to determine their status before the improvement (Table 12.3). With these four importance and satisfaction indices, we constructed a performance matrix of the major causal categories before the improvement (Fig. 12.3).

12.4.2 Defective Service

To analyze the defective service, we consider three items: 11, 14 (which need more resources), and 6 (the resources devoted to which should be reduced). In the following section, we analyze all three of these defective service items.

There are two ways to address a defective service: with a focus on each service as a unit or on the defective service as a whole. Determining which to use depends on whether each defective service item can be measured or is related to overall service quality. In this study, as we used 15 items to measure the service of the products consumers bought, the three defective service items should be analyzed together. We developed a cause-and-effect diagram in which the three defective service items served as the three major causal categories. We again identified minor causes within these three major causal categories, namely, the specific service strategies that could enhance overall service quality. The minor causes included ten items: (A) conducting data mining about customer's purchases of products, (B) providing active customer consultations, (C) providing active customer advice about purchasing products, (D) offering a 30-day full refund warranty, (E) offering home service for purchases and repair, (F) providing a 1-year repair warranty, (G) making prices of products and service competitive, (H) allowing customers to choose product delivery times and place, (I) leaving servers on and available 24 h a day year round, and (J) providing weekly updated Web pages and updated product catalogs.

12.5 Improving Defective Products and Services

12.5.1 Defective Products

Our cause-and-effect diagram pertaining to the Tyan Tiger motherboard indicates four major causal categories and several minor causes, which we used to define two critical causal categories: I/O speed and computer crashes. Engineers responsible for the motherboard designed improvements based on the three minor I/O causes: (1) insufficient flash memory capacity, (2) overly complex logic circuitry, and (3) plethoric function design. At the same time, they addressed the five minor causes of computer crashes: (1) insufficient capacitance refractory ability, (2) inadequate multi-power project design, (3) inadequate electronic insulation, (4) inadequate graphing function, and (5) inadequate cooling system. They improved the product by decreasing the LSL and standard deviation (δ) of the process capabilities of the two critical causal categories. In turn, we considered the device parameters, tolerance, and best manufacturing conditions as criteria for enhancing process capability. After a period of manufacturing process stability, we collected our data again and determined the I/O speed to be 1.25, according to Eq. 12.4, which corresponds to a 5σ quality level. The process capability increased to 1.96, a 7σ quality level as show in Fig. 12.4.

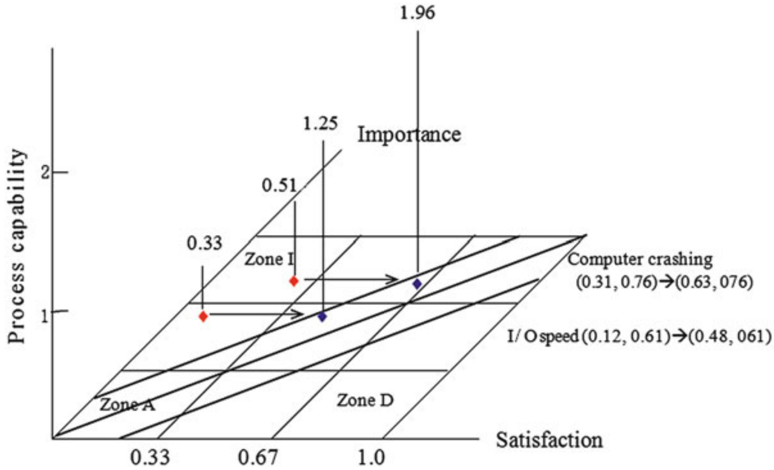


Fig. 12.4 Process capability performance matrix of two critical causal categories before and after improvement

12.5.2 Defective Service

Our holistic analysis with a cause-and-effect diagram for the three defective service items identifies ten specific service strategies (A–J) that could enhance overall service quality. These ten strategies are interdependent with the three defective service items and overall service quality; therefore, we undertake quality function development and use the relationship between the ten strategies and the three defective service items to determine critical service strategies. The quality function development process reveals three large positive values and two large negative values, as we show in Table 12.4.

After we identified these critical service strategies, we required the internet marketing company to implement improvements in them. The company announced the improvements on its Web site and through e-mails to customers who participated in our survey. After the improvement, 286 customers who provided effective return questionnaires again applied the same criteria to measure their satisfaction with item 6, providing customized service; item 11, warranty, service, and repair of the product; and item 14, overall quality and flow of electronic commerce. The resulting service quality satisfaction index shows that the value of Item 6 decreased from .85 to .38 and that of item 11 from .20 to .44, whereas the value of Item 14 increased from .33 to .69. We used the importance and satisfaction indices of these items after their improvements to compose the performance matrix of defective service quality after improvements. As shown in Fig. 12.5, the three defective service items improved and have reached the ideal target zones.

Table 12.4 Critical service strategies for enhancing overall service quality through quality function development

Defective service items \ Specific service strategies (Weight w_i: Very strong 4.0, Strong 3.0, Fair 2.0, Weak 1.0, Not relevant 0.0)	Review/examination on index δ_i	A conducting data mining on customer's purchasing products	B commissioners actively providing customers consultation	C commissioners actively providing customers advice for purchasing products	D 30-day warranty of full-refund with customer dissatisfaction	E home service for purchase and repair	F one-year warranty of prompt repair	G competitive prices of products and service	H customers choosing product-delivery time and place	I servers never shut down 24 h a day year round	J weekly updated web pages with updated products catalogues
6. Providing customized service	−.58	4	4	4	2	1	2	0	2	3	1
11. Warranty, service, and repair of the product	.53	2	1	1	4	4	4	4	3	3	3
14. Overall quality and flow of electronic commerce	.45	3	2	1	3	4	2	4	4	4	4
Total weight (TW_i)		.09	−.9	−1.3	2.31	3.34	1.86	3.92	2.23	1.65	2.81
Critical service strategies			−2	−1		2		1			3

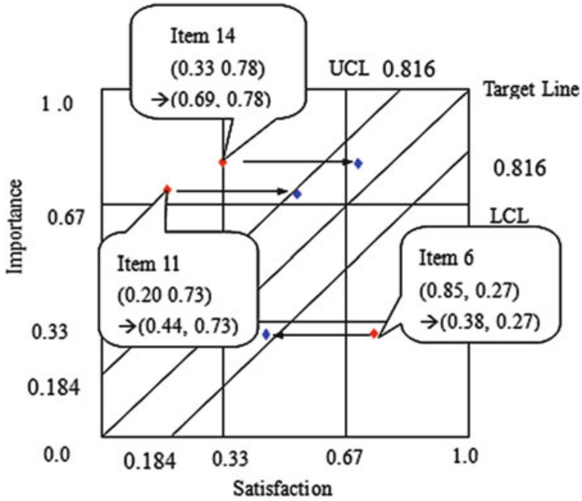


Fig. 12.5 Service quality performance matrix of after improvement

12.6 Control of Product and Service Quality

After corroborating the effective improvements in product and service quality, we moved to the five stages of DMAIC, which entail controlling defective products and services so that they remain within the ideal zones. We considered the concept of control from both narrow and broad perspectives.

12.6.1 Narrow Perspective

The narrow perspective of control focuses on activities. Specifically, such control entails corroborating the enhanced process capability of the critical causal categories of the defective product, ensuring satisfaction matches importance within the target zones, implementing critical service strategies, and finally confirming consumer acceptance of these levels of product and service quality. After control has been implemented, standard operating procedures and a knowledge management system may be established, with a particular focus on the R&D, manufacturing, marketing, and maintenance departments to ensure the effectiveness of long-term controls.

12.6.2 Broad Perspective

The broad perspective of control integrates all the elements of DMAIC as a form of control itself and focuses on appropriate controls of each DMAIC step to enhance the company's product and service quality.

12.7 Conclusion

Internet marketing represents a growing trend whose success is critically determined by product and service quality. Most previous researches on customer satisfaction in the domain of internet marketing analyzes and attempts to improve customer satisfaction through a single aspect, whereas we used Six Sigma DMAIC approach to establish a series of systems that can enhance overall consumer satisfaction with product and service quality.

With increasing technology capabilities and expanded global networks, many companies have developed complex logistic management systems that depend more and more on the internet to locate and purchase needed products and services. Thus, a high-quality internet marketing system can enhance a company's competitiveness in international markets. This study provides a procedure to measure, analyze, improve, and control product and service quality through a step-by-step approach and posits a practical method by which firms may holistically consider R&D, manufacturing, marketing, and maintenance to enhance their quality and customer satisfaction. In turn, their efforts will likely improve consumers' inclination to purchase, as well as their own business performance.

References

Chen KS, Chen SC, Li RK (2002) Process quality analysis of capacitors. Int J Adv Manuf Technol 19:623–628

Chen SC, Chen KS, Hsia TC (2005) Promoting customer satisfactions by applying six sigma example of automobile industry process. Qual Manage J 12(4):21–33

De Vellis RF (1991) Scale development theory and applications. Sage Publications, London

Gay P, Airasian W (1992) Educational research competencies for analysis and application. Macmillan, New York

Hsia TC, Chen SC, Chen KS (2009) Enhancement of service quality in internet-marketing through application of the six sigma process. J Chin Inst Ind Eng 26(1):11–21

Hung YH, Huang ML, Chen KS (2003) Service quality evaluation by service quality performance matrix. Total Qual Manage 14(1):79–89

Lambert M, Sharma A (1990) A customer-based competitive analysis for logistics decisions. Int J Phys Distrib Logist Manage 20(1):23

Michael LG (2002) Lean six sigma. McGraw-Hill, New York

Nunnally JC, Ira B (1978) Psychometric theory, vol 2. McGraw-Hill, New York

Parasuraman A, Valarie Z, Leonard LB (1985) A conceptual model of service quality and its implications for future research. J Market 49(Fall):41–50

Parasuraman A, Valarie Z, Leonard LB (1991) Understanding customer expectation of service. Sloan Manage Rev 33:1–28

Pearn WL, Chen KS (1997) Multi-process performance analysis: a case study. Qual Eng 10(1):1–8

Pyzdek T (2001) The six sigma handbook. McGraw-Hill, New York

Rosen LD, Karwan K, Scribner L (2003) Service quality measurement and the disconfirmation model: taking care in interpretation. Total Qual Manage 14(1):3–14

Thomas P (2003) The six sigma project planner. McGraw-Hill, New York

Chapter 13
Production Process Improvement Using the Six Sigma DMAIC Methodology: A Case Study of a Laser Computer Mouse Production Process

S. Pimsakul, N. Somsuk, W. Junboon, and T. Laosirihongthong

Abstract This paper aims to improve a production process of a laser computer mouse by using the Six Sigma DMAIC Methodology. This study focuses on the functional test procedure because of its lowest yield. The regression analysis and two-level factorial design of experiments is employed in order to determine the optimal conditions of parameters. By operating under these resulting conditions, yield of the functional test procedure increases from 96.2 to 98.6 %.

Keywords Design of experiments • Production process improvement • Six Sigma DMAIC methodology

13.1 Introduction

With the advent of a laser computer mouse and the increasingly accurate technology behind it, the laser computer mouse based on vertical-cavity surface-emitting laser (VCSEL) chips needs a high overall yield, including epitaxy, processing, dicing, mounting, and testing (Jäger and Riedl 2011).

Production process should be beneficial to both quality control and yield enhancement in order to be competitive for manufacturers (Hung et al. 2009).

S. Pimsakul (✉)
Department of Industrial Engineering, Faculty of Engineering, King Mongkut's Institute of Technology Ladkrabang, Bangkok, Thailand
e-mail: pimsakuls@yahoo.com

N. Somsuk
Department of Industrial Engineering, Faculty of Engineering, Eastern Asia University, Pathumthani, Thailand

W. Junboon • T. Laosirihongthong
Department of Industrial Engineering, Faculty of Engineering, Thammasat University, Pathumthani, Thailand

E. Qi et al. (eds.), *The 19th International Conference on Industrial Engineering and Engineering Management*, DOI 10.1007/978-3-642-37270-4_13,

As you know Six Sigma, a quality improvement method, is the most well known quality paradigm for zero-defect quality excellence in the corporate world today (Bhote 2003; Shanmugaraja et al. 2011). The DMAIC method through five steps – including Define, Measure, Analyze, Improve, and Control – is the prescribed improvement process for the Six Sigma methodology that focuses on understanding root causes of defects prior to applying solutions (Leaphart et al. 2012; Tague 2004; Pyzdek and Keller 2009).

In this study, a case study of the laser computer mouse production process in a computer parts manufacturing plant located in Thailand, its functional test procedure has the lowest yield among all procedures, in which the value of the loss cost is about 4.8 million baht per year. Therefore, the primary purpose of this case study is to improve the production process by using the Six Sigma DMAIC Methodology.

The remainder of the paper is organized as follows. Section 13.2 is literature review on the Six Sigma DMAIC approach. The methodology in Sect. 13.3 describes the application of Six Sigma DMAIC process improvement tools. The results and discussions are presented in Sect. 13.4. Finally, Sect. 13.5 concludes the improvements in this case study.

13.2 Literature Review

Six Sigma is a process improvement method by reducing variation and eliminating defects, target to achieve zero defect (Bhote 2003). The implementation of Six Sigma processes is designed to maximize yield and performance and minimize process variation, which leads to reduction in defects and increase in profits (Pickrell et al. 2005; Kumar et al. 2008). In this study, the DMAIC methodology of Six Sigma is applied. In general, Six Sigma with its DMAIC roadmap focuses on improving an existing process using the five phases: Define Measure, Analyze, Improve, and Control.

According to the references Bhote (2003), Pyzdek and Keller (2009), and Kumar et al. (2008), each phase of the DMAIC methodology can be summarized as follows. In the "Define" phase, the problem and/or process that needs to be solved and critical-to-quality characteristics (CTQs) are identified. In the "Measure" phase, it must be determined how to measure the process and its capability or performance, and then they are measured. In the "Analyze" phase, the most likely causes of defects are determined by identifying the potential root causes for the problem and then confirming the actual root causes with historical data. In the "Improve" phase, it focuses on improving the process by addressing the root causes of problem, modifying the process to stay within an acceptable range, and eliminating the defects. Lastly, "Control" phase focuses on ensuring that gains from the improve phase are maintained by determining how to maintain the improvements.

13.3 Methodology

In this study, the Six Sigma DMAIC methodology is applied to identify root causes of existing defects and to determine optimal solutions providing significant business impact as well as improving an existing process in five phases: Define, Measure, Analyze, Improve, and Control.

13.3.1 Define Phase

Initially, the case study adopts Thailand Quality Award (TQA) criteria (Baldridge-based) as the Six-Sigma project selection criteria. Hence, the project selection criteria of this case study are as follows:

- Consistency with the company's quality policy,
- Expected monetary value of the project (or significant business impact),
- Possible problem solving methods within a reasonable project budget, and
- A project relating to responsibility of project team members.

During the define phase, various tools – including Pareto, Cause and Effect, and SIPOC diagram – are used. The SIPOC diagram is used to identify all relevant elements of the process improvement project before work begins while the Pareto diagram is used to prioritize types of problem to improve the production process. The Cause and Effect diagram is used to logically organize possible causes of defects.

Based on the data collection, the Pareto diagram as shown in Fig. 13.1 indicates that functional test failure is the most frequently occurring problem and also the first type of the problem to be studied and improved in the production process.

The Cause and Effect diagram shown in Fig. 13.2 is developed by brainstorming and discussing among all stakeholders to identify the potential root causes of functional test failure.

The SIPOC diagram as shown in Table 13.1 is used to determine the focus requirements of the project. The SIPOC helps to prepare the business case and project charter to drive the business case.

In the define phase, based on an analysis through the project selection criteria mentioned above, the functional test procedure of the laser computer mouse production process is selected to be improved by the project team members. Its target yield is set about 2 % increasing from the existing yield, e.g., from 96 to 98 %.

The project team members comprise a quality manager, a quality engineer, a quality assurance staff, a project engineer, a production supervisor, a process engineer, and a test engineer.

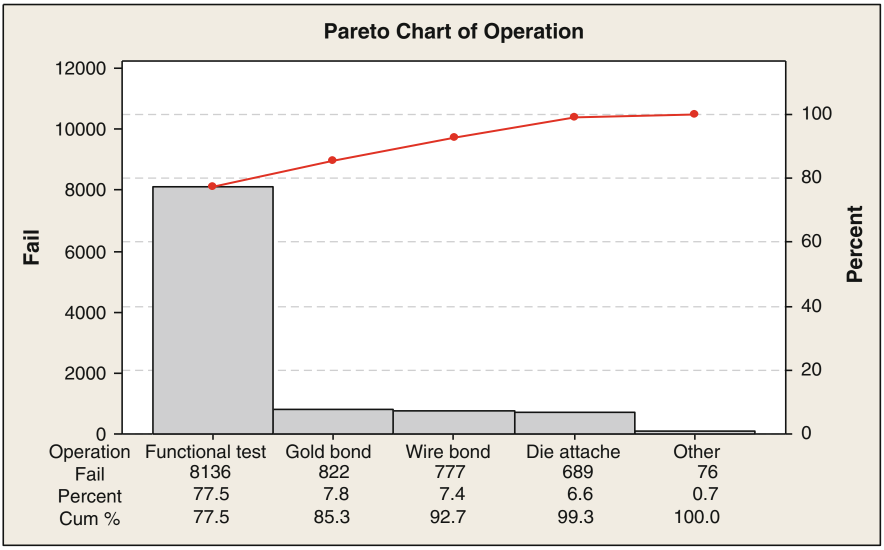


Fig. 13.1 Pareto diagram

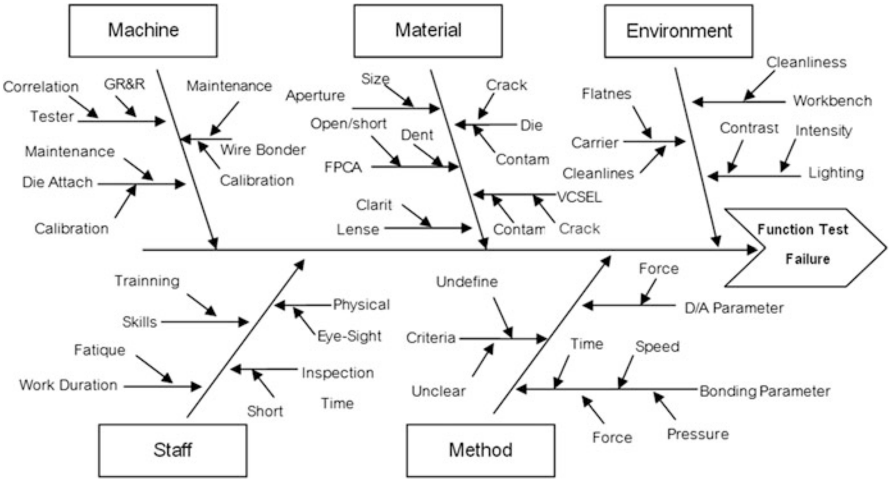


Fig. 13.2 Cause and effect diagram

13.3.2 *Measure Phase*

In the measure phase, the project team members operationally define Critical-To-Quality (CTQ) variables, determine the validity of the measurement system for each CTQ, and establish baseline capabilities for each CTQ.

Table 13.1 The SIPOC diagram

Suppliers	Inputs	Processes	Outputs	Customers
SMT	FPCA		E module	Logitect
EMM	Die	Die attach		
3M	D/A glue for die			
ULM	VCSELs	VCSELs attach		
	D/A glue for VCSELs			
	Al wire	Al bonding		
	Au wire	Au bonding		
	Lense	Aperture assembly		
	Functional tester	Functional test		

In order to understand the whole situation of the production process and establish baseline capabilities, various activities, including process mapping, CTQs identification, process capability analysis (PCA), and measurement system analysis (MSA), are carried out by the project team members.

– Yield throughout an entire process is measured. The use of process flow diagram can help the project team members to better understand how the process flows and how productivity can be improved by using work study.
– Cause and Effect Matrix is conducted to prioritize input variables based on their impacts to outputs. It can also be used to screen important input variables.
– Process capability analysis is conducted to measure CTQs of the process, to understand the ability of the process, to produce a very high proportion of outputs within specification, and to determine the inherent variation of the process.
– MSA is conducted for all gages, measurements, and test equipment to determine the capability of measuring system and to ensure the validity of the measurements. MSA – gage repeatability and reproducibility (GR&R) – can consequently provide reliable and accurate analysis.

Results of the cause and effect matrix (Table 13.2) in the laser computer mouse production process indicate that key process input variables can be found in Die-VCSELs attachment process and Aluminum bonding process.

13.3.3 Analyze Phase

In the analyze phase, hypotheses are confirmed with historical data. The project team members confirm that whether the factors obtained from the measure phase, are present or not, and they also confirm that changes in these factors (or variables) substantially impact the outcomes.

Statistical tools, including hypothesis tests, analysis of variance (ANOVA), and design of experiments (DOE), are employed by the Six Sigma project team members to identify potential root causes of problems.

Table 13.2 Cause and effect matrix

		51	21	10	7	4	2	
Rating of importance to customer		1	2	32	4	5	6	
Process step	Process input	Pixel map	Regulator	Validation	Power	Laser	Open/Short	Total
Die attach	Die attach machine: force	2	2	2	2	1	2	186
	Die attach machine: syringe pressure	4	1	1	1	1	1	248
	Silver epoxy viscosity	2	1	1	1	1	1	146
	Silver epoxy particle	2	1	1	1	1	1	146
	Attached position	9	1	1	1	2	2	509
	Clean room particle	1	1	1	1	1	1	95
	FPCA defect	1	3	1	3	3	8	173
	Noise: supervisor	1	1	1	1	1	1	95
	Skill: training	2	2	2	2	2	2	190
	Working duration: fatigue	1	1	1	1	1	1	95
VCSELs attach	Die attach machine: force	2	1	1	2	2	1	179
	Die attach machine: syringe pressure	2	1	1	2	2	1	157
	Silver epoxy viscosity	2	1	1	1	1	1	146
	Silver epoxy particle	2	1	1	1	1	1	146
	Attached position	9	1	1	1	1	1	503
	Clean room particle	1	1	1	1	1	1	95
	FPCA defect	1	3	1	3	3	7	171
	Noise: supervisor	1	1	1	1	1	1	95
	Skill: training	2	2	2	2	2	2	190
	Working duration: fatigue	1	1	1	1	1	1	95
Aluminium wire bond	Bond pad cleanness	2	1	1	1	2	9	166
	Bond force	2	7	7	7	7	9	414
	Bond time	2	7	7	7	7	9	414
	Bond power	2	7	7	7	7	9	414

	Bond speed	2	7	7	7	7	9	414
	Loop high	2	7	7	7	7	9	414
	Clean room particle	1	1	1	1	1	1	95
	Operator skill	2	2	2	2	2	2	190
Gold bond	Bond pad cleanness	2	2	2	7	7	1	243
	Bond force	4	1	1	1	5	1	264
	Bond time	4	1	1	1	5	1	264
	Bond power	4	1	1	1	5	1	264
	Clean room particle	1	1	1	1	1	1	95
	Operator Skill	2	2	2	2	2	2	190
Aperture assembly	Glue contamination	3	1	1	1	1	1	197
	Glue viscosity	2	1	1	1	1	1	146
	Air pressure	2	1	1	1	1	1	146
	Aperture defects	3	1	1	1	1	1	197
	Operator skill	1	1	1	1	1	1	95
FPCA separation	Mechanism stress	1	1	1	1	1	6	105
	Operator skill	1	1	1	1	1	1	95
	Working duration	1	1	1	1	1	1	95
Lens assembly	Mechanism stress	2	1	1	1	1	1	146
	Lens defects	4	1	1	1	1	1	248
	Operator skill	1	1	1	1	1	1	95
	Working duration	1	1	1	1	1	1	95
Functional test	Test program limit	3	1	1	1	1	1	197
	Electrical over stress (EOS) of M/C	2	3	3	3	3	3	234
	Module defects	2	2	2	2	2	2	190
	Testing time	1	1	1	1	1	1	95
	Operation skill	1	1	1	1	1	1	95

Table 13.3 Factors tested using two sample t-test

Factors	Level		Outcomes	p-value
	Low	High		
Die-VCSELs pitch	3.1 mm	3.3 mm	Map pixel value	<0.05
Bonding force	20 g	30 g	Pull force	<0.05
Bonding power	60 mW	90 mW	Pull force	<0.05
Bonding time	20 mS	30 mS	Pull force	<0.05
Loop height	Low	High	Pull force	0.444
Bonding speed	700 mS/wire	200 mS/wire	Pull force	0.058

According to the results of various Six Sigma tools applied in the measure phase, the possible factor affecting the functional test results can be the distance of the bonding between Die and VCSELs (or Die-VCSELs pitch) and the possible factors affecting the strength of aluminum wire can be various setup parameters for a wire bonder machine, including bonding force, bonding power, bonding time, and bonding speed, and a loop height.

Firstly, an experiment strategy that changes one factor at a time (OFAT) to evaluate its effects on an outcome is conducted. The differences between means for low and high levels of each factor are tested using two sample t-test to determine whether these means have a significantly different effect on an outcome (either map pixel value or pull force). Null hypothesis is $H_0 : \mu_1 = \mu_2$ with sample size of 30 for each t-test. Summary of factors and their low and high levels, outcomes, and p-value of the t-test is shown in Table 13.3.

Based on the t-test results, the four null hypotheses (with p-value < 0.05) are rejected at the 95 % confidence level. Therefore, only these four factors (e.g., Die-VCSELs pitch, bonding force, bonding power, and bonding time) will be considered to improve their either map pixel value or pull force in the improve phase.

13.3.4 Improve Phase

In the improve phase, a solution based on the exposed problem in the first three phases is determined. It begins with brainstorming to create solutions to the problem of functional test failure. After that testing the solutions is executed in order to ensure that the solutions meet the requirements defined, and then the outcomes of the executed solutions are assessed.

- Brainstorming by the project owners and the project team members is conducted to identify factors affecting test failure and eventually find out how to improve the process. Solutions to reduce the failure rate (or improve the process) are identified, and the prevention plan to prevent the problems from occurring is also defined.

Regression Analysis: Pixel Map versus Pitch

```
The regression equation is
Pixel Map = 359 - 107 Pitch


Predictor       Coef   SE Coef        T      P
Constant     359.467     1.833   196.09  0.000
Pitch       -107.011     0.573  -186.84  0.000


S = 0.495996   R-Sq = 99.6%   R-Sq(adj) = 99.6%


Analysis of Variance

Source           DF      SS      MS         F      P
Regression        1  8588.5  8588.5  34910.79  0.000
Residual Error  148    36.4     0.2
Total           149  8624.9
```

Fig. 13.3 Linear regression analysis results

- The identified process improvement solutions which are obtained based on brainstorming are also tested. The yield of the executed solutions is assessed, and then the solutions are slightly modified to meet the target yield before embarking on full scale implementation.
- The above activities are done to determine the finest settings for implementation. However, the most effective setting depending on company's resources, procedures, and policies has to be rationally chosen.

Map pixel value improvement: According to the analyze phase, Die-VCSELs pitch significantly affects the Map pixel value. In order to improve the Map pixel value approaching to the laser index of 15, the optimal value of Die-VCSELs pitch has to be determined. Hence, linear regression analysis is carried out to determine the optimal condition of the Die-VCSELs pitch that produces the specific laser index of 15.

From the regression analysis, correlation coefficient of 99.6 % indicates an acceptable linearity (Darlington 1990). A linear regression model is stated in Map pixel value = 359 – 107 × Pitch. Based on the model, to achieve the Map pixel value of the laser index of 15, therefore the Die-VCSELs pitch should be approximately determined at 3.22 mm (Fig. 13.3).

Pull force improvement: Based on the OFAT results, the setup parameters – including bonding force, bonding power, and bonding time – significantly affect the pull force. These three factors (or setup parameters) at two levels each are studied using DOE. In order to improve the pull force, the optimal conditions of these parameters have to be determined. Therefore, a three-factor two-level factorial

Factorial Fit: Pull strength versus Force, Power, Time

Term	Effect	Coef	SE Coef	T	P
Constant		6.0828	0.01957	310.85	0.000
Force	-0.2781	-0.1391	0.01957	-7.11	0.000
Power	0.8331	0.4166	0.01957	21.29	0.000
Time	-0.2856	-0.1428	0.01957	-7.30	0.000
Force*Power	-0.1881	-0.0941	0.01957	-4.81	0.000
Force*Time	-1.0394	-0.5197	0.01957	-26.56	0.000
Power*Time	-0.7106	-0.3553	0.01957	-18.16	0.000
Force*Power*Time	0.3731	0.1866	0.01957	9.53	0.000
Ct Pt		1.8532	0.05323	34.81	0.000

S = 0.110695 PRESS = 18.8505
R-Sq = 99.05% R-Sq(pred) = 47.78% R-Sq(adj) = 98.78%

Analysis of Variance for Pull strength (coded units)

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Main Effects	3	6.8243	6.8243	2.2748	185.64	0.000
2-Way Interactions	3	12.9654	12.9654	4.3218	352.70	0.000
3-Way Interactions	1	1.1138	1.1138	1.1138	90.90	0.000
Curvature	1	14.8510	14.8510	14.8510	1211.99	0.000
Residual Error	28	0.3431	0.3431	0.0123		
Pure Error	28	0.3431	0.3431	0.0123		
Total	36	36.0976				

Fig. 13.4 Factorial design results

design (a 2^3 factorial design) with three replicates of the central point is applied to determine their optimal conditions.

Results of the 2^3 factorial design as shown in Fig. 13.4 are observed that all main effects and all interactions among those parameters are significant at the 95 % confidence level (p-value < 0.05). Based on the response optimization results as shown in Fig. 13.5, they indicate that the optimal values of the three parameters for maximizing Pull force are as follows: a bonding force of 25 g, a bonding power of 90 mW, and a bonding time of 20 mS.

13.3.5 Control Phase

In control phase, standardizing and documenting the new processes across the project, training operators, and creating a plan to monitor the process are executed. Six Sigma tools, including control charts, process maps, and monitoring plans, are used.

In order to ensure that the yields obtained during the improve phase are maintained even after the project ends. Therefore, it is necessary to standardize and document procedures and make sure that all operators are trained.

The project team members create a control plan (Table 13.4) for ongoing monitoring of the processes. Moreover, the application of information technology used to control the functional test procedure of the process is provided. By linking

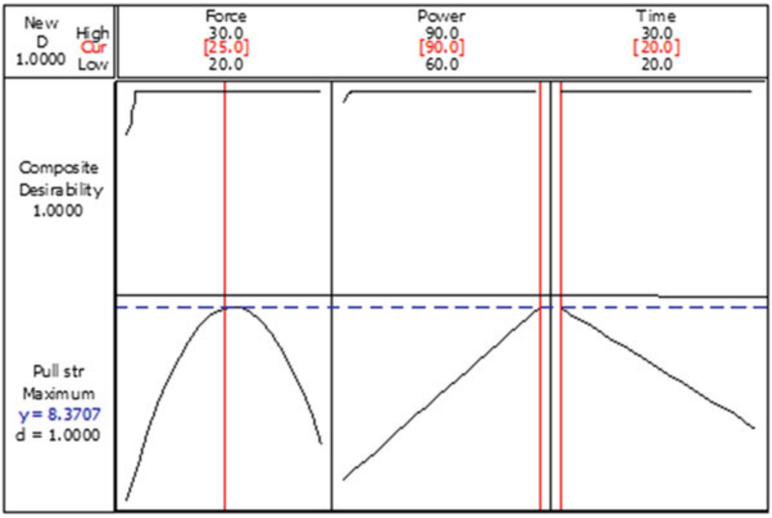


Fig. 13.5 Response optimization results

test results to an automatic data transfer system, therefore quick response to any problems can help team members solve these problems promptly.

13.4 Results and Discussions

The findings on how Six Sigma tools and other advanced statistical tools can be used to help in improving the production process of a laser computer mouse in each phase of DMAIC methodology are discussed as follows.

13.4.1 Define Phase

The aim of the define phase is to determine the process requirements and define the scope and goal(s) of the improvement project (Antony et al. 2012).

According to the project selection analysis based on Thailand Quality Award (TQA) criteria, a functional test procedure of the laser computer mouse production process is selected by the Six Sigma project team members after brainstorming and discussion with the stakeholders. Based on the data collection and brainstorming, the Pareto diagram and the cause and effect diagram are used to find out the root causes of functional test failure, in order to eliminate them. Final step of the define phase, the SIPOC diagram is developed. It not only helps to determine the focus requirements of the project but also to prepare the project charter to drive the business case.

Table 13.4 A control plan

Operation	Equipment/ Fixture used	Chemical material used	Control parameter	Control tools	Action	Frequency	Responsibility	Reference document No.
Al wire bonding	Al wire bonder	Al wire ∅1.0 mm.	Machine parameters: File name; E Force = 25, Power = 90, Time = 20	Wire bonding check list	Inform tech, Eng	Every start shift	Production	OS04031
	Bond pull tester	Finger cots						
	High power scope							
	Microscope 40 ×		Bond pull force $\geq$ 4 g.	$\bar{X} - R$ chart	Issue IRR to production	10 wires at start of shift	QA	WI-03215 T
			Loop height 3–8 mm.	$\bar{X} - R$ chart	Issue IRR to production	10 wires at start of shift	QA	
Functional test	Tester	Finger cots	Reject criteria per test software	Functional test record		100 %	Production	
		IPA						
		Cotton swap						
	Functional tester		Pre test procedure for triggering 98 %		Inform production supervisor/TE to verify the defect Issue IRR to production		PD, QA, TE	TP-00138

13.4.2 Measure Phase

The measure phase involves identifying the key input and output variables, determining operational measures and definition for each CTQ, and performing a GR&R analysis for each CTQ. The cause and effect matrix as shown in Table 13.2 indicates the key input variables of processes.

Moreover, the MSA of the functional test procedure is conducted. The map pixel measurement system is found to be reliable. Therefore, the use of the existing staffs, equipment, and the functional test system is reliable and repeatable (not reported in this paper).

13.4.3 Analyze Phase

The analyze phase involves identifying and confirming potential root causes of problems.

In this study, the OFAT experimentation using two sample t-test is used to evaluate the effect of a factor on an outcome in order to screen the significant factors by finding the influential factors. Only four factors, including Die-VCSELs pitch, bonding force, power, and time, are considered to be improved in the improve phase.

13.4.4 Improve Phase

The potential root causes of defects identified during the measure phase provide inputs to the improve phase. The improve phase involves selecting solutions to eliminate these causes of defects.

The map pixel value improvement is conducted. The optimal Die-VCSELs pitch is determined using a linear regression analysis. And the pull force improvement is conducted. A 2^3 factorial design with three replicates is applied to determine the optimal conditions of the three factors, including bonding force, power, and time. The results as shown in Fig. 13.4 indicate that all main effects and interactions among these factors are significant at the 95 % confidence level, and the response optimization results also indicate the optimal values of these factors.

13.4.5 Control Phase

The control phase involves establishing measures to standardize, monitor, and integrate the changes within the existing framework. The project team members develop a control plan as shown in Table 13.4 to maintain the desired improvements.

Control tools such as the $\overline{X} - R$ chart for pull strength and check list are applied. Besides, certain activities to monitor and control the improvements are applied during this control phase, including establishing procedure standardization and documentation.

13.5 Conclusion

In order to improve the production process of a laser computer mouse, a case study, the DMAIC methodology of Six Sigma is applied. According to the study, the functional test procedure has the lowest yield of 96.2 %. After the implementation of the DMAIC methodology, the yield of the functional test procedure increases from 96.2 to 98.6 %, higher than the target yield of 98 %.

Contributions of this study include demonstrating how Six Sigma tools and other advanced statistical tools can be applied to help in improving the production process of a laser computer mouse in each phase of DMAIC methodology.

References

Antony J, Bhuller AS, Kumar M, Mendibil K, Montgomery DC (2012) Application of Six Sigma DMAIC methodology in a transactional environment. Int J Qual Reliab Manage 29(1):31–53

Bhote K (2003) The power of ultimate Six Sigma. Amacom, New York

Darlington RB (1990) Regression and linear models. McGraw-Hill, New York

Hung YY, Chen YS, Ng SP, Liu L, Huang YH, Luk BL, Ip RWL, Wu CML, Chung PS (2009) Review and comparison of shearography and active thermography for nondestructive evaluation. Mater Sci Eng R Rep 64(5–6):73–112

Jäger R, Riedl MC (2011) MBE growth of VCSELs for high volume applications. J Cryst Growth 323(1):434–437

Kumar S, Strandlund E, Thomas D (2008) Improved service system design using Six Sigma DMAIC for a major US consumer electronics and appliance retailer. Int J Retail Distrib Manage 36(12):970–994

Leaphart CL, Gonwa TA, Mai ML, Prendergast MB, Wadei HM, Tepas JJ III, Taner CB (2012) Formal quality improvement curriculum and DMAIC method results in interdisciplinary collaboration and process improvement in renal transplant patients. J Surg Res 177(1):7–13

Pickrell G, Lyons HJ, Shaver J (2005) Lean Six Sigma implementation case studies. IntJ Six Sigma Compet Adv 1(4):369–379

Pyzdek T, Keller PA (2009) The Six Sigma handbook, 3rd edn. McGraw-Hill, New York

Shanmugaraja M, Nataraj M, Gunasekaran N (2011) Defect control analysis for improving quality and productivity: an innovative Six Sigma case study. Int J Qual Innov 1(3):259–282

Tague NR (2004) Seven basic quality tools. American Society for Quality, Milwaukee

Chapter 14
Machine Vision-Aided Quality Decision System for Solder Joint Defect Evaluation

Chien-Chih Wang

Abstract To improve the printed circuit board (PCB) manufacturing process, it is important to have an automatic inspection system that classifies information regarding defects in solder joints. This paper proposes a quality decision system for solder joint defect classification on a PCB. An experiment was conducted to demonstrate the application of this technique. The results showed that the inspection accuracy reached 94 %, which is superior to the results achieved by other methods. The results of this study provide an effective solution for the inspection of the solder joint quality.

Keywords Automatic visual inspection • Multivariate analysis • Defect classification

14.1 Introduction

Solder Joint quality is the key factor in the assembly of electronic components. In practice, the main testing methods are electric test, function test and appearance test. Among, the inspection results of appearance test were difficult assessment. The main reason is that product design tends to be ‘small, thin, and light’, as in cellular phones. Therefore, although modern technology has enabled the traditional semi-automatic production in electronic industries to be gradually replaced by fully

---

This research was supported by the National Science Council of Taiwan, Project Number NSC 90-2218-E-131-012 and NSC 91-2218-E-131-005.

C.-C. Wang (✉)
Department of Industrial Engineering and Management, Ming Chi University of Technology, Taipei County 243, Taiwan
e-mail: ieccwang@mail.mcut.edu.tw

E. Qi et al. (eds.), *The 19th International Conference on Industrial Engineering and Engineering Management*, DOI 10.1007/978-3-642-37270-4_14,

automatic production, most solder joint defect inspections are still performed manually. Therefore, it is crucial to develop optimal defect detection and classification methodology using machine vision for quality improvement in order to ensure the quality of mass production manufacturing (Wang and Jiang 2001; Wang et al. 2011; Jiang et al. 2002, 2004, 2007, 2010; Chu et al. 2008). The mission of traditional solder joint inspection was focused on go/no-go quality judgement. This inspection result can provide improvement information is limited. Therefore, there is an urgent need for inspection and classification methods to be integrated into the detection of defects in a quality decision system.

Various statistical and non-statistical methods have been used for defect classification. Some of the most popular methods are the K-nearest neighbor classifier, the Linear discriminate function classifier, tree classifier, neural network classifier, and the Bayes classifier (Ripley 1996).

Oyeleye and Lehtihet (1999) used a three-dimensional solder joint model to inspect and classify actual solder joints. Liao et al. proposed a welding flaw detection method based on a fuzzy k-NN classifier. Ko and Cho (2000) proposed a classification method based on the neural network and fuzzy rule for inspection of the solder joints of printed circuit boards. The results of their experiment have proven effective in the classification of solder-joint images, with a high success rate compared to two neural network methods: the LVQ self-clustering algorithm, and the Kohonen neural network classifier. Jiang et al. (2001) proposed bootstrap sampling techniques for the generation of enough samples to determine population parameters and a tree classifier for golden finger defect classification. Jiang et al. (2007) proposed a novel two-stage machine vision inspection process to segment and classify solder joints for PCBs, for a monitor manufacturer in Taiwan.

Among known classifiers, the Bayesian classifier is the optimal system, producing a minimum error rate if the underlying posterior distribution of the data is known. When Bayesian classification is applied in practice, it is often assumed that the prior distribution does not exist, or equal and that the posterior distribution is a multi-normal distribution. With regard to the posterior distribution, multivariate normal distribution is often assumed, with no further consideration of the characteristics of the data. Before the Bayesian classification is used, the data are assumed to display a multivariate normal distribution, regardless of whether this is actually the case; the result of this is ‘GIGO’ (Garbage In, Garbage Out). In theory, the more the parameters, the more are the required samples; this further complicates the deduction of the Bayesian decision formula. Hence, the Bayesian classifier under multivariate normal distribution is a commonly used method in practical application. In this paper, therefore, to preserve the minimum error rate property of the Bayesian classifier, we propose a quality decision system to improve the Bayesian classifier with variables that represent non-normal features. To demonstrate the proposed procedures, we also provide a defect classification experiment conducted upon solder joints on a PCB board obtained from a monitor manufactured in Taiwan. The PCB board contained examples of no-solder, short, and open-soldered joint defects.

14.2 Methodology

The Bayesian classifier is a classification method based on a probability distribution. In practice, the Bayesian classifier relies on the assumption of the feature variables of multivariate normality. To address the problem of defect classification, the basic approach of Bayesian classification addresses the historical data of defect types (called prior information) and the feature variables of the known defects information (called posterior information). The rules for decision-making in classification can be derived from these two sets of data. Theoretically, if the prior and posterior information for defects are known, an optimal decision rule can then be derived. Assume that there are c types of defect, denoted as $G_1, G_2, \cdots, G_c$. All defect types can be selected as p feature variables quantification its information. The data structure of defect G_i can then be expressed as follows:

$$G_i = \begin{bmatrix} x_{i11} & x_{i12} & \cdots & x_{i1p} \\ x_{i21} & x_{i22} & \cdots & x_{i2p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{im_i1} & x_{im_i2} & \cdots & x_{im_ip} \end{bmatrix} \tag{14.1}$$

where x_{ijk} denotes the i-th defect, the j-th measurement value under the k-th feature variable.

Let $p\,(G_i)\,, i = 1, 2\ldots, c$ denote the c types of defect that could appear with some probability, referred to as priori probability, and let $f(\vec{\mathrm{x}}|G_i), i = 1, 2, \ldots, c$ denote the i-th class conditioned probability density for feature variables vector $\vec{\mathrm{x}}$. According to the Bayes theorem, the posterior probability $p(G_i|\vec{\mathrm{x}})$ is

$$p\left(G_i|\vec{\mathrm{x}}\right) = \frac{f_i\left(\vec{\mathrm{x}}|G_i\right) \times p\,(G_i)}{p\left(\vec{\mathrm{x}}\right)}, p\left(\vec{\mathrm{x}}\right) = \sum_{i=1}^{c} f_i\left(\vec{\mathrm{x}}|G_i\right) \times p\,(G_i)$$

In Bayesian classification, the decision rule is based on the minimum loss principle, with less $e^i\left(\vec{x}\right)$ defined as classification error probability $1 - p\left(G_i|\vec{\mathrm{x}}\right)$. The optimal Bayesian decision rule for c types of defect (Mardai 1970) is then

$$e^*\left(\vec{\mathrm{x}}\right) = \underset{i=1,2,\ldots,c}{\mathrm{Min}}\, e^i\left(\vec{\mathrm{x}}\right) = \underset{i=1,2,\ldots,c}{\mathrm{Min}}\left[1 - p\left(G_i|\vec{\mathrm{x}}\right)\right] \tag{14.2}$$

We can simplify the above equation to obtain the decision rule $\hat{W}\left(\vec{\mathrm{x}}\right)$ of Bayesian classification for two-class product defect classification, as

$$\hat{W}\left(\vec{\mathrm{x}}\right)=\begin{cases}G_i,\ \text{if}\ p\left(G_i\right)\times p\left(\vec{\mathrm{x}}|G_i\right)>p\left(G_j\right)\times p\left(\vec{\mathrm{x}}|G_j\right)\\ G_j,\ \text{if}\ p\left(G_i\right)\times p\left(\vec{\mathrm{x}}|G_i\right)<p\left(G_j\right)\times p\left(\vec{\mathrm{x}}|G_j\right)\end{cases},$$

$$p\left(\vec{\mathrm{x}}|G_i\right)\times p\left(G_i\right)=p\left(\vec{\mathrm{x}}|G_j\right)\times p\left(G_j\right) \tag{14.3}$$

We solve $\vec{\mathrm{x}}$ to obtain the optimal Bayesian decision rule formula. This type of decision formula is related to the number of feature variables. For one feature variable, the decision formula is rectilinear; for two feature variables, it is linear.

A Bayesian classifier that uses multivariate normal distribution to approximate multivariate non-normal distribution will increase the classification error rate (Wang and Jiang 2001). Therefore, in this paper, we propose a multivariate transformation Bayesian classifier (MTBC). To develop this classification algorithm, we first executed a multi-normal test for feature variables to confirm whether there is multivariate normal posterior distribution. If the posterior distribution was multi-normal, we then applied the traditional Bayesian classifier. If it was not, we applied a transformation procedure to convert the multi-non-normal distribution in multi-normal distribution, after implementation of the Bayesian classification. The MTBC involves two key procedures, which include multivariate normal test and multivariate transformation. Many researchers have proposed multivariate normal test methods (Smith and Jain 1988). In this paper, we used the multivariate skewness and kurtosis test for Multivariate Normality (Mardai 1970). This procedure makes use of calculations as follows:

- Step 1: Calculate all the variables of coefficient $r_{i,j}$; $\sum$ is the variance-covariance matrix of G.

$$r_{i,j}=\left(X_i-\overline{X}\right)^T\Sigma^{-1}\left(X_i-\bar{X}\right) \tag{14.4}$$

- Step 2: Calculate Mardia's skewness measure value $b_{1,p}=\sum\limits_{i=1}^{n}\sum\limits_{j=1}^{n}r_{i,j}^3/n^2$
- Step 3: Calculate test statistic $A=nb_{1,p}/6$
- Step 4: If test statistic $A>\chi^2_{1-\alpha}(f)$, then reject the assumption of multivariate normality, and $f=p\,(p+1)\,(p+2)/6$, p is the number of feature variables, with n as the sample size.

In this paper, multivariate Box-Cox transformation was applied to remove non-normality. The multivariate Box-Cox transformation is defined by

$$X_j^{(\lambda)}=\left[\frac{x_{ij1}^{\lambda_1}-1}{\lambda_1},\frac{x_{ij2}^{\lambda_2}-1}{\lambda_2},\ldots,\frac{x_{ijp}^{\lambda_p}-1}{\lambda_p}\right],j=1,2\ldots,m,$$

$$if\ \lambda_i\neq 0=\left[\ln x_{ij1},\ln x_{ij2},\ldots,\ln x_{ijp}\right],j=1,2\ldots,m,if\ \lambda_i=0 \tag{14.5}$$

We must determine $\vec{\lambda} = (\lambda_1, \lambda_2, \ldots, \lambda_p)$, such that the transformation variables $X_j^{(\lambda)}$ are closer to multivariate normality. The optimal estimate $\vec{\lambda} = (\lambda_1, \lambda_2, \ldots, \lambda_p)$ is calculated by maximizing the likelihood function $L(\vec{\lambda}) = -\frac{n}{2} \ln \det \left(\hat{\Sigma}_{\vec{\lambda}} \right) + \sum_{j=1}^{p} (\lambda_j - 1) \sum_{k=1}^{m} \ln x_{ijk}$, where $\Sigma_{\vec{\lambda}} = \sum \left(X_j^{(\lambda)} - \bar{X}^{\lambda} \right) \left(X_j^{(\lambda)} - \bar{X}^{\lambda} \right)^T / n$.

14.3 Results

In this section, we use solder joint defect data to demonstrate the proposed procedure. The solder joint data contained 285 samples that included normal solder joints and no-solder, short, and open types of defects (Table 14.1). Among our samples, 85 solder joints served as training samples and the rest as testing samples.

For the defect detection of solder joints, we used a two-stage procedure to segment solder joints from the inspected image (Wang and Jiang 2001). The first stage defines segments on the bare copper region of the PCB (Fig. 14.1a). To do this, we first acquired a color image of the bare PCB. Masking was then performed using a median filter to remove the noise and fill in the areas reserved for the placement of pins. The final image represents the copper foil area depicted in white with the remaining area in black, thus obtaining a solder area image (Fig. 14.1b) to serve as a reference image.

Table 14.1 The defect types used in this study, along with their appearance

Defect type	Open	No solder	Short, Bridge
Appearance			

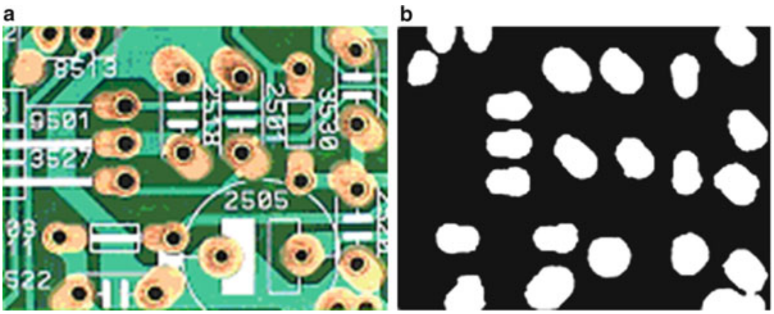


Fig. 14.1 (**a**) Original image; (**b**) Bare PCB copper region

Fig. 14.2 (**a**) Inspection image; (**b**) Reference image; (**c**) Solder joint segment

The second stage involves comparing the divided copper area with the inspected image from the first stage. This is done by means of a pixel-by-pixel comparison. Because the grey in the copper area in the reference image has a pixel value of 0, a 'minimum' function calculation is performed in the inspection (Fig. 14.2a) and reference images (Fig. 14.2b). As shown below, the solder joint in the image is then further articulated for complete inspection (Fig. 14.2c)

Next, we selected 16 candidate feature variables that take into account the geometric and statistical features of solder joints. These features were selected on the basis of their ease of calculation and their capacity to discriminate among the three types of defects. The seven statistical features based on the grey image include: mean (F_1), mode (F_2), median (F_3), variance (F_4), standard deviation (F_5), skewness (F_6), and kurtosis (F_7). The nine geometric features based on a binary image include: variance in the X-axis direction (F_8), variance in the Y-axis direction (F_9), covariance of the X and Y axis (F_{10}), area (F_{11}), perimeter (F_{12}), circularity (F_{13}), major axis length (F_{14}), minor axis length (F_{15}), and rectangularity (F_{16}). In this paper we implemented the following analysis procedures to perform solder joint defect classification.

- Step 1: Eliminate the feature variables with poor discriminative capacity.

ANOVA technology was applied to eliminate the feature variables with poor discriminative capacity. ANOVA is a statistical method used to test the differences between two or more population means. The results of the *p*-value are presented in Table 14.2. Because the *p*-value is smaller than 0.05, this indicates that the seven feature variables (F_2, F_8, F_9, F_{11}, F_{14}, F_{15}, and F_{16}) can be used to distinguish the differences among four solder joints.

- Step 2: Construct the experimental design for the remaining feature variables from Step 1.

Step 2 involves determining an optimal feature subset. From the results of Step 1, we can dispose of the 2_{III}^{k-p} design.

- Step 3: Calculate the Pillai statistics to evaluate the multiple group differences.

Table 14.2 Feature selection analysis result

Feature variables	*p*-value
F_1	0.17
F_2	***0.01***
F_3	0.12
F_4	0.22
F_5	0.50
F_6	0.47
F_7	0.47
F_8	***0.02***
F_9	***0.00***
F_{10}	0.14
F_{11}	***0.00***
F_{12}	0.78
F_{13}	0.81
F_{14}	***0.00***
F_{15}	***0.00***
F_{16}	***0.00***

For each experimental combination, we executed MANOVA and calculated the Pillai statistics in response. The Pillai statistic (V) is defined as follows

$$V = trace\left(\mathrm{B}(\mathrm{W} + \mathrm{B})^{-1}\right) = \sum_{i=1}^{p} \frac{\lambda_i}{1 + \lambda_i} \tag{14.6}$$

where λ_i is the eigenvalue of the $\mathrm{B}(\mathrm{B} + \mathrm{W})^{-1}$ matrix, $i = 1, 2, \ldots, p$. **B** is the between sum of the squares and cross-products matrix, and **W** is the within sum of the squares and cross-products matrix.

- Step 4: Optimal feature variables

With a type I error α of 0.05, we consulted a Pareto chart that shows the estimated effects of feature variables. This Pareto chart displays the absolute value of these effects, and draws a reference line that allows users to see both the magnitude and importance of a given effect. Any effect that extends past this reference line is potentially important. Applying this to Fig. 14.3, we can confirm that the F_2, F_{16}, F_{11}, and F_9 constitute an optimal feature subset.

- Step 5: Fitted optimal Bayes classification model

From Step 4, we obtained the optimal feature variables and performed the MTBC procedure. First, we tested the optimal feature variables to confirm their compliance with our assumption of multi-normality distribution. We then plotted the multi-normality probability plot and calculated the Mardia skewness coefficient value for four types of solder joints. From Table 14.3, with the optimal feature variables, the normal solder joints and open solder joints are not in compliance with the assumption of multi-normality distribution, while the no-solder joints and short

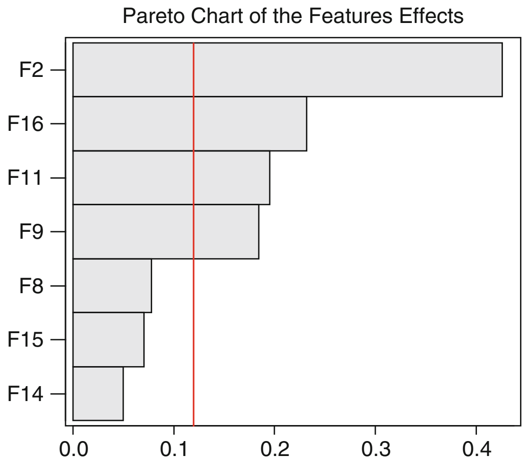


Fig. 14.3 Pareto chart of the effects of feature variables when $\alpha = 0.05$

solder joints comply with that assumption. No-solder joints and short solder joints are therefore used for the multi-normality Bayesian classifier, and normal solder joints and open solder joints are used for the MTBC.

According to a maximum likelihood estimate, we can affirm that the optimal Box-Cox transformation λ_s for multivariate feature variables F_2, F_{16}, F_{11}, and F_9 of normal solder joints are 0.3, 0.8, 0.9, and 0.6, respectively, and that for open solder joints they are 0.1, 1.6, 1.5, and 0.3, respectively. Next, we can determine the MTBC for normal and open solder joints.

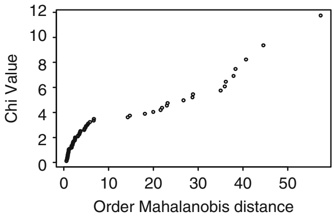
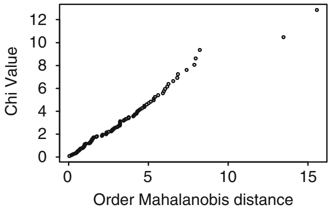
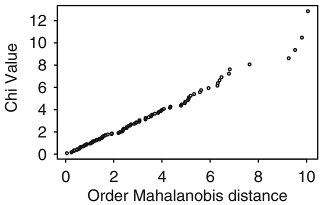
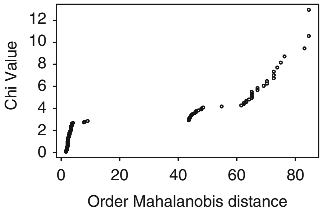
- Step 6: Estimate the rate of correct classification.

A testing sample of 200 solder joints was used to evaluate the proposed analysis procedure. The results are shown in Table 14.4. The correct classification rate was as follows: 93 % for normal solders, 90 % for no-solders 100 % for short solders, and 94 % for open solders. The average correct classification rate was 94 %. Next, we compare these results with those of the other three methods, including the multivariate approach Bayes classifier, linear discriminant function classifier, and 1-NN. The results are shown in Table 14.4. The multivariate approach Bayes classifier has a correct classification rate of 89 %, the linear discriminant function classifier rate is 85 %, and the 1-NN classifier rate is 78 %. As shown in Table 14.4, the proposed MTBC was superior to the multi-normal approach Bayes classifier, the linear discriminate function classifier, and 1-NN.

14.4 Discussion

This study proposes a quality decision system for solder joint defect classification to improve the results obtained by the traditional Bayes classifier. The traditional multivariate Bayesian classifier is a widely used method for defect classification that

Table 14.3 Optimal feature variables for multivariate normal test

Type	Multivariate normal probability plot	Test statistic
Normal		33.09946
No-solder		18.674075
Short		12.00483
Open		37.6055

produces a minimum error rate. Because these methods rely on the assumption of multivariate normality, the results of their practical application cannot be considered optimal. In practice, the assumption of multivariate normality is ignored, and the result is an increase in the rate of incorrect classification. To address this problem, in this paper, we proposed a quality decision system. In order to validate this procedure, it was applied to a large sample of solder joint defects, and the result achieved a classification correctness of 94 %. Our experimental analysis results showed that that in terms of classification accuracy, the MTBC is superior to the traditional Bayesian classifier, the linear discriminate function classifier, and the 1-NN classifier.

The proposed method also offers the following advantages: (1) It can handle multivariate data classification problems, (2) It produces a minimum classification

Table 14.4 A comparison of the multivariate transformation Bayesian classifier and other classifiers

Classification method	Defect type	Corrected classification	Average correct %
MTBC (proposed in this paper)	I	0.93	0.94
	II	0.9	
	III	1.00	
	IV	0.94	
Multinormal approach Bayes classifier	I	0.84	0.89
	II	0.9	
	III	1.00	
	IV	0.87	
Linear discriminate function classifier	I	0.81	0.85
	II	0.87	
	III	1.00	
	IV	0.75	
1-NN	I	0.83	0.78
	II	0.68	
	III	0.86	
	IV	0.73	

I: Normal, II: No-solder, III: Short, IV: Open

error rate, and (3) It can be readily applied to practical classification problems. These research results provide an effective new alternative that resolves a weakness in the traditional detection and classification of solder joint defects.

References

Chu CC, Jiang BC, Wang CC (2008) Modified gamma correlation method to enhance BGA surface image for surface defect inspection. Int J Prod Res 46(8):2165–2178

Jiang BC, Wang YM, Wang CC (2001) Bootstrap sampling techniques applied to PCB golden fingers defect classification study. Int J Prod Res 39(10):2215–2230

Jiang BC, Tasi SL, Wang CC (2002) Machine-vision based grey relational theory applied to IC marking inspection. IEEE Trans Semicond Manuf 15(4):531–539

Jiang BC, Wang CC, Chen PL (2004) Logistic regression tree applied to classify PCB golden finger defects. Int J Adv Manuf Technol 24(7):496–502

Jiang BC, Wang CC, Hau YN (2007) Machine vision and background remover-based approach for PCB solder joints inspection. Int J Prod Res 45(2):451–464

Jiang BC, Wang CC, Chen HJ, Chu CC (2010) Automatic bubble defect inspection for microwave communication substrates using multi-threshold technique based co-occurrence matrix. Int J Prod Res 48(8):2361–2371

Ko KW, Cho HS (2000) Solder joints inspection using a neural network and fuzzy rule-based classification method. IEEE Trans Electron Pack Manuf 23(2):93–103

Mardai KV (1970) Measures of multivariate skewness and kurtosis with applications. Biometrika 36:519–530

Oyeleye OU, Lehtihet EA (1999) Automatic visual inspection of surface mount solder joint defects. Int J Prod Res 37:1217–1242

Ripley BD (1996) Pattern recognition and neural networks. Cambridge University Press, New York

Smith SP, Jain AK (1988) A test to determine the multivariate normality of a dataset. IEEE Trans Pattern Anal Mach Intell 10:757–761

Wang CC, Jiang BC (2001) PCB solder joints defects detection and classification using machine vision. Int J Ind Eng 8(4):359–368

Wang CC, Jiang BC, Chou YS, Chu CC (2011) Multivariate analysis techniques-based image enhancement model for electronic products inspection. Int J Prod Res 49(15):2999–3021

Chapter 15
Demand Analysis and Framework of Theory About Chinese Events Sports Logistics

Peng-hui Liu and Li Gao

Abstract As a special kind of logistics activities, sports logistics affect deeply the rapid development of logistics industry in China. In this article, the author analyzes the present requirement of events sports logistics in China using literature and logical reasoning, the article analyzes the problems of sports logistics in present development, and on this basis to explore the frame structure of events sports theory research, so as to perfect further sports logistics theory.

Keywords Demand analysis • Events • Sports logistics • Theory system

15.1 Introduction

"We should emphasize the modern logistics, and broaden our horizons, and look for new economic growth point, and broaden our living space" (The state council 2008), this is that The Deputy Prime Minister Li LanQing introduces solemnly in teleconference of the national commodity circulation work in the December 25, in 1997, and that emphasized mainly the importance of developing modern logistics. In the twenty-first century, the overall size of the Chinese logistics industry realized the fast growth. Builded the logistics distribution center in all over the country, logistics park, It raised the level of service of the logistics industry and improved environment condition, and laid even more the solid foundation to develop logistics industry's in the future.

P. Liu (✉)
PE School, Beihua University, Jilin, China

Yongin University, Yongin, South Korea
e-mail: z7811232005@yahoo.com.cn

L. Gao
Qingbei Nine Years schools, Piaohe town Jiaohe, China

E. Qi et al. (eds.), *The 19th International Conference on Industrial Engineering and Engineering Management*, DOI 10.1007/978-3-642-37270-4_15,

China has entered the world sports country ranks, the Olympic Games, the Asian games such as domestic and international well-known events held a grand now, it put forward higher request for the domestic logistics industry, promoting the development of “sports logistics”. Sports logistics encompasses all the economic significance of the things flow that it is sports-related; is a kind of large and complex system of social economic activity, it not only relates with the success or failure of sports activities but also plays a very important role for the development of local economy (Yongsheng Dong 2006). Nowadays, only by researching continuously the opportunities and challenges that sports activities bring for logistics development, and researching the theoretical framework of the sports logistics structure, and praising scientifically the relationship between the physical activity and the logistics economic, we can promote our country sports logistics business scale expansion (Qinlin Deng 2006).

15.2 Demand Analysis About Events Sport Logistics in China

15.2.1 Demand Status

In 1996, the freight turnover is 3.6271 trillion ton-km that is 1.29 times that of 1990; in 2010 the freight turnover is about seven trillion ton-km. According to relevant statistics, physical logistics is around 5 % percent of total logistics in China, and raised to around 10 % during the Beijing Olympic Games, the Olympic Games generate physical logistics business and is also a period of rapid growth of logistics in China (Danqing Liang and Hongyi Jia 2008).

The sports event convenes grandly, has provided the infinite opportunity to the physical distribution industry, following is a bigger test. According to the actual needs of competition, transport, storage, handling, processing, distribution, handling, packing, the circulation of information processing, and other basic function implementation organic combination, and provide extension services as needed (Xianwu Li 2005). In 2008 the Beijing Olympic Games, the participative country and the area had achieved 204, the participative athlete have only achieved 11,483 people. In 2010 Guangzhou Asian Games, also has collected the Asian more than 46 participation countries and the local tens of thousands of athletes, reporter, the audience, the concerned personnel communicates, the journey, the traveling and around the consumable allocation and the competition the equipment, the sporting goods ships, aspects and so on storage, packing, information processing all needs the physical distribution system the coordination and the control. In addition, but also has the sports to equip the market, the competition performs the market, the sports intermediary market, the sports traveling market, the sports insures aspects and so on market to need the sports physical distribution to provide the service, has brought the huge pressure for the sports physical distribution.

15.2.2 Event Sports Physical Distribution Type Division

During sports activities, there are a large number of resource allocations and flow of logistics to implement events in sports consumption needed flow of goods from suppliers to the receiving entity (Zhongfu Cui pay 2007). Transport of personnel, equipment guarantee before the match, racing sports, successfully carried out as scheduled, ordered after the game control, evacuation, and more. Depending on the angle, racing sports Logistics Division.

Physical logistics of fluids can be divided into means of use, means of subsistence and waste logistics. Most sports logistics information, broken down by time, racing, racing back to before the match logistics; but regardless of what stage the race, cannot be separated from the waste material logistics and production materials and supplies, so you can divide the two consider integrated (Yong Hu and Jing Jia 2010). From the perspective of the workload, the three logistics are large.

Logistics can be divided into local and off-site logistics on the flow from geographical and physical activities. Local logistics is a division of logistics, off-site logistics is the logistics in offsite and international logistics due to content rich events activities, there will be more games, spatial transfer of logistics activities will be realized, the logistics activities before the game flows to areas of logistics, and after the game has left the logistics of organized sports activities.

In addition, the service object also can be under logistics activities, logistics movement patterns to divide, not described here.

15.2.3 Main Features of Sports Logistics

Racing sports logistics arising from physical activity, not only throughout the process there is a demand for the movement of persons, but also a great deal of material flows of the entity, logistics entity of the larger differences between individuals. Therefore, unlike most of the logistics, showing their unique characteristics:

1. Sports logistics spaces span, involving wide geographical range.

Whether international matches, domestic competitions, there is competition from different domain, may be all over the country, or it may be all over the world, it had the necessary material to the competition venues, this space of large-span, region wide, increasing difficulty for logistics activities.

2. Sports logistics time limits strictly.

Competition sports activities are generally after the organizing committee determined between many times, every link content, each time arrangement, in addition to very special circumstances are is not change characteristics, which requires the sports activities required prepared materials must be on time in position. Therefore, the distribution of the sports logistics to arrange orderly according to this time, to ensure that every logistics links on time, smoothly and efficiently.

3. Sports total logistics business, logistics cost is high, the phase change there.

Sports activity is ready to period before the construction long, game activity concentration, the time is short, involve wide; After the game the features of the derived function, so need to various material is enough, the equipment is perfect, in order to meet the needs of the sports activities. In the Beijing Olympic Games in 2008, the investment of the logistics total cost is more than 400 yuan, the staff participating supplies logistics spending to nearly 600 million yuan, after the meeting spending of sports activities consumables, waste 1.2 billion multivariate logistic expenses, sports activity, the logistics business decision is very big, the organizer of sports activities it need a lot of logistics investment, and for logistics enterprise is brought a lot of benefits.

From the sports event sports physical distribution total business volume looked quantity is astonishing, but from the competition different time, the business volume has presented the gradual change. For instance, in front of the match we must complete the facility the construction, the competition facility, the equipment are transported massively, Physical distribution quantity is quite centralized; But after the match, we must realize abandons the physical distribution and the reverse logistics, the logistics service increase again.

4. Sports logistics professional level requirement is high, and safety is strong, and always pays attention to the green, environmental protection.

Competition sports event is various, shapes of required equipment are respectively different; The sports media needs the spare parts excellent, needs the physical distribution enterprise to carry on the different packing to it, in view of the logistics entity characteristic, carries on the transportation and loading and unloading, guarantees the transportation commodity the imperfection. The logistics enterprise is very strict, not only to have the logistics personnel of high quality, strict management system, more to have professional delivery ability, to ensure that the whole logistics process safety items.

At the same time, in society's all the various trades and occupations all advocated the green environmental protection, the physical distribution industry is not exceptional, also initiates the green logistics.

All the world has started a green logistics for the corresponding technology research, from the packaging and distribution path choice to solve the energy and protect the environment (Pingping Zhang and Yanrong Cui 2011). Because sports activities to be able to reflect a country, a nationality spirit civilization, culture, therefore, not only to ensure a successful event, more important is to put the "energy saving, environmental protection, green" campaign theme it down and implement the green logistics process.

15.2.4 The Status of Implementation of Chinese Events Sports Logistics

Presently, our country sports obtained the country and the local authority take highly, the sports increased unceasingly, cause the sports logistics also to raise the important program. Although, our country's sports logistics industry obtained the fast development, but compared with the international on sports logistics still has certain disparity, exposed some questions.

1. Because our country sports logistics lacks the plan which should have, causes the sports logistics establishment layout not to be unreasonable, the logistics equipment is imperfect, supply and demand contradictory pressure high.
2. The lack of sports logistics professionals leads to enterprise management level not high, logistics power shortage, logistics technology update lags behind, unable to adapt to the needs of the sports logistics.

The logistics enterprise in the sports logistics market demand, lacks the management, the technology and the management talented person, that cause the logistics business management aspect to be in the inferiority directly. The sports logistics flow is relatively complex, the logistics entity difference is big, the spatial transportation distance is long, needing advanced logistics technical and so on management technology, intelligence transmission to support, the short period logistics business volume is big, that request that logistics enterprise has certain strength, satisfying the need of the sports logistics. Our country logistics company is at present multitudinous, but the scale, the technology and the management level are irregular, the level not one, cause some large amount service not to dare to meet the list, the logistics enterprise loses the opportunity of creating the brand.

3. Lack of the strength of the logistics outsourcing companies, the logistics company failed to fully realize the importance of logistics alliance.

Sports general logistics outsourcing to have actual strength is the international logistics companies, domestic logistics enterprise business is less; At the same time sports logistics is often the major logistics activities, needing to implement multiple, complicated logistics activities, and sometimes is finished difficultly a logistics business by a logistics company alone. Modern logistics in China the enterprise did not attach importance to the logistics companies in the form of alliance, and realize the logistics enterprise work division between the cooperation, to finish the.

15.3 Framework of Theory About Chinese Events Sports Logistics

15.3.1 The Conditions That Sports Logistics Theory System Should Satisfy

Sports logistics theory comes from sports practice, and many times repeatedly is through the combining theory with practice, and finally form a system. Because the theory system is sports logistics theory according to certain elements of the logical relationship of the formation of the combination of the whole, which in theory structure must meet certain requirements.

1. Sports logistics theory system should be complete, comprehensive content, cover the entire sports logistics activities.
2. With a scientific perspective to study the theory, ensuring scientific in the physical theory of logistics system planning, and logic.
3. Due to the high physical activity requirements, complexity of the stage, so to clear a level, features a clear theoretical system.
4. Because the entire sport activity is dynamic, so pay attention to flexible logistics system of physical education.
5. Sports logistics theory system is in order to solve the real problems in sports activities, so better maneuverability.

15.3.2 Service Object

The establishment sports physical distribution theory system is wants better to serve for the sports sports event, looked from the concrete range of service, the demand object mainly has the sports physical distribution highest policy-maker, the operation superintendent, physical distribution operation business, the corresponding commodity supplier and the direct user (Peifeng Zhang and Xiaoling Zeng 2010). The concrete content see Table 15.1:

15.3.3 Competition Sports Logistics Theory System Framework

1. Sports logistics theory frame

The sports event sports physical distribution theory frame mainly by the basic theory, the application theory and the expansion theory is composed, as shown in Fig. 15.1:

The basic theory. the basic theory contains to the sports physical distribution concept, the characteristic and so on the elementary knowledge summary and the

Table 15.1 Service object of our country sports logistics

Service object	The main contents
Top decision-makers	Event the most senior decision-makers, responsible for overall planning and control of information
Operations managers	Events related to department managers, the plan for supplies and equipment, procurement, organization and control the implementation
Logistics operators	Domestic and foreign logistics companies to provide logistics services
Material suppliers	Suppliers, sponsors, and related businesses
Users	Athletes, coaches, medical personnel, sports teams leadership

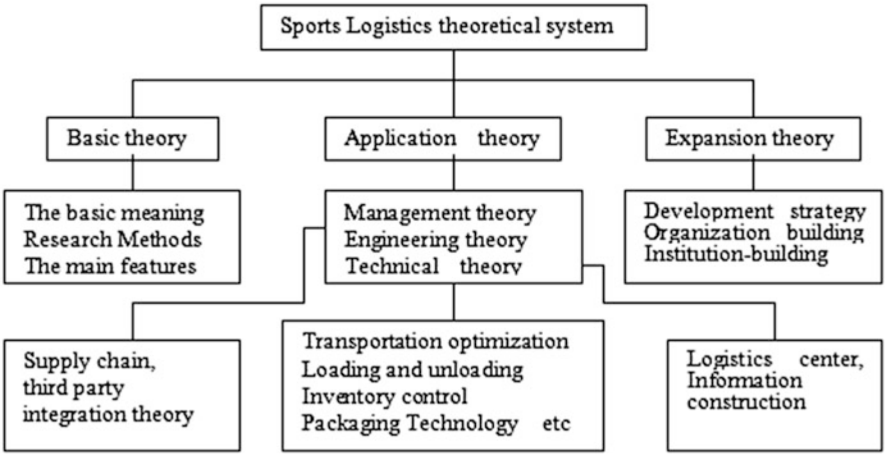


Fig. 15.1 Sports logistics theory frame system

introduction, makes the support for the better understanding other branch theory, only then the basic theory consummates unceasingly, can guarantee the sports physical distribution develops with steady steps.

Applied theory. sports logistics theory, is the core of guiding physical logistics activities properly, is the main means to translate theory into practice, through the practice of implementation of specific laws (Huanyu Qin et al. 2010). From the three aspects of management, technical, engineering, using various advanced technologies of logistics activities. This state of the art technology: supply chain management, integrated management, and third-party management theory; optimizing inventory control, packaging, transportation, handling, network planning, technical theory; there are physical logistics projects such as the construction of logistics center construction, information theory.

Extension theory. sports logistics expansion theory is to follow the future development trend of logistics and technology law, made to the physical logistics of long-term scientific planning; is based on the premise of basic theory and applied

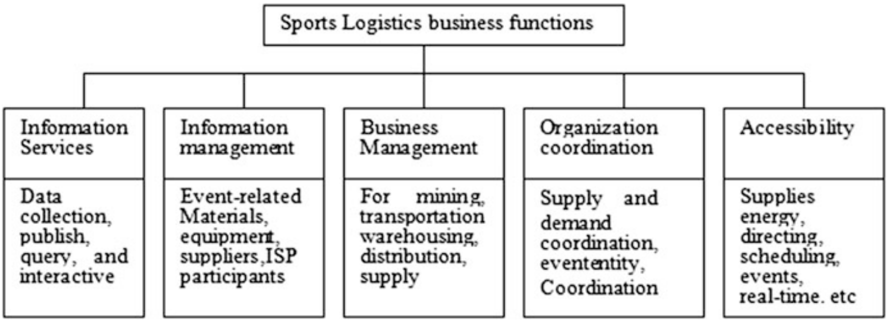


Fig. 15.2 Sports logistics business functions

theory of in-depth study and exploration on sports logistics through time and practice to verify the correctness of technology eventually accumulate into basic theory and applied theory.

2. The framework of sports logistics business functions

By determining the physical object structure and theory system frame of logistics services, summing up the physical logistics should implement business functions including: physical logistics information service, information management, business management, organization, coordination, and other decision of five functions (see Fig. 15.2), only these five functions coordinate the work, in order to achieve maximum physical logistics functions and benefits.

15.4 Suggestions for Developing Sports Logistics

Chinese sports logistics must face up to the present situation of logistics enterprise, to overcome difficulties, and carry actively out effective strategies to adapt quickly to the market demand of the sports events,

1. Continue to benefit from science and technology, improving the logistics management information system.

Sport events logistics information uncertainty requires enterprises to establish information management systems for timely response and continued use of the network, information, and other advanced technological innovations in science and technology and equipment. Through the introduction of foreign talents and advanced equipment, means of strengthening technical training talent, to raise the level of modernization of logistics enterprise's core technology and management, logistics, logistics platforms, regional sports information communication and interaction, information sharing, ensure sport smoother, coordination of logistics tasks.

2. To develop high quality talents, and satisfy the demand of logistics industry

The physical distribution enterprise is the emerging profession which our country develops, serious lack of professional talent, must speed up the cultivation of logistics personnel, meeting logistics needs. By continuing to strengthen knowledge and working knowledge of logistics theory to in-depth research, fully mobilize the logistics enterprises, universities, research institutions to work together. On the personnel training in colleges and universities and training in logistics enterprises, should aim at improving the practical ability, scientific and rational development of training programmes, logistics theory to the total rejection of backward technology and teaching methods and teaching means, emancipate the mind, change attitudes, promote various types of logistics education, improve quality of logistics management and logistics technicians.

3. Establish physical distribution strategy organization, make the whole plan to the entire sports event sports physical distribution, strengthen appraisal to the physical distribution organization.

The sports physical distribution to the large-scale sports event, is responsible for the entire sports commodity and the equipment transportation, the storage, the allocation and so on many links, the sports physical distribution whether can guarantee safe, on time, carried on smoothly, decides the sports event convened success or failure, therefore, to establish the powerful sports physical distribution strategy organization that was responsible for the entire sports event sports physical distribution the whole plan and the unification mixes. Because the sports physical distribution service total quantity is big, the physical distribution entity is diverse, is complex, that needs tender and so on many kinds of forms, to choose the best physical distribution enterprise to participate in the cooperation, and carries on the appraisal to the physical distribution organization; Like this already can guarantee the physical distribution enterprise splendid completes the task, also can promote the domestic physical distribution through the physical distribution enterprise competition the level of development (Haihui Xiao and Jinchun Wu 2010).

4. The country must strengthen the infrastructure construction unceasingly, consummate various local physical distribution space layout.

The domestic public logistics information service can guarantee that logistics information flow, but the logistics enterprise information system that is perfect, strength and strong, also can not get away from national infrastructure support. Sports activities is the host city for the layout of a space, sports activities all the necessary materials, to realize through space and time on the logistics of displacement, distribution is the sea, land, air many of the means of transportation comprehensively. Therefore, only then physical distribution infrastructure in and so on conformity existing transportation, warehousing foundations, rests on various local population quantity, the city scale carries on the whole plan to the physical distribution transportation network, consummates the transportation network layout, realizes the multi-transport mode combined transport; Simultaneously strengthens the physical distribution allocation center, the physical distribution garden area relay combined transport infrastructure construction, guarantees in sports event period each item of facility to arrive punctually, and the sports event to carry on normally.

15.5 Conclusion

Nowadays, the sports event large scale, modernized as well as the internationalization trend of development, needs modern physical distribution better integration to the sports. The sports event has provided the good turning point for the modern physical distribution development, needs National All levels of Governments to give the high value, the enlarge domestic physical distribution development support dynamics, at in the sports event plan, in the national physical distribution facility plan and the construction pays great attention to both to coordinate. The physical distribution enterprise speeds up the physical distribution informationization plat form the construction, the synthesis utilizes each kind of vanguard technology, enhances the sports physical distribution the service horizontal and the international competitive power, impels the sports physical distribution the fast development.

References

Danqing Liang, Hongyi Jia (2008) For our big sports events of logistics problems [J]. Chin Mark 1(19):23–24

Haihui Xiao, Jinchun Wu (2010) On the strategy of the logistics operation in the major sports events-to the 2010 Guangzhou Asia n games as an example [J]. Logist Technol 1(03):104–106

Huanyu Qin, Yihong Ru, Xiuquan Xu (2010) Large-scale comprehensive sporting events logistics system analysis of research [J]. Logist Technol 29(05):43–45

Peifeng Zhang, Xiaoling Zeng (2010) Based on large-scale event sport logistics theory [J]. Sports Adult Educ 26(02):11–12

Pingping Zhang, Yanrong Cui (2011) Sports logistics development opportunities and challenges [J]. China Bus 1(06):107–108

Qinlin Deng (2006) The Olympic logistics to Guangzhou Asian games the enlightenment of the logistics [J]. J Guagzhou Inst Phys Educ 25(6):20–22

The State Council (2008) The logistics industry adjust and revitalize the planning of the notice [N]. Modern Logist Newsp, pp 03–17

Xianwu Li (2005) The modern logistics theory analysis of the application of the sports events in the large [J]. J Xian Sports Inst 23(3):25–28

Yong Hu, Jing Jia (2010) Sporting events in the logistics characteristics and logistics planning [J]. China Bus 1(28):139–140

Yongsheng Dong (2006) Beijing Olympic Games to the sports industry in China with the opportunities and challenges [J]. Zhanjiang Norm Univ 27(6):100–103

Zhongfu Cui pay (2007) Our modern logistics management mechanism of the research and exploration [J]. China Logist Procure 1(15):14–16

Chapter 16
Dual Selection of Competing Channels Based on Internet Channel and Power Structures

Jun Liu and De-qing Tan

Abstract In competing channels with a dominant retailer, the article first calculates equilibrium outcomes under three power structures. Then it analyzes the effects of the reservation price difference and internet channel on channel members' profits. Finally, it investigates dual selection equilibriums of competing channels based on internet channel and power structures. The study shows that with the enhancement of consumer preferences for internet channel, the possibility of opening an internet channel will increase for manufacturer. However, the manufacturer's profits may not increase when it opens an internet channel.

Keywords Competing channels • Dominant retailer • Game theory • Internet channel • Power structures

16.1 Introduction

In recent years, news reports about the conflict between dominant retailers and manufacturers not only often have appeared in various media, but also have received the extensive attention from marketing scholars. Some recent research has focused on the effect of dominant retailers on other channel members. For example, Bloom and Perry used Compustat data to study nonfoods manufacturers, and found that manufacturers may not perform relatively as well financially when they have Wal-Mart as a primary customer (Bloom and Perry 2001). In a bargaining frame, Iyer and Villas-Boas found that a dominant retailer might be beneficial to other channel members (Iyer and Villas-Boas 2003). Raju and Zhang developed a

---

J. Liu (✉) • D. Tan
School of Economics and Management, Southwest Jiaotong University, Chengdu, China
e-mail: liu_jun79@126.com

E. Qi et al. (eds.), *The 19th International Conference on Industrial Engineering and Engineering Management*, DOI 10.1007/978-3-642-37270-4_16,

channel model in the presence of a dominant retailer to examine how manufacturer coordinate the channel (Raju and Zhang 2005). In a channel model of competing manufacturers and competing multiproduct retailers, Dukes, Gal-Or and Srinivasan found that manufacturers may get increased profits when a retailer gains an exogenous cost advantage over another retailer (Dukes et al. 2006). Jerath studied the effect of two dominance strategies of power retailers (market dominance and channel dominant) on other channel members' profits (Jerath 2008). Dukes, Geylani and Liu investigated the diverging incentives for product quality in a channel of two asymmetric retailers, and found that the low service dominant retailer benefits from quality reduction that is detrimental for the other members of the channel (Dukes et al. 2010).

Another stream has focused on how other channel members strategically react to the emergence of dominant retailers. For instance, Geylani, Dukes and Sinivasan characterized a theoretical model with a dominant retailer and a weak retailer, in which the dominant retailer has the power to dictate the wholesale price while the manufacturer sets the wholesale price for the weak retailer, and illustrated a strategic manufacturer response to a dominant retailer. They found that the dominant retailer can gain more profits than the weak retailer by engaging in joint promotions and advertising (Geylani et al. 2007). In a channel model with a large retailer, Pu, Shi and Ling investigated the effect of the direct marketing on the retailing channels and the condition that the manufacturer use direct marketing to gain more profits (Pu et al. 2007). Dukes, Geylani and Srinivasan showed that if the dominant retailer has the channel power to determine its assortment, it can choose strategic assortment reduction which will lead to lower consumer surplus (Dukes et al. 2009). Ailawadi et al. examined incumbent retailers' reactions to a Wal-Mart entry and the impact of their reactions on the retailers' sales, and found that incumbent retailers will suffer significant sales losses when Wal-Mart enters the local market. They showed that incumbent retailers' sales outcomes are significantly affected by their reactions (Ailawadi et al. 2010).

However, game sequence is supposed mostly Manufacturer Stackeberg in these literatures. As in Choi, Lee et al., Trivedi and Dong et al., according to the pricing order of distribution channel, power structure is divided into three categories: Manufacturer Stackeberg (MS), Retailer Stackeberg (RS) and Vertical Nash (VN) (Choi 1991, 1996; Lee and Staelin 1997; Trivedi 1998; Dong et al. 2009). At this point, what kind of power structures should channel members choose? In addition, with the rapid developments of internet technology and logistics industry, it is more common among manufacturers selling their products through internet channel. Then, will the manufacturer open an internet channel if it doesn't consider the construction cost of internet channel under three power structures? In this paper, we will analyze the effects of the reservation price difference and internet channel on channel members' profits, and investigate dual selection of competing channels under internet channel and power structures.

The rest of this paper is organized as follows: we describe the basic model and assumptions in Sect. 16.2. Then, Sect. 16.3 is devoted to calculate equilibrium outcomes under three power structures. In Sect. 16.4, we explore the effects of the reservation price difference and internet channel on members' profits. Section 16.5 investigates dual selection equilibrium of competing channels. We finally summarize the main conclusions in Sect. 16.6.

16.2 Basic Model and Assumptions

We suppose that a manufacturer M sells their products to the local market in a region through a dominant retailer R_1 and a weak retailer R_2. In the distribution channel, dominant retailer is dominant to the manufacturer with its huge store size. So the manufacturer loses the pricing power of wholesale price w_1 for dominant retailer. Similar to Pu et al. and Geylani et al., we suppose w_1 is exogenous while the manufacturer directly decides wholesale price w_2 for weak retailer. It is assumed that the local consumers are uniformly distributed on [0, 1] with dominant retailer and weak retailer located at 0 and 1 respectively. Let v_1 and v_2 represent consumer reservation price to traditional retail channel and internet channel respectively. It is assumed that $v_i(i = 1, 2)$ is sufficiently large so that every consumer can buy one unit of the product. Let $v = v_1 - v_2$ $(v_2 \leq v_1)$ represents the reservation price difference. The total market demand is normalized to 1, without loss of generality.

Because consumers prefer to go shopping at closer retail stores, t represents consumer loyalty for retailer and $x \in [0, 1]$ represents the distance from a consumer to dominant retailer. Given retail prices p_1 and p_2, consumer's surplus utility of buying one unit product from both dominant retailer and weak retailer is $U_1 = v_1 - tx - p_1$ and $U_2 = v_1 - t(1 - x) - p_2$ respectively. The indifferent consumer is located at $\bar{x}$ such as $U_1(\bar{x}) = U_2(\bar{x})$. Therefore, when the manufacturer doesn't open an internet channel, both retailers' market demands are expressed by

$$q_1 = \bar{x} = (p_2 - p_1 + t)/(2t)$$
$$q_2 = 1 - \bar{x} = (p_1 - p_2 + t)/(2t) \tag{16.1}$$

Manufacture and retailers profits are $\pi_m = w_1 q_1 + w_2 q_2$ and $\pi_{ri} = (p_i - w_i)q_i$.

When an internet channel is opened, given retail price p_3 which the manufacturer chooses, consumers' surplus utility of buying one unit product from it is $U_3 = v_2 - p_3$. The indifferent consumers are located at x_1 and x_2 respectively such as $U_1(x_1) = U_3(x_1)$ and $U_2(x_2) = U_3(x_2)$, or

$$x_1 = (p_3 - p_1 + v)/t$$
$$x_2 = (p_2 - p_3 - v + t)/t \tag{16.2}$$

Therefore, market demands of retailers and internet channel can be written as respectively

$$
\begin{aligned}
q_1 &= x_1 = (p_3 - p_1 + v)/t \\
q_2 &= 1 - x_2 = (p_3 - p_2 + v)/t \\
q_3 &= x_2 - x_1 = (p_1 + p_2 - 2p_3 - 2v + t)/t
\end{aligned} \tag{16.3}
$$

Manufacture and retailers' profits are $\pi_m = w_1 q_1 + w_2 q_2 + p_3 q_3$ and $\pi_{ri} = (p_i - w_i) q_i$.

Because a weak retailer whose store size is small doesn't have the right to choose power structure, we suppose that both manufacturer and dominant retailer choose competing channels' power structure. For consumers, the reservation price v_2 to internet channel doesn't exceed the reservation price v_1 to traditional retail channel. Let $0 < v < t$ so that competing channels has pure strategy Nash equilibrium outcomes under three power structures. In the following section, we calculate equilibrium outcomes under three power structures with opening an internet channel.

16.3 Three Power Structures and Equilibriums

16.3.1 Power Structure MS

Under power structure MS, the manufacturer takes both retailers' retail price reaction function into consideration for its pricing decision in order to maximize its profits. Based on this, both retailers determinate retail prices of their products. Firstly, both retailers' reaction functions can be derived from first-order optimality conditions $\partial \pi_{ri}^{MS} / \partial p_i^{MS} = 0 (i = 1, 2)$:

$$
\begin{aligned}
p_1^{MS} &= (w_1 + p_3^{MS} + v)/2 \\
p_2^{MS} &= (w_2^{MS} + p_3^{MS} + v)/2
\end{aligned} \tag{16.4}
$$

Then, the manufacturer's profits function can be written as from (16.4):

$$
\begin{aligned}
\pi_m^{MS}(w_2^{MS}, p_3^{MS}) = \big[& w_2^{MS}(v + 2p_3^{MS} - w_2^{MS}) \\
& + 2p_3^{MS}(t - v + w_1 - p_3^{MS}) + w_1(v - w_1) \big] / (2\mathrm{t})
\end{aligned} \tag{16.5}
$$

From first-order optimality conditions, the manufacturer's retail price and wholesale price can be obtained $w_2^{MS} = w_1 + t$ and $p_3^{MS} = w_1 + (2t - v)/2$. The corresponding Hessian matrix is negative definite after an inspection. Therefore, channel members' profits can be derived:

$$
\begin{aligned}
\pi_m^{MS} &= w_1 + \left(v^2 - 2tv + 2t^2\right)/(4\mathrm{t}) \\
\pi_{r1}^{MS} &= (2t + v)^2/(16\mathrm{t}) \\
\pi_{r2}^{MS} &= v^2/(16\mathrm{t})
\end{aligned} \tag{16.6}
$$

16.3.2 Power Structure RS

Under power structure RS, both retailers choose retail prices of their products given the manufacturer's price reaction functions so as to maximize their profits. Based on this, the manufacturer chooses wholesale and retail prices of the internet channel. Specifically, given retail margins (rm_1^{RS}, rm_2^{RS}), the manufacturer's price reaction functions firstly can be derived from their first-order optimality conditions $\partial\pi_m^{RS}/\partial w_2^{RS} = 0$ and $\partial\pi_m^{RS}/\partial p_3^{RS} = 0$:

$$
\begin{aligned}
w_2^{RS} &= t + w_1 + p_1^{RS} - p_2^{RS} \\
p_3^{RS} &= \left(t - v + w_1 + p_1^{RS}\right)/2
\end{aligned} \tag{16.7}
$$

Then both retailers' profits functions can be written as from (16.6):

$$
\begin{aligned}
\pi_{r1}^{RS} &= \left(p_1^{RS} - w_1\right)\left(t + v + w_1 - p_1^{RS}\right)/(2t) \\
\pi_{r2}^{RS} &= \left(2p_2^{RS} - p_1^{RS} - t - w_1\right)\left(p_1^{RS} - 2p_2^{RS} + t + v + w_1\right)/(2t)
\end{aligned} \tag{16.8}
$$

From first-order optimality conditions $\partial\pi_{ri}^{RS}/\partial p_i^{RS} = 0(i = 1, 2)$, both retailers' retail prices can be obtained $p_1^{RS} = w_1 + (t + v)/2$ and $p_2^{RS} = w_1 + (3t + 2v)/4$: The corresponding Hessian matrix is negative definite after an inspection. Therefore, channel members' profits can be derived:

$$
\begin{aligned}
\pi_m^{RS} &= w_1 + \left(2v^2 - 6tv + 9t^2\right)/(16\mathrm{t}) \\
\pi_{r1}^{RS} &= (t + v)^2/(8\mathrm{t}) \\
\pi_{r2}^{RS} &= v^2/(8\mathrm{t})
\end{aligned} \tag{16.9}
$$

16.3.3 Power Structure VN

Under power structure VN, all channel members make decisions at the same time. On one hand, the manufacturer determinates wholesale price w_2^{VN} and retail price p_3^{VN} conditional on both the observed retail price of the competitor's product and the observed retailer's margin on its own product so as to maximize its profits.

On the other hand, given the manufacturer's price (w_2^{VN}, p_3^{VN}), both retailers choose retail price (p_1^{VN}, p_2^{VN}) in order to maximize their profits. From price reaction functions (16.4) and (16.7), the corresponding wholesale and retail price are

$$
\begin{aligned}
w_2^{VN} &= w_1 + 2t/3 \\
p_1^{VN} &= w_1 + (t + v)/3 \\
p_2^{VN} &= w_1 + (2t + v)/3 \\
p_3^{VN} &= w_1 + (2t - v)/3
\end{aligned}
\tag{16.10}
$$

Therefore, channel members' profits can be derived:

$$
\begin{aligned}
\pi_m^{VN} &= w_1 + 2(v^2 - 2tv + 2t^2)/(9\mathrm{t}) \\
\pi_{r1}^{VN} &= (t + v)^2/(9\mathrm{t}) \\
\pi_{r2}^{VN} &= v^2/(9\mathrm{t})
\end{aligned}
\tag{16.11}
$$

16.4 Equilibriums Analysis

In this section, we investigate the effect of the reservation price difference and internet channel on channel members' profits under three power structures respectively.

Obviously, the smaller the reserve price difference v, the higher the evaluation of consumers' internet channel shopping. So, the effect of opening internet channel on traditional retail channels is higher than that of not opening it. The analysis gives the following result.

Proposition 16.1 $\partial\pi_m/\partial v < 0, \partial\pi_{ri}/\partial v > 0 (i = 1, 2)$.

Proposition 16.1 shows that manufacturer's profits will increase while retailers suffer losses with reduction of the reservation price difference. The reason for this result is as follows: the reduction of the reservation price difference means the enhancement of consumer preferences for internet channel, and more consumers will shift in purchasing products from retailer channels to internet channel. As this time, the manufacturer can set a higher retail price, and get a bigger market share and higher profits. However, both retailers will set lower retail margins and prices so as to reduce the losses of their market share, and get lower profits.

When an internet channel isn't opened, the manufacturer whose distribution channels are highly subject to dominant retailer loses the pricing power of wholesale price for dominant retailer. With the opening of internet channel, the manufacturer is able to accurately grasp its product demand information and strengthen the control of its distribution channels. However, can the manufacturer gain more profits through opening an internet channel? Will the manufacturer open internet channel if it doesn't consider the construction cost of internet channel under three power structures? The analysis gives the following result.

Proposition 16.2 *Let* Δ *represents channel members' profits difference through opening an internet channel*. (1) $\Delta\pi_{ri} < 0 (i = 1, 2)$. (2) a. $\Delta\pi_{ri}^{RS} > 0$. b. *When* $0 < v < 0.293t (0.293t < v < t)$, $\Delta\pi_m^{MS} > 0 (< 0)$. c. *When* $0 < v < 0.485t (0.485t < v < t)$, $\Delta\pi_m^{VN} > 0 (< 0)$.

Proposition 16.2 shows that opening an internet channel will intensify price competition among retailers and reduce their profits. The manufacturer gets higher profits under power structure RS while it may not get higher profits under both power structure MS and VN. When $0.293t < v < t$, the manufacturer will get lower profits under power structure MS. When $0.485t < v < t$, it will get lower profits under power structure VN.

16.5 Dual Selection Equilibrium

Our analysis shows that the manufacturer will open an internet channel under power structure RS while it may not open it under both power structure MS and VN. Based on this, what kind of power structure do both the manufacturer and dominant retailer choose? From two angles of internet channel and power channel, it is analyzed and derived:

Proposition 16.3 (1) *When* $0 < v < 0.485t$, $\pi_m^{VN} < \pi_m^{MS} < \pi_m^{RS}$, $\pi_{r1}^{VN} < \pi_{r1}^{RS} < \pi_{r1}^{MS}$. (2) *When* $0.485t < v < t$, $\pi_{r1}^{RS} < \pi_{r1}^{VN} < \pi_{r1}^{MS}$. *Meanwhile, if* $0.485t < v < 0.634t$, *then* $\pi_m^{VN} < \pi_m^{MS} < \pi_m^{RS}$. *Otherwise*, $\pi_m^{VN} < \pi_m^{RS} < \pi_m^{MS}$.

Proposition 16.3 states that under the influence of the reservation price difference v and consumer loyalty t, when $0 < v < 0.293t$, dual selection equilibrium is (open, MS) and (open, RS). When $0.293t < v < 0.485t$, dual selection equilibrium is (no open, MS) and (open, RS). When $0.485t < v < t$, dual selection equilibrium is (no open, MS). Obviously, with the enhancement of consumer preferences for internet channel, the possibility of opening an internet channel will increases for manufacturer.

16.6 Conclusions

In competing channels with a dominant retailer, our analysis obtains the following results: (1) With the enhancement of consumer preferences for internet channel, both retailers will get lower profits, and the possibility of opening an internet channel will increase for manufacturer. (2) Opening an internet channel for the manufacturer will reduce both retailers' profits while it may not increase the manufacturer's profits.

Acknowledgment This work is supported by Science and Technology Support Program of Sichuan Province (2011SZ0228).

References

Ailawadi KL, Zhang J, Krishna A, Kruger MW (2010) When Wal-Mart enters: how incumbent retailers react and how this affects their sales outcomes. J Market Res 47(4):577–593

Bloom PN, Perry VG (2001) Retailer power and supplier welfare: the case of Wal-Mart. J Retail 77(3):379–396

Choi SC (1991) Price competition in a channel structure with a common retailer. Market Sci 10(4):271–296

Choi SC (1996) Price competition in a duopoly common retailer channel. J Retail 72(2):117–134

Dong L, Narasimhan C, Zhu K (2009) Product line pricing in a supply chain. Manage Sci 55(10):1704–1717

Dukes AJ, Gal-Or E, Srinivasan K (2006) Channel bargaining with retailer asymmetry. J Market Res 43(1):84–97

Dukes AJ, Geylani T, Srinivasan K (2009) Strategic assortment reduction by a dominant retailer. Market Sci 28(2):309–319

Dukes AJ, Geylani T, Liu YC (2010) Dominant retailers' incentives for product quality. Working Paper, University of Southern California

Geylani T, Dukes AJ, Sinivasan K (2007) Strategic manufacturer response to a dominant retailer. Market Sci 26(2):164–178

Iyer G, Villas-Boas JM (2003) A bargaining theory of distribution channels. J Market Res 40(1):80–100

Jerath K (2008) Essays on the impact of power retailers on distribution channels, PhD dissertation, University of Pennsylvania

Lee E, Staelin R (1997) Vertical strategic interaction: implications for channel pricing strategy. Manage Sci 16(3):185–207

Pu XJ, Shi Q, Ling LY (2007) Effect of direct marketing on retailing channels where large retailers exist. J Manage Sci China 10(6):49–56 (in Chinese)

Raju J, Zhang ZJ (2005) Channel coordination in the presence of dominant retailer. Market Sci 24(2):254–262

Trivedi M (1998) Distribution channel: an extension of exclusive retailer ship. Manage Sci 44(7):896–909

Chapter 17
Supply-Chain Ripple Effect Under Duopoly Retailers' Operating Uncertainty with Two Different Contracts

Qian Zhao, Hong Chen, and Zhong-he Wu

Abstract In the past few years, the problem of supply chain uncertainty attracts more and more people's attention. However, people place almost all their focus on demand and supply uncertainties and seldom study operating uncertainties of supply chain. This paper investigates the problem of operating uncertainty from the aspect of oligarch competition. First, we examine the supply chain, one supplier and two competing retailers, coordination mechanism with a revenue sharing contract and a quantity-discount contract under a deterministic environment. Second, we assume the retailers' uncertain behaviors, price competition, occur. Then, we discuss the supply-chain ripple effect due to duopoly retailers' price competition with the two different contracts. By comparing the intensity of the supply-chain ripple effect with the different contracts, we find out that the quantity-discount contract can be well suited to the uncertain environment.

Keywords Coordination • Operating uncertainty • Quantity-discount contract • Revenue sharing contract • Supply-chain ripple effect

17.1 Introduction

As the market competition growing, the uncertainties and risks that confront firms become more and more. Therefore, the problems of how to deal with supply chain uncertainties and limit the risks due to those uncertainties attract many scholars' attention. Tsiakis et al. (2001) considered a two-stage stochastic programming model for supply chain network design under demand uncertainty. Qi et al. (2004) introduced a model to coordinate an one-supplier-one-retailer supply chain that

Q. Zhao (✉) • H. Chen • Z. Wu
School of Economics and Management, University of Electronic Science and Technology of China, Chengdu, China
e-mail: cliffhf88@163.com

E. Qi et al. (eds.), *The 19th International Conference on Industrial Engineering and Engineering Management*, DOI 10.1007/978-3-642-37270-4_17,

experiences the demand disruption once the production plan had been made. Yu et al. (2005, 2006) applied a modified buy back contract and a modified wholesale price contract to coordinate a supply chain under the demand disruption. Hult et al. (2010) used real options theory to investigate supply chain investment decisions when facing high levels of risk uncertainty. Khan et al. (2009) developed framework for design-led risk management and thus presents a case for recognizing design as more than a creative function in the supply chain but as a platform to manage risk in supply chains. Xiao et al. (2010) developed a game theoretic model of a three-stage supply chain consisting of one retailer, one manufacturer and one subcontractor to study ordering, wholesale pricing and lead-time decisions, where the manufacturer produces a seasonal/perishable product. Anupindi and Akella (1993) addressed the operational issue of quantity allocation between two uncertain suppliers and its effects on the inventory policies of the buyer. Based on the type of delivery contract a buyer has with the suppliers, they suggest three models for the supply process. Refik et al. (1999) analyzed a periodic review, single-item inventory model under supply uncertainty. Kouvelis and Milner (2002) studied the interplay of demand and supply uncertainty in capacity and outsourcing decisions in multi-stage supply chains. The authors above pay more attention to supply-chain demand or supply uncertainty. You et al. (2009) considered the risk management for mid-term planning of a global multi-product chemical supply chain under demand and freight rate uncertainty.

According to the literature above, we find that people pay more attention about demand or supply uncertainties, but seldom concentrate on operating uncertainties (Zhang and Sun 2005). However, operating uncertainties also deserve people's attention. Conventional operating uncertainties are due to supply chain members' uncertain behaviors. For example, a retailer cuts his retail price for sales promotion. His behavior will definitely stimulate customer demand. Then, he will enlarge his order quantity. Finally, his supplier faces the fluctuations of order quantity and takes on its risk. Chen (Chen et al. 2004; Wei et al. 2009; Zhao et al. 2011) calls this phenomenon in supply chain, one's influence on the others through a supply chain, as supply-chain ripple effect.

In this paper, we will discuss the supply-chain ripple effect due to retailers' operating uncertainties.

17.2 Basic Model

In our model there are three firms consisting a two-echelon supply chain, a supplier and duopoly retailers denoted by 1 and 2, respectively. The supplier has enough capacity to meet the retailers' demand. Before a selling season, the supplier offers the retailers a contract respectively; retailer i submits his order quantity q_i to the supplier at the wholesale price w_i which is the supplier offered. The retailers face a Bertrand problem as:

$$q_i = a - bp_i + \gamma p_j \quad i, j = 1, 2 \quad i \neq j \tag{17.1}$$

where $a, a > 0$, is the market scale. p_i represents retail price determined by retailer i in one-shot game. The difference $(b - \gamma), b > \gamma$, is negative correlated to the degree of product substitutable between the two retailers' products. That is, the smaller the difference, the more substitutable their products, and the more intense price competition. Let c_s denote the supplier's unit production cost, $c_s < w_i$. The retailer i's marginal cost per unit is c_s. Assume all unit costs and revenue functions is common knowledge for every member of the supply chain and all firms are risk neutral. Then, the channel revenue of retailer i is:

$$R_{r_i} = p_i q_i \quad i, j = 1, 2 \quad i \neq j \tag{17.2}$$

The profit of retailer i is

$$\Pi_{r_i} = (p_i - w_i - c_i) q_i \tag{17.3}$$

The profit of supplier is

$$\Pi_s = \sum_{i=1}^{2} (w_i - c_s) q_i \tag{17.4}$$

The supply chain's profit is

$$\Pi_{sc} = \sum_{i=1}^{2} (p_i - c_s - c_i) q_i \tag{17.5}$$

For the centralized problem, the optimal retail price for retailer i is

$$p_i^c = \frac{1}{2} \cdot \frac{a}{b - \gamma} + \frac{c_s + c_i}{2} \tag{17.6}$$

The optimal order quantity for retailer i is

$$q_i^c = \frac{1}{2} \left[a - b(c_i + c_s) + \gamma (c_j + c_s) \right], i, j = 2, i \neq j \tag{17.7}$$

For the decentralized problem, we derive the supplier's maximum profit

$$\Pi_s^d = \sum_{i=1}^{2} \left[\frac{1}{2} \cdot \frac{a}{b - \gamma} - \frac{c_s + c_i}{2} \right] q_i^d \tag{17.8}$$

and retailer i's maximum profit

$$\Pi_{r_i}^d = \left[p_i^d - \frac{1}{2} \cdot \frac{a}{b - \gamma} - \frac{c_s + c_i}{2} \right] q_i^d \tag{17.9}$$

where p_i^d is the retailer i's optimal retail price for the decentralized problems and $q_i^d = a - bp_i^d + \gamma p_j^d \quad i, j = 1, 2 \quad i \neq j$.

Before discussing the supply-chain ripple effect under an uncertain environment, we are going to address the supply chain coordination condition under a deterministic environment.

17.3 The Coordination of the Supply Chain

We introduce two different contracts, revenue sharing contracts and all-unit quantity-discount contracts, and check if the contracts can coordinate the supply chain or not.

17.3.1 Coordination with a Revenue Sharing Contract

Theorem 17.1 *Let $\phi_i, 0 < \phi_i < 1$, be the fraction of channel revenue the retailer i keeps, then the supplier earns $(1 - \phi_i)$ of channel revenue. For any given $(w_i, \phi_i), i = 1, 2$, if w_i satisfies*

$$w_i = \frac{(1 - \phi_i)}{b}(q_i^c - p_i^c) + \frac{\gamma}{b}(p_j^c - c_s + c_j) + c_s, \tag{17.10}$$

then supply chain can be coordinated. In addition, the retailer i's profit is

$$\Pi_{r_i} = \frac{q_i^c}{b}\left[b(\phi_i p_i^c - c_s - c_i) - (1 - \phi_i)(q_i^c - p_i^c) - \gamma(p_j^c - c_s + c_j)\right] \tag{17.11}$$

the supplier's profit is

$$\Pi_s = \sum_{i \neq j}^{2} \left[(1 - \phi_i)p_i^c - c_s\right] q_i^c \tag{17.12}$$

Proof With the revenue sharing contract, the first derivative of retailer i's profit function with respect to the retail price p_i is

$$\Pi_i^{r_i} = \phi_i(a - 2bp_i + \gamma p_j) + b(w_i + c_i), \tag{17.13}$$

where $\Pi_i^{r_i} = \partial \Pi_{r_i} / \partial p_i$. Substituting (17.10) into (17.13) then yields $\Pi_i^{r_i}(p_i^c, p_j^c) = a - 2bp_i^c + 2\gamma p_j^c + b(c_s + c_i) - \gamma(c_s + c_j)$ In centralized decision, when $p_i = p_i^c$, the supply chain's profit function satisfies the first order condition

$$\Pi_i^{sc} = a - 2bp_i^c + 2\gamma p_j^c + b(c_s + c_i) - \gamma(c_s + c_j) = 0, \tag{17.14}$$

where $\Pi_i^{sc} = \partial\Pi_{sc}/\partial p_i$. Form (17.14), we derive

$$\Pi_i^{r_i}(p_i^c, p_j^c) = 0 \tag{17.15}$$

(17.15) shows that the optimal retail price of the centralized decision satisfies the first order conditions of (17.3). It means that retailer i's optimal price decision of the decentralized decision is identical with the optimal price of the centralized decision, when Eq. (17.10) is satisfied. In conclusion, the revenue sharing contract can coordinate the supply chain. According to (17.10), (17.11) and (17.12) can be derived from (17.3) and (17.4).

17.3.2 Coordination with a All-Unit Quantity-Discount Contract

Theorem 17.2 *Consider a kind of all-unit quantity-discount contracts as*

$$w_i = \omega_i - kq_i \quad i = 1, 2 \tag{17.16}$$

where ω_i is initial wholesale price for retailer i and k, $k > 0$, is quantity-discount rate. For any given ω_i, if ω_i satisfies

$$\omega_i = 2k(a - bp_i^c + \gamma p_j^c) + \frac{\gamma}{b}(p_j^c - c_s - c_j) + c_s, \tag{17.17}$$

then supply chain can be coordinated. In addition, the retailer i's profit is

$$\Pi_{r_i} = \left[p_i^c - kq_i^c - \frac{\gamma}{b}(p_j^c - c_s - c_j) - c_i - c_s\right] q_i^c \tag{17.18}$$

the supplier's profit is

$$\Pi_s = \sum_{i \neq j}^{2} \left[kq_i^c + \frac{\gamma}{b}(p_j^c - c_s - c_j)\right] q_i^c \tag{17.19}$$

Proof Similarly as the proof of Theorem 17.1, the first derivative of retailer i's profit function with respect to the retail price p_i is

$$\Pi_i^{r_i} = [p_i - c_i - \omega_i + kq_i]q_i \tag{17.20}$$

Substituting (17.17) into (17.20) then yields

$$\Pi_i^{r_i}(p_i^c, p_j^c) = a - 2bp_i^c + 2\gamma p_j^c + b(c_s + c_i) - \gamma(c_s + c_j)$$

Form (17.14), we also derive

$$\Pi_i^{r_i}(p_i^c, p_j^c) = 0 \tag{17.21}$$

It means that retailer i's optimal price decision of the decentralized decision is identical with the optimal price of the centralized decision, when Eq. (17.17) is satisfied. In conclusion, the all-unit quantity-discount contract can coordinate the supply chain. According to (17.17), (17.18) and (17.19) can be derived from (17.3) and (17.4).

So, we address the coordination conditions of the supply chain with two contracts, respectively.

17.4 The Discussion of Supply-Chain Ripple Effect Under Retailers' Operating Uncertainty

The previous section shows the supply chain can be coordinated by the revenue sharing contract and the all-unit quantity-discount contract under a deterministic environment. But it is ideal conditions, almost not happening in the real world. Instead, in the real operational environment, there are a lot of uncertainties that members of supply chain need to cope with, such as the fluctuations of raw material price, marketing promotion, the destruction of logistics system, the risk of exchange rate, and so on. Among those uncertainties, the uncertainties due to firms' uncertain operating behaviors are called operating uncertainties. In market competitions, the most common uncertain operating behavior is adjusting product price arbitrarily due to firms' self-interests. Surely, this arbitrary behavior causes uncertainties of the supply chain. In fact, in a supply chain, when one member changes his operating behaviors arbitrarily, his change will not only affect itself, but also affect the other members. So, the effect of uncertainty will transmit along a supply chain from upstream to downstream or from downstream to upstream. This phenomenon of effect transmission is called supply-chain ripple effect (Chen et al. 2004). Supply-chain ripple effect usually results in firms' interest loss. The following, we will discuss the supply-chain ripple effect when operating uncertainty of the supply chain due to retailers' price competition occurs.

Based on the coordination we discuss in the previous section, we assume the following sequence of events occurs: One retailer adjusts his retail price arbitrarily due to his self-interests, the other retailer makes quick response to his competitor's action, the fluctuations of market demand occurs and the change of retailers order commitment to the supplier occurs. Furthermore, if interest improvement is less than the retailers' expected, they will accelerate price competition. It would result in even more aggravated levels of volatility of market demand. This process ultimately results in the fluctuations of retailers' order quantity. In addition, the fluctuations of the retailers' order quantity will bring more operating cost to the supplier.

In many cases, the supplier always becomes the victim of the retailers' market promotion, such as price discounts. The process, the retailers' competition affecting the supplier's performance finally, is called the supply-chain ripple effect.

To measure the intensity of supply-chain ripple effect, we need to introduce a concept in oligopoly theory, strategic complements and substitutes, first. The theory can be concluded as: in an oligopoly competition, if firm A's strategy adjustment raises firm B's marginal profitability, firm B will respond by a more aggressive strategy, in the time firm B regards his products as strategic complements. On the other hand, if firm A's strategy adjustment lowers firm B's marginal profitability, firm B will respond by a less aggressive strategy, in the time firm B regards his products as strategic substitutes (Bulow et al. 1985). In our model, the more aggressive strategy means that the retailers will adjust their retail price in the same direction and the less aggressive strategy means that the retailers will adjust their retail price in the opposite direction. Then, we will examine the intensity of supply-chain ripple effect with the two contracts, respectively.

Let Δp_i denote the range of retailers' price adjustment. From (17.1), the fluctuations of the retailers' order quantity is

$$|\Delta q| = \left|\sum_{i=1}^{2} \Delta q_i\right| = (b-\gamma)\left|\sum_{i=1}^{2} \Delta p_i\right| \tag{17.22}$$

Case 17.1 The leader of the supply chain, the supplier, adopts the revenue sharing contract to coordinate the supply chain.

The second derivative of retailer i's profit function with respect to the retail price p_i and p_j is

$$\Pi_{ij}^{r_i} = \phi_i \gamma$$

where $\Pi_i^{r_i} = \partial \Pi_{r_i}/(\partial p_i \partial p_j)$. When $\phi_i > 0, \gamma > 0$, then $\Pi_{ij}^{r_i} > 0$. According to Bulow's definition (Bulow et al. 1985), the two retailers regard their goods as strategic complements and take more aggressive strategies in market competition. They raise or cut their retail price at the same time, which results in the fluctuations of retailers order quantity as

$$|\Delta q|^* = \sum_{i=1}^{2} |\Delta q_i| = (b-\gamma)\sum_{i=1}^{2} |\Delta p_i| \tag{17.23}$$

From the perspective of supply-chain ripple effect, the retailers' behaviors caused the higher intensity of supply-chain ripple effect. This higher intensity of supply-chain ripple effect is called superposing supply-chain ripple effect (Chen et al. 2004). Obviously, the supplier takes on more order quantity risk, when he faces the superposing supply-chain ripple effect.

Case 17.2 The leader of the supply chain, the supplier, adopts the revenue sharing contract to coordinate the supply chain.

The second derivative of retailer i's profit function with respect to the retail price p_i and p_j is

$$\Pi_{ij}^{r_i} = \gamma - 2kb\gamma$$

The sign of $\Pi_{ij}^{r_i}$ is associated with the value of quantity-discount rate k. So, there are two following situations.

1°

$$\Pi_{ij}^{r_i} > 0 \Rightarrow 0 < k < \frac{1}{2b}$$

In this situation, when the value of k is set within the range $(0, 1/2/b)$, the two retailers regard their goods as strategic complements and take more aggressive strategies in market competition. They raise or cut their retail price at the same time, which results in the fluctuations of retailers order quantity as (17.23). The supplier faces the superposing supply-chain ripple effect.

2°

$$\Pi_{ij}^{r_i} < 0 \Rightarrow k > \frac{1}{2b}$$

In this situation, when the value of k is set within the range $(1/2/b, +\infty)$, the two retailers regard their goods as strategic and take less aggressive strategies in market competition. They adjust their retail price in the opposite direction, which results in the fluctuations of retailers order quantity as

$$|\Delta q|^{**} = \left|\sum_{i=1}^{2} \Delta q_i\right| = (b - \gamma)(|\Delta p_i| - |\Delta p_j|)^+ \tag{17.24}$$

With the same range of retailers' price adjustment, we derive from (17.23) and (17.24)

$$|\Delta q|^* > |\Delta q|^{**} \tag{17.25}$$

From (17.25), in contrast to the case of strategic complements, the retailers' price competition leads lower fluctuations of retailers order quantity in the case of strategic substitutes. In other words, the retailers' behaviors results in the lower intensity of the supply-chain ripple effect in the case of strategic substitutes than in the case of strategic complements. This lower intensity of supply-chain ripple effect is called offsetting supply-chain ripple effect (Chen et al. 2004). Comparing with the superposing supply-chain ripple effect, in the face of offsetting supply-chain ripple effect, the supplier will take less order quantity risk, because he faces lower fluctuations of retailers order quantity. It is low risk because the retailers' behaviors forms a mechanism for order quantity risk hedging as (17.24).

As a result, when the retailers' price competition occurs, the supplier will face the superposing supply-chain ripple effect, if he chooses the revenue sharing contract before. However, if the supplier chooses the all-unit quantity-discount contract before, he will face the offsetting supply-chain ripple effect, when the value of quantity-discount coefficient, k, is set within the range $(1/2/b, +\infty)$. Since the superposing supply-chain ripple effect cause high order quantity risk and the offsetting supply-chain ripple effect cause low order quantity risk, the all-unit quantity-discount contract is more suitable for the supplier, when facing retailers' price competition, comparing with the revenue sharing contract. In addition, the supply-chain ripple effect can be weakened by all-unit quantity-discount contract, but cannot be eliminated.

17.5 Conclusion

In this paper, we discuss the supply-chain ripple effect and contracts selection under the duopoly retailers' operating uncertainty: price competition. In the absence of price competition, revenue sharing contracts and all-unit quantity-discount contracts can both coordinate the supply chain in which there are one supplier and duopoly retailers. However, when two retailers compete with each other on price for their self-interests, the situation is becoming different. If the supplier adopts the revenue sharing contract, no mater what value the revenue sharing coefficient is, the retailers' price competition will lead higher fluctuations of retailers' order quantity, which means that the supplier always faces the superposing supply-chain ripple effect. The supplier will take on more order quantity risk. Yet, with the all-unit quantity-discount contract, the supplier can limit his risk by setting the value of quantity-discount coefficient within some specific range, because the supplier will face the offsetting supply-chain ripple effect. Hence, in the face of retailers' operating uncertainties, the supplier is best to choose the all-unit quantity-discount contract to coordinate the supply chain and limit his risk due to the supply-chain ripple effect.

Acknowledgment This research was partly supported by the National Natural Science Foundation of China (Grant No. 70773017).

References

Anupindi R, Akella R (1993) Diversification under supply uncertainty. Manage Sci 39(8):944–963

Bulow J, Geanakoplos J, Klemperer P (1985) Multimarket oligopoly: strategic substitutes and complements. J Pol Econ 93(3):488–511

Chen H, Li L, Wang J (2004) Quantifying the ripple effect in supply chain and generally describing the bullwhip effect. In: Proceedings of 2004 international conference on management science & engineering, Harbin, 2004, pp 967–972

Hult GTM, Craighead CW, Ketchen DJ Jr (2010) Risk uncertainty and supply chain decisions: a real options perspective. Decis Sci 41(3):435–458

Khan O, Christopher M, Burnes B (2009) The role of product design in global supply chain risk management. Int Ser Oper Res Manage Sci 124(2):137–153

Kouvelis P, Milner JM (2002) Supply chain capacity and outsourcing decisions: the dynamic interplay of demand and supply uncertainty. IIE Trans 34:717–728

Qi X, Bard JF, Yu G (2004) Supply chain coordination with demand disruptions. Omega 32(4):301–312

Refik G, Ebru O, Erkip N (1999) Analysis of an inventory system under supply uncertainty. Int J Prod Econ 59:377–385

Tsiakis P, Shah N, Pantelides CC (2001) Design of multi-echelon supply chain networks under demand uncertainty. Ind Eng Chem Res 40(16):3585–3604

Wei L, Chen H, Zhang J (2009) Asymmetric transmission of industry chain price ripple effect. Syst Eng Theory Pract 29(7):1–7

Xiao T, Jin J, Chen G, Shi J, Xie M (2010) Ordering, wholesale pricing and lead-time decisions in a three-stage supply chain under demand uncertainty. Comput Ind Eng 59(4):840–852

You F, Wassick JM, Grossmann IE (2009) Risk management for a global supply chain planning under uncertainty: models and algorithms. AIChE J 55(4):931–946

Yu H, Chen J, Yu G (2005) Supply chain coordination under disruptions with buy back contract. Syst Eng Theory Pract 25(8):38–43

Yu H, Chen J, Yu G (2006) Managing wholesale price contract in the supply chain under disruptions. Syst Eng Theory Pract 26(8):33–41

Zhang T, Sun L (2005) Managing supply chain uncertainty: techniques and strategies. Tsinghua University Press, Beijing, pp 90–91

Zhao Q, Chen H, Wu Z, Xie D (2011) Coordination and ripple effect of a coopetition supply chain with quantity-discount contract. In: Proceedings of 2011 international conference on management science & engineering, Rome, 2011, pp 222–226

Chapter 18
The Influence of Inter-enterprise Value Co-creation on Innovation Based on Resource Theories

Steven Ji-fan Ren, Qiong Bu, Ming-jian Zhou, and Cai-hong Hu

Abstract In today's globalized competitive environment, realizing the fact that the individual resources are insufficient, more and more companies are gradually deepening the understanding and awareness of the win-win business model and alliance. Under the influence of the concept of service-dominate logic, and value co-creation, a relatively new concept, has attracted the attention of scholars. This study focuses on the inter-enterprise value co-creation and its mechanisms. Basing on resource theories, we analyzed the relationships among asset co-specialization, technical dependence, inter-enterprise value co-creation, partnership quality and innovation. A conceptual model is developed.

Keywords Asset co-specialization • Innovation • Inter-enterprise value co-creation • Partnership quality • Technical dependence • Value co-creation

18.1 Introduction

Facing fiercer and fiercer price competition in the industry, especially in the manufacturing industry, entrepreneurs and scholars are looking for a new way for the firms. More and more entrepreneurs and scholars realize that ignoring the cooperation and the value co-creation leads to the failure to get advantage in future competition (Zhang and Chen 2009).

Foundation Item: The National Natural Science Funds (71172158)

S.J. Ren (✉) • Q. Bu • M. Zhou • C. Hu
Center of Business Administration, Harbin Institute of Technology Shenzhen Graduate School, Shenzhen, China
e-mail: renjifan@gmail.com; buqiong.doc@gmail.com; mngzmj@gmail.com; caihong11@yahoo.cn

E. Qi et al. (eds.), *The 19th International Conference on Industrial Engineering and Engineering Management*, DOI 10.1007/978-3-642-37270-4_18,

The theory of value co-creation (VCC) and related service dominate logic (S-D Logic) is the hot spot in the recent management, marketing and information field. Top journals in the marketing and information system field like Journal of Academy of Marketing Science (2008 Vol. 36), MIS Quarterly (2011, Special Issues on Service Innovation in Digital Age) have published or are publishing the special journal about this theory.

The research of VCC has two streams. Some of them focus on the relationship between firms and individual (e.g. Andreu et al. 2010; Nambisan and Baron 2009), the other study the VCC between enterprises (Forsstrom 2005; Ho et al. 2010). More literature is related to the former one, less is founded about the VCC between enterprises. In this paper, we focus on the latter.

The resource theory (resource based-theory and resource dependence theory) has been proved a useful tool to analysis the strategic alliance (Eisenhardt and Schoonhoven 1996; Gulati 1995), which is a meaning reference for this study.

This study firstly summarized the VCC theories systematically. Then we focus on the VCC on the enterprise level, discuss the effect of VCC to innovation performance and cooperation quality basing the resource theory. This plays important role for figuring out the conceptualization and the mechanism of VCC.

18.2 Literature Review

18.2.1 Value Co-creation Theory

The concept of VCC is introduced by Normann & Ramirez in 1998, they suggested that the interaction between producer and consumer is the basic part of value creation (Andersson et al. 2007). Existing literature share a common internal logic named S-D Logic. S-D Logic consider the service as consumer's potential or existing right which is inalienable, that is quite different from G-D Logic that focuses on the product exchange.

Prahalad and Ramaswamy are the most important scholars for the research of VCC. They suggest that the clients are not passive but active in their paper, and the core of VCC is the interaction between enterprise and customer (Prahalad and Ramaswamy 2004). Besides, they also consider that the dialogue, access, risk assessment and transparency are the base of VCC, named DART model. Table 18.1 list the factors in DART model.

When Prahalad and Ramaswamy studied this, they didn't clarify the subject of VCC. That is, these four factors can be applied no matter between enterprise and customer or between enterprises.

Table 18.1 Factors in DRAT model

DART	Description (organizational level)
Dialogue	The communication between upstream and downstream enterprise to know each other's wish, perceive the cooperative attribution, solute problems equally and share the knowledge
Access	Built an information platform for the cooperation to elicit service information timely. More attention on the right of common use but not own
Risk assessment	Two side of cooperation analysis the risk of product or service, and offer reliable method to assess the risk
Transparency	Make sure the symmetric of information, improve the transparency of resource, and enhance the trust and willing to cooperate

18.2.2 Resource Theories

Resource theories are quite important for the organizational level research. Among resource theories literatures which related to VCC, Vargo and Lusch (2004, 2008) point out that in the concept of S-D Logic, clients are no longer the operand resources but the more important operant resources. Baron and Warnaby (2011) analysis the cases of English library to figure out how to transform the operant resources of clients to the usable, physical and cultural social resource. Forsström (2005) analyze the importance of technical dependence during the value co-creation.

This study discusses the VCC between enterprises basing on the resource based theory and resource dependence theory.

1. *Resource based theory*: resources-based view of the firm discusses the reason of performance difference from the inner view of firms, basing which people can figure out the method to gain the continuous competitive advantage (Wernerfelt 1984). In the resource based theory, the process of enterprises to improve their competitiveness is the process of enlarging the diversity and enhancing the rareness of inner resource. However, in the current environment, no existing firms can hold all superior resource. That stimulated the emergence of VCC. Basing the inner advanced resource, the best way for the enterprise to achieve a wider market is using the outer (the other organizations') resource, and cooperate. Owning the resource which is owned by the other is the motive power and the base of cooperation.
2. *Resource dependence theory*: the basic hypothesis of resource dependence theory is no organization is self-sufficient, an all of them exchange with the environment, and basing this to survive and develop. During the exchange, the environment offers the critical resource which is rare and not contained in the organization. For the VCC activity, the dependency with partner is necessary. The firm depends on the partner shows that the partners own the critical scare

resource. But if the partner doesn't depend on the firm, the motive power of cooperation may disappear; the co-created value also can't come out. Thus, the dependency of resource is the base of practical cooperation and the necessary condition of two sides of VCC.

Resource based theory and resource dependence theory provide condition and foundation for the VCC activities. However, it doesn't mean owning the heterogeneous resources which has dependence relationship with the enterprise leads to high quality co-created value.

18.3 Hypotheses

18.3.1 Conceptual Model

Resource based theory emphasize that there is the heterogeneous resource in the enterprise to enhance the firm's competitiveness, which is a foundation of an enterprise. But with continuous improvement of inter resource, a single enterprise can't hold more resources to get a higher market goal in the condition of maintain exiting scarce resource. As a single unit is limited by the resource, people realize the fact of dependent resource and enhance the competitiveness by cooperation.

In the process of cooperation, the phenomenon of joint development and ownership of product or service is emerging, which is an expression of asset co-specialization. With the lasting and development of cooperation, the technical dependence is promoted. Asset co-specialization and technical dependence ask for the better communication between partners and take the risk together, all of which form the foundation of VCC. The trust with each other during the VCC is beneficial to maintain cooperation, develop the relationship. VCC and good relationship with partner can lead to positive effect to the innovation (Fig. 18.1).

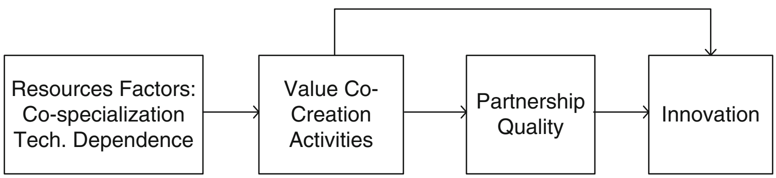


Fig. 18.1 Conceptual model

18.3.2 The Hypothesis About Asset Co-specialization

Among the alliance, some specific agreements may need asset co-specialization which exists in each activity and developing stage. In the mechanism of co-development between firms, different kinds of partner's advantage can help the enterprise figure out how to build long-term cooperation and integration. In the agreement of co-manufacturing, the united manufacture basing the strength of each side can enhance the alliance's ability to resistant the examination and monitoring. Co-Promote (products selling with a common brand) and co market (a product selling in the partners' brand) need the union of organizations to cooperatively face the all kinds of sharing and developing activities (Simonet 2002). Kogut and Zander suggest that co-specialization can be considered as a quasi-organizational structure. In this mechanism each partner builds relationship agreement and exchange information by transferring their superiorities. However, this kind of organizational structure, which will help each other to transfer the superiority, may cause the investing surplus sometimes. Oxley (1997) points out that when one unit tries to enhance the superiority transfer, it is possible that more information than what is demanded is conveyed, which cause damage to transfer process. Besides, co-specialization builds a routine relationship with alliance and non-transactional character, thus the partner has the speculative opportunity to hold-up and bargaining with each other.

Teece (1992) points out asset co-specialization is a kind of mutual investing which can reflect the specific cooperating activities or the organizational structure to facilitate the exchange of information and knowledge, which is decided by its own character. And one of the aims of asset co-specialization is promoting the information exchange among firms and transfer advantage. This kind of activity can improve the transparency to each other because the building of communication platform or channel must exist in the process of information exchange. Asset co-specialization undoubtedly base on the contract, no matter co-development, the co-manufacturing, co-promote or co-market. There is clause that rules joint adventure to clarify the risk contribution in the contract. In the most situations, sharing the risk is the potential premise of asset co-specialization. Once the asset co-specialization is chose, what is mostly concerned is the question about the sharing of risk. Besides, from the beginning of plan to the final practical operating, the dialogue and communication is essential, which will mature with the whole process. The character of asset co-specialization promotes the dialogue, the channel building, the risk assessment and the transparency, which has positive effect to the VCC activity. Basing the discussion above, we suggest the hypothesis 1:

H1 Asset co-specialization positively influences VCC activities.

18.3.3 The Hypotheses About Technical Dependence

Forsström (2005) points out the technical dependence between firms is an essential condition of VCC. One-side dependence usually leads to the other side lose the motive power of VCC.

Recent research shows that the strength of a single company is a part of the whole strength (Gaski and Nevin 1985). The overall strength depends on the accumulation, asymmetry and diversity of two sides' power (Bacharach and Lawler 1981). After the focus moves to the overall strength and asymmetrical power, the study analyzing a single firm becomes not so convincing. Thus, most studies about these are related to the technical dependence.

Among the literature about technical dependence and asymmetrical dependence, two common theories are bilateral deterrent theory (BDT) and conflict spiral theory (CST). Bilateral deterrent theory suggest that because two side of cooperation face the potential loss and they have strong motivation to avoid these loss, when the overall strength increases, the punishing activity reduces. Oppositely, conflict spiral theory points out that more powerful overall strength leads to the increasing allure and implement of punishing activity (Lawler et al. 1988). The motivation of BDT is reducing the threat, and CST is basing the assumption that the company will use all its power to achieve established aims. As the dependence shows the potentiality of value loss, that is, the value will disappear if the relationship breaks up; DT is more suitable to be applied to the study of technical dependence. Because the punishing ability is the storage of the power that may cause damage, the CST is only related to the partner's punishing ability.

According to BDT, we can inter that if a company's strength bases on the increasing dependence, the cooperating relationship will improve, then the expectation of attack become lower. Then the company and his partner will share the technical dependence and be afraid to break this relationship, hope the other side also realizes this. Kinds of benefit can be attributed to the high-level dependence including no punishment and so on (Gundlach and Cadotte 1994).

BDT suggests that an asymmetric relationship contains inner instability and only symmetry can obstruct the punishing behavior. The more powerful the overall strength is, this obstruction is stronger. Many scholars have proved that the symmetrical relationship is more stable and beneficial than an asymmetrical one (Kumar et al. 1995). More asymmetry lead to more possibility of punishing behavior (Lacity and Hirschhein 1993). According to BDT, the dependence is the main motivational factors for the company with a lower-level demand for the dependence. Because companies of this kind get relatively less benefit from the relationship, they won't worry about breaking the relationship or intentionally control the punishing behavior. Though the other side which more depends on the partner will loss more, it still may be attacked by some activities, thus it has a strong motive to conduct preemptive punishing strategy. The relationship power theory, which may lead to a different conclusion, suggests that continuous improvement of asymmetric dependence will influence the company and its supply chain cooperator. Especially, the side which has a weaker power is less willing to use the punishment, while the stronger one is more willing.

Forsström points out that the independence relationship shows that each side has the resource that the other hasn't hold (Ho et al. 2010). The demand of complementary resource forms the technical dependence and makes it tighter. For example, firm A designs a system for B basing B's existing operating system to improve the performance. The technical dependence will get stronger with the continuous updating of system. This established relationship will lead to great cost to break, and causes a long-term cooperation which is necessary for the VCC between firms. Basing the statement above, we can get hypothesis 2 followed:

H2 Technical dependence positively influences VCC activities.

18.3.4 The Hypothesis About Partnership Quality

The view point of strategic alliances has more and more accepted, people pay more attention on the successful cooperating relationships between service customer and provider. A lot of companies in the world has built intimate relationship with their service provide, for example, the cooperative partner relationship among Kodak-IBM-DEC, USAA-IBM and Xerox-EDS. In the academic field, more scholars come to realize the importance of partner relationship.

In this paper, the partner relationship is defined as an inter-organizational relationship to achieve the common goal of alliance members. In the marking field, the partner relationship is built on the social exchange theory instead of economic viewpoint like transaction cost theory and agent cost theory. Social exchange theory consider the relationship is a powerful process in which one has common ordered behaviors with partner, and exchange the valuable resource. This assumption shows the continuous trustful process lasts through the whole time period of cooperation, and the relationship also need the economic aspect (Lacity and Hirschhein 1993).

Blau suggests that social exchange theory explain the exchange relationship basing the concept of trust (Blau 1964). For the firms, the trust is the other firms have favorable behavior instead of harmful. Trust is a critical factor for the long-term relationship and promotion of exchange. Besides, trust can evolve more satisfactory interaction and then increase the trust in each other.

Transparency is an essential condition of building trust, and a well-working dialogue mechanism as well as the communication platform is the important approach to improve the trust. Sharing the risk is also the expression of trust. Assessment and solutions to the risk is a critical indicator of trust. Thus VCC activities are beneficial to improve the trust between firms. According to the statement above, we can get hypothesis 3:

H3 VCC activities positively influence partnership quality.

18.3.5 *The Hypotheses About Innovation*

Creativity and innovation is the critical resource to prevent the organization from stagnate (Cameron 1996). In this paper, as other scholars do, we define creativity as any new product and useful thought in this field; meanwhile, innovation is defined as the phenomenon or behavior in the organization which manages to apply the creative thought. From this viewpoint, organizational creativity is from the individual in the organization or the entire organization. However, innovation is not limited to this, and it is influenced by other factors, like technology transformation.

Traditional literature about innovation largely focuses on the individual character. However, recent research of innovation began to study in the organizational level that is organizational creativity or organizational innovation. Woodman et al. define the organizational creativity like this: organization is a complex system where the organization member's ability of cooperating to create useful and valuable new technology, product/service and flow/procedure (Woodman et al. 1993). Organizational creativity is indispensably related to creative context, individual with creative ability, procedure of innovation and creative product or service. Mohamed and Rickard put forward an assessment of company's innovation. They suggests organizational innovation is shown in the improvement of technology/product/service, the introduction and implement of new technology (like equipment or the procedure), interaction with outside and its effect, and the inner training about innovation (Mohamed and Rickards 1996).

VCC activity changes the inner or outer environment of company. The dialogue with each other about the diversity of enterprise culture or concerned benefit actually is a process of improving the enterprise's openness which is beneficial to organizational innovation (Service and Boockholdt 1998). The change the inner or outer environment which usually introduce new creative individual is also good for the formation of innovation. With the improvement of transparency, the open, accuracy demand and interaction is good for building the innovational circumstance. Besides, because of the scarcity of technical resource, the newness of related information can expand the employee's thought, which has promoting effect for the individual innovation. Basing the discussion above, we suggest hypothesis 4 followed:

H4 VCC activities positively influence innovation.

Interaction with outside itself is beneficial to organizational innovation. Furthermore, establishment of partner relationship can change the environmental factors, facilitate establishment of innovational context, and patch the shortage of inner innovation. To the partner, trust is the foundation, and the improvement of performance is the source force of following persistent cooperation. The transparency is helpful to develop the targeted cooperative innovation, basing a common goal. The innovation performance basing the partner's cooperation can increase the identification with each other firm. This attitude of identity and affirmation can promote the development of relationship and then encourage the organizational innovation. Basing the discussion above, we can get hypothesis 5:

H5 Partnership quality positively influences innovation.

18.4 Discussion

This paper introduces the resource theory to study the relationship among the asset co-specialization, technical dependence, VCC activities, partnership quality and innovation. Previous researches about VCC mostly focus on the relation between enterprise and customers, this paper can be used for reference for the scholars who are willing to study VCC between enterprises. Besides, most scholars are still discussing the concept of VCC. Few of them connect this with the other factors to analysis the enterprise's operation and figure out which factors can affect the VCC activities as well as what benefit can be brought by VCC. This paper provides a new viewpoint to VCC theory by building the research model which studies VCC as a mediator. Another innovation of this paper is the introduction of resource theory, which provides the theoretical foundation for the future research.

For the practical implication, this study shows the necessity and importance of understanding and conducting the cooperation. The factors in the model are the common problems which current companies are facing or concerning, which is beneficial for the enterprises which is in their bottleneck. Besides, the research also offers a new way to improve the innovation performance by cooperation. It may be an implicational viewpoint for the companies which are still fighting alone.

The paper still has some inevitable limitations. The background theory is limited to the resource theory, which is just a viewpoint among numerous theories that may relate. Besides, we didn't specify an industry, which may ignore has some unpredictable effect. In addition, VCC activities also have different forms, which are not involved in this paper. All of these can provide the future directions for the scholars who are interested.

References

Andersson P et al (2007) Mobile innovations in healthcare: customer involvement and the co-creation of value. Int J Mob Commun 5:371–388

Andreu L et al (2010) Value co-creation among retailers and consumers: new insights into the furniture market. J Retail Consum Serv 17:241–250

Bacharach SB, Lawler EJ (1981) Bargaining: power, tactics, and outcome. Jossey-Bass, San Francisco

Baron S, Warnaby G (2011) Individual customers' use and integration of resources: empirical findings and organizational implications in the context of value co-creation. Ind Market Manage 40:211–218

Blau PM (1964) Exchange and power in social life. Wiley, New York
Cameron MF (1996) A theory of individual creative action in multiple social domains. Acad Manage Rev 21:1112–1142
Eisenhardt KM, Schoonhoven CB (1996) Resource-based view of strategic alliance formation: strategic and social effects in entrepreneurial firms. Org Sci 7:136–150
Forsström B (2005) Value co-creation in industrial buyer-seller partnerships– creating and exploiting interdependencies. Åbo Akademi University Press, Åbo
Gaski JF, Nevin JR (1985) The differential effects of exercised and unexercised power sources in a marketing channel. J Market Res 22:130–142
Gulati R (1995) Social structure and alliance formation patterns: a longitudinal analysis. Adm Sci Q 40:619–652
Gundlach GT, Cadotte ER (1994) Exchange interdependencea nd interfirm interaction: research in a simulated channel setting. J Market Res 31:516–532
Ho YC et al (2010) Value co-creation in business models: evidence from three cases analysis in Taiwan. Bus Rev Camb 15:171–178
Kumar N et al (1995) The effects of perceived interdependence on dealer attitudes. J Market Res 32:348–356
Lacity MC, Hirschhein R (1993) Information systems outsourcing. Wiley, New York
Lawler EJ et al (1988) Coercive capability in conflict: a test of bilateral deterrence versus conflict spiral theory. Soc Psychol Q 51:93–107
Mohamed MZ, Rickards T (1996) Assessing and comparing the innovativeness and creative climate of firms. Scand J Manage 12:109–121
Nambisan S, Baron RA (2009) Virtual customer environments: testing a model of voluntary participation in value co-creation activities. J Prod Innov Manage 26:388–406
Oxley J (1997) Appropriability hazards and governance in strategic alliances: a transaction cost approach. J Law Econ Org 13:387–409
Prahalad CK, Ramaswamy V (2004) Co-creating unique value with customers. Strategy Leadersh 32:4–9
Service RW, Boockholdt JL (1998) Factors leading to innovation: a study of managers' perspectives. Creativity Res J 11:295–307
Simonet D (2002) Licensing agreements in the pharmaceutical industry. Int J Med Market 2: 329–341
Teece DJ (1992) Competition, cooperation, and innovation: organizational arrangements for regimes of rapid technological progress. J Econ Behav Org 18:1–25
Vargo SL, Lusch RF (2004) Evolving to a new dominant logic for marketing. J Market 68:1–17
Vargo S, Lusch R (2008) Service-dominant logic: continuing the evolution. J Acad Market Sci 36:1–10
Wernerfelt B (1984) A resource-based view of the firm. Strategic Manage J 5:171–180
Woodman RW et al (1993) Toward a theory of organizational creativity. Acad Manage Rev 18: 293–321
Zhang X, Chen R (2009) New sources of competitive advantage, and customers to create value. Manage Eng J 23:14–19 (in Chinese)

Chapter 19
The Analysis of Fresh Food Safety Risks from the Cold Chain Logistics System

Fang Yang and Jian-ming Cai

Abstract The cold chain management plays a strategic role for the quality of fresh food arriving to the consumers. This research describes the structure of fresh food cold chain logistics system and identifies the safety risk factors impacting to fresh food at every link of the cold chain, and then analyzes the fresh food safe risks from the cold chain logistics system using the Analytic Hierarchy Process (AHP) and the Risk Priority Number (RPN) methodology, and considering the time-varying of the system. The methodology can find out the key risk factors and the weakest link in the fresh food cold chain logistics system so as to improve the performance of the cold chain system.

Keywords Cold chain • Fresh food • Logistics • Risk Priority Number • Safety risk

---

Supported by the Philosophy and Social Science Fund of Hunan Province (Grant No.11YBA025), and the Promoting Talents Program of Two Sides of the Straits (Grant No.11F8120105).

F. Yang (✉)
Department of Business Administration, Changsha University, Changsha 410003, People's Republic of China

School of Traffic and Transportation Engineering, Central South University, Changsha 410075, People's Republic of China
e-mail: ylzhuo@163.com

J. Cai
School of Traffic and Transportation Engineering, Central South University, Changsha 410075, People's Republic of China
e-mail: jmcai@csu.edu.cn

E. Qi et al. (eds.), *The 19th International Conference on Industrial Engineering and Engineering Management*, DOI 10.1007/978-3-642-37270-4_19,

19.1 Introduction

In today's society, the safety of foods is closely connected to issues like public health, commercial interests and general development of the society as well as the health and well-being of the individual. Food's qualities, i.e. its look, smell, taste, safety and health benefits attract all consumers' attention. Therefore, any new event regarding safety of the food, immediately create interests light in the mind of the public. Once unsafe or contaminated food enters the food chain, it may be distributed more rapidly and to a greater number of consumers—hence the risks are higher. The economic consequences of contaminated food and farm products can be devastating (Burlingame and Pineiro 2007).

Many foods need to be refrigerated for preservation. These foods include fruit and vegetables, meat, fish, dairy products, seafood and so on. Refrigeration would make it possible to partially reduce post-harvest losses. With the rapidly growing urban populations, the cold chain logistics is a kind of primary means to ensure food quality and safety (Coulomb 2008).

For these reasons, the authorities are now requesting more and more guarantees concerning product quality and traceability throughout the cold chain. These have need of ongoing technological improvements and optimizing the cold chain.

Many food safety issues have highlighted the need for early identification of food safety hazards and the subsequent mitigation, control, and/or prevention of the associated risks (Marvin and Kleter 2009).

The SAFE FOODS framework describes an iterative decision-making process with an integrated five-step procedure for the risk analysis, i.e. framing, risk (risk-benefit) assessment, evaluation, risk management, and review. The framework is designed to address both risks and benefits, and improve the transparency, openness, and accountability of the risk analysis process (Knudsen 2010; König et al. 2010; Kuiper and Davies 2010; König 2010). Notermans and Teunis (1996) studied the application of quantitative risk analysis (QRA) in the production of microbiologically safe food products (Notermans and Teunis 1996). Doménech et al. (2007) proposed a QRA model which allows the evaluation of the risk for consumers' health and the induced economic risk for the company as a result of the occurrence of failures in food safety. An integrated way is considered in this model that all the agents are involved in the decision-making on food quality and safety, and all the links of the food chain, from the farm to the table (Doménech et al. 2007). Within the work package 3 of the EU-funded project SAFE FOODS, van Klaveren and Boon (2009) developed an electronic platform of food consumption and chemical concentration databases harmonized at raw agricultural commodity level base on probabilistic risk assessment of single and multiple exposure to pesticide residues or contaminants, and an integrated probabilistic risk assessment (IPRA) model was developed to compare health risks between multiple chemicals in complex risk assessment situation (van Klaveren and Boon 2009). Smith and König (2010) analyzed various environmental risk assessment practices that were at that time in place for food-related substances in both the EU and the USA (Smith and

König 2010). A system dynamic model to support cold chain management in food supply chain was presented (Oliva and Revetria 2008).

This paper will describe the structure of fresh food cold chain logistics system and identify the risk factors impacting to fresh food safety at every link in the cold chain, and analyzes the fresh food safety risks of the cold chain logistics system using the Analytic Hierarchy Process (AHP) and the Risk Priority Number (RPN) methodology.

19.2 Fresh Food Cold Chain Logistics System

The cold chain logistics system is the system which deals with the produce, packaging, transport, storage, distribution and selling of perishable products within the safe temperature range from the place of primary production to the point of final preparation and consumption ensuring people receive an effective products that has retained its potency and has not been affected, and minimizes the economic losses.

The fresh food cold chain system refers to the incessant refrigerated handling of fresh food from the farm to the market. For the system, the desired temperature level is to achieve and maintain the quality of fresh food at every link of the supply chain.

The cold chain system needs the players to participate in the handling, storage and transport of the perishable produce. These include farmers, packers, transport service suppliers, and workers and staff.

The quality of fresh foods needs professional and specialized facilities to maintain the right temperatures from production to delivery for consumption. The facilities in a cold chain logistics system consist of pre-coolers, packing houses, refrigerated chambers and transportation tools (trucks, cargo ships, freight cars, etc.) for refrigerated transport. Pre-coolers are used to remove the heat quickly after harvest to obtain the acquired conditions. Packing houses are absolutely necessary to prepare the fresh foods (e.g. fruits and vegetables) before entering the market such as trimming and cleaning, sorting out the deficient products, among others. Refrigerated chambers maintain the required storage temperature in production of high quality fresh food. Refrigerated container vans/trucks gather the fresh food from the refrigerated chambers/packing houses and transport them.

The complexity of cold chain logistics system may vary from cold chain to cold chain due to different products demand, producer and consumer. Figure 19.1 shows the cold chain system in new agro-food marketing.

19.3 Identification of Fresh Food Safety Risks in the Cold Chain

The safety for fresh food generally refers to the content of the various chemical and microorganism elements in food, i.e. toxicants, pollutants and other potentially risk elements.

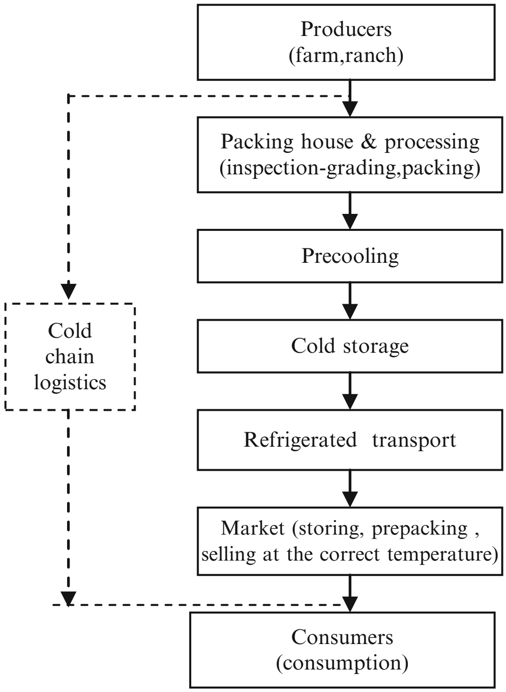


Fig. 19.1 The cold chain system for new agro-food marketing

In fresh foods, the factors which affect growth and inactivation parameters comprise combined effects of intrinsic (food-related) factors and extrinsic (environmental) factors. The food-related factors contain nutrients, water activity (aw), pH and microbial inhibitory factors, and the environmental factors include gaseous atmosphere, storage temperature and relative humidity.

Fresh foods, unless the appropriateness of the processing, packaging, distribution and storage, can be destroyed in relatively short time and cause a potential hazard for the consumer when temperature abuses.

The fresh degree of fresh foods is connection with many logistic attributes, such as storage & transportation time, surrounding temperature, relative humidity and the gas content in containers or packing receptacle. In the process of logistics, many factors can cause random variations of these logistics attributes, e.g. packaging damage, distribution centers capacity, transport conditions, failures of facilities & equipment, operational errors, and others.

In the cold chain logistics system, the risk factors that affect the safety of fresh food mainly include poor environment, technical factors, human error and organizational factors. These factors are the common performance influencing factors at every link in the fresh food cold chain logistics system.

1. *Environmental factors*: Poor working conditions (improper temperature, improper humidity and improper working space), natural disasters, etc.

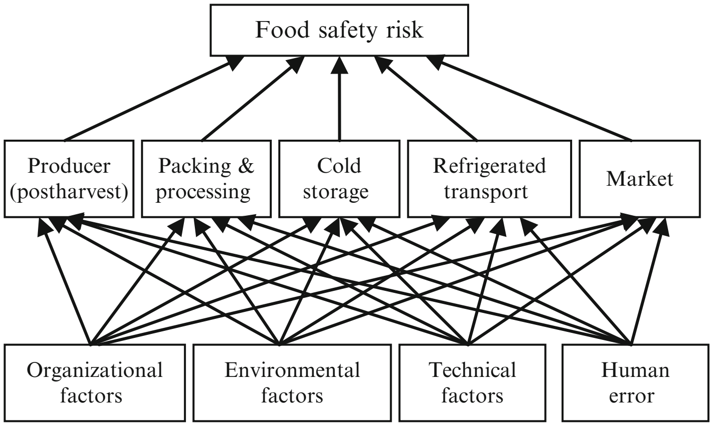


Fig. 19.2 A three-levels hierarchy for fresh food safety risks from the cold chain logistics system

2. *Technical factors*: Failures of temperature control & test equipment, quality testing equipment, transportation equipment, operation technology; inadequacy of man machine interface and operational support.
3. *Human error*: Inadequacy of training and experience, deficient crew collaboration quality, operational errors.
4. *Organization factors*: Deficient or inefficient organization, inappropriate procedures/plans.

In this respect, the Analytic Hierarchy Process (AHP) (Saaty 1980) is here employed to structure the hierarchy of the fresh food safety risks from the cold chain system. Figure 19.2 shows a three-level hierarchy for fresh food safety risks from the cold chain system.

19.4 Assessment of Fresh Food Safety Risks from the Cold Chain Logistics System

19.4.1 Overview of the Risk Priority Number

A FMECA can be used to identify the potential failure modes for a product or process. The Risk Priority Number (RPN), which is a well-known methodology used to rank the order of priority of alternatives, is used to analyze the risk associated with potential problems identified during a Failure Mode Effect and Criticality Analysis (FMECA) and to prioritize problems for corrective action. There is a large amount of variation among FMECA practitioners in the case of the specific analysis procedure and some analyses may include alternative calculation methods (Gionata 2009; Cai et al. 2011). The parameters of the traditional RPN are defined as:

- Severity (S): Represents the seriousness of the effects of a failure which affect the consumer or system that uses the product or component.
- Occurrence (O): Represents the probability that the failure occur.
- Detection (D): Represents the likehood that the failure mode can be identified.

The RPN is the product of these three parameters:

$$\mathrm{RPN} = \mathrm{S} \times \mathrm{O} \times \mathrm{D}, \tag{19.1}$$

with the higher number indicating the higher risk or hazard. An analysis team is usually required to use past experience and engineering judgment for rating each potential problem base on these three rating scales (Severity, Occurrence and Detection) of the RPN methodology. Rating scales usually range from 1 to 5 or from 1 to 10, the higher number indicates the higher risk or seriousness.

19.4.2 Revised RPN for the Assessment of the Risks from the Cold Chain Logistics System

Fresh food production is a dynamic activity. The quality of fresh food may change owing to seasonal effects. The composition of product may vary due to different consumer demand. Also, processing conditions will change continuously. In addition, processing procedures may be different from plant to plant. Thus, there is the time-varying characteristic of the risks for fresh food safety from cold chain logistics system.

For the above reasons, the three parameters of the RPN are redefined, i.e. $S_{ij}^k(t)$ represents the severity rank of the event j of the category i at the link k in a given time interval t, $O_{ij}^k(t)$ represents the occurrence probability of the event j of the category i at the link k in a given time interval t, $D_{ij}^k(t)$ represents the detection difficulty rank of the event j of the category i at the link k in a given time interval t.

Therefore, we can easily obtain the following results:

For the RPN of the event j of the category i at the link k in a given time interval t, the calculation formula is

$$RPN^j(t) = S_{ij}^k(t) \times O_{ij}^k(t) \times D_{ij}^k(t). \tag{19.2}$$

For the RPN of the category i at the link k in a given time interval t, the calculation formula is

$$RPN_j^i(t) = \sum_{j=1}^{m_{ki}} S_{ij}^k(t) \times O_{ij}^k(t) \times D_{ij}^k(t). \tag{19.3}$$

Table 19.1 Severity levels for the risk factors

Severity level	Severity description
1	None or very low
2	Minor: a failure that may be overcome with an
3	unscheduled measure and/or little economic impact
4	Medium: a failure that may cause infection, the state or
5	quality of being incredulous and/or some economic
6	losses
7	Critical: a failure that may cause severe food-borne
8	diseases and/or great economic losses
9	Catastrophic: a failure that may cause casualty and/or
10	huge economic losses

For the RPN of the link k in a given time interval t, the calculation formula is

$$RPN_{ij}^{k}(t) = \sum_{i}^{n_k} \sum_{j=1}^{m_{ki}} S_{ij}^{k}(t) \times O_{ij}^{k}(t) \times D_{ij}^{k}(t). \tag{19.4}$$

For the RPN of the cold chain logistics system in a given time interval t, the calculation formula is

$$RPN(t) = \sum_{k=1}^{N} \sum_{i=1}^{n_k} \sum_{j=1}^{m_{ki}} S_{ij}^{k}(t) \times O_{ij}^{k}(t) \times D_{ij}^{k}(t). \tag{19.5}$$

Where, m_{ki} represents the number of the event j of the category i at the link k, n_k represents the number of the category i at the link k, N represents the number of the links in the cold chain logistics system.

19.4.3 Example

Here is a cold chain logistics system for vegetables and four main links are considered. The four main links of cold chain logistics are cold processing & packing, cold storage, refrigerated transport and selling in the right temperature. Where, Table 19.1 describes the rating scale of the Severity with a ten point for vegetables safety risks, Table 19.2 describes the Occurrence scale with a ten point. Table 19.3 describes the RPNs and indexes of risk factors.

To simplify the problem, let $D_{ij}^{k}(t) = 1$.

As shown in Table 19.3, the severities are the average value of a set of similar events in a given time interval t, the first number in parentheses represents the RPN in the given time interval $t = 1$, while the second number represents the RPN in the given time interval $t = 2$.

Table 19.2 Occurrence levels for the risk factors

Occurrence level	Probability for reference (Pr)	Description
1	$Pr < 0.00001$	Extremely rare: a failure whose probability of occurrence is near to zero
2	$0.00001 \leq Pr < 0.0001$	Rare: an unlikely probability of occurrence
3	$0.0001 \leq Pr < 0.005$	Occasional: an infrequency probability of occurrence
4	$0.005 \leq Pr < 0.01$	Reasonably common: a moderate probability of occurrence
5	$0.01 \leq Pr < 0.10$	
6	$0.10 \leq Pr < 0.20$	
7	$0.20 \leq Pr < 0.30$	Frequent: a high probability of occurrence
8	$0.30 \leq Pr < 0.40$	
9	$0.40 \leq Pr < 0.50$	Very frequent: a very high probability of occurrence
10	$Pr \geq 0.50$	

According to Table 19.3, the RPN of the cold chain logistics system is 1,077 when $t = 1$, and 969 when $t = 2$. Therefore, the vegetables safety risk from the cold chain logistics system is higher in the given time interval $t = 1$ than the risk in the given time interval $t = 2$. Apparently, the highest risk is from cold storage, followed by market, refrigerated transport and picking house. Environmental factors contribute higher risk to packing house and cold storage, whilst risk at the links of refrigerated transport and market is mainly from the organizational factors.

19.5 Conclusion

The RPN methodology is an available technique for analyzing the food safety risk from the cold chain logistics system. It can not only find out the key risk factors and the weakest link in the fresh food cold chain logistics system so as to improve the performance of the cold chain system, but also choose the optimal one from different cold chain systems.

Because the RPN is the product of three ratings, various sets of Severity, Detection and Occurrence values may produce a same value of RPN, however, the risk meaning may be totally different. Thus, basing decisions solely on the RPN may result in inefficiency and/or increased risk. To close this gap, a method to assign a set of weights to the RPN is available. For example, issues with higher severity and/or higher occurrence ratings may be considered as a higher risk than issues with higher detection ratings. Usually, the rank ordering of risk may be determined by Severity, Occurrence or Detection.

Table 19.3 The levels and categories of risk factors from the cold chain logistics system

	First-level index (Risks from links)	Second-level index (Categories of risk factors)	Third-level index (Risk factors)	$t = 1$			$t = 2$		
				$S_{ij}^{k}(t)$	$O_{ij}^{k}(t)$	$RPN^{j}(t)$	$S_{ij}^{k}(t)$	$O_{ij}^{k}(t)$	$RPN^{j}(t)$
Vegetables safety risks (1077,969)	Risks from packing house (219,190)	Organizational factors (60,54)	Deficient or inefficient organization	6	5	30	6	4	24
			Inappropriate procedures/ plans	6	5	30	6	5	30
		Environmental factors (75,70)	Destruction by heat or irradiation	9	5	45	9	5	45
			Poor working conditions	5	6	30	5	5	25
		Technical factors (40,26)	Spoiled by packing tools	4	3	12	4	3	12
			Improper packing materials	7	4	28	7	2	14
		Human error (44,40)	Packers error	6	5	30	6	4	24
			Load and unload worker error	4	4	14	4	4	16
	Risks from cold storage (330,267)	Organizational factors (89,73)	Deficient or inefficient organization	6	3	18	6	3	18
			Inappropriate procedures/ plans	5	3	15	5	3	15
			Too long time	8	7	56	8	5	40
		Environmental factors (125,86)	Destruction by heat or irradiation	9	5	45	9	4	36
			Cross contamination	10	8	80	10	5	50
		Technical factors (76,72)	Refrigerating battery failure	10	4	40	10	3	30
			Inappropriate placement	6	6	36	6	7	42
		Human error (40,36)	Warehouseman error	4	5	20	4	4	16
			Load and unload worker error	4	5	20	4	5	20

(continued)

Table 19.3 (continued)

First-level index (Risks from links)	Second-level index (Categories of risk factors)	Third-level index (Risk factors)	$t = 1$			$t = 2$		
			$S_{ij}^k(t)$	$O_{ij}^k(t)$	$RPN^j(t)$	$S_{ij}^k(t)$	$O_{ij}^k(t)$	$RPN^j(t)$
Risks from refrigerated transport (239,248)	Organizational factors (96,94)	Too long time en route	8	6	48	8	7	56
		Inappropriate route	6	3	18	6	3	18
		Inappropriate conveyance	5	6	30	5	4	20
	Environmental factors (40,40)	Natural disasters	10	1	10	10	1	10
		Accident	6	5	30	6	5	30
	Technical factors (46,54)	Conveyance failure	8	2	16	8	3	24
		Refrigerating battery failure	10	3	30	10	3	30
	Human error (57,60)	Drivers failure	5	9	45	6	8	48
		Load and unload worker error	3	4	12	3	4	12
Risks from market (289,264)	Organizational factors (138,117)	Deficient or inefficient organization	6	7	42	6	5	30
		Inappropriate procedures/plans	6	7	42	6	4	24
		Too long time	9	6	54	9	7	63
	Environmental factors (81,79)	Improper temperature	10	3	30	10	2	20
		Improper humidity	7	3	21	8	3	24
		Poor gaseous atmosphere	6	5	30	7	5	35
	Technical factors (20,20)	Failures of temperature control & test equipment	10	2	20	10	2	20
	Human error (50,48)	Sellers error	6	5	30	6	5	30
		Load and unload worker error	5	4	20	6	3	18

References

Burlingame B, Pineiro M (2007) The essential balance: risks and benefits in food safety and quality. J Food Compos Anal 20(3–4):139–146

Cai JM, Li X.-M, Yang GH (2011) The risk priority number methodology for distribution priority of emergency logistics after earthquake disasters. In: Proceedings of the 2011 international conference on management science and industrial engineering, Harbin, China, 2011, pp 560–562

Coulomb D (2008) Refrigeration and cold chain serving the global food industry and creating a better future: two key IIR challenges for improved health and environment. Trends Food Sci Technol 19(8):413–417

Doménech E, Escriche I, Martorell S (2007) Quantification of risks to consumers' health and to company's incomes due to failures in food safety. Food Control 18(11):1419–1427

Gionata C (2009) An integrated structural framework to cost-based FMECA: the priority-cost FMECA. Reliab Eng Syst Saf 94:861–871

Knudsen I (2010) The SAFE FOODS framework for integrated risk analysis of food: an approach designed for science-based, transparent, open and participatory management of food safety. Food Control 21(12):1653–1661

König A (2010) Compatibility of the SAFE FOODS risk analysis framework with the legal and institutional settings of the EU and the WTO. Food Control 21(12):1638–1652

König A, Kuiper HA, Marvin HJP, Boon PE, Busk L, Cnudde F et al (2010) The SAFE FOODS framework for improved risk analysis of foods. Food Control 21(12):1566–1587

Kuiper HA, Davies HV (2010) The SAFE FOODS risk analysis framework suitable for GMOs? A case study. Food Control 21(12):1662–1676

Marvin HJP, Kleter GA (2009) Early awareness of emerging risks associated with food and feed production: synopsis of pertinent work carried out within the SAFE FOODS project. Food Chem Toxicol 47(5):911–914

Notermans S, Teunis P (1996) Quantitative risk analysis and the production of microbiologically safe food: an introduction. Int J Food Microbiol 30(1–2):3–7

Oliva F, Revetria R (2008) A system dynamic model to support cold chain management in food supply chain. In: Proceedings of the 12th WSEAS international conference on SYSTEMS, Heraklion, Greece, July 2008, pp 361–365

Saaty TL (1980) The analytic hierarchy process. McGraw-Hill International, New York

Smith MR, König A (2010) Environmental risk assessment for food-related substances. Food Control 21(12):1588–1600

van Klaveren JD, Boon PE (2009) Probabilistic risk assessment of dietary exposure to single and multiple pesticide residues or contaminants: summary of the work performed within the SAFE FOODS project. Food Chem Toxicol 47(12):2879–2882

Chapter 20
Literate Review on Firm Network Capability: A Process Perspective

Yi-bo Lv, Qing-yang Mao, and Yu-wen Ding

Abstract Networking growth based on outsourcing has become an important organizational mode of firm growth, firm network capability concentrates on reflecting elemental basis and managerial efficiency of outsourcing. This paper reviews relative literature of firm network capability at home and abroad, based on "process perspective" of network activities, it proposes a "process perspective" basic framework of firm network capability by refining content and feature, identifying constituent elements and dividing dimensions, combing influential factors and mechanism of action.

Keywords Constituent elements • Firm network capability • Influential factors • Process perspective

20.1 Introduction

With the development of global value chain driven by network economy, networking growth based on outsourcing has become an important organizational mode; firm network capability reflects elemental basis and managerial efficiency

---

Foundation item: National Natural Science Foundation of China (70902032), Doctoral Fund of Ministry of Education of China (20090041120014), Science Foundation of Ministry of Education (09YJC630015).

Y. Lv (✉) • Q. Mao
Faculty of Management and Economics, Dalian University of Technology Dalian, Liaoning, China
e-mail: luyibo@dlut.edu.cn; mydreamqingyang@126.com

Y. Ding
School of Accountancy, Shanghai University of Finance and Economics, Shanghai, China
e-mail: excellentmqy@hotmail.com

E. Qi et al. (eds.), *The 19th International Conference on Industrial Engineering and Engineering Management*, DOI 10.1007/978-3-642-37270-4_20,

of outsourcing. Since Håkansson (1987) proposed the concept of firm network capability firstly, network capability has gradually become an important theme of firm growth and innovation. Many scholars have done a large number of theoretical constructions and empirical researches.

Early studies focused on the definition and characteristics of network capability. From the view of organizational competitive advantage, Ritter (1999) identified and analyzed organizational elements of firm network capability, measuring the constituent elements of firm network capability from two dimensions of task execution and qualifications (Ritter et al. 2002) which become basic dimensions of follow-up studies of network capability. Zhu Xiumei et al. (2010) defined the structure of firm network capability from three dimensions of network orientation, network building and network managing. Ren Shenggang et al. (2011) further developed a measurement scale of firm network capability from four levels including network vision, network building, relationship management and portfolio management. Recently more and more studies have begun to pay attention to the practical significance of network capability. Ritter and Gemünden (2003) confirmed that network capability positively affected business innovation success in follow-up studies. Walter et al. (2006) found that the network capability had a positive regulatory role between entrepreneurial orientation and organizational performance. The important role that network capability plays in firm innovation and performance improvement has been widely confirmed by different researches.

Originality, composition and influential factors of network capability have been the cornerstone of studies. In the context of firm network capability, there is always integration of skill and action view. Based on "process perspective" of network activities, it proposes a basic framework of network capability by combing connotation features, constituent elements and influential factors.

20.2 Connotation and Characteristic of Firm Network Capability

From different definitions of firm network capability, the process of network activities is common concerns. Therefore, based on "process perspective" of network activities, firm network capability is defined as a dynamic capability that firms can take the initiative to implement a series of network planning, building, and managing activities to get scarce network resources and competitive advantages based on internal resources and resources leverage. From definitions of existing researches, network capability has three typical features including process-oriented, resource-dependent and dynamic evolution:

1. *Process-oriented*: Firm network activities consist of activities of planning, building, and managing network, so process-oriented is the basic feature of firm network capability. Dyer and Singh (1998) from the process perspective, regarded network capability as collections of a series of capabilities by creating,

developing, managing and using relationships. Zhao Shuang and Xiao Hongjun (2010) from overall network planning process, proposed network capability was to obtain scarce network resources by identifying strategic opportunities, dealing with network relations, and managing network locations. So firm network capability is obviously process-oriented.

2. *Resource-dependent*: Firm network capability belongs to scope of resource-based view in the connotation. Amit and Schoemaker (1993) pointed out that resources played a basic role in firm ability. Wang Xiayang and Chen Honghui (2002) showed that resources were starting point, incentives and conditions of firm network construction activities. Therefore, resources are the basic element of networking activities, cultivating and developing firm network capability show obvious resource-dependent.
3. *Dynamic evolution*: As an external oriented capability, firm network capability makes changes and adjustments with changing external environment. Gulati (1999) pointed out that firm network capability changed constantly with external environment. Xing Xiaoqiang and Tong Yunhuan (2006) stressed self-renewing characteristics which reflected dynamic properties of network capability. Therefore, firm network capability is dynamic with external environment.

20.3 Constituent Elements of Firm Network Capability Under "Process Perspective"

Based on process prospective of network activities, firm network capability is divided into: network planning capability, network building capability, and network managing capability.

20.3.1 Network Planning Capability

From existing researches on network planning capability, this paper considers that network planning capability is the ability that enterprises plan and design network from strategic perspective based on current situation and trend of external environment. And it belongs to strategic level capability. This article deems network planning capability consists of five abilities:

1. *Network requirements analysis capability*
   Network requirements analysis capability is an ability that enterprises confirm their network status and requirements by analyzing own resources and network environment (Ritter 1999; Zhu Xiu-mei et al. 2010). Including: internal analysis, environmental analysis, and match degree analysis.
2. *Network vision shaping capability*
   Network vision shaping ability is the ability that enterprises make network vision and goal combined with their development strategies based on understanding

network trends. Xing Xiaoqiang and Tong Yunhuan (2006) proposed that enterprises should adjust and reduce specific network locking effect to achieve best benefits; Zhao Shuang and Xiao Hongjun (2010) put forward the ability to ensure network activity goal, including cognizing network activities objective and designing network activities mode.

3. *Network strategy formulation capability*
   Network strategy formulation capability refers to a capability to make network strategy for shaping network vision and achieving competitive strategy. Von Friedrichs Grängsjö and Gummesson (2006) pointed out an ability to make specific network action plan, including choosing partners, developing relations, using interactive methods. Zhao Shuang and Xiao Hongjun (2010) proposed an ability to make network action strategy, including selecting appropriate mode to ally with partners and searching partners.
4. *Network opportunity recognition capability*
   Network opportunity recognition capability refers to a capability to find and identify potential opportunities in network environment. Möller and Halinen (1999) pointed out that key challenges of network opportunity recognition capability were how to identify network and find opportunities effectively; Fang Gang (2011) proposed the ability to choose the best networking opportunity to make enterprises exploit and use strategic opportunities in network, this capability was an important dimension of network planning capability.
5. *Network evolution prediction capability*
   Network evolution prediction capability is the ability to forecast network trend and evolution based on understanding, network environment. Möller and Halinen (1999) proposed that predicting network evolution trend was key of network vision capability. Xing Xiaoqiang and Tong Yunhuan (2006) considered that forecasting network structure and changes in magnitude and direction could reduce locking effect of specific network and get the best benefits.

20.3.2 Network Building Capability

From existing views on network building capability, this article thinks network building capability is a capability to seek and identify potential network partners, and build effective network contacts. And it belongs to executive level capability. This paper believes that network building capability comprises five areas:

1. *Network information gathering capability*
   Network information gathering capability refers to a capability that enterprises collect valuable information of potential partners through various channels. Ritter et al. (2002) pointed out that collecting information of potential partners by various channels was an important part to build network relations; Fang Gang (2011) proposed that discovery activities of network partners were to search and attract companies that can provide valuable information and make them become partners.

2. *Network resources evaluation capability*
   Network resources evaluation capability is a capability that enterprises assess values of potential partners' resources. Burt (1992) deemed that enterprises might avoid continuing to maintain redundant links which only provided less valuable resources and information; Möller and Halinen (1999) pointed out that assessing future values of relations was key challenge of single relation management.
3. *Network partners choice capability*
   Network partners' choice capability is a capability that enterprises select potential partners who are beneficial to network development and value maximization. Hagedoorn et al. (2006) indicated selecting valuable partners could improve efficiency of overall network; Ma Hongjia et al. (2010) pointed out that enterprises with strong ability of choosing partners might promote their network positions, and have an important effect on development of self-centric network.
4. *Network relations launched capability*
   Network relations launched capability is the ability that enterprises approach potential partners, exchange information, assess possibilities of establishing network relations. Zhu Xiumei et al. (2010) considered that enterprises launched network relations including visiting, getting to know potential partners and disclosing related information; Ren Shenggang et al. (2011) considered that enterprises took the initiative to launch cooperation with external actors.
5. *Network relations building capability*
   Network relations building capability refers to a capability that enterprises establish relations with potential partners using different strategies and roles. Zhu Xiumei et al. (2010) proposed that strengthening contact and building relationships with potential partners by using different strategies were important parts of network building activities. Äyväri and Jyrämä (2007) pointed out abilities in establishing network relationships included communication skills, ways of cooperation and so on.

20.3.3 Network Managing Capability

From existing concepts on network managing capability, the study considers that network managing capability refers to the ability enterprises coordinate single network relationship and deal with multiple network relations. Network managing capability in this article consists of the following six components:

1. *Internal resources allocation capability*
   Internal resources allocation capability is the ability that enterprises configured resources and put them into different contacts by assessing future values of network relations. Network as a resource, it need managerial and financial resources (Xu Jin-fa et al. 2001). Network managing activities included confirming the contribution of each network relation, specifying responsible persons, assigning public relations fees, holding meetings regularly (Zhu Xiu-mei et al. 2010).

2. *External resources acquisition capability*
   External resources acquisition capability is a capability that enterprises acquire resources from external organizations by network. Enterprises used other organizations' resources and capabilities to obtain scarce network resources through network management tasks of initiation, exchange, coordination and control (Xing Xiao-qiang and Tong Yun-huan 2006). Song Tiebo and Kong Lingcai (2008) considered that this capability was to get related resources from other network subjects based on trust.
3. *External resources integration capability*
   External resources integration capability is the ability that enterprises integrate network resources from different channels and internal resources to achieve strategic goals. Firms needed to integrate information, knowledge and other resources in the network, transforming them into enterprise's unique knowledge and abilities (Xing Xiao-qiang and Tong Yun-huan 2006); while the enterprise also needed to regroup various resources in the network to produce a synergistic effect (Ren Sheng-gang et al. 2011).
4. *Network relationship optimization capability*
   Network relationship optimization capability is a capability that enterprises form the optimal combination of multiple relationships by recognizing, developing and adjusting specific relationship based on strategic goals, environment changes and integration of network resources. Enterprises needed to continue to improve and develop network relations to achieve ideal combination between inputs and outputs (Ren Sheng-gang et al. 2011); the key challenge of enterprises in relationship portfolio management was how to optimize the relationship portfolio (Möller and Halinen 1999).
5. *Network relationship coordination capability*
   Network relationship coordination capability refers to the ability that enterprises coordinate effectively various network relationships, deal with problems in specific relationships and conflicts in multiple relationships to conduct network activities. On the one hand, organizations needed to coordinate specific network relationships, handle relationships with partners, solve problems and conflicts in cooperation and monitor implementation of commitments (Zhu Xiu-mei et al. 2010); On the other hand, enterprises needed to resolve conflicts in cooperation between different network relationships (Lumpkin and Dess 1996).
6. *Network relationship control capability*
   Network relationship control capability refers to a capability that enterprises implement control activities of strengthening, weakening or terminating relations by comparing and evaluating costs, contributions and future values of different relationships. Ritter (1999) indicated control activities of multiple relations including results of network cooperation, network personnel performance, contributions of partners; Ren Shenggang et al. (2011) proposed the concept of relationship reconstruction, describing enterprises adjusted multiple cooperation relations to create, manage and terminate partnerships by evaluating future value of different network relations.

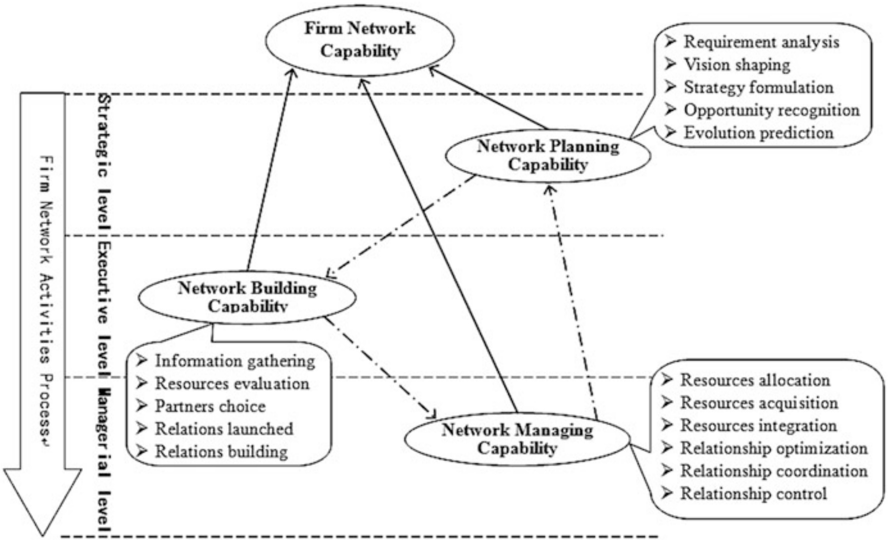


Fig. 20.1 Constituent dimension of firm network capability under "process perspective"

20.3.4 Constituent Dimension of Firm Network Capability Under "Process Perspective" (Fig. 20.1)

20.4 Literature Review of Influential Factors on Firm Network Capability

Summarizing existing researches, this paper divides influential factors of firm network capability into indemnificatory factors and promotion factors; they affect the formation and improvement of network capability jointly.

20.4.1 Indemnificatory Factors

Indemnificatory factors are prerequisites and foundational elements of improving network capability, including entrepreneurial orientation, internal resources and corporate culture.

Entrepreneurial orientation affected network capability from three dimensions of risk, innovativeness and initiative (SongTie-bo and Kong Ling-cai 2008). Firstly, risk tendency can promote enterprises to construct network for acquiring external resources and upgrade network capability purposefully; secondly, innovation-oriented enterprises could plan their network behaviors and improve their network capabilities by identifying potential needs and opportunities in the market, strengthening linkages with research institutions. Enterprises with strategic initiative orientation need use strategic foresight to anticipate future needs and seize market opportunities (Renzulli et al. 2000), and strategic vision comes from understanding and perception of the external environment, it can stimulate enterprises to extend network relations and strengthen cooperation and exchange.

Adequate internal resources would attract potential partners, they were beneficial to carry out network activities and enhance network capability (Wang Xia-yang and Chen Hong-hui 2002). Specifically, adequacy of financial resources and other material resources can ensure that enterprises invest necessary physical resources to network activities; the number and quality of personnel who implemented network cooperation task could also affect intensity of network capability (Xu Jin-fa et al. 2001). Timely, rich and high-quality information resources made enterprises better understand customer needs and more quickly identify network opportunities; therefore, information resources for firm network capability are also important.

The openness of corporate culture was also fundamental factor of network capability development (Ritter 1999). Flexible, spontaneous and encouraging individuality open corporate culture encouraged employees to take responsibility and stimulated their enthusiasm and innovative spirit to seek and create new external relations and opportunities (Xu Jin-fa et al. 2001). Open corporate culture also could motivate staff to coordinate and balance various relations and maintain benign development of contacts (Li Zhen and Ti-Qin Zhang 2010).

20.4.2 Promotion Factors

Promotion factors emphasize specific methods and means of improving network capability, including communication structure, network orientation of human resource management and organizational learning.

Integration of communication structure focuses on reflecting internal exchange efficiency within the enterprise. Efficient integration of communication structure made network staff obtains sufficient information to establish and optimize network relations; smooth information exchange of cross-sector could raise interdepartmental collaboration to promote formation of common vision (Calantone et al. 2002). Open communication can ensure roles of departments, adjust expectations and avoid misunderstandings, it affects construction and operation of business relations.

Enterprise network personnel need to have certain qualifications, such as skills, knowledge, communication, cooperation, conflict management and so on. Network orientation of personnel selection regarded capabilities and experience in network activities as key indicators of job description and personnel selection (Li Zhen and Ti-Qin Zhang 2010), it could help to deploy suitable personnel on network relations. Network orientation of personnel training is to improve qualifications and skills of staff by training their job skills, to ensure cooperation smoothly. Network orientation of personnel assessment is to intensify evaluation standards of network executive effect in individual performance.

Network capability, as an organizational ability, organizational learning structure was key element (Möller and Halinen 1999). Organizational learning structure consisted of team orientation, learning orientation, system orientation and memory orientation (Hult and Ferrell 1997). Team-oriented features affected establishment and development of individual knowledge network; they play a role in collaboration and improve the efficiency of relationships. Important significance of learning orientation lay in stimulating employees' learning enthusiasm and motivation by commitment, open mind and shared vision (Calantone et al. 2002), enhancing understanding of organization environment, network status and requirement to launch and establish new relations. System orientation from overall perspective enhances network capability by identifying existing problems, adjusting network strategic goals and optimizing network portfolio. Memory orientation could accelerate accumulation and routinization of experiences which was beneficial to accumulation and diffusion of knowledge and experience within the firm, learning effect caused by experience accumulation can enhance network managing capability.

20.5 Conclusion

Based on "process perspective" of network activities, the paper proposes a "process perspective" basic framework of firm network capability. The main conclusions are drawn as follows:

1. From "process perspective" of network activities, firm network capability is defined as a dynamic capability that firms can take the initiative to implement a series of network planning, building, and managing activities to obtain scarce network resources and competitive advantages based on internal resources and resources leverage. It has three typical features including process-oriented, resource-dependent and dynamic evolution.
2. Based on process prospective of network activities, firm network capability is divided into: network planning capability, network building capability, and network managing capability. Network planning capability is the ability that enterprises plan and design network from strategic perspective based on understanding entire external network, including five abilities of requirements analysis, vision shaping, strategy formulation, opportunity recognition and evolution prediction; network building capability is the ability to seek and identify potential network partners, and build effective network contacts, including five areas of information gathering, resources evaluation, partners choice, relations launched and relations building; Network managing capability refers to the ability that enterprises coordinate single network relationship and deal with multiple network relations, including six components of resources allocation, resources acquisition, resources integration, relationship optimization, relationship coordination and relationship control.

3. From "process perspective", the paper proposes influential factors' analysis framework of firm network capability. Indemnificatory factors are prerequisites and foundational elements, including entrepreneurial orientation, internal resources and corporate culture; promotion factors emphasize specific methods and means of improving network capability, including communication structure, network orientation of human resource management and organizational learning. They affect the formation and improvement of network capability jointly.

References

Amit R, Schoemaker PJH (1993) Strategic assets and organizational rent. Strateg Manage J 14(1):33–46

Äyväri A, Jyrämä A (2007) The networking abilities of craft entrepreneurs. In: 23rd IMP conference, University of Manchester, Manchester, 2007

Burt RS (1992) Structural holes: the social structure of competition. Harvard University Press, Cambridge

Calantone RJ, Cavusgil ST, Zhao Y (2002) Learning orientation, firm innovation capability and firm performance. Ind Market Manage 31(6):515–524

Dyer JH, Singh H (1998) The relational view: cooperative strategy and sources of inter organizational competitive advantage. Acad Manage Rev 23(4):660–679

Fang Gang (2011) The component of network capability and its effect on innovation performance. Stud Sci Sci 29(3):461–470

Gulati R (1999) Network location and learning: the influence of network resources and firm capabilities on alliance formation. Strateg Manage J 20:397–420

Hagedoorn J, Roijakkers N, Kranenburg HV (2006) Inter-firm R&D networks: the importance of strategic network capabilities for high-tech partnership formation. British J Manage 17(1):39–53

Håkansson H (1987) Industrial technological development: network approach. Croom Helm, London

Hult GTM, Ferrell OC (1997) Global organizational learning capacity in purchasing: construct and measurement. J Bus Res 40(2):97–111

Li Zhen, Ti-Qin Zhang (2010) Theoretical framework and constituent dimensions of corporate knowledge network competence. China Ind Econ 1(10):107–116

Lumpkin GT, Dess GG (1996) Clarifying the entrepreneurial orientation construct and linking it to performance. Acad Manage Rev 21(1):135–172

Ma Hong-jia, Dong Bao-guo, Chang Guan-qun (2010) Network capability and entrepreneurial capability: the empirical study of the start-ups in the northeast China. Stud Sci Sci 28(7):1008–1014

Möller KK, Halinen A (1999) Business relationships and networks: Managerial challenge of network era. Ind Market Manage 28(5):413–427

Ren Sheng-gang, Yu Meng, Wang Long-wei (2011) The measurement of firm network competence and empirical research. Chin J Manage 8(4):531–538

Renzulli LA, Aldrich H, Moody J (2000) Family matters: gender, networks and entrepreneurial outcomes. Soc Forces 79(2):523–546

Ritter T (1999) The networking company: antecedents for coping with relationships and networks effectively. Ind Market Manage 28(5):467–479

Ritter T, Gemünden HG (2003) Network competence: its impact on innovation success and its antecedents. J Bus Res 56(9):745–755

Ritter T, Wilkinson IF, Johnston WJ (2002) Measuring network competence: some international evidence. J Bus Ind Market 17(23):119–138
SongTie-bo, Kong Ling-cai (2008) The research on the relationship between enterprises' network competence and discontinuous technological innovation. Stud Sci Sci 26(11):224–229
Von Friedrichs Grängsjö Y, Gummesson E (2006) Hotel networks and social capital in destination marketing. Int J Serv Ind Manage 17(1):58–75
Walter A, Auer M, Ritter T (2006) The impact of network capabilities and entrepreneurial orientation on university spin-off performance. J Bus Ventur 21(4):541–567
Wang Xia-yang, Chen Hong-hui (2002) Internationalization of SMEs based on the resource base and network capability. Foreign Econ Manage 24(6):23–28
Xing Xiao-qiang, Tong Yun-huan (2006) Network competence: concept, structure and influential factors. Stud Sci Sci 24(12):558–563
Xu Jin-fa, Xu Qiang, Wang Yong (2001) Firm network capability analysis. Foreign Econ Manage 23(11):21–25
Zhao Shuang, Xiao Hong-jun (2010) Study on the increasing paths of firm performance based on network capability. Sci Technol Prog Policy 27(6):71–75
Zhu Xiu-mei, Chen Chen, Yang Jun-ping (2010) Dimension examination of new firm's network competence and conceptual construction. Stud Sci Sci 28(8):1222–1229

Chapter 21
The Shortest Path Algorithm Based on Petri Net

Yu-jie Zheng, Kai-hu Hou, Wei-zhen Liao, and Lin Yang

Abstract The Dijkstra algorithm is the classic algorithm to solve the shortest path problem, but the solving process is relatively complicated. As the visual graphics ability and good computer skills of Petri Net, it is used to solve the shortest path problem, and according to the thought of directed Petri Net and transition enabled rules, Petri Net algorithm of solving the shortest path problem is designed. Compared to the Dijkstra algorithm, this algorithm which omits the P, T tabs and λ, S functions of the Dijkstra algorithm, can make the solution of the shortest path simpler and more convenient, improve the solution efficiency, and at the same time provide convenience for achieving algorithm objectively using computer.

Keywords Dijkstra algorithm • Petri Net algorithm • Petri Net of directed network • The shortest path problem

21.1 Introduction

The shortest path problem is one of the most important optimization problems. It can be directly applied to solve many of the problems in the actual production, such as pipe laying, circuit arrangement, factory layout, equipment updating, and it is often as one of the basic tools, used to solve the problems of the optimization of the others (Sun Zhixin et al. 2007). Theoretically, the question has many algorithms, the classic Dijkstra algorithm is put forward in 1959, namely Dijkstra method. This algorithm is a kind of excellent shortest path algorithm, which has a wide range of applications in the network calculation and optimization (Noshita 1985). But there are heavy steps and complicated calculation, its availability is not

Y. Zheng (✉) • K. Hou • W. Liao • L. Yang
Department of Industrial Engineering, Kunming University of Science and Technology, Kunming, People's Republic of China
e-mail: zhengyujiesmile@sina.com; kaihu_kmustie@163.com

E. Qi et al. (eds.), *The 19th International Conference on Industrial Engineering and Engineering Management*, DOI 10.1007/978-3-642-37270-4_21,

satisfactory. In recent years, some scholars have put forward many new methods to solve this problem. In the first literature, an adjacency list storage model with heuristic information was proposed, the process of initial population with heuristic information and chromosome coding, and the strategy of dynamic adjustment of the fitness function were presented (Li Xiongfei et al. 2011). In the second literature, a path nodes-driven idea was introduced, which derived from the destination does-drive idea. Using the idea, an algorithm called least-cost shortest path tree (LCAPT) was designed (Zhou Ling and Wang Jianxin 2011). In the third literature, Dijkstra algorithm was revised, and all shortest paths can be given according to the number of edges by applying Yen algorithm (Wang Zhijian et al. 2010). In the fourth literature, the solving random network of the shortest circuit simulation method based on STPN was proposed, the shortest circuit of distribution function from starting to finishing the random network was presented (Hu Xiongying et al. 2010). In the fifth literature, the traffic network weighted S-figure of nets model was put forward, and the weighted S-figure of the shortest path algorithm was proposed (Li Shuju et al. 2011).

Petri Net was firstly suggested by Germany's CA. Dr Petri in 1962, it is a kind of system modeling and analysis tool, suitable for describing parallel, a synchronization, distribution, uncertainty and randomness of information processing system (Jong et al. 1999). Petri Net takes graphical modeling method, which makes model intuition and more easy to understand. It has a profound connotation of mathematics and foundation, and the good formalism description method. Because of the need to depict system, some experts define and research all the factors of Petri Net to make the application field of Petri Nets further expanded (Ding Yijie and Li Xiaozhong 2011; Yao Yulin 1994).

In view of Petri Net's simple, intuitive and accurate graphical modeling ability and logical ability, the paper imitates the shortest path problem with Petri Net. On the base of this model, the shortest path problem is solved, and the shortest path problem of Petri Nets algorithm is put forward, which makes the shortest path algorithm more convenient.

21.2 The Petri Net of Directed Network Presentation

21.2.1 Definition of the Fundamental Petri Net

The fundamental Petri Net is a triple $N = \{P, T, F\}$, among:

1. $P = \{p_1, p_2, \ldots, p_n\}$ is place set, n is place number;
2. $T = \{t_1, t_2, \ldots, t_m\}$ is transition set, m is transition number;
3. $P \cup T \neq \emptyset$
4. $P \cap T = \emptyset$

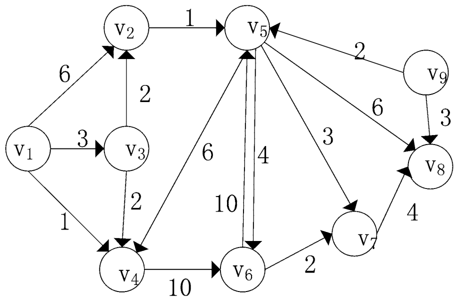


Fig. 21.1 The directed network $G = (V, A, W)$

5. $F \in (P \times T) \cup (T \times P)$, the single direction of the contact, which from place to transition, and from transition to place is set up. And similar provisions between elements can't directly connected (Song Yubo et al. 2011; Wang Guangqiu and Zhang Bin 2009).

21.2.2 The Petri Net's of Directed Network Definition

In order that the weight in G can be signified by PN, the Petri Net of directed network is defined. The Petri Net of directed network is a septet, $PN = \{P, T, F, D, I, W, M_0\}$, among:

1. $\{P, T, F\}$ is fundamental Petri Net;
2. D is the least time which tokens take from place p_1 to place p_i;
3. I is that the former place of p_n is p_i, in the shortest cut from place p_1 to p_n;
4. M_0 is the initial mark.

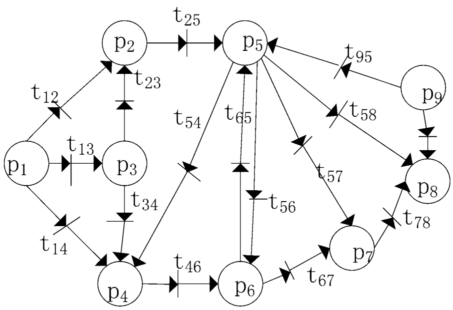
Given the directed network $G = (V, A, W)$, among $V = \{v_1, v_2, \ldots, v_n\}$ is vertex set, $A = \{a_1, a_2, \ldots, a_m\}$ is directed arc set, a is a random edge in A and strictly ordered elements in $V = \left(v_i, v_j\right)$. W is nonnegative function set, w_{ij} is the weigh from v_i to v_j (Dijkstra 1999; Noshita et al. 1978).

The vertex v_i in Fig. 21.1 (Gan Ying'ai 2005) is signified by place p_i, the transition which from place p_i to p_j is signified by t_{ij}, also the directed arc a_{ij}, the time that transition is trigged need to take is signified by the value of t_{ij}, then the structure and logical relationship of the directed network G are well signified by Petri Net, as Fig. 21.2.

21.3 The Shortest Algorithm Based on Petri Net

The algorithm of Petri Net marks the trigger of transition and the time of tokens reach to the place, looking for the shortest path from the origin to terminal point. $L(p_i)$ and $L\left(t_{ij}\right)$ are the marks of place and transition. $L(p_i) = 0, L(p_i) = 1$

Fig. 21.2 Petri Net



mean that there is no tokens in place p_i and tokens reached to the place p_i respectively, $L(p_i) = 2$ is a token which takes the least time to reach to the place p_i. $L(t_{ij}) = 0$ and $L(t_{ij}) = 1$ are the transition trigged or not. Initially $L(p_1) = 1, L(p_i) = 0$, among $i \neq 1, D(p_1) = 0$.

1. The algorithm is finished when $L(p_i) \neq 1$, by now, there is $D(p_i) = d(p_1, p_i)$ to every p_i.
2. Choosing $\min\{D(p_i^*)\}$ from all $L(p_i) = 1$, transition t_{ij}^* related to the place is superior to be trigged, then $L(p_i^*) = 2, D(p_j) = D(p_j^*) + t_{ij}^*, L(t_{ij}^*) = 2$, $L(p_j) = 1$.

   Attention $D(p_i^*) = \min\{D(p_i^k), D(p_i^m)\}$, m and k are iterations.

21.4 Application Example of the Shortest Path Algorithm

The directed network is shown in Fig. 21.1, vertex v_1 is the initial, vertex v_9 is terminal point. The digital w_{ij} next to arc is the weight, please to solve the shortest path of the network.

According to the algorithm in this paper, firstly the directed network of Petri Net is modeled. The place is marked as shown in Fig. 21.2. Iterative processes are as follows:

1. As $L(p_1) = 1$, it means that there is a token in place $p_1, D(p_1) = 0$, so $L(p_1) = 2$, transition t_{12}, t_{13} and t_{14} are trigged.

$$D(p_2) = D(p_1) + t_{12} = 0 + 6 = 6$$

$$L(t_{12}) = 1, L(p_2) = 1$$

$$D(p_3) = D(p_1) + t_{13} = 0 + 3 = 3$$

$$L(t_{13}) = 1, L(p_3) = 1$$

$$D(p_4) = D(p_1) + t_{14} = 0 + 1 = 1$$

$$L(t_{14}) = 1, L(p_4) = 1$$

2. As $L(p_2) = 1, L(p_3) = 1, L(p_4) = 1$, the least one p_4 among p_2, p_3 and p_4 is chosen, the transition t_{46} related to p_4 is prior to be triggered.

$$L(p_4) = 2$$
$$D(p_6) = D(p_4) + t_{46} = 1 + 10 = 11$$
$$L(t_{46}) = 1$$
$$L(p_6) = 1$$

3. As $L(p_2) = 1, L(p_3) = 1, L(p_6) = 1$, the least one p_3 among p_2, p_3 and p_6 is chosen, the transition t_{32} related to p_3 is prior to be triggered.

$$L(p_3) = 2$$
$$D(p_2) = D(p_3) + t_{32} = 3 + 2 = 5$$

According to

$$D\left(p_i^*\right) = \min\left\{D\left(p_i^k\right), D\left(p_i^m\right)\right\}$$
$$D(p_2) = \min\left\{D\left(p_2^1\right), D\left(p_2^2\right)\right\} = 5$$

4. As $L(p_2) = 1, L(p_6) = 1$, the least one p_2 among p_2 and p_6 is chosen, the transition t_{25} related to p_2 is prior to be triggered.

$$L(p_2) = 2$$
$$D(p_5) = D(p_2) + t_{25} = 5 + 1 = 6$$
$$L(t_{32}) = 1$$
$$L(p_5) = 1$$

5. As $L(p_5) = 1, L(p_6) = 1$, the least one p_5 among p_5 and p_6 is chosen, the transition t_{56}, t_{57}, t_{58} are prior to be triggered.

$$L(p_5) = 2$$
$$D(p_6) = D(p_5) + t_{56} = 6 + 4 = 10$$

According to

$$D\left(p_i^*\right) = \min\left\{D\left(p_i^k\right), D\left(p_i^m\right)\right\},$$
$$D(p_6) = \min\left\{D\left(p_6^2\right), D\left(p_6^4\right)\right\} = 10$$
$$L(t_{56}) = 1$$
$$L(p_6) = 1$$
$$D(p_7) = D(p_5) + t_{57} = 6 + 3 = 9$$

$$L\left(t_{57}\right)=1$$

$$L\left(p_{7}\right)=1$$

$$D\left(p_{8}\right)=D\left(p_{5}\right)+t_{58}=6+6=12$$

$$L\left(t_{58}\right)=1$$

$$L\left(p_{8}\right)=1$$

6. As $L\left(p_{6}\right)=1, L\left(p_{7}\right)=1, L\left(p_{8}\right)=1$, the least one p_7 among p_6, p_7 and p_8 is chosen, the transition t_{78} related to p_7 is prior to be triggered.

$$L\left(p_{7}\right)=2$$

$$D\left(p_{8}\right)=D\left(p_{7}\right)+t_{78}=9+4=13$$

According to

$$D\left(p_{i}^{*}\right)=\min\left\{D\left(p_{i}^{k}\right), D\left(p_{i}^{m}\right)\right\}$$

$$D\left(p_{8}\right)=\min\left\{D\left(p_{8}^{5}\right), D\left(p_{8}^{6}\right)\right\}=12$$

$L\left(t_{78}\right)=1, L\left(p_{8}\right)=1$ are also achieved.

7. As $L\left(p_{6}\right)=1, L\left(p_{8}\right)=1$ the least one p_6 among p_6 and p_8 is chosen, $L\left(p_{6}\right)=2$, but $L\left(p_{5}\right)=2, L\left(p_{7}\right)=2$, so the transition t_{65} and t_{65} are not triggered.
8. Just because $L\left(p_{8}\right)=1$, so $L\left(p_{8}\right)=2$, and no transition can be triggered. So $D\left(p_{8}\right)=12$.

By now, only place p_9 remains no token, the algorithm is finished.

There are the shortest path for $i=2,3,4,5,6,7,8$ and no path from p_1 to p_9. According to the reversed track of the triggered transition sequence, the shortest path from p_1 to p_9 can be gained $i=2,3,4,5,6,7,8$.

21.5 The Procedure of the Shortest Path Algorithm Based on Petri Net

```
#include <stdio.h>
#include "Conio.h"
#define I  9999
#define N  9
int cost[N][N] = {
   {0,6,3,1,I,I,I,I,I},
   {I,0,I,I,1,I,I,I,I},
   {I,2,0,2,I,I,I,I,I},
```

```
    {I,I,I,0,I,10,I,I,I},
    {I,I,I,6,0,4,3,6,I},
    {I,I,I,I,10,0,2,I,I},
    {I,I,I,I,I,I,0,4,I},
    {I,I,I,I,I,I,I,0,I},
    {I,I,I,I,2,I,I,3,0}
};
int dist[N];
int v0 = 'A' - 65;
int main()
{
    int P[N],i,v,w,min,k;
        printf("\n 任意两个定点之间的最短路径如下 :\n\n");
    for(k=0;k<N;k++)
    { for (v = 0; v < N; v++)
            {
              { P[v] = 0;
         dist[v] = cost[v0][v];
                  }
                }
        P[v0] = 2;
        for (i = 0; i < N-1; i++)
        { min = I;
          for (w = 0; w < N; w++)
         { if (P[w]=1 && dist[w] < min)
           {min = dist[w];
                 v = w;
           }
         }
            P[v] = 2;
            for (w = 0; w < N; w++)
            { if (P[w]=1 && dist[v] + cost[v][w] < dist[w])
              dist[w] = dist[v] + cost[v][w];
            }
        }
         for (i = 0; i < N; i++)
         { printf("%c->%c: %2d\t", v0 + 65, i + 65, dist[i]);
           }
         printf("\n");
            v0++;
    }
      return 0;
    }
```

The result of procedure is showed in Fig. 21.3.

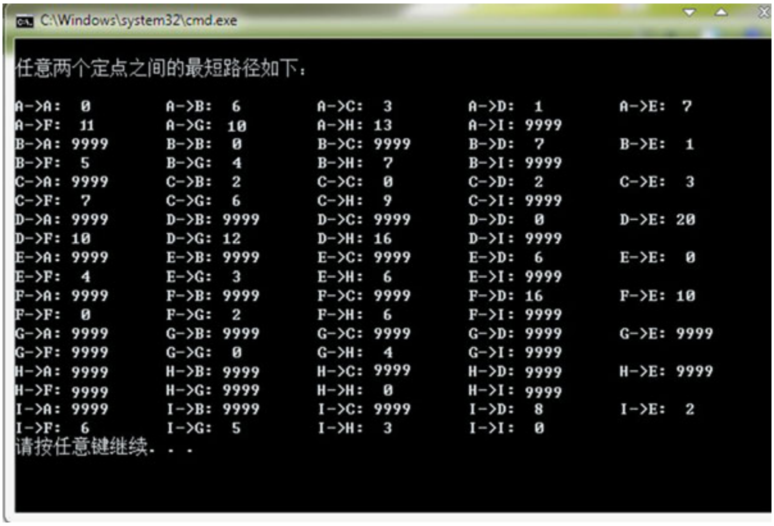


Fig. 21.3 The result of procedure

21.6 Conclusion

Petri Net has been proved to be an effective tool for system modeling. Petri Net has a great value in system static and dynamic characteristics by using network graphics to describe objects of the relationship between input and output (Yu Ping et al. 2011). In this paper, Petri Net method is applied to solve the shortest path problem, according to the thought of directed Petri Net and transition enabled rules, the Petri Net algorithm of the shortest path problem is designed. The algorithm simplifies the Dijkstra method, uses the reversed trace of transition instead of the marks of P and T, omits the functions of λ and S, improves the solution efficiency, at the same time provides convenience for the computer to achieve objectively algorithm.

References

Dijkstra EW (1999) A nite on two problem in connexion with graphs. Numerische Mathematik 1(1):269–271

Ding Yijie, Li Xiaozhong (2011) Modeling and analysis of knowledge flow based on colored Petri Nets. J Syst Simul 23(Suppl 1):287–290 (in Chinese)

Gan Ying'ai (2005) The operational research. Tsinghua University Press, Beijing, pp 260–264 (in Chinese)

Hu Xiongying, Hu Bin, Zhang Jinling (2010) Solving the random network shortest circuit simulation method based on STPN. J Wuhan Univ Sci Technol (Inf Manag Eng Ed) 32(2):197–201 (in Chinese)

Jong WT, Shiau YS, Horng YJ et al (1999) Temporal knowledge representation and reasoning techniques using time Petri Nets. IEEE Trans Syst Man Cybern Part B 29(4):541–545

Li Shuju, Li Wenjing, Wang Rujing (2011) Study of shortest path based on weighted indentification S-graph. J Comput Eng 37(15):46–48 (in Chinese)

Li Xiongfei, Zhang Hailong, Liu Zhaojun (2011) Problem for shortest path problem based on heuristic algorithm. J Jilin Univ (Eng Technol Ed) 41(1):182–187 (in Chinese)

Noshita K (1985) A theorem on the expected complexity of Dijkstra's shortest paths algorithm. J Algorithm 6(3):400–408

Noshita K, Masuda E, Machida H (1978) On the expected behaviors of the Dijkstra's shortest paths algorithm for complete graph. Inf Process Lett 7(5):237–243

Song Yubo, Jiang Zhaoyuan, Mu Haibo (2011) Algorithm for seeking the minimal cost and maximal flow based on Petri net. J Lanzhou Jiaotong Univ 30(3):67–70 (in Chinese)

Sun Zhixin, Gao Yanjuan, Wang Wending (2007) Complete dynamic algorithm for updating shortest path tree. J Jilin Univ (Eng Technol Ed) 37(4):860–864 (in Chinese)

Wang Guangqiu, Zhang Bin (2009) Optimum route sequence search in stochastic Petri Net based on reliability of transition. J Syst Eng Electron 31(1):187–194 (in Chinese)

Wang Zhijian, Han Weiyi, Li Yijun (2010) Shortest path problem with multiple shortest paths. J Harbin Eng Univ 42(9):1428–1431 (in Chinese)

Yao Yulin (1994) A Petri Net modal for temporal knowledge representation and reasoning. IEEE Trans Syst Man Cybern 24(9):1374–1382

Yu Ping, Cao Jiping, Song Jianshe (2011) Improved fuzzy Petri Net algorithm and its application in fault diagnosis. Ordnance Ind Autom 30(6):20–22 (in Chinese)

Zhou Ling, Wang Jianxin (2011) Path nodes-driven least-cost shortest path tree algorithm. J Comput Res Dev 48(5):721–728 (in Chinese)

Chapter 22
Measuring Satisfaction of Two Level Employees in a Representative Chinese Steel Plant: A Structure Equation Modeling Approach

Guo-wen Huang and Hong-wei Liu

Abstract Although employee satisfaction has been well explored in academic literature, there exists little research on this subject with respect to steel plants, especially to different level employees in China. In this paper, the latent variables including employee relationships, salary, career development, organization provision, and employee satisfaction are constructed based on previous literatures. The objective of this paper is to explore the different relations and effects of these latent variables between production workers and managers in C steel plant which is a representative Chinese steel plant. Data obtained from the questionnaires was analyzed by structure equation modeling by AMOS 17.0. The analytical results show that: production workers are more care about their salary while managers pay more attention to organization provision and relationships; for both production workers and managers, the covariance between salary and achievements is negative, which indicates that employees feel their salaries are not enough to fit their efforts.

Keywords AMOS • Employee satisfaction • Structure equation modeling • Steel plant management

22.1 Introduction

Employee satisfaction is defined as the satisfaction of employees with their jobs or the degree to which employees like their jobs. Employee satisfaction can reduce staff turnover, enhance creativity and commitment, and improve productivity and quality (Halkos and Bousinakis 2010). The employee satisfaction has been well explored in academic literature in last several decades. Hartline and Ferrell (1996)

G. Huang (✉) • H. Liu
College of Management and Economics, Tianjin University, Tianjin, China
e-mail: hgw0610209@gmail.com

E. Qi et al. (eds.), *The 19th International Conference on Industrial Engineering and Engineering Management*, DOI 10.1007/978-3-642-37270-4_22,

have argued that service quality is influenced by job satisfaction of employees. They found evidence that employee satisfaction is associated with service quality. Dubrovski (2001) pointed out that improving customer satisfaction not only raises company profits, but also facilitates company development. Oshagbemi (1997) indicated that employee satisfaction had been found to be as important as customer satisfaction. Nebeker et al. (2001) pointed out that employees are the greatest assets of a company, and that customer satisfaction is based on satisfying employee requirements. In a word, employees, who are the internal customers of the business, their satisfaction can influence organizational performance as much as customer satisfaction. They are willing to cooperate with the business to accomplish business goals when they satisfy the current working environment.

Employee satisfaction is a subject that has appeared many times over the past decade in academic research and various employee satisfaction-related elements exist. Antoncic and Antoncic (2011) mentioned that job satisfaction factors can be classified accordingly to the well-known Herzberg's (1964) two-factor theory into hygienes (supervision, working conditions, co-workers, pay, policies/procedures and job security), which lead to dissatisfaction, and motivators (achievement, recognition, the work itself, responsibility, advancement and growth), which lead to satisfaction. In order to discusses the relationship between staff satisfaction and loyalty to the employing organization, Becker and Kaerkes (2006) presented a table of employee satisfaction factors and lists parameters affecting satisfaction, including co-operation with colleagues, actual work content, environment, working time, management attitudes, in-company information, payment, personal development and further education opportunities. Xuejun (2011) put forward 20 variable factors which affect employee satisfaction achieved by analyzing employee satisfaction of Changjiang hotel and five factors are extracted by factor analysis based on SPSS, including work environment, incentives, work intensity, career planning, job stress. According to previous literatures, important elements that affect employee satisfaction include:

Work conditions (Turkyilmaz et al. 2011; Androniceanu et al. 2010; Roelen et al. 2008; Togia et al. 2004), working time (Becker and Kaerkes 2006; Jelacic et al. 2007; Pierce and Newstrom 1980), reputation of the company (Jelacic et al. 2007), Employee relationships (Becker and Kaerkes 2006; Jelacic et al. 2007; Bowen et al. 2008; Westover and Taylor 2010; Yee et al. 2008), salary (Becker and Kaerkes 2006; Togia et al. 2004; Bowen et al. 2008; Yee et al. 2008; Hanneman and Schwab 1985), promotion (Androniceanu et al. 2010; Togia et al. 2004; Bowen et al. 2008), further education (Becker and Kaerkes 2006; Schmidt 2010), career development (Becker and Kaerkes 2006; Androniceanu et al. 2010), incentives (Westover and Taylor 2010), working intensity (Roelen et al. 2008), environment (Becker and Kaerkes 2006; Jelacic et al. 2007), work content (Becker and Kaerkes 2006), employment assurance (Jelacic et al. 2007), benefits (Williams 1995).

Basing on the analysis of the previous literatures, in this paper, we assume that employee satisfaction is affected by four factors including employee relationships, salary, career development, and organization provision (Fig. 22.1). We classify the

Fig. 22.1 Employee satisfaction model

employees into two levels: production workers and managers. Our objective is to examine a structural model of the relationships of both two level employees between employee satisfaction and its affected factors in a representative Chinese steel plant. And then show the differences of satisfaction relationships between production workers and managers.

22.2 Methodology

22.2.1 Questionnaire Design and Structure

The purpose of this study is to explore the differences of employee satisfaction between production workers and managers in steel plants. To achieve the objectives of the present study, a questionnaire is designed (Table 22.1). Respondents were asked to rate on each item on a seven-point Likert-type scale anchored at 1 = "totally disagree" and 7 = "totally agree".

22.2.2 Data Collection and Analysis

Data for this study were gathered using a questionnaire that were distributed to 3,500 employees (including 2,500 production workers and 1,000 managers) of C steel plant which is a representative steel plant in China. Herein, we want to explain a little more why C steel plant is a representative steel plant in China. In total, the response rate was 73 and 87 % for production workers and managers, respectively. Next, we made a pre-dispose of the return questionnaires that eliminated the blank ones and identical ones. Two questionnaires were identical if they were answered with the same handwritings. Finally, there were 2,426 useable questionnaires (1,674 from production workers and 752 from managers), which forms a satisfactory dataset for subsequent analysis.

Table 22.1 Employee satisfaction questionnaire

Scales	Items
Employee satisfaction	You are committed to the organization
	You are proud to tell others that you are part of the organization
	For you, this is the best of all possible organizations for which to work
	You feel a great sense of personal satisfaction when you do the job well
	You feel the respect of your position inside the company
Relationship	Your fellow workers are the kind you would like to have around
	You get along well with your coworkers
	You are happy with the relationship with your fellow workers
Salary	You are satisfied with the pay you receive for your job
	Your salary is higher while comparing to the same position in other organizations
Organization provision	Employees are relatively well-rewarded financially for their work
	You are satisfied with the fairness of promotion system in the company
	You are satisfied with your working conditions
Career development	You feel satisfied with your career development
	You feel satisfied with potential promotion opportunities

Table 22.2 Cronbach's alpha-value of latent variables

Scale	Cronbach's alpha (production workers)	Cronbach's alpha (managers)
Employee satisfaction	0.90	0.88
Relationship	0.89	0.81
Salary	0.78	0.78
Career development	0.70	0.79
Organization provision	0.82	0.93

22.2.3 Analysis of Reliability and Validity

Before evaluating the conceptual relationships of the proposed model, an assessment of internal consistency and reliability of the measurement scale were conducted. Cronbach's alpha is a coefficient of consistency that measures how well a set of variables or items measures a single, unidirectional latent construct. In this study, Cronbach's alpha was applied to evaluate the internal consistency of the five latent factors. Ma et al. (2010) pointed out that a Cronbach's alpha of 0.7 or more indicates acceptable reliability. As shown in Table 22.2, the reliability of the scales is generally acceptable which implies that the survey approach is valid.

22.3 Methodology

22.3.1 *The Employee Satisfaction Model*

The research model was tested using the maximum likelihood methods of AMOS version 17.0, a structural equation modeling software. Relationships between the different factors and their impact on employee satisfaction were explored. The standardized path coefficients for production workers and managers models are reported in Figs. 22.2 and 22.3, respectively. All the standardized factor loadings were greater than 0.50 with statistically significant (all t-values are greater than 2.0). Therefore, convergent validity is achieved. In order to assess discriminant validity, a constrained CFA (Confirmatory Factor Analysis) model was used for each possible pair of constructs in which the correlations among this pair of constructs were fixed to 1. This model was subsequently compared to the original unconstrained model, in which the correlations among constructs are freely estimated. A significant difference of the two statistics between the fixed and unconstrained models indicates high discriminant validity. In our study, all the differences of two between the fixed and unconstrained model were significant at the 0.05 level. Therefore, discriminant validity is ensured. To estimate the accuracy of the SEM, five evaluation indexes are calculated and compared with ideal values. As demonstrated in Table 22.3, the comparison result shows that the models fit well with the ideal values. Based on the above results, the postulated models were considered tenable.

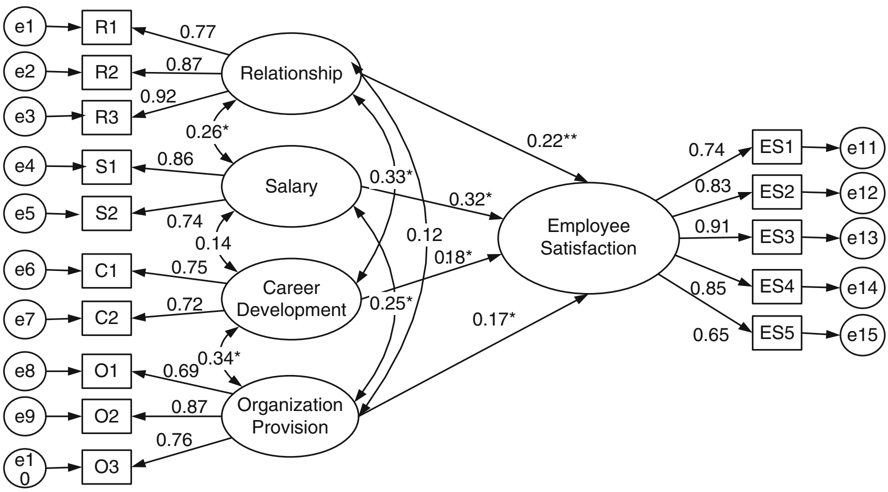


Fig. 22.2 Structural equation model (production workers). *Notes*: *Significance at 0.05; **Significance at 0.01

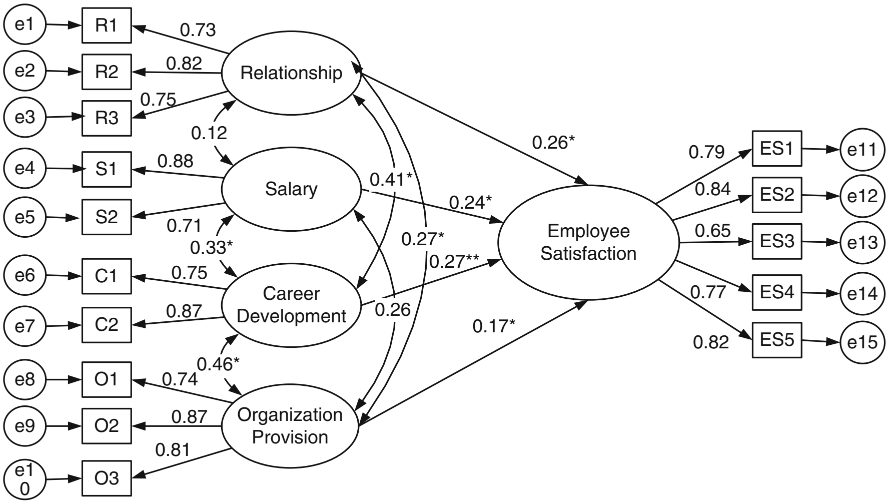


Fig. 22.3 Structural equation model (managers). *Notes*: *Significance at 0.05; **Significance at 0.01

Table 22.3 Model fitting degree testing

Evaluation index	Index calculation (production workers)	Index calculation (managers)	Ideal value
χ^2/degree of freedom	5.562	6.273	Range with sample scale
Goodness of fit index (GFI)	0.952	0.948	>0.9
Adjusted Goodness of fit index (AGFI)	0.926	0.911	>0.9
Root mean residual (RMR)	0.027	0.042	<0.05
Root mean square error of approximation (RMSEA)	0.058	0.063	<0.08

22.3.2 Discussion

The primary objective of this study is to explore the differences of employee satisfaction between production workers and managers in a representative Chinese steel plant. From the structural equation modeling results (see Figs. 22.2 and 22.3), it is apparent that employee satisfaction is highly related to employee relationship, salary, career development and organization provision for both production workers and managers. Furthermore, several important findings are generalized while taking a deep insight to the details of the path coefficients for the two models:

1. All of the four factors (employee relationship, salary, and career development and organization provision) have positive relationship to employee satisfaction for both production workers and managers.

2. Production workers pay more attention to salary while managers care more about their career development and employee relationship.
3. For production workers and managers, organization provision affects employee satisfaction in the same degree.

22.4 Conclusion

Finally, this paper takes the employees of C steel plant as the studying object, which company is a large-scale state-owned enterprise by the steel and iron production as the primarily business, and it is a representative example of Chinese steel plants. By reviewing the predecessors' researches, defining the frame of reference, and reviewing the theories about employee satisfaction, then developing a questionnaire to test the employee satisfaction of two-level employees (production workers and managers) of this company and studying the influence by the different factors on the employee satisfaction, finally, we get some findings about the differences of employee satisfaction between production workers and managers. The result shows: (1) All of the four factors (employee relationship, salary, career development and organization provision) have positive relationship to employee satisfaction for both production workers and managers; (2) Production workers pay more attention to salary while managers care more about their career development and employee relationship; (3) For production workers and managers, organization provision affects employee satisfaction in the same degree.

References

Androniceanu A, Sora S, Paun D (2010) Employee satisfaction in the public sector. In: 5th WSEAS international conference on economy and management transformation, West University of Timisoara, Romania, pp 272–277

Antoncic JA, Antoncic B (2011) Employee satisfaction, intrapreneurship and firm growth: a model. Ind Manag Data Syst 111(4):589–607

Becker R, Kaerkes W (2006) Employee satisfaction and loyalty. Qualitaet und Zuverlaessigkeit 51(3):18–22 (in German)

Bowen P, Cattell K, Mechell K, Edwards P (2008) Job satisfaction of South Africa quantity surveyors: are employers happier than employees? J Eng Des Technol 6(2):124–144

Dubrovski D (2001) The role of customer satisfaction in achieving business excellence. Total Qual Manag 12(7):920–925

Halkos G, Bousinakis D (2010) The effect of stress and satisfaction on productivity. Int J Product Perform Manag 59(5):415–431

Hanneman HG, Schwab DP (1985) Pay satisfaction: its multidimensional nature and measurement. Int J Psychol 20(2):129–141

Hartline MD, Ferrell OC (1996) The management of customer contact service employees: an empirical investigation. J Market 60(4):52–70

Herzberg F (1964) The motivation-hygiene concept and problems of manpower. Pers Adm 27:3–7

Jelacic D, Grladinovic T, Pirc A, Oblak L (2007) Motivation factors analysis in industrial plants. Strojarstvo 49(3):137–148
Ma M, Yan X, Huang H, Abdel-Aty M (2010) Safety of public transportation occupational drivers. Transp Res Rec 21(45):72–79
Nebeker D, Busso L, Werenfels PD, Diallo H, Czekajewski A, Ferdman B (2001) Airline station performance as a function of employee satisfaction. J Qual Manag 6:29–45
Oshagbemi T (1997) Job satisfaction profiles of university professors. J Manag Psychol 12(1):27–39
Pierce JL, Newstrom JW (1980) Toward a conceptual clarification of employee response to flexible working hours: a work adjustment approach. J Manag 6(2):117–134
Roelen CAM, Koopmans PC, Groothoff JW (2008) Which work factors determine job satisfaction? WORK: J Prev Assess Rehabil 30(4):433–439
Schmidt SW (2010) The relationship between job training and job satisfaction: a review of literature. Int J Adult Vocat Educ Technol 1(2):19–28
Togia A, Koustelios A, Tsigilis N (2004) Job satisfaction among Greek academic librarians. Libr Inf Sci Res 26(3):373–383
Turkyilmaz A, Akman G, Ozkan C, Patuszak Z (2011) Empirical study of public sector employee loyalty and satisfaction. Ind Manag Data Syst 111(5):675–696
Westover JH, Taylor J (2010) International differences in job satisfaction: the effects of public service motivation, rewards and work relations. Int J Product Perform Manag 59(8):811–828
Williams ML (1995) Antecedents of employee benefit level satisfaction. J Manag 21:1097–1128
Xuejun Z (2011) An empirical research on hotel employee satisfaction based on factor analysis – a case study of Changjiang hotel. In: International conference on management and service science, MASS, Wuhan
Yee RWY, Yeung ACL, Cheng TCE (2008) The impact of employee satisfaction on quality and profitability in high-contact service industries. J Oper Manag 26:651–668

Chapter 23
Performance Evaluation of CRM in Securities Investment Consulting Business Based on Fuzzy-AHP

Qing-guo Zhao and Song-tao Jiang

Abstract Performance evaluation plays an important role in the customer relationship management (CRM) process. This paper is to evaluate CRM performance of investment consulting business of securities companies based on analytic hierarchy process and fuzzy comprehensive algorithm through the establishment of evaluation indicators. At the end of the paper, a securities company is taken as an example to analyze the performance of securities investment consulting business customer relationship management and come to the corresponding conclusion.

Keywords Analytic hierarchy process • Balance score card • Customer relationship management • Fuzzy comprehensive algorithm

23.1 Introduction

With the development of China's securities industry, customer relationship management system has been given extensive attention and widely used in the securities investment consulting business. The role of CRM system is to help investment consultant data mining and customer segmentation, which is conducive to investment consultant to provide customers with one-on-one personalized service to increase customer satisfaction and loyalty, and ultimately securities companies to maximize profits (Zhao 2011).

Performance evaluation of CRM is an extremely important work. It is necessary for the enterprise to evaluate the performance of CRM scientifically and objectively to improve it when the enterprise carries out the CRM (Zhou et al. 2008). Many scholars have conducted studies on it. Curry and Kkolou (2004) studied the CRM

Q. Zhao (✉) • S. Jiang
Department of Economics and Management, Shenyang Aerospace University, Shenyang, China
e-mail: zhqg1000@163.com; jsten163@163.com

E. Qi et al. (eds.), *The 19th International Conference on Industrial Engineering and Engineering Management*, DOI 10.1007/978-3-642-37270-4_23,

evaluation to contribute to TQM improvement through a cross-case comparison. Antonio et al. (2007) researched the evaluation of CRM practices among agribusiness firms. Zhou et al. (2008) used fuzzy comprehensive algorithm in order to do a scientific evaluation on CRM performance. Kong et al. (2005) evaluated the effectiveness of CRM based on the BP algorithm. Wang (2008) evaluated the CRM performance of the third-party logistics from money value, image and fame three dimensions. Hu (2010) evaluates the performance of enterprise CRM system based on the grey-fuzzy comprehensive evaluation method. Chang et al. (2009) used analytic hierarchy process evaluating collaborative CRM System. Ullah and Al-Mudimigh (2009) proposed a CRM Scorecard based on Balanced Scorecard for the evaluation and enhancement of CRM systems of the Saudi Arabian Banks to improve the customer satisfactions and loyalty. Hong and Kim (2007) researched the selection of CRM systems in financial institutes through the analytic hierarchy. Tang and Liang (2009) studied the performance evaluation of CRM under supply chain integration through the fuzzy comprehensive evaluation method.

However, most of their researches are based on the corporate level, and few are focused on the business of a company. In this paper, we selected appropriate performance evaluation indicators to build CRM performance evaluation system based on the securities investment consulting business which is also the innovation of this paper. Finally, we hope that our research can provide some valuable advice for securities companies to carry out securities investment consulting business (SICB).

23.2 Performance Evaluation Index System for Securities Investment Consulting Business CRM

23.2.1 The Process of Securities Investment Consulting Business CRM

Investment consultant is the core of securities investment consulting business customer relationship management process. First of all, as a knowledge-intensive profession, learning and development is very important for investment consultant, it is the foundation for investment consultant to continue to enhance the consulting capacity. Secondly, customer data mining and management is the core of CRM, which is the prerequisite for securities investment consulting business. Through data mining and analysis of customer information,financial consultant take customer segmented and provide customers with targeted, personalized service and assist in the marketing department to do targeted marketing, ultimately to achieve the purpose of improving financial indicators and enhancing customer satisfaction. Meanwhile, the indicators such as financial and customer satisfaction feedback the

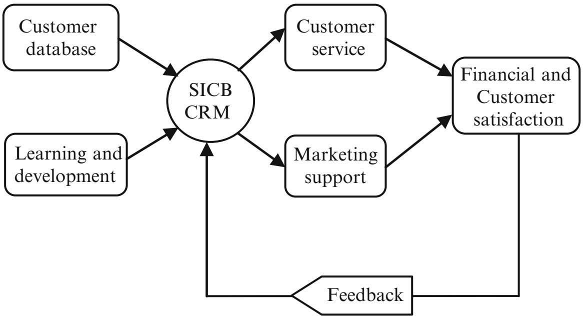


Fig. 23.1 The process of securities consulting business CRM

performance of CRM to investment consultant, show gaps and deficiencies of CRM and investment consultant take measures to improve CRM. Finally, a constantly upward of CRM cycle process is created. The specific process is shown in Fig. 23.1.

23.2.2 Build Performance Evaluation Index System of Securities Investment Consulting Business CRM Based on Balanced Scorecard

Currently, the scholars mostly choose Balanced Scorecard theory as a method to select evaluation indicators when they study the performance evaluation of CRM. Balanced Scorecard links the organization's strategic objectives and implementation process, the current performance and future profitability, which breaks the traditional evaluation system that takes the finance as the core (Meng et al. 2006). The Balanced Scorecard clears the customer-oriented design principle of performance evaluation system and this coincides with the concept of "customer-centric" of the CRM. The amazing consistency of this value orientation has laid a solid theoretical foundation for the Balanced Scorecard in the performance evaluation of the CRM (Lu 2011). As mentioned above,we choose Balanced Scorecard as the method to build customer relationship management performance evaluation indicator system.

Combining with the characteristics of the securities investment consulting business and customer relationship management process system, the paper divides securities investment consulting business CRM performance evaluation system into financial, customer service, marketing support, learning and development four dimensions, as the first-level indicators, then identifies the four first-level indicators of their main influencing factors by consulting industry experts, as the second-level indicators. After the above division we get the performance evaluation system table as the following Table 23.1 show which is consisted of 4 first-level indicators and 18 second-level indicators, and constitute the basis of CRM performance evaluation.

Table 23.1 The performance evaluation system of SICB

	First-level indicators (B_i)	Second-level indicators (C_{ij})
The performance of investment securities consulting business CRM A	Financial B_1	Commission growth rate C_{11}
		Trading volume C_{12}
		Financial product sales C_{13}
		Average cost of service C_{14}
		Profit margin C_{15}
	Customer service B_2	Customer asset growth rate C_{21}
		Customer profit proportion C_{22}
		Customer satisfaction C_{23}
		Service coverage C_{24}
		Loss of customers C_{25}
	Marketing support B_3	Customer segmentation C_{31}
		Marketing success rate C_{32}
		Customer signing rate C_{33}
		Number of new customers C_{34}
	Learning and development B_4	Comprehensive service capabilities C_{41}
		Staff training rate C_{42}
		Staff satisfaction C_{43}
		Staff maintains rate C_{44}

23.3 Performance Evaluation of Securities Investment Consulting Business CRM Based on a Securities Company

The authors employ the method of combining the analytic hierarchy process (AHP) and fuzzy comprehensive evaluation in this paper to evaluate the securities investment consulting business CRM, and get the conclusion of the evaluation results. The process is as follows:

23.3.1 Construct Pair Wise Comparative Judgment Matrix

We invited a number of experts who are familiar with securities investment consulting business CRM and have some theoretical or practical experience on it, to build an expert team to seek expert opinion on the relative importance of the same level indicators. Structure comparison matrix Y, by "scaling by 1–9". $Y = (a_{ij})$, a_{ij} represents the hierarchy of importance compared to the indicator i and indicator j. Table 23.2 lists the nine hierarchy of importance and their assignment.

According to the above principles, construct the judgment matrix of the first-level indicators A through expert consultation method.

Table 23.2 The importance hierarchy and their assignment scale of the indicators when pair wise comparisons

Assignment scale	Importance hierarchy
1	Two indicators compared with the same importance
3	Two indicators compare, the former is slightly more important than the latter
5	Two indicators compare, the former is obvious more important than the latter
7	Two indicators compare, the former is intense more important than the latter
9	Two indicators compare, the former is extremely more important than the latter
2, 4, 6, 8	Hierarchy of importance between the above
Reciprocal	If the importance ratio of indicators i and j is a_{ij}, the importance ratio of indicators j and i is $a_{ji} = \frac{1}{a_{ij}}$

$$A = \begin{bmatrix} 1 & 1/3 & 3 & 2 \\ 3 & 1 & 5 & 3 \\ 1/3 & 1/5 & 1 & 1/2 \\ 1/2 & 1/3 & 2 & 1 \end{bmatrix}$$

Similarly, construct the judgment matrix of second-level indicators B_i, $(i = 1, 2, 3, 4.)$

$$B_1 = \begin{bmatrix} 1 & 1 & 5 & 3 & 5 \\ 1 & 1 & 5 & 3 & 5 \\ 1/5 & 1/5 & 1 & 1/2 & 1 \\ 1/3 & 1/3 & 2 & 1 & 2 \\ 1/5 & 1/5 & 1 & 1/2 & 1 \end{bmatrix} B_2 = \begin{bmatrix} 1 & 1/5 & 1/2 & 1/4 & 1/2 \\ 5 & 1 & 2 & 1 & 2 \\ 2 & 1/2 & 1 & 1/3 & 1 \\ 4 & 1 & 3 & 1 & 2 \\ 2 & 1/2 & 1 & 1/2 & 1 \end{bmatrix}$$

$$B_3 = \begin{bmatrix} 1 & 3 & 3 & 5 \\ 1/3 & 1 & 1 & 3 \\ 1/3 & 1 & 1 & 3 \\ 1/5 & 1/3 & 1/3 & 1 \end{bmatrix} \quad B_4 = \begin{bmatrix} 1 & 2 & 3 & 5 \\ 1/2 & 1 & 2 & 3 \\ 1/3 & 1/2 & 1 & 2 \\ 1/5 & 1/3 & 1/2 & 1 \end{bmatrix}$$

23.3.2 Calculate the Indicators Weight

According to the principle of hierarchy analysis method, use matlab7.0 software to derive the largest eigenvalues λ_{max} of above five matrixes and their corresponding eigenvectors ω. Finally, the normalized eigenvector value is the indicator on upper indicator weights. The result is as shown in Table 23.3.

Table 23.3 Indicators weight of the performance evaluation system

	First-level indicators B_i	Weight ω_i	Second-level indicators C_{ij}	Weights ω_{ij}
The performance of securities investment consulting business CRM A	Financial B_1	0.2398	Commission growth rate C_{11}	0.3643
			Trading volume C_{12}	0.3643
			Financial product sales C_{13}	0.0703
			Average cost of service C_{14}	0.1308
			Profit margin C_{15}	0.0703
	Customer service B_2	0.5232	Customer asset growth rate C_{21}	0.0717
			Customer profit proportion C_{22}	0.3134
			Customer satisfaction C_{23}	0.1387
			Service coverage C_{24}	0.3267
			Loss of customers C_{25}	0.1496
	Marketing support B_3	0.0851	Customer segmentation C_{31}	0.5222
			Marketing success rate C_{32}	0.1998
			Customer signing rate C_{33}	0.1998
			Number of new customers C_{34}	0.0781
	Learning and development B_4	0.1519	Comprehensive service capabilities C_{41}	0.4829
			Staff training rate C_{42}	0.2720
			Staff satisfaction C_{43}	0.1570
			Staff maintains rate C_{44}	0.0882

23.3.3 Consistency Test

The consistency index (CI) and consistency ratio (CR) are used to verify the consistency of the comparison matrix. CI and CR are defined as follows:

$$\mathrm{CI} = \frac{\lambda_{max} - n}{n - 1}$$

$$\mathrm{CR} = \frac{\mathrm{CI}}{\mathrm{RI}}$$

The RI represents the average consistency index over numerous random entries of same order reciprocal matrices. If the $\mathrm{CR} < 0.1$, the estimate is accepted; otherwise, a new comparison matrix is solicited until $\mathrm{CR} < 0.1$ (Chang et al. 2009). After the computation, all levels of indicators have the satisfactory consistency.

23.3.4 Fuzzy Comprehensive Evaluation

1. Establish indicators set and evaluation set, the set of indicators W = (financial indicators, customer service, marketing support, learning and development), the evaluation set R = {excellent, good, medium, poor}.

2. Determine the set of evaluation indicators weights. For convenience, the first-level indicators weights set is denoted as $W1 = \{\omega_1, \omega_2, \omega_3, \omega_4\}$, the second-level indicators weights set is denoted as $W_{ij} = \{\omega_{i1}, \omega_{i2}, \ldots, \omega_{ij}\}$ $(i, j = 1, 2, 3, 4)$
3. Establish the evaluation matrix R. Invite a number of experts to evaluate the various indicators and to determine the degree of membership of each indicator. The single factor fuzzy evaluation matrix is as follows:

$$R_{B1} = \begin{bmatrix} 0.4 & 0.3 & 0.2 & 0.1 & 0 \\ 0.5 & 0.3 & 0.2 & 0 & 0 \\ 0.6 & 0.3 & 0.1 & 0 & 0 \\ 0.7 & 0.2 & 0.1 & 0 & 0 \\ 0.7 & 0.3 & 0 & 0 & 0 \end{bmatrix} R_{B3} = \begin{bmatrix} 0.5 & 0.3 & 0.2 & 0 & 0 \\ 0.6 & 0.4 & 0 & 0 & 0 \\ 0.4 & 0.3 & 0.2 & 0.1 & 0 \\ 0.7 & 0.3 & 0 & 0 & 0 \end{bmatrix}$$

$$R_{B2} = \begin{bmatrix} 0.2 & 0.2 & 0.3 & 0.2 & 0.1 \\ 0.3 & 0.2 & 0.1 & 0.3 & 0.1 \\ 0.5 & 0.3 & 0.2 & 0 & 0 \\ 0.7 & 0.3 & 0 & 0 & 0 \\ 0.7 & 0.2 & 0.1 & 0 & 0 \end{bmatrix} R_{B4} = \begin{bmatrix} 0.5 & 0.3 & 0.1 & 0.1 & 0 \\ 0.4 & 0.3 & 0.2 & 0.1 & 0 \\ 0.7 & 0.3 & 0 & 0 & 0 \\ 0.8 & 0.2 & 0 & 0 & 0 \end{bmatrix}$$

4. Single-factor evaluation based on formula

$$P = W \bullet R$$

$$P_{B1} = W_{1j} \bullet R_{B1} = (0.5108\ 0.2869\ 0.1658\ 0.0364\ 0)$$
$$P_{B2} = W_{2j} \bullet R_{B2} = (0.5111\ 0.2466\ 0.0955\ 0.1084\ 0.0385)$$
$$P_{B3} = W_{3j} \bullet R_{B3} = (0.5156\ 0.3200\ 0.1444\ 0.0200\ 0)$$
$$P_{B4} = W_{4j} \bullet R_{B4} = (0.5307\ 0.2912\ 0.1027\ 0.0755\ 0)$$

5. Conduct comprehensive evaluation based on single-factor evaluation

$$\begin{aligned} P_A &= W_1 \bullet R_A \\ &= (0.2398\ 0.5352\ 0.0851\ 0.1519) \bullet \\ &\quad \begin{bmatrix} 0.5108 & 0.2869 & 0.1658 & 0.0364 & 0 \\ 0.5111 & 0.2466 & 0.0955 & 0.1084 & 0.0385 \\ 0.5156 & 0.3200 & 0.1444 & 0.0200 & 0 \\ 0.5307 & 0.2912 & 0.1027 & 0.0755 & 0 \end{bmatrix} \\ &= (0.5205\ 0.2722\ 0.1188\ 0.0799\ 0.0206) \end{aligned}$$

23.4 Results

According to the principle of maximum membership degree, compare with the evaluation standard, the performance evaluation result of securities investment consulting business CRM is “excellent”. But through the single-factor evaluation results we can found that several second-level indicators get low score, such as customer asset growth rate and customer profit proportion, comprehensive service capabilities and staff training rate, which means that the quality of customer service and the service capacity of financial consultants need to be improved to improve the performance of the overall customer relationship management.

23.5 Conclusion

Results of the performance evaluation of securities consulting business customer relationship management by analytic hierarchy process (AHP) and fuzzy comprehensive evaluation show that the evaluation indicators can be quantified effectively. It should be noted that the effectiveness of the method and application depends primarily on the performance evaluation indicators selection and design of securities consulting business CRM.

References

Antonio T, Akridge J, Gray A, Boehlje M, Widdows R (2007) An evaluation of CRM practices among agribusiness firms. Int Food Agribus Manag Rev 10:36–60

Chang C, Wu C, Liao C (2009) Using analytic hierarchy process evaluating collaborative customer relationship management system. In: Proceedings of the international conference on management and service science, MASS, Beijing

Curry A, Kkolou E (2004) Evaluating CRM to contribute to TQM improvement-a cross-case comparison. TQM Mag 16:314–324

Hong T, Kim E (2007) The selection of CRM systems in financial institutes using the analytic hierarchy. In: 2nd international conference on digital information management, ICDIM, vol 1. IEEE, Piscataway

Hu G (2010) Performance evaluation of enterprise CRM system based on grey-fuzzy theory. In: International conference on networking and digital society, vol 1. ICNDS 2010, Piscataway, pp 32–35

Kong Q, Han Y, Meng Q (2005) Research on evaluation for the effectiveness of CRM based on BP algorithm. Oper Res Manag Sci 14:155–160

Lu H (2011) The construction of CRM performance evaluation system of commercial banks based on the balanced scorecard. Huaihai Inst Technol 9:87–90

Meng Q, Han Y, Wu Z (2006) A study on CRM effectiveness evaluation model based on balanced scorecard. J Ind Eng Eng Manag 4:46–50

Tang B, Liang H (2009) Performance evaluation of customer relationship management under supply chain integration: a fuzzy comprehensive evaluation approach. In: 1st international conference on information science and engineering, ICISE 2009, IEEE, Piscataway, pp 2983–2986

Ullah Z, Al- Mudimigh AS (2009) CRM scorecard measurement: the case for the banking sector in Saudi Arabia. In: EMS 2009 – UKSim 3rd European modelling symposium on computer modelling and simulation, IEEE Conference Publications, Athens, pp 269–273

Wang J (2008) Performance evaluation for CRM of the third-party logistics. Donghua University, Shanghai, pp 35–42

Zhao Q (2011) Securities companies customer relationship management operation evaluation. MS thesis, Ocean University of China, Qingdao, pp 19–21 (Chinese)

Zhou J, Yu T, Zhang Y, Dong S, Wang W (2008) System of CRM performance evaluation based on fuzzy comprehensive algorithm. In: Proceedings of the international conference on information management, vol 1. Innovation Management and Industrial Engineering, ICIII 2008, Sunnyvale, pp 382–385

Chapter 24
A Study on the Model of Selecting the Loyalty Card Issuance Method Based on Fuzzy Set Pair Analysis

Yu Su and Jin Song

Abstract With the aggravation of market competition, more and more enterprises are using loyalty card targeted promotion as an important means to cultivate a stable group of consumers and build competitive advantages. In this paper, we mainly discuss the popular ways of issuing loyalty card such as "purchase", "buy gifts", "buy + 10" and "complimentary". What's more, we use membership base, the number of repurchase, the issuing income as the evaluation indexes to establish the model of selecting loyalty card issuance method based on fuzzy set pair analysis. The model can provide method to issue loyalty card for enterprises.

Keywords Evaluation index • Fuzzy set pair analysis • Issuance method • Loyalty card

With the aggravation of market competition, the promotion means that enterprises adopted are continuously renovated and the enterprises' promotion costs are growing rapidly. Especially since the 1990s, with the increasingly intense of market competition and the increasing of consumer differences, the importance and advantages of targeted promotion become increasingly prominent. Loyalty card targeted promotion is one of means the enterprises use in targeted promotion and

Y. Su
The TUC-FIU Cooperation School of Tianjin University of Commerce, Tianjin, China

J. Song (✉)
The Business School of Tianjin University of Commerce, Tianjin, China
e-mail: happysongjin@163.com

E. Qi et al. (eds.), *The 19th International Conference on Industrial Engineering and Engineering Management*, DOI 10.1007/978-3-642-37270-4_24,

in recent years is adopted by more and more enterprises as an important means to cultivate a stable group of consumers and build competitive advantages.[1]

24.1 Literature Review

The current research is focusing on the effectiveness of loyalty card, implementation of loyalty card and factors that impact the effect of implementation of loyalty card. The research of loyalty card's effectiveness was around consumer loyalty: Wright and Sparks (1999) pointed out that the retail loyalty development programs had become the most significant feature of retail marketing in the 1990s. Some retailers had begun to question the limit of loyalty and consumers were starting to become more vigilance and more selectivity for loyalty card and loyalty programs (Wright and Sparks 1999); Mauri (2003) studied whether the card holder was loyalty (Mauri 2003); Bellizzi and Bristol (2004) studied whether the loyalty card issuance was really associated with consumer loyalty by surveying a big city in the western United States (Bellizzi and Bristol 2004); Van Heerde and Bijmolt (2005) studied the differences and reasons of the spending by members and non-members, the partly reasons about the differences was loyalty consumers choose to participate in loyalty programs (Van Heerde and Bijmolt 2005); Demoulin and Zidda (2008) studied on the differences about loyalty for store and price sensitivity between the satisfied and dissatisfied consumers on loyalty card return (Demoulin and Zidda 2008). In addition, Meyer-Waarden (2007) proposed that loyalty program had a positive role in the consumer life cycle and consumer wallet share (Meyer-Waarden 2007); Cortin et al. (2008) and other scholars proposed database generated by consumer loyalty program can help enterprises make decisions on pricing policies, promotions, inventory and category management (Cortin et al. 2008). The studies on how to implement loyalty card include: Worthington (1998) through qualitative and quantitative analysis studied the development of loyalty card in town's center district, proposed the standard that loyalty card must consider in other towns' central area and discussed the options can be used for members' card upgrade in accordance with the development of technology (Worthington 1998); Rowley (2000) described a new applications of loyalty card based on the self-service equipment. This new applications can enhance the effectiveness of loyalty card (Rowley 2000); Leenheer et al. (2007) and other scholars established a model that can determine membership, measure loyalty behavior and the share of wallet through a study of loyalty programs' impact on consumer loyalty (Leenheer et al. 2007). The studies on factors that impact the effect of implementation of loyalty card include: Demoulin and Zidda (2009) studied a new loyalty card adoption process in grocery retailing and found that attitude, behavior and socio-demographic variables had varying degrees impact on the likelihood and using times (Demoulin and Zidda 2009); Smith et al.

[1]Tianjin Colleges and Universities Humanities.
Social Science Funded Project (20112115).

(2003) and other scholars pointed out that the effects of loyalty card are related with many consumers factors associated with loyalty programs, including the subordinate degree to retailers, the operation of the plan in practice and the attitude of consumers without cards (Smith et al. 2003); Keh and Lee (2006) studied the return time of loyalty card (now vs. delay) and return type (direct vs. indirect) under the two service states (Satisfactory or unsatisfactory) that can produce different effects (Keh and Lee 2006). Yu Su (2010a, b) studied the factors that influence the value and risk of loyalty card targeted promotion (Yu Su 2010a) and studied the effect of enterprises' mass promotion (Yu Su 2010b) and competitors' promotion (Yu Su and Yi Zong 2010) on loyalty card targeted promotion using empirical methods.

From the above, there are not scholars to study on how enterprises to choose loyalty card issuance method. In this paper, we use fuzzy set pair analysis to establish a model of loyalty card issuance method, in order to provide the best method to choose loyalty card issuance method.

24.2 Benefits and Costs of the Loyalty Card Issuance Method

There are mainly four ways to issue loyalty card that are "purchase", "buy gifts", "buy + 10" and "complimentary". The method of "purchasing" can get a certain income through the issuance of loyalty card, no matter the consumer will repurchase or not. But consumers are generally reluctant to get loyalty card through this way, so the membership base is smaller than other ways. As members who use the way to get loyalty card must pay money, therefore members absorbed by this method often favorite the enterprise's products and are more likely to repurchase; the method of "buy gifts" can provide a certain incentives to consumers for the first time, but there is no income of issuing cards. Members absorbed by this way are less likely to repurchase than members absorbed by the method of "purchasing"; the method of "Buy + 10" also provides a certain income for enterprises to issue cards. As the cost of "¥10", it has smaller member base than "buy gifts". But members repurchase probability is high compared with "buy gifts"; "Complimentary" attracts more consumers to apply for loyalty card. Consumers do not have the cost of getting cards and members absorbed by this method have small probability to repurchase. The enterprise only has cost and does not have immediate revenue by issuing cards.

For the four loyalty card issuance methods, the main cost comes from the cost of fabrication, except for a small amount of management cost.

24.3 The Model of Loyalty Card Issuance Method Based on Fuzzy Set Pair Analysis

When we select the issuance method of loyalty card, we should not only consider its benefits, but also take into account its costs. We can use fuzzy set pair analysis to evaluate the different methods in order to choose a better one.

24.3.1 Select Evaluation Index of the Loyalty Card Issuance Method

There are four loyalty card issuance methods that are A_1 (purchase), A_2 (buy gifts), A_3 (buy + 10), A_4 (complimentary); three evaluation indexes are P_1 (membership base), P_2 (the number of repurchase), P_3 (the issuance income of loyalty card) and the cost of every method are C_1, C_2, C_3, C_4.

24.3.2 Calculate the Relatively Value r_i of Each Loyalty Card Issuance Method Based on the Evaluation Method of Fuzzy Set Pair Analysis Closeness Degree

There are four issuance schemes of loyalty card A_1, A_2, A_3, A_4 and each scheme has three indexes P_1, P_2, P_3, the dimension of each index can be the same or different and each index has a value recorded as $f_{ij} \geq 0 (i = 1, 2, 3, 4, j = 1, 2, 3)$ that represents the value of the index j in the loyalty card issuance method i. The three indexes are efficiency indexes, that is to say it is excellent to take a large value of the index. The steps of evaluation method based on fuzzy set pair analysis closeness degree are as following:

1. Make the best scheme A_0 and the worst scheme A_s according to the four schemes. The value of each index in the best scheme A_0 should be the best value of each index in the four schemes, namely efficiency index is the maximum one. f_{0j} is the value of index j in the best scheme A_0. If f_{ij} is the efficiency index, $f_{0j} = \max_i \{f_{ij}\}$; the value of each index in the worst scheme A_s should be the worst value of each index in the four schemes, namely efficiency index is the minimum one. f_{sj} is the value of index j in the worst scheme A_s. If f_{ij} is the efficiency index, $f_{sj} = \min_i \{f_{ij}\}$.
2. Calculate the same membership degree a_{i0j} and opposite membership degree c_{i0j} of the index pair $\{f_{ij}, f_{0j}\}$ of set pair (A_i, A_0) consisted by the loyalty card issuance method A_i and the best scheme A_0. If f_{ij} is efficiency index, the compare interval of the index is $[f_{sj}, f_{0j}]$. Then

$$a_{i0j} = \frac{f_{ij}}{f_{0j} + f_{sj}} \tag{24.1}$$

$$c_{i0j} = \frac{f_{0j} f_{sj}}{(f_{0j} + f_{sj}) f_{ij}} \tag{24.2}$$

3. The closeness degree of index f_{ij} and the correspond index f_{0j} in the best scheme is

$$r_{i0j} = \frac{a_{i0j}}{a_{i0j} + c_{i0j}} \tag{24.3}$$

Establish the matrix of closeness degree that evaluates loyalty card issuance indexes and the ideal loyalty card issuance indexes:

$$H = \begin{pmatrix} r_{101} & r_{102} & r_{103} & r_{104} \\ r_{201} & r_{202} & r_{203} & r_{204} \\ r_{301} & r_{302} & r_{303} & r_{304} \\ r_{401} & r_{402} & r_{403} & r_{404} \end{pmatrix}$$

4. Determine the weights of each index, then using matrix of indexes' weights $\omega = [\omega_1, \omega_2, \omega_3, \omega_4]$ to calculate connection matrix of loyalty card issuance scheme A_i and the best scheme A_0.

$$R = H^* \omega^T \tag{24.4}$$

The element r_i in R is the same membership degree of the evaluate scheme A_i and the best scheme A_0.

5. Regard r_i as the relatively value of loyalty card issuance method A_i.

24.3.3 Determine the Loyalty Card Issuance Method

Compare method i and method j. The selection method is as follows:

1. When $r_i > r_j, C_i < C_j$, the scheme j is eliminated;
2. When $r_i < r_j, C_i < C_j$, definite the effect better degree of loyalty card issuance method j than loyalty card issuance method i.

$$R_{ji} = \frac{\Delta r_{ji}}{r_i} \tag{24.5}$$

$\Delta r_{ji} = r_j - r_i$ is the relatively increment of loyalty card issuance method j compared with loyalty card issuance method i.

The cost better degree of loyalty card issuance method i than loyalty card issuance method j:

$$c_{ij} = \frac{\Delta C_{ij}}{C_i} \tag{24.6}$$

$\Delta C_{ij} = C_j - C_i$ is the cost increment of loyalty card issuance method j compared with loyalty card issuance method i.

When $R_{ji} > c_{ij}$, the scheme i is eliminated; when $R_{ji} < c_{ij}$, the scheme j is eliminated.

3. Using (24.1) and (24.2) gradually eliminate some of the inferior schemes and select the best scheme.

Table 24.1 The evaluate table of membership base, re-purchase number issuing income and issuing costs

Loyalty card issuance methods	P_1	P_2	P_3	C
A_1	20	18	400	20
A_2	30	15	1,500	30
A_3	25	18	250	25
A_4	100	20	0	100

24.4 Example

A Company issued loyalty card. The data of the number of members absorbed by each type of loyalty card, of repurchasing members, of issuing income and of issuing costs is shown in the Table 24.1 (UNITE 1000).

We can get $\omega = [0.3, 0.5, 0.2]$.

$$a_{10} = \left[\frac{1}{6}, \frac{18}{35}, \frac{4}{5}\right], \quad a_{20} = \left[\frac{1}{4}, \frac{3}{7}, 1\right], \quad a_{30} = \left[\frac{5}{24}, \frac{18}{35}, \frac{1}{6}\right], a_{40} = \left[\frac{5}{6}, \frac{4}{7}, 0\right]$$

$$c_{10} = \left[\frac{5}{6}, \frac{10}{21}, 0\right], \quad c_{20} = \left[\frac{5}{9}, \frac{4}{7}, 0\right], \quad c_{30} = \left[\frac{2}{3}, \frac{10}{21}, 0\right], \quad c_{40} = \left[\frac{1}{6}, \frac{3}{7}, 0\right]$$

$$r_{10} = \left[\frac{1}{6}, \frac{104}{105}, 1\right], \quad r_{20} = \left[\frac{9}{29}, \frac{3}{7}, 1\right], r_{30} = \left[\frac{7}{8}, \frac{104}{105}, 1\right], \quad r_{40} = \left[\frac{6}{6}, \frac{4}{7}, 0\right]$$

R = (0.75, 0.5, 0.96, 0.54)

The best loyalty card issuance method is “buy + 10”.

24.5 Conclusion

This paper establishes the model of selecting the loyalty card issuance method based on fuzzy set pair analysis, the methods of issuing loyalty card are “purchase”, “buy gifts”, “buy + 10” and “complimentary”. What’s more, the evaluation indexes are membership base, the number of repurchase, the issuing income. Using this model companies can choose their own loyalty card issuance method according to its own actual situation.

References

Bellizzi JA, Bristol T (2004) An assessment of supermarket loyalty cards in one major US market. J Consum Market 21(2):144–154

Cortin M, Elorz M, Múgica JM (2008) The use of loyalty-cards databases: differences in regular price and discount sensitivity in the brand choice decision between card and non-card holders. J Retail Consum Ser 15:52–62

Demoulin NTM, Zidda P (2008) On the impact of loyalty cards on store loyalty: does the customers' satisfaction with the reward scheme matter? J Retail Consum Serv 15:386–398

Demoulin NTM, Zidda P (2009) Drivers of customers' adoption and adoption timing of a new loyalty card in the grocery retail market. J Retail 85(3):391–405

Keh HT, Lee YH (2006) Do reward programs build loyalty for services? The moderating effect of satisfaction on type and timing of rewards. J Retail 82(2):127–136

Leenheer J, van Heerde HJ, Bijmolt THA, Smidts A (2007) Do loyalty programs really enhance behavioral loyalty? An empirical analysis accounting for self-selecting members. Int J Res Market 24:31–47

Mauri C (2003) Card loyalty. "A new emerging issue in grocery retailing". J Retail Consum Serv 10(1):13–25

Meyer-Waarden L (2007) The effects of loyalty programs on customer lifetime duration and share of wallet. J Retail 83(2):223–236

Rowley J (2000) Loyalty kiosks: making loyalty cards work. Br Food J 102(5/6):390–398

Smith A, Sparks L, Hart S, Tzokas N (2003) Retail loyalty schemes: results from a consumer diary study. J Retail Consum Serv 10(2):109–119

Van Heerde HJ, Bijmolt THA (2005) Decomposing the promotional revenue bump for loyalty program members versus non-members. J Market Res 42(4):443–457

Worthington S (1998) Loyalty cards and the revitalization of the town center. Int J Retail Distrib Manag 26(2):68–77

Wright C, Sparks L (1999) Loyalty saturation in retailing: exploring the end of retail loyalty cards? Int J Retail Distrib Manag 27(10):429–440

Yu Su (2010a) A study of value and risk of loyalty card targeted promotion. Jiangsu Commer Theor 6:6–8 (Chinese)

Yu Su (2010b) A study on the effect of public promotion on members' perceived value of loyalty card. Ind Technol Econ 29(10):25–32 (Chinese)

Yu Su, Yi Zong (2010) A study on the effect of competitors promotion on loyalty card target promotion. Stat Decis 10:137–140 (Chinese)

Chapter 25
Complex Network Community Detection Algorithm Based on Genetic Algorithm

Yun Li, Gang Liu, and Song-yang Lao

Abstract For the problem of complex network community detection, propose a new algorithm based on genetic algorithm to solve it. This algorithm sets network modularity function as target function and fitness function, uses matrix encoding to describe individuals, and generates initial population using nodes similarity. The crossover operation is based on the quality of individuals' genes, in this process, all nodes that weren't partitioned into any communities make up a new one together, and the nodes that were partitioned into more than one community are placed into the community to which most of their neighbors belong. The mutation operation is non-uniform, which splits the mutation gene into two new genes or fuses it into the others randomly. The experiment proved that this algorithm could effectively detect communities in complex networks.

Keywords Complex network • Community detection • Genetic algorithm

25.1 Introduction

In 2002, Newman and Girvan opened up a new field of complex network research – complex network community detection (Girvan and Newman 2002). It uses the information contained in network topology to get the community structure of complex network, which has some connotations. The research of this problem, which is of great significance, helps to study the whole network's module, function and its evolution in a divide and conquer way so that to understand complex network's organization principle, topology and dynamics characteristic more correctly (Luo et al. 2011).

Y. Li (✉) • G. Liu • S. Lao
Information System and Management College, National University of Defense Technology, Changsha, China
e-mail: Lee2003100@163.com

E. Qi et al. (eds.), *The 19th International Conference on Industrial Engineering and Engineering Management*, DOI 10.1007/978-3-642-37270-4_25,

In 2004, they proposed network modularity (Newman and Girvan 2004) to quantitatively evaluate the result's quality of complex network community detection, and its function is as follows:

$$Q = \frac{1}{2m}\sum_{ij}\left[a_{ij} - \frac{k_i k_j}{2m}\right]\delta\left(c_i, c_j\right) \tag{25.1}$$

Where a_{ij} represents an element of network adjacency matrix $A = \left(a_{ij}\right)_{n\times n}$, if nodes i and j are connected by an edge, then $a_{ij} = 1$, or $a_{ij} = 0$; c_i and c_j respectively stand for the communities that nodes i and j belongs to; if $c_i = c_j$, $\delta\left(c_i, c_j\right) = 1$, or $\delta\left(c_i, c_j\right) = 0$; k_i and k_j respectively represent the degrees of nodes i and j, $k_i = \sum_{j=1}^{n} a_{ij}, k_j = \sum_{i=1}^{n} a_{ij}, i, j = 1, 2, \ldots, n$; m and n are respectively the amount of network's edges and that of nodes. $|Q| \leq 1$, the greater it is, the better the result of complex network community detection is.

25.2 Methodology

To assume each node in complex network $G(V, E)$ belongs to only one community, that's to say, communities are not intersected or overlapped, and:

$$V = \{v_i | i = 1, 2, \ldots, n\},\ E = \{e_i | i = 1, 2, \ldots, m\}$$

Where V and E respectively represent the node set and edge set; n is the amount of nodes, and m edges.

Then complex network community detection is to partition n nodes into t communities that are not intersected, so get $c_1, c_2, \ldots, c_t$ as a partition of complex network, $1 < t < n$, which means:

$$c_i \neq \varnothing;\ c_i \subset V;\ \bigcup_{i=1}^{t} c_i = V;$$

$$\forall i \neq j, c_i \cap c_j = \varnothing, i, j = 1, 2, \cdots, t$$

Thus, complex network community detection is a typical NP combinatorial optimization problem, and genetic algorithm can effectively solve such problem. Therefore, by setting network modularity function as target function and fitness function, we propose a new Complex Network Community Detection algorithm based on Genetic algorithm (CNCDG).

25.2.1 Individual Encoding

In the field of complex network community detection, string encoding (Jin et al. 2011; Li et al. 2010; Liu et al. 2007; Gog et al. 2007; He et al. 2010) and graph-based encoding (Pizzuti 2008a, b, 2009; Shi et al. 2010) are being widely used at present, and they all have shortcomings, the former fails traditional crossover operator, while the latter requires additional decoding, which are specifically explained in Jin et al. (2011) and He et al. (2010). To avoid above shortcomings, we adopt binary matrix to encode individuals, which means, however complex network is partitioned, it could be always represented by a binary matrix M:

$$M = \begin{pmatrix} m_{11} & m_{12} & \cdots & m_{1t} \\ m_{21} & m_{22} & \cdots & m_{2t} \\ \vdots & \vdots & \ddots & \vdots \\ m_{n1} & m_{n2} & \cdots & m_{nt} \end{pmatrix}$$

M is a $n \times t$ matrix, t $(1 < t < n)$ is the amount of communities after partitioning complex network. Row i $(1 \le i \le n)$ of M corresponds to the partition result of v_i, and column j $(1 \le j \le t)$ corresponds to community c_j. If v_i belongs to c_j, then $m_{ij} = 1$; otherwise, $m_{ij} = 0$.

Since each node in complex network must and can be only partitioned into one community, encoding matrix M must follow the constraints of (25.2) and (25.3):

$$\sum_{j=1}^{t} m_{ij} = 1 \tag{25.2}$$

$$\sum_{i=1}^{n} m_{ij} > 0 \tag{25.3}$$

Different partitions of complex network brings different amounts of communities, thus the amount of columns of encoding matrix M is variable. And regardless of columns' order, it represents the same partition of complex network unless any column's value changes.

25.2.2 Population Initialization

By referring to the quantitative description regarding nodes similarity of complex network in the reference Leicht et al. (2006), we propose a new population initialization method based on traditional clustering method.

Leicht et al. (2006) with other scholars took into account the similar relation between nodes in both long and short paths and derived a formula from that, as (25.4), to calculate the nodes similarity of complex network by using network topology.

$$D \times S \times D = \frac{\alpha}{\lambda_{\max}} \times A \times (D \times S \times D) + I \tag{25.4}$$

Where A represents network adjacency matrix; $\lambda_{\max}$ is the maximum eigenvalue of A; D is the n order diagonal square matrix which takes degrees of nodes as diagonal elements, its diagonal element $d_{ii} = \sum_{j=1}^{n} a_{ij}$, $i, j = 1, 2, \ldots, n$; I represents identity matrix; S is the nodes similarity matrix. Another notable parameter α $(0 < \alpha < 1)$ controls the relative weight of long and short paths; a smaller α value makes the calculation focus more on short paths, whereas a bigger one takes long and short paths into consideration at the same time. Generally, it's optimal to set α as 0.97 (Leicht et al. 2006).

Then steps of population initialization are as follows:

Step 1: Iterate (25.4) until the result is convergent to get S, the nodes similarity matrix.

Repeat steps 2 and 3 based on population scale P_n:

Step 2: Randomly select several nodes $v_1, v_2, \ldots, v_t$ as centers of communities, $1 < t < n$.

Step 3: By following the max-similarity principle which means a node has greater similarity with one community center than with the others, use nodes similarity matrix S to partition non-community-center nodes into each community.

Time complexity of step 1 is $O\left(ln^3\right)$, l is the iteration times. Each implementation of step 2 and 3 obtains a community partition of complex network. In step 2, the community centers chosen each time are different, which makes initial individuals have strong diversity, and its time complexity could be ignored. Step 3 greatly reduces the possibility that nodes with low similarity or no link between each other are divided into the same community, this provides initial individuals with a certain accuracy, lessens algorithm's search space and speeds up the convergence; its time complexity is $O\left(P_n\left(n - \bar{t}\right)\bar{t}\right)$, $\bar{t}$ is the average number of communities after each partition. Therefore, the time complexity of population initialization is $O(n^3 + (n - \bar{t})\bar{t})$.

25.2.3 Crossover Operator

To adopt traditional single-point crossover operator to cross individuals, before that, first, we quantitatively describe the quality of individuals' genes. According to *Individual Encoding*, a column of encoding matrix M corresponds to one gene of an individual, which also represents a complex network community. And generally, the larger the average nodes similarity of a community is, the better the community structure is. Then crossover operation based on the quality of individuals' genes is as follows:

Step 1: Use network modularity function to calculate all individuals' fitness, and arrange them in descending order according to the fitness. Then select the top $P_n \times P_c$ individuals which have optimal fitness and pair the head and the tail ones to cross. P_c is a fixed constant, $0 < P_c < 1$, and $(P_n \times P_c) \mod 2 = 0$. For example, select the top six individuals $I_1, I_2, I_3, I_4, I_5, I_6$, which have optimal fitness that are arranged in descending order, and pair the head and the tail ones to cross; then I_1 and I_6, I_2 and I_5, I_3 and I_4 crosses with each other respectively.

Step 2: Measure the quality of crossover individuals' genes. Suppose in an individual encoding matrix M, column M_i is expressed as $(m_{1i}, m_{2i}, \ldots, m_{ni})^{-1}$, and there are r nonzero elements, $m_{u_1 i}, m_{u_2 i}, \ldots, m_{u_r i}$, in M_i, $1 \leq u_p \leq n$, $1 \leq p \leq r$, $1 \leq r < n$, then use nodes similarity matrix S and (25.5) to calculate the average similarity $\bar{s}_i$ among nodes $v_{u_1}, v_{u_2}, \ldots, v_{u_r}$, which also indicates the quality of gene i of the individual. In (25.5), $s_{u_p u_q}$ is the similarity between nodes v_{u_p} and v_{u_q}.

$$\bar{s}_i = \frac{\sum_{p=1}^{r} \sum_{p<q}^{r} s_{u_p u_q}}{r(r-1)/2} \tag{25.5}$$

Arrange each crossover individual's genes in descending order according to $\bar{s}_i$, whose value is 0 when a community has only one node.

Step 3: Implement traditional single-point crossover. Exchange two crossover individuals' best genes which correspond to the first columns of their encoding matrixes, then two new individuals are generated.

It's worth noting that the new individuals generated by step 3 may be illegal, which means some new individuals' encoding matrixes don't satisfy (25.2) and (25.3). In this case, some nodes may belong to either none of communities or more than one community, which contradicts the presupposition. Therefore, it needs to revise these invalid solutions, the method are as follows:

1. Form a new community made up of nodes that belong to none of communities. It means that if individual encoding matrix has some rows with only 0 which are called 0-rows for short, then add a new column with only 0 to the matrix, and set the elements in both the new column and the 0-rows as 1;
2. Partition those nodes that belong to more than one community into the communities that their most neighbors belong to, which is called neighbor-most principle (He et al. 2010). It means that if individual encoding matrix has a row with several 1 s which is called 1 s-row for short, then find out the column which has most neighbors of the node that the 1 s-row corresponds to, and set all of elements in the 1 s-row as 0, except the one also in that column as 1.

In the above crossover operation, all arrangements using bubble sort. Suppose the average numbers of nodes in each community and those which are partitioned into more than one community after each partition, are respectively αn and βn, $0 < \alpha < 1$, $0 < \beta \leq 1$. Then all above steps' time complexities are respectively $O\left(n^2 + {P_n}^2\right)$, $O\left((P_n \times P_c)\left((\alpha n)^2\bar{t} + \bar{t}^2\right)\right)$, $O\left((P_n \times P_c)\,\beta n\bar{t}\right)$, thus the time complexity of crossover operation is $O\left(n^2\bar{t} + \bar{t}^2\right)$.

25.2.4 Mutation Operator

For matrix encoding, we use non-uniform mutation on the basis of individuals' genes arranged in descending order to change the mutation gene through split or fusion operation. In an individual encoding matrix, split operation means randomly split the last column whose elements' sum is greater than 1, into two columns to replace the original one; whereas fusion operation is moving those none-zero elements of the last column to other ones according to the neighbor-most principle and delete the last one. Below are the detailed steps:

Step 1: Select the last $P_n \times P_m$ individuals which have minimum fitness, then measure and arrange their genes according to the method mentioned in the 2nd step of crossover operation. P_m is a fixed constant, $0 < P_m < 1$, and $P_n \times P_m$ is an integer.

Step 2: In an individual encoding matrix, if the amount of columns is 2, implement the split operation; if it's greater than 2, implement split or fusion operation randomly. In fusion operation, if one node's most neighbors are in the community corresponding to the last column, partition it into the community which has the second most neighbors of it.

Suppose the average number of nodes which involve in fusion operation after each partition, is γn, $0 < \gamma < 1$. Then above two steps' time complexities are respectively $O\left((P_n \times P_m)\left((\alpha n)^2\bar{t} + \bar{t}^2\right)\right)$ and $O\left((P_n \times P_m)\,\gamma n\bar{t}\right)$, thus the time complexity of mutation operation is $O\left(n^2\bar{t} + \bar{t}^2\right)$.

25.2.5 Selection Operator and Description of CNCDG

We use $\mu + \lambda$ strategy (He et al. 2010) to select the top P_n individuals which have optimal fitness as the progeny generation from the parent generation and the new population generated by crossover and mutation. Then description of CNCDG is as follows:

Table 25.1 Parameters of CNCDG

Parameter name	Parameter value	Parameter description
α	0.97	Control parameter used to calculate nodes similarity
P_n	100	Population scale
P_c	0.8	Ratio of the crossover individuals to all individuals of population
P_m	0.2	Ratio of the mutation individuals to all individuals of population
N_{max}	100	Algorithm's maximum iteration times

Algorithm Input: The adjacency matrix A of complex network and the parameters of CNCDG shown in Table 25.1;

Algorithm Output: Encoding matrix M that represents the community partition of complex network;

Terminal Condition to End Iteration: Algorithm's iteration times achieve $N_{\max}$;

Algorithm Pseudo-code:

1. Generate initial population $Population_{original}$;
2. $for\ i = 1 : N_{max}$
3. Calculate all individuals' fitness using (25.1), and arrange them in descending order according to their fitness;
   //Generate $P_n \times (P_c + P_m)$ new individuals through//crossover and mutation to form new population// $Population_{new}$
4. Select the top $P_n \times P_c$ individuals which have optimal fitness to cross, and calculate their genes' quality using (25.5), then arrange each crossover individual's genes in descending order according to their quality;
5. Pair the head and the tail crossover individuals and exchange their best genes;
6. Revise invalid individuals generated by the crossover operation;
7. Select the last $P_n \times P_m$ individuals which have minimum fitness to mutate, and calculate their genes' quality using (25.5), then arrange each mutation individual's genes in descending order according to their quality;
8. Make mutation individuals mutate non-uniformly;
   //Finish step 4–8, then get a new population// $Population_{new}$
9. Using (25.1) to calculate all individuals' fitness in $Population_{new}$, integrate $Population_{new}$ and $Population_{original}$, and select the top P_n individuals which have optimal fitness as the progeny generation;
   end
10. Select the individual with maximum fitness as the result of complex network community detection.

Refer to above, and $\bar{t}$ is usually far less than n, thus the time complexity of CNCDG is $O\left(n^3\right)$.

25.3 Results

We choose two real networks, Zachary's karate club network (Zachary 1977) and Dolphin social network (Lusseau et al. 2003), which are widely used, to verify the effectiveness of CNCDG.

25.3.1 Zachary's Karate Cub Network

Zachary's karate club, as a karate club in an American university, is a social network which has 34 nodes and 78 edges. Each node represents a club member, and each edge means that there's social interaction between two club members. This club splits into two independent clubs due to internal divergence, which are respectively led by the original coach and director. Figure 25.1 shows a result of CNCDG randomly running on this network once. In the figure, the black and white nodes respectively represent two independent communities after the split, while the four shapes of nodes respectively represent four small communities with different sizes, and nodes 1 and 34 respectively represent the coach and director. It shows that CNCDG not only accurately discovers the real community structure of Zachary's karate club network, but also small communities contained in those two real ones. The average network modularity of CNCDG running on this network 50 times is 0.4198, which is greater than 0.3715 corresponding to the real community structure of this network.

25.3.2 Dolphin Social Network

Dolphin social network, containing two communities of male and female dolphins, is a dolphin contact network which has 62 nodes and 159 edges. Each node

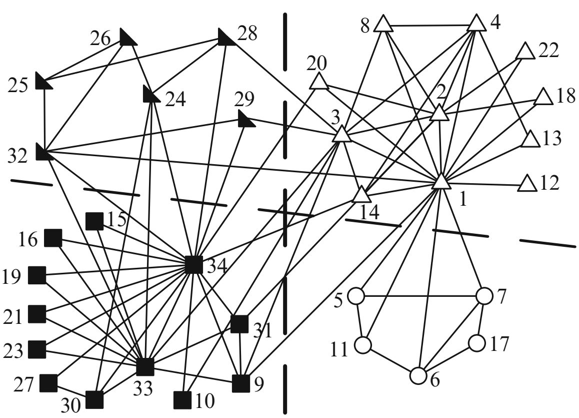


Fig. 25.1 Community detection result of Zachary's karate club network using CNCDG

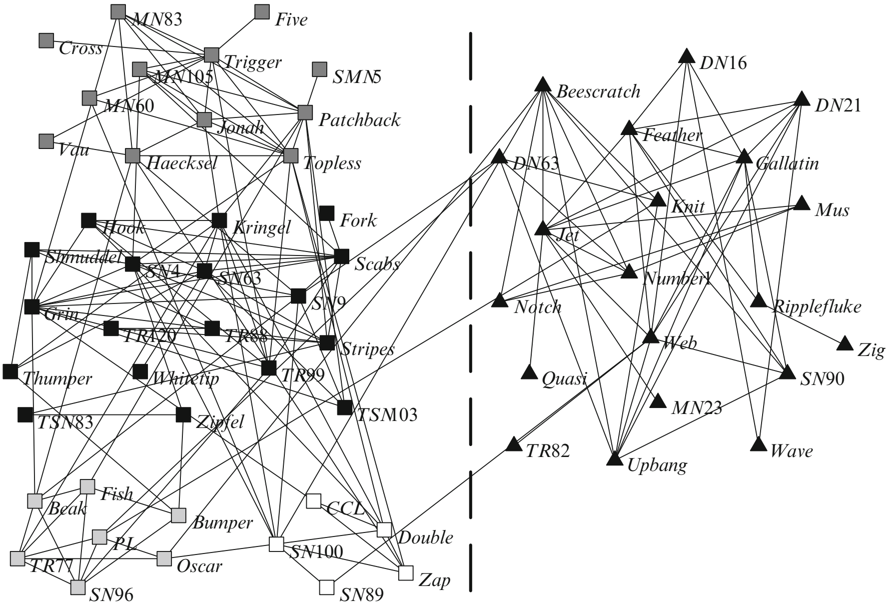


Fig. 25.2 Community detection result of dolphin social network using CNCDG

represents a dolphin, and each edge means that two dolphins contact frequently. CNCDG randomly runs on this network once and the result is shown in Fig. 25.2, in which square nodes represent female dolphins, and different colored square nodes represent the small communities in female dolphins, while triangular nodes represent male dolphins. It shows that CNCDG not only accurately discovers the real community structure of Dolphin social network, but also smaller communities in female dolphins. The average network modularity of CNCDG running on this network 50 times is 0.5272, which is greater than 0.3722 corresponding to the real community structure of this network.

25.4 Discussion

Hereto, we described CNCDG in detail and tested its effectiveness, but we think there are still some works needed to be further improved.

First, the analysis of the time complexity of CNCDG shows that its time complexity suffers from the calculation of nodes similarity. Because of the method to calculate nodes similarity, in Leicht et al. (2006), involves matrix multiplication, it increases the time complexity of CNCDG to $O\left(n^3\right)$. Therefore, it needs to find a new method to calculate nodes similarity which has smaller time complexity, to decrease the time complexity of CNCDG.

Second, algorithms, proposed in Jin et al. (2011) and He et al. (2010), have many optimizing measures which complicate the algorithms' steps in a certain extent, but they really improve the algorithms' accuracy. In view of this, it is useful to import other simple but efficient optimizing measures to further improve the accuracy of CNCDG.

Third, to completely test the performance of CNCDG, it still needs to test the initial individuals' accuracy and diversity, and its own convergence, by using artificial random networks and other real networks. And it's necessary to compare its accuracy with that of some classical algorithms in the field of complex network community detection.

25.5 Conclusion

By setting network modularity function as target function and fitness function, we propose a new complex network community detection algorithm based on genetic algorithm. It uses matrix encoding to describe individuals, which is concise; and generates initial population using nodes similarity, which has both randomness and selectivity, and makes initial individuals have a certain accuracy and strong diversity. Revising invalid solutions in crossover operation insures new individuals are valid; and the split operation and fusion operation in mutation operation makes new individuals have stronger diversity. The results of experimental test using two real networks show that this algorithm could effectively solve the problem of complex network community detection.

References

Girvan M, Newman MEJ (2002) Community structure in social and biological networks. Proc Natl Acad Sci U S A 99(12):7821–7826

Gog A, Dumitrescu D, Hirsbrunner B (2007) Community detection in complex networks using collaborative evolutionary algorithms. In: Proceedings of the 9th European conference on advanced in artificial life, Lisbon, 2007, pp 886–894

He DX, Zhou X, Wang Z, Zhou CG, Wang Z, Jin D (2010) Community mining in complex networks clustering combination based genetic algorithm. Acta Automatica Sinica 36(8):1160–1170 (in Chinese)

Jin D, Liu J, Yang B, He DX, Liu DY (2011) Genetic algorithm with local search for community detection in large-scale complex networks. Acta Automatic Sinica 37(7):873–882 (in Chinese)

Leicht EA, Holme P, Newman MEJ (2006) Vertex similarity in networks. Phys Rev E 73(2):026120

Li SZ, Chen YH, Du HF, Feldman MW (2010) A genetic algorithm with local search strategy for improved detection of community structure. Complexity 15(4):53–60

Liu X, Li DY, Wang SL, Tao ZW (2007) Effective algorithm for detecting community structure in complex networks based on GA and clustering. In: Proceedings of the 7th international conference on computational science, part II, Beijing, 2007, pp 657–664

Luo ZG, Ding F, Jiang XF, Shi JL (2011) New progress on community detection in complex network. J Natl Univ Def Technol 33(1):47–52 (in Chinese)

Lusseau D, Schneider K, Boisseau OJ, Haase P, Slooten E, Dawson SM (2003) The bottlenose dolphin community of doubtful sound features a large proportion of long-lasting associations. Behav Ecol Sociobiol 54(4):396–405

Newman MEJ, Girvan M (2004) Finding and evaluating community structure in networks. Phys Rev E 69(2):026113

Pizzuti C (2008) GA-net: a genetic algorithm for community detection in social networks. In: Proceedings of the 10th international conference on parallel problem solving from nature, Dortmund, 2008, pp 1081–1090

Pizzuti C (2008) Community detection in social networks with genetic algorithms. In: Proceedings of the 10th annual conference on genetic and evolutionary computation, Atlanta, 2008, pp 1137–1138

Pizzuti C (2009) A multi-objective genetic algorithm for community detection in networks. In: Proceedings of the 21st IEEE international conference on tools with artificial intelligence, Newark, 2009, pp 379–386

Shi C, Yan ZY, Wang Y, Cai YN, Wu B (2010) A genetic algorithm for detecting communities in large-scale complex networks. Adv Complex Syst 13(1):3–17

Zachary WW (1977) An information flow model for conflict and fission in small groups. J Anthropol Res 33(4):452–473

Chapter 26
Complexity Analysis of Software Based on Function-Call Graph

Yang Guo, Zheng-xu Zhao, and Wei Wang

Abstract The scientific study of networks, such as social networks, biological networks and computer networks, is an emerging filed. A large number of recent works have been devoted to the study of community structure and topology in networks. In this paper, we use a function-call graph reconstruction algorithm and present a complex network approach to the study of software engineering. We have examined a function-call graph of software system, and found it to reveal small-world, scale-free features which are similar to those identified in other technological, biological, and sociological systems. The result can contribute to the maintaining of software systems.

Keywords Connected component • Function-call graph • Power-law distributions • Scale-free network • Small world effect

26.1 Introduction

Complex network is being considered as an important interdisciplinary approach to representing complex systems these years. Small-world phenomenon and power-law distributions have been found in many instances of natural and man-made systems. As the basis for understanding the behavior of many large complex systems, networks have become a promising research area in different fields of sociology and biology, such as the World Wide Web, the Internet, the proteins

---

Y. Guo
School of Mechanical Engineering, Shandong University, Jinan, Shandong 250061, China

Z. Zhao (✉) • W. Wang
School of Information Science and Technology, Shijiazhuang Tiedao University, Shijiazhuang, Hebei 050043, China
e-mail: zhaozx@stdu.edu.cn

E. Qi et al. (eds.), *The 19th International Conference on Industrial Engineering and Engineering Management*, DOI 10.1007/978-3-642-37270-4_26,

and genes, and also the power grid of US. These networks are considered to show small-world and scale-free behavior. Small-world here means that characteristic path length is very short in a network, regardless of the clustering property. Scale-free behavior implies that the network lacks a “characteristic length scale”. The distribution of their network graphs, where the vertices represent the instances of system, and the edges represent the relationships between them, follows a power-law. Power-law distribution means that small values are greatly common, and greater values are extremely rare (Newman 2003; Clauset et al. 2009; Guo et al. 2011; Zhao and Lee 2008). Community structure is another common feature of networks. It implies the trend of dividing vertices into groups, where dense links are shown in groups and infrequent links between them (Newman 2004, 2006a).

Recently, researchers have paid close attention to software systems, expecting for finding some features similar to the findings of complex systems. Treating software systems as networks is an emerging field. The exploration of them may contribute to solving the problems found in software engineering, for example, managing the complexity of software growth, or maintaining codes for software engineers.

The remainder of this article is structured as follows. Section 26.2 expresses related works of other researchers. Section 26.3 gives the definition of function-call graph and introduces small-world effect and power-law distributions. We present our experimental results and relevant analysis in Sect. 26.4. Finally, the conclusion is given in Sect. 26.5.

26.2 Related Work

Over recent years, there have been substantial works for discussing power-law distributions or community structure in natural phenomena and the web. Only recently researchers have turned attentions to studying the scale-free behavior in Java or C++ software programs.

Valverde et al. (2002) used the method of complex network first to analyze the topological structure of the Java Development Framework. They had come to a conclusion that the software architecture graphs, where nodes represented classes and edges means relationship between them, revealed small-world and scale-free behavior.

Myers (2003) had analyzed software collaboration graphs of some software systems written in C/C++, and found the same properties. He had also proposed some measures of network topological structure to discuss the relationship to software engineering practices.

Potanin et al. (2005) had examined numerous different systems written in C++, Java, and Smalltalk languages. They studied the graphs where objects as nodes and references as edges. Finally they had drawn a conclusion that the graphs obeyed a power-law distribution.

Concas et al. (2007) had presented a study which used an implementation of an object-oriented system in order to search for scaling laws. They studied about ten system properties, such as the distributions of method or variable names, method and class sizes and inheritance hierarchies. They had also made it clear that why the out-degree distribution showed log-normal or power-low behavior in class graphs.

Louridas et al. (2008) provided evidences to discuss that in software systems the distribution with long, fat tails appearing at different levels of abstraction in diverse languages or systems were much more common than that known before.

Girvan and Newman (Girvan and Newman 2002; Newman and Girvan 2004; Newman 2006b) had proposed an algorithm which used the method of iterative removing edges with high "betweenness" scores and appeared to identify this structure with high sensitivity.

Similar statistical features by other authors have also been identified in networks where nodes represent code files and edges mean relationships between them, for example in C or C++ programs one source file may include other header files. They had also found power-law distributions in networks.

26.3 Networks

We can use various levels of organization, which are ranging from small and simple functions to large and complex modules, to analyze the structure of software systems. The function-call graph, which describes the functions' calling relationships, is a useful representation of programs for software engineers to understand codes for maintaining them effectively.

26.3.1 Function-Call Graph

A function-call graph is a directed graph that represents calling relationships between functions (or methods in object-oriented language) in software systems. Specifically, every node stands for a function and each edge (f_1, f_2) indicates that function f_1 calls function f_2. Thus, a cycle in the graph indicates recursive function calls.

Function-call graphs describe a significant information space of object-oriented software systems. A prerequisite for software maintenance and evolution is that engineers understand the functions performed in systems. As a basic program analytical result, function-call graphs can be used for human understanding of programs, or as a basis for further analyses, such as an analysis that tracks the flow of values between functions, finding functions that are never called, etc.

26.3.2 Small World Effect

One of the most widely discussed of network phenomena is the small world effect, which means the typical network distances between vertices are surprisingly small. This finding may data back to the 1960s, from the Stanley Milgram's letter-passing experiment. In mathematical terms, small world effect is a hypothesis that the mean distance is small, in a sense that will be defined shortly.

Watts and Strogatz have studied lots of real networks, finding the average path length shorter and clustering coefficient higher. The network with these features is called a small-world network. Clustering coefficient is inversely proportional to the size of network, which is significantly larger in a small-world network than that for the random network. The average path length in a small-world network is approximate with that of the random network, where the average path length in random networks is proportional to the logarithm of their size (Newman 2010).

26.3.3 Power Laws

Researchers have paid particular attentions to the mathematical properties of power laws over these years. As the degree distribution of a network obeys the form of power laws, we can consider that the network has a scale-free phenomenon. Mathematically, a quantity x obeys power laws if it is described from a statistical distribution

$$p(x) \propto x^{-\alpha} \tag{26.1}$$

Where α is a constant parameter of the distribution known as the exponent or scaling parameter. However, there're not many empirical phenomena which obey power laws with all values of x. The power-law distribution usually applies only for values larger than some minimum $x_{\min}$ and in such case it can conclude that the tail of distribution follows a power law.

Power laws can be also defined as a complementary cumulative definition under some circumstances, where the probability that a random variable takes at least a value is proportional to a negative power of that value, that is:

$$P(X \geq x) \propto cx^{-k'} \tag{26.2}$$

Scale-free networks appear everywhere and may be able to describe all sorts of different phenomena, so researchers analyze the ubiquity of power laws for a reasonable explanation and have proposed some models, for example, the preferential attachment model. In this paper, we studied the existence of scale-free networks at the function level, expecting of optimization models for code design.

26.3.4 Experiment Database

Celestia is a typical application of interactive and real-time visualization for space. It contains over 10,000 galaxies and more than 100,000 stars of the solar system in detail. Not only that, this program provides an extensibility mechanism for users to add more objects. Users can run Celestia which is based on the Hipparcos Catalogue for travelling through extensive universe at any direction or speed, especially at any time in history. The application also offers a unique perspective with objects ranging from single spacecraft to entire galaxies in scale, which other ground-based display can't achieve.

26.4 Statistical Analysis

26.4.1 Topology and Modularity

We have elaborated a reconstruction algorithm based on Windows systems for obtaining the function-call graph from software systems. Figure 26.1 shows the topology of function-call graph for Celestia. This network with 2,990 functions and 22,435 calls can be divided into 59 modules. There is a clear modular organization and nodes naturally cluster in different modules which are distinguished by node size. The clustering coefficient C of this network is about 0.079, much larger than

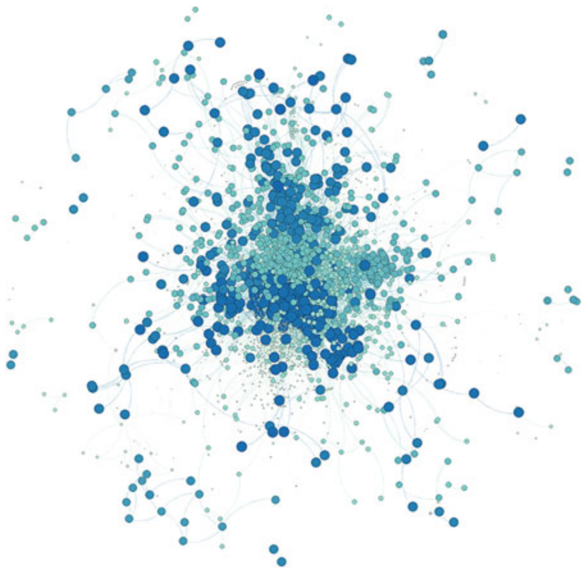


Fig. 26.1 Directed function-call graph for the Celestia

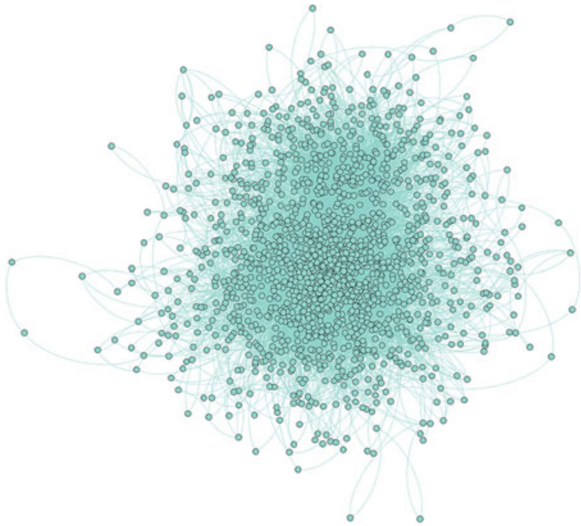


Fig. 26.2 Largest strong connected component for Celestia

that of random network. The average path length 3.862 is also proportional to the logarithm of network size. That is to say the clustering coefficient in function-call graphs is well above expectations of random graphs while the average path length is so small.

26.4.2 Connected Component

Directed graphs have weakly and strongly connected components (SCC). A strongly connected component is a maximal subset of vertices in a network such that each can reach and is reachable from all of the others along a directed path. There is typically one large strongly connected component in a directed network and a selection of small ones.

Figure 26.2 shows the strongly connected component for Celestia, comprising a large fraction of total nodes in the system (more than 50 %). This strong membership in SCC is in accordance with that found in other directed complex networks.

26.4.3 Degree Distributions

It is often illuminating to make a plot of the degree distribution of a large network. Figure 26.3 shows an example for Celestia system at function level. From it we can see that most of the vertices in the network have low degree, however, there is a significant "tail" to the distribution, corresponding to vertices with substantially higher degree.

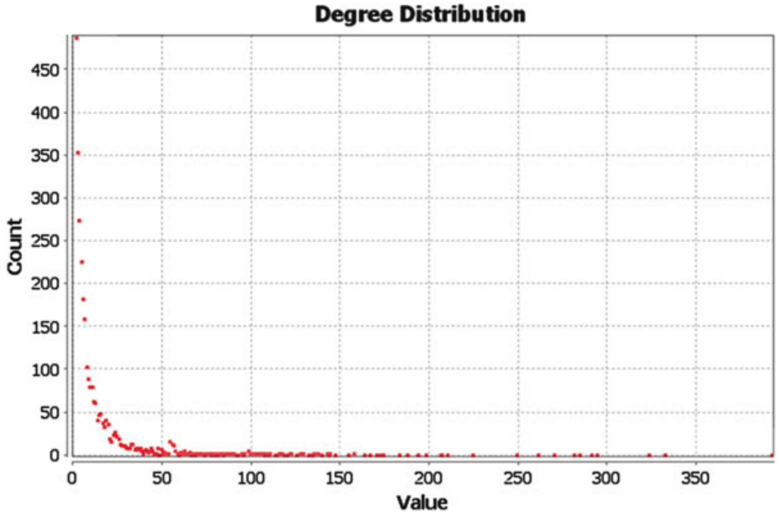


Fig. 26.3 The degree distributions of Celestia

For the particular data set of Celestia, a total of 2,990 vertices in the network, the highest degree vertex in the network has degree 392. The average degree of this network is about 7.503 and the network diameter is about 14. It means that the most highly connected vertex is connected to about 13 % of all other vertices in the network. We often call such a well-connected vertex a hub. In fact, it turns out that almost all real-world networks have degree distributions with a tail of high-degree hubs like this.

We have studied the distribution and plotted the cumulative degree distributions in log-log scale. Figure 26.4 shows the cumulative degree distributions for Celestia. Longitudinal coordinate indicates the number of nodes in the function-call graph of Celestia with degree greater than or equal to *k* and the dotted line expresses power-law fits on log-log scales. As the figure shows, when viewed in this way, the degree distribution follows a straight line.

26.5 Conclusion

According to recent studies in the research of complex networks, we have analyzed the function-call graph contained within Celestia systems. Small-world effect could have substantial implications for networked systems. We have paid attentions to static analyses of source code at function level by using a one open-source visual simulation platforms, and concluded that the Celestia system analyzed here have displayed the small-world and power-law behavior. The average distance between any pair of functions is very small and the structure of them is characterized by

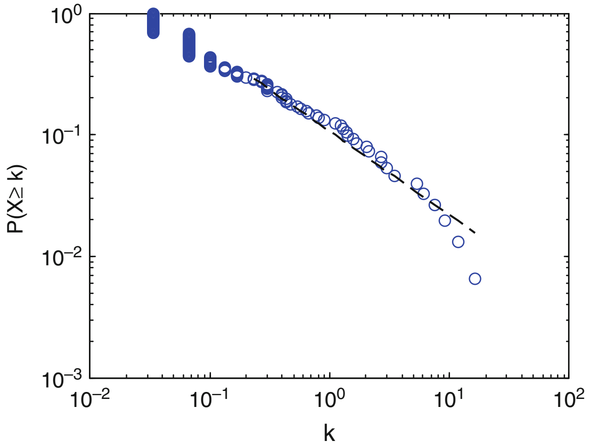


Fig. 26.4 Cumulative degree distribution for Celestia

a degree distribution following a power-law with similar exponents. The largest strongly connected component of the network fills about a half of network. The implication of this study will contribute to predict the dimensions of future systems and evaluate the complexity of maintaining and developing those systems.

Acknowledgment The work is funded by the NSFC Funding Council, the Peoples' Republic of China, under the grant number 60873208.

References

Clauset A, Shalizi CR, Newman MEJ (2009) Power-law distributions in empirical data. Siam Rev 51(4):661–703

Concas G, Marchesi M, Pinna S, Serra N (2007) Power-laws in a large object-oriented software system. IEEE Trans Softw Eng 33(10):687–708

Girvan M, Newman MEJ (2002) Community structure in social and biological networks. Proc Natl Acad Sci 99(12):7821–7826

Guo Y, Liu ZW, Zhao ZX (2011) Complexity analysis on link structure of World Wide Web. Comput Eng 37(23):105–106 (Chinese)

Louridas P, Spinellis D, Vlachos V (2008) Power laws in software. ACM Trans Softw Eng Methodol 18(1):1–26

Myers CR (2003) Software systems as complex networks: structure, function, and evolvability of software collaboration graphs. Phys Rev E 68(4):61–76

Newman MEJ (2003) The structure and function of complex networks. Siam Rev 45(2):167–256

Newman MEJ (2004) Detecting community structure in networks. Eur Phys J B 38(2):321–330

Newman MEJ (2006a) Modularity and community structure in networks. Proc Natl Acad Sci 103(23):8577–8582

Newman MEJ (2006b) Finding community structure in networks using the eigenvectors of matrices. Phys Rev E 74(3):6104–6125

Newman MEJ (2010) Networks: an introduction. Oxford University Press, New York

Newman MEJ, Girvan M (2004) Finding and evaluating community structure in networks. Phys Rev E 69(2):6113–6128
Potanin A, Noble J, Frean M, Biddle R (2005) Scale-free geometry in object-oriented programs. Commun ACM 48(5):99–103
Valverde S, Ferrer-Cancho R, Sole R (2002) Scale-free networks from optimal design. Eur Phys Lett 60(4):512–517
Zhao Z, Lee ZZ (2008) Small-world phenomenon: toward an analytical model for data exchange in product lifecycle management. Int J Internet Manuf Serv 1(3):213–230

Chapter 27
Fault Diagnosis of Wet Flue Gas Desulphurization System Based on KPCA

Yu-ping Zheng and Li-ping Zhang

Abstract Fault detection and diagnosis for sensor are necessary, which affect the performance of the thermal power plant of wet flue gas desulphurization system seriously. A fault diagnosis method using kernel principal component analysis (KPCA) is proposed to affectively capture the nonlinear relationship of the process variables, which computes principal component in high dimensional feature space by means of integral operators and nonlinear kernel functions. The faults are detected by calculating the statistics of the square prediction error (SPE) and identified by calculating the change diagram of contribution percentage of Hostelling T^2. At last, employing the actual data from wet flue gas desulphurization system of Huaneng Fuzhou power plant, it's proved effectively to detect and identify four kinds of faults, which is the complete invalidation fault, fixed bias fault, drift bias fault and precision degradation fault. The result shows the KPCA method has a good performance in fault detection and diagnosis.

Keywords Fault detect and diagnosis • Gas desulphurization • KPCA • Wet flue sensors

27.1 Introduction

In recently the focus of controlling the SO_2 emissions at home is still dropped on the reduction of the SO_2 emissions in the thermal power plant. The thermal generator mainly through the way of desulphurization of coal fired boiler flue gas to reduce the SO_2 emissions. Compared all kinds of gas desulphurization, wet flue gas desulphurization has became the main gas desulphurization technology

Y. Zheng (✉) • L. Zhang
College of Mechanical Engineering and Automation, Fuzhou University, Fuzhou, China
e-mail: 1034656760@qq.com

E. Qi et al. (eds.), *The 19th International Conference on Industrial Engineering and Engineering Management*, DOI 10.1007/978-3-642-37270-4_27,

as for the advantages of the process maturity and high effect desulphurization at home and abroad. It is a must to monitor closely the process parameter and the states of some important machinery equipment to guarantee the wet flue gas desulphurization operation normally. The accuracy of the gain monitoring data rely on the reliability of all kinds of sensors, which will influence the desulphurization system operated normally and not effectively gas desulphurization as to discharge SO_2 to the atmosphere if the sensors occur the faults (Zhou Zhixiang et al. 2006; Zhong et al. 2008). So it is necessary to diagnosis and research the question on desulphurization system sensors faults.

Wet flue gas desulphurization technology (FGD) does not use the accurate mathematical model to diagnosis the faults, due to huge project, system complexity, process mechanism complexity, numerous of variables, and difficultly relying on accurate mathematical model. Principal Component Analysis (PCA) is a way of data driven for process monitoring widely, but it is only a way of linear transformation and can't use to nonlinear process. Kernel Principal Component Analysis (KPCA) can handle the nonlinear data very good, not only has the simplicity of PCA, but also possess the practical value (Scholkopf et al. 1988; Lee and Yoo 2003).

Therefore the paper proposes a way of fault diagnosis of wet flue gas desulphurization system based on KPCA, namely use KPCA to built the sensor fault diagnosis model, and use SPE statistic to diagnosis the sensor faults. At last, the data acquainted from wet flue gas desulphurization system of Huaneng Fuzhou power plant proves that the way can effectively monitor sensor faults.

27.2 KPCA

27.2.1 KPCA Algorithm

The key idea of KPCA is through the nonlinear transform $\Phi(\cdot)$ to map the sample data from the input space to high dimensional space F, and carry PCA analysis and extraction of characteristic variables out in the feature space (Mika et al. 1999; Cho et al. 2005; Maestri and Cassanello 2009).

Given original sample $x_k \in R^N$ $(k = 1, 2, \cdots, M)$, if they meet the zero mean condition, the covariance matrix can be expressed in the feature space:

$$C = \frac{1}{M} \sum_{i=1}^{M} \Phi(x_i)\Phi(x_i)^T \tag{27.1}$$

The covariance matrix characteristic equation

$$\lambda \nu = C \nu \tag{27.2}$$

Where the eigenvalue of matrix C is λ, ν is the eigenvector of matrix C.

Eigenvector ν can be expressed:

$$\nu = \sum_{j=1}^{M} \alpha_j \Phi\left(x_j\right) \tag{27.3}$$

There is a mapping vector $\Phi\left(x_k\right)$, thus

$$\lambda\left(\Phi\left(x_k\right) \cdot \nu\right) = \left(\Phi\left(x_k\right) \cdot C\nu\right), k = 1, 2, \cdots, M \tag{27.4}$$

Let's define a $M \times M$ matrix K:

$$K_{ij} = k\left(x_i, x_j\right) = \left(\Phi\left(x_i\right) \Phi\left(x_j\right)\right)$$

Combining Eqs. (27.3) and (27.4), we can obtain

$$M\lambda\alpha = K\alpha \tag{27.5}$$

$M\lambda$ is the eigenvalue of K, coefficient vector $\alpha = (\alpha_1, \alpha_2, \cdots, \alpha_M)^T$ is the eigenvalue $M\lambda$.

Normalizing the vector ν:

$$\nu^k \cdot \nu^k = \langle \nu^k \cdot \nu^k \rangle = 1, k = 1, 2, \cdots, M$$

We can obtain PCs, where $k = 1, 2, \cdots, M$.

$$t_k = \langle \nu^k \cdot \Phi(x) \rangle = \sum_{i=1}^{M} \alpha_i^k K\left(x, x_i\right) \tag{27.6}$$

27.2.2 Choose Kernel Function

The requirement on the kernel function is that it satisfies Mercer's theorem. In general, kernel functions are as follows:

1. Polynomial kernel

$$k(x, y) = < x, y >^d,$$

2. Sigmoid kernel

$$k(x, y) = \tanh\left(\beta_0 < x, y > +\beta_1\right),$$

3. Radial basis kernel

$$k(x, y) = \exp\left(-\frac{\|x - y\|^2}{c}\right),$$

Up to the present there is not a united theory to guide the choice of kernel function and fewer papers refer to the choice of KPCA parameter. So we usually use it by experience. This paper chooses the sigmoid kernel function.

27.2.3 Fault Detect

Fault detect based on KPCA is realized through monitoring the transform between statistic Hostelling T2 and SPE which is similar to PCA. T^2 changes in the PCs and is standard square of PCs vector, which is on behalf of change trend and degree of deviation in amplitude each sample, expresses a kind of prediction in internal model. SPE statistic changes in the residual space, is another important statistic indexes for fault detect, which expresses every change trend and error of building statistic model and a measurement outside the model. Both variables are as follows (Choi et al. 2005; Niu Zhen 2006; Rongyu 2007; Zhang et al. 2012; Cui et al. 2008):

$$T^2 = [t_1, t_2, \cdots, t_k] \wedge [t_1, t_2, \cdots, t_k]^{\mathrm{T}} \tag{27.7}$$

Where t_k is obtained by Eq. (27.6), $\wedge$ is diagonal matrix formed by eigenvalue, and confidence limit of T^2 is get by F distribute.

$$SPE = \|\Phi(x) - \Phi_b(x)\|^2 = \sum_{i=1}^{M} t_i^2 - \sum_{i=1}^{b} t_i^2 \tag{27.8}$$

The confidence limit of SPE can calculate approximately:

$$SPE_a \sim gx_h^2$$

Where g and h are weight parameters and degree of freedom of SPE, if a and c are evaluating mean and variance of SPE, g can be expressed as $g = c/2a, h = 2a^2/c$, b is the number of PCs.

27.2.4 Fault Diagnosis

The faults are detected, it need to find the reason of fault emerge and fault variables in time. Because fault variables have a linear relation to detected variables, using

PCA can easily calculate variable contribution and draw contribution diagram. As the nonlinear transform process has not use explicit nonlinear transform function, adding that kernel function doesn't provide the relation between measures variables and monitor variables, now many papers consider the way of contribution diagram using in PCA can't fit in KPCA (Xiaogang and Xuemin 2005).

This paper employs a way of contribution diagram to solve original measure variables relate to faults, and compares by the change percentage of the contribution before and after faults happened to separate fault variables (Nguyen and Golinval 2010; Shao et al. 2009).

The number j original measure variable has a contribution to T^2 statistic:

$$cntr_{j,k} = \sum_{i=k}^{b} \left| t_k^T x_j / \lambda_k \right| \tag{27.9}$$

The percentage of contribution is as follows:

$$cper_{t_1/t_2,j} = \frac{cntr_{j,k}}{\sum_{j=1}^{n} cntr_{j,k}}, \Delta cper_j = cper_{t_2,j} - cper_{t_1,j} \tag{27.10}$$

Where b is number of PCs, t_k is number k nonlinear PCs, x_j is number j measure variable, t_1 and t_2 are the time of sensor fault accursed before and after, and n is number of sensor.

27.3 Fault Diagnosis Simulation Experience

Wet flue gas desulphurization system use low-priced limestone as desulphurization absorbent and the absorption liquid injected into absorber by nozzle atomization disperse tiny drops to cover all absorber cross section. These drops counter-current contact with gas in tower, occur mass transfer and absorb reaction.

For example, using the sensor operation data of wet flue gas desulphurization to built KPCA model by Fuzhou Huaneng power plant third unit. The sample times is per minute and choose fourteen variables 500 groups of data operated normally to built sensor fault detect model in steady working condition, such as boiled load, FGD entry signal, FGD entry flow, FGD entry gas temperature, clean gas exit pressure, clean gas entry pressure, boiled chimney entry signal, boiled chimney entry humidity, boiled chimney entry temperature, booster fan exit pressure, clean chimney signal, booster fan power, mist eliminator entry pressure and mist eliminator exit pressure.

Choose another 500 groups of data operated normally to sensor fault diagnosis and analysis, and start from 150 min to join zero mean complete failure fault, 3 °C bias fault, 0.05 drift fault and precision degradation fault of one variance into faults. The figures are as follows:

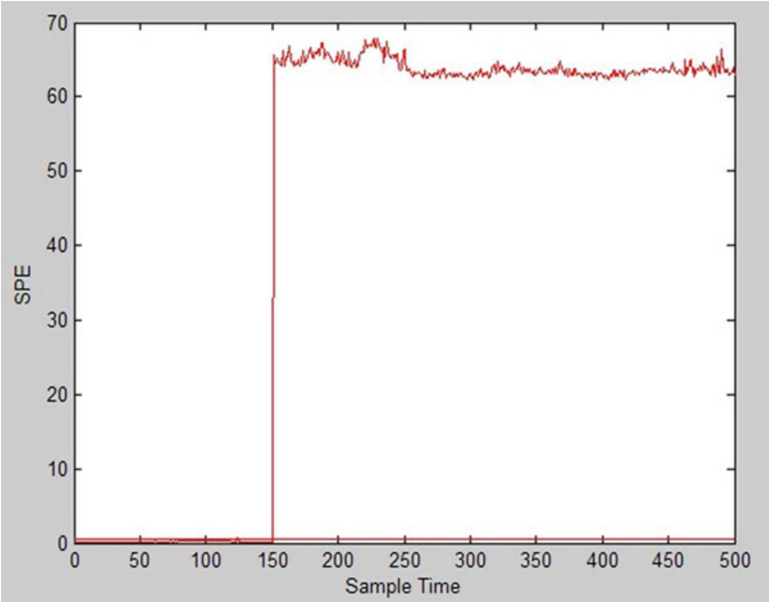


Fig. 27.1 SPE of complete failure fault

In Fig. 27.1, as the SPE statistic in fault is too large to exceed the control limit, leading the control limit and SPE statistic curve close together. In Fig. 27.2, the SPE statistic exceeds the control limit from 151 min. In Fig. 27.3, the SPE statistic grows rapidly, because the fault and the SPE are small at the beginning. After the time grows, the fault deviates from actual value which corresponding to SPE grows more. In Fig. 27.4, the SPE fluctuates up and down the control limit but the most are up the control limit after sensor occur precision degradation fault. The above figures see KPCA can detect four normal faults, and KPCA is better than linear PCA in precision degradation fault.

When the faults detected, it needs to confirm which sensor is the fault sensor, and estimates by the percentage of contribution. As the ninth variable's contribution is the maximum, it's found that the fault occur in the ninth sensor (Figs. 27.5 and 27.6, here it only gives two kinds of faults such as complete failure fault and drift fault):

27.4 Conclusion

In this paper, aiming at the question of wet flue gas desulphurization sensor fault to influence all the system normal operation, diagnosis and research faults may occur based on KPCA is proposed. An experience is simulated by Matlab, the result proves KPCA can diagnosis wet flue gas desulphurization system fault, and affectively detect whether the system occurs fault or not. It has practical significance.

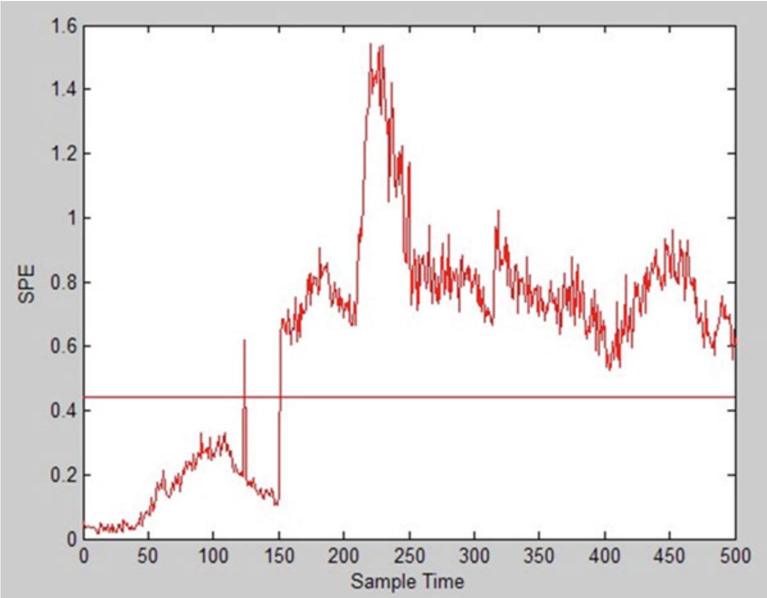


Fig. 27.2 SPE of bias fault

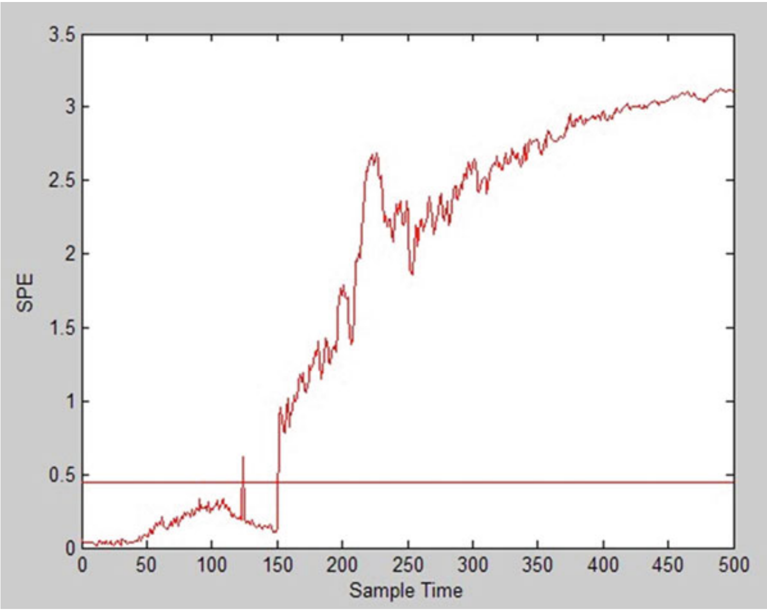


Fig. 27.3 SPE of drift fault

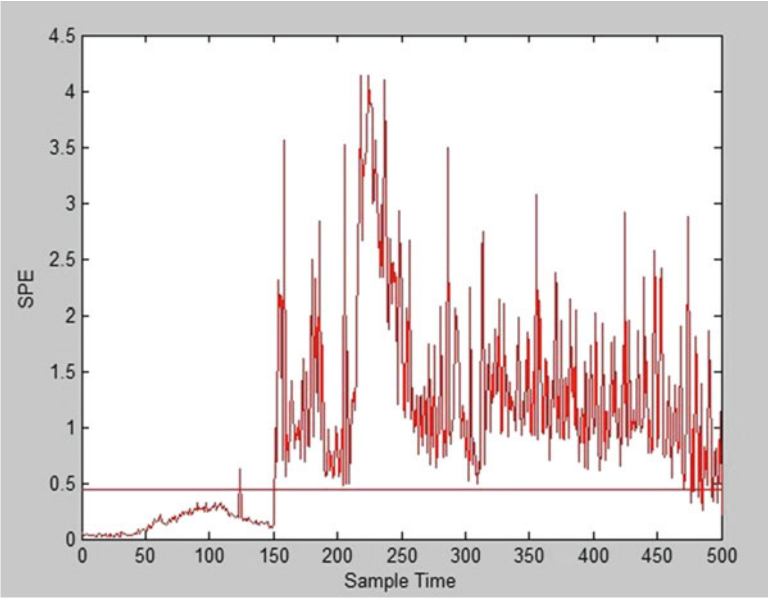


Fig. 27.4 SPE of precision degradation fault

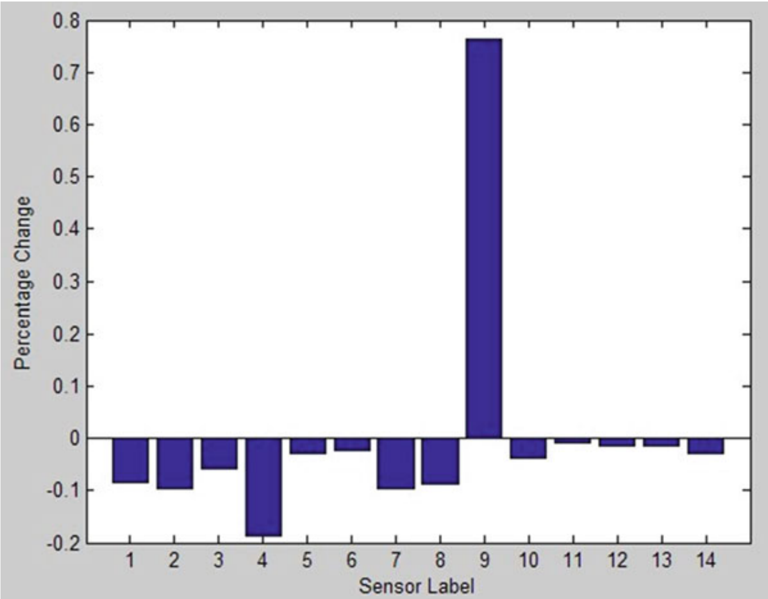


Fig. 27.5 Change diagram of contribution percentage of complete failure fault

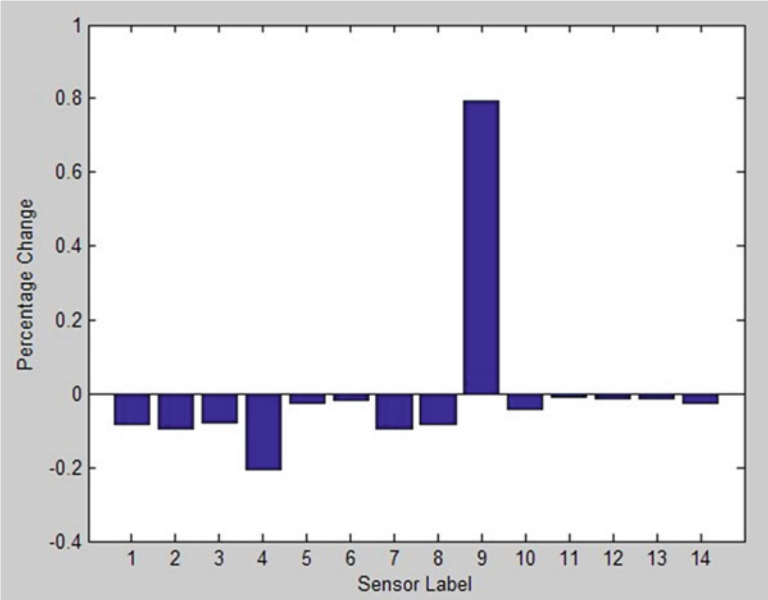


Fig. 27.6 Change diagram of contribution percentage of drift fault

Acknowledgment This paper is done by the careful guidance of professor Zhang. In the process of the paper, tutor and senior sister pay for great effort. I thank my tutor and senior sister great help and deep respect.

References

Cho JH, Lee JM, Choi SW, Lee D, Lee IB (2005) Fault identification for process monitoring using kernel principal component analysis. Chem Eng Sci 60(1):279–288

Choi SW, Lee C, Lee JM, Park JH, Lee IB (2005) Fault detection and identification of nonlinear process based on kernel PCA. Chemom Intell Lab Syst 75(1):55–67

Cui P, Li J, Wang G (2008) Improved kernel principal component analysis for fault detection. Expert Syst Appl 34:1210–1219

Lee J-M, Yoo CK (2003) Nonlinear process monitoring using kernel principal component analysis. Chem Eng Sci 7(59):223–234

Maestri ML, Cassanello MC (2009) Kernel PCA performance in processes with multiple operation modes. Chem Prod Process Model 4(5):1934–2659

Mika S, Scholkopf B, Smola AJ (1999) Kernel PCA and de-nosing in feature spaces. Adv Neural Inf Process Syst 11:536–542

Nguyen VH, Golinval J-C (2010) Fault detection based on kernel principal component analysis. Eng Struct 32:3683–3691

Niu Zhen (2006) Fault diagnosis for control systems of thermal power plant based on multivariate statistical analysis. North China Electric Power University, Beijing

Rongyu L (2007) Research on statistical process monitoring based on PCA. Zhejiang University, Zhejiang, China

Scholkopf B, Smola A, Muller KR (1988) Nonlinear component analysis as a kernel eigenvalue problem. Neural Comput 10(5):1299–1319

Shao J-D, Rong G, Lee JM (2009) Learning a data-dependent kernel function for KPCA-based nonlinear process monitoring. Chem Eng Res Des 87:1471–1480

Xiaogang D, Xuemin T (2005) Nonlinear process fault diagnosis method using kernel principal component analysis. J Shandong Univ (Eng Sci) 35(3):103–106

Zhang Y, Li S, Teng Y (2012) Dynamic processes monitoring using recursive kernel principal component analysis. Chem Eng Sci 72:78–86

Zhong Y, Gao X, Huo W et al (2008) A model for performance optimization of wet flue gas desulphurization system of powers plants. Flu Process Technol 89(11):1025–1032

Zhou Zhixiang, Duan Jianzhong, Xie Jianming (2006) Wet flue gas desulphurization technology handbook. China Electric Power Press, Beijing

Chapter 28
The Research on Negative Perfectionism Personality of Engineering Students from the Perspective of Game Theory

Ying Sun, Ying-ying Yang, Chen-xu Feng, and Hui Zhang

Abstract With the development of industrial engineering, it is the bounden responsibility of colleges to develop engineers which can meet the needs of the society. Mental health level of engineering talents has a profound impact on engineering talent structure. Many of "perfectionism" personality traits of the individual resulted in many serious physical and psychological problems that are eroding physical and mental health of people which has become a serious social problem. In this paper, the perfectionism personality of individual acts are shown on the principal-agent theory, namely through the establishment of a two-stage dynamic game model to explain the behavioral choice between "ideal self" and the "actual self" and other possible incentive problems caused by repeated games ("ratchet effect"). The key point is tried to be found which cause individual strange thinking model. The internal mechanism is further analyzed through the model in order to promote engineering talents' growth and development.

Keywords Negative perfectionist • Ratchet effect • Ideal self • Actual self • Engineering students

---

Foundation programs: State Social Sciences Fund project in 2008. The research of availability of promoting the level of college student's psychological well-being by flexible ideological and political education (08BSH047); The Ministry of Education of Humanities and social science research planning fund projects in 2011 the interpretation of college student's meaning in life and research of the way of promoting. (11YJA190014); The key research topic of Tianjin Institute of higher education twelve-five plan for education science and planning "The talent interpretation of Education Excellence Engineer training plan under the perspective of Engineering Ethics" (125z005).

Y. Sun (✉) • Y. Yang • C. Feng • H. Zhang
School of Vocational and Technical Education, Tianjin University, Tianjin, China
e-mail: wyssying@126.com

E. Qi et al. (eds.), *The 19th International Conference on Industrial Engineering and Engineering Management*, DOI 10.1007/978-3-642-37270-4_28,

Human resources are the strategic resources of modern enterprise. Human beings, in the core position, are the most critical factor of the development of the enterprise in the enterprise's personnel, financial and material resources elements and production-supply-marketing links. So that is to say, some competitions of modern enterprise in the final analysis are the competitions for talent. With the development of industrial engineering, it is the bounden responsibility of colleges to develop engineers which can meet the needs of the industrial areas, the world and the future. Currently, the concepts of training talents in the industrial engineering are constantly updated; meanwhile, many colleges have praised the training mode of paying attention to the practice ability, innovation ability and leadership. However, in the society many people with "negative perfectionist" personality have appeared especially in the excellent elite group. As a result, it has become a very serious social problem (Zhang et al. 2010). People with "negative perfectionist" personality always feel that their behaviors do not reach their requirements, although they usually do better than around, they cannot feel happy (Fang et al. 2007). Currently, negative perfectionism is a common psychological phenomenon in college students and middle school students, especially in the excellent students (Sun and Feng 2012). Because the way of thinking of engineering discipline itself is in pursuit of rigorous, accurate and strict logic reasoning (Yang et al. 2004), and engineering which permits no dishonesty, needs careful calculation and thinking, and its design thought strives for excellence, and then put the design into practice perfectly. This kind of thinking mode and strategy of doing things are advocated in the "scientific world" and all of the engineering students should accept the strict training. But in this kind of training of thinking mode, if we lack of proper intervention and concept of macro-control and do not pay attention to counseling, the engineering students will confound the "scientific world" with "the world we live in" (Husserl 1997) easily, and they don't know the dangers which brought by confusion. Under this precondition, engineering students also bring the strict requirements to the world they live in, finally when some of the engineering students think about the problems about life, their thinking has get ossified, rigid and has no idea how to be flexible, and they often deal with problems in real-life just only according to one principle. Furthermore, they insist on wasting time on some unworthy issues and excessively put their energy on some inadequately things, and the evaluative standard is rigid in their life, and they blindly regardless of objective reality, so it leads to many psychological problems. So it is not surprising that a series of researches show that perfectionism has a close relationship with some psychological distresses and mental disorders (Zi and Zhou 2005). Take engineering students for example, for many engineering students with personality trait of perfectionism, their "actual self" and "ideal self" have a delicate internal contact. On the one hand, they give themselves a high standard; on the other hand, they show some "lazy" in the real life. After some time, because ideal self of high standards is dissatisfied with actual self of making no effort to seek progress and mental deterioration, it is very easy to cause all kinds of psychological entanglement, more seriously it might make some perfectionists terrible depression and anxiety because of their demotivation, in the end engineering students become a kind of people who often experience psychological problems (Sun 2012).

In the previous researches for industrial engineering talents, the main focus is practice ability, innovation ability and leadership ability. However, with the development of times, increasingly, educators find that the mental health level of engineering talents has a profound impact on engineering talent structure. So this paper thinks the research of "negative perfectionist" personality which easily results in mental health problems for engineering students under the interdisciplinary perspective is meaningful. Meanwhile, as an aspect of the research of engineering talents training it deserves special attention.

28.1 Perfectionism and Negative Perfectionism

Perfectionist refers to those people who set high individual standard, which is beyond their own ability, force themselves to some impossible goals and evaluate self-value according to the quantity of finished task. Their quest for perfect can only lead to their self-defeating (Burns 1980). Psychologists find two kinds of perfectionism, they are positive (normal) perfectionism and negative (nervous) perfectionism (Dunkey et al. 2003; Slade and Owens 1998; Bergman et al. 2007). Negative perfectionist set too high and ideal work standard for themselves with strict self-evaluation, and they always tend to adopt an "all or nothing" way of thinking, which is either success or failure (Pacht 1984). The needs of reaching perfection of negative perfectionism make them feel that they are in a predicament of self-defeating (Weisinger and Lobsenz 1981; Xie 2004).

The gap between negative perfectionists' "real self" and excessive, unrealistic "ideal self" is too large, resulting in that negative perfectionist may feel more psychological frustration in reality. And when the psychological frustration continue to accumulate, negative perfectionist is prone to distress and psychological and behavioral problems (Nahoko 2004). We can imagine an engineering practitioners suffering from such psychological and behavioral problems, which not only hinder his personal career development, but also influence their particular field of industrial engineering development. So research on the internal psychological mechanism of negative perfectionism has become an important topic of industrial engineering education. Explain clearly the internal psychological mechanism of negative perfectionism, which is helpful to provide better psychological assistance initiatives to the prospective engineers who are puzzled by negative perfectionism, and indirectly promote the progress of industrial engineering.

It is the interaction mechanism between the negative perfectionist "ideal self" and "real self" that initiate the ratcheting effect during the development of negative perfectionist, that is, the "ideal self" hope the "real self" act as perfect standard, and try to evaluate the "real self" according to the perfect standard and past outstanding performance. The harder the "real self" works, the more possibly it gets good performance, and the higher the standard will be, so when "real self" realize his effort will improve the standard, it will low its positive will. So we can analysis the perfectionists' ratcheting effect under the framework of principal-agent theory, and try to explain the internal psychological mechanism of negative perfectionists according to the ratcheting effect in economics. Here we consider a two-stage dynamic model and analysis the ratcheting effect that negative perfectionists perform and the internal mechanism lead to the individual psychological distress through the model. We analysis how the ratcheting effect weaken the incentive mechanism, find the key point cause the thinking circle, and then seek the aid method to negative perfectionist.

28.2 Engineering Students' Negative Perfectionism Under Two-Stage Dynamic Game Model

Analyzing the negative perfectionist's behavior under the framework of principal agent theory, "ideal self" is the principal, while the "real self" is an agent. The "ideal self" cannot directly observe what actions the "real self" chosen, but it can not observe some variables such as individual achievements, honor and success. These variables is decided by the effort level of "real self", the intrinsic ability of "real self" and other exogenous random factors, they are incomplete information for the action of "real self". So The general reputation model is applied to the "principal-agent" model, namely, explain the commissioned the behavior choice through the establishment of a two stage dynamic game model, as well as many game may bring another incentive problems – "ratcheting effect" (Zhang 2003).

Assumed that the game has two stages, it is expressed as $t = 1, 2$, the Single-stage output function is,

$$\pi_t = a_t + \theta + u_t \quad t = 1, 2$$

Among them, the individual output is π_t, and π_t is common knowledge, that is, the information of output that "ideal self" and "real self" all know; a_t is effort level of "real self", is the private information of "real self"; θ is the individual intrinsic productivity, assumed the intrinsic production capacity obey the distributed random variables with the mean is the $\mathrm{E}(\theta) = \bar{\theta} > 0$ and the variance is σ_θ^2; u_t is exogenous random variables, assumed u_t obey normal distribution with mean is 0 and variance σ_u^2, θ and u_t are independent.

And we assume that u_1 and u_2 are independent, so $\mathrm{cov}(u_1, u_2) = 0$. We also assume that "real self" is neutral risky and the discount rate is zero (the discount factor is 1). Thus the utility function as follows:

$$U = w - c(a_1) - c(a_2)$$

Among them, w is the total reward that "real self" obtains. According to the hypothesis, $w = w_1 + w_2$, w_1 and w_2 are the reward that "real self" obtains in stage 1 and stage 2.

So, the utility function can also express as follows:

$$U = w_1 - c(a_1) + w_2 - c(a_2)$$

Here, $c(a)$ is the cost function of "disutility" or effort of "real self" single stage, assumed $c(a)$ is a strictly increasing convex function, and $c'(a) > 0$, $c''(a) > 0$; $c''(a) > 0$; $c'(a) > 0$ means that the marginal utility of the "real self" to is increasing, $c'(a) > 0$ means that "real self" hope make less effort, which is just conflict with the high demand of "ideal self".

Under the above assumptions, if the "ideal self" and "real self" to sign an incentive contract: $w_t = \pi_t -$ y0, among them, y0 doesn't rely on π_t, represents a fixed payment that "real self" gives to "ideal self" for the control for individual thought and the individual ownership. The "ideal self" cannot commitment a long-term contracts to the "real self", that means the stage 2 contract can be made after observe π_1, which is completely compatible with the negative perfectionist personality traits. "Ideal self" always hope "be better", so it will think that once the commitment to a long-term contract, "standard" will be fixed, while in the eye of "ideal self", fixing "standard" means that "real self" will stay in this "standard" without the guidelines of high standard. It is this fear that make the "ideal self" easy to go back, the specific performance is, when it learned the output that "real self" in the first period, "ideal self" will want to modify the contract and promote "standard". At the same time, because the "ideal self" want to pretend the pride of "real self", always demand "real self" perfect, in this situation, after experience the t1 and t2, the hard-working "real self" doesn't get more reword from "ideal self", which makes the "real self" realize that their efforts did not get due incentive.

On the other hand, usually the realistic environment also cannot separate θ and u_1, that is to say, realistic environment don't know π_1 is the result of management ability of "actual self" or the result of exogenous u_1 which is uncertain factor. Therefore realistic environment need to deduce θ according to π_1. So, order:

$$\tau = \frac{Var(\theta)}{Var(\theta) + Var(u_1)} = \frac{\sigma_\theta^2}{\sigma_\theta^2 + \sigma_u^2}$$

Var means variance.

τ is the ratio of the variance between θ and π_1. The larger σ_θ^2 is, the larger τ is. That is to say, the less ensure about someone's inner productivity, the larger τ is. "Ideal self" could observe π_t when each stage finished, but it could not observe θ and u_t. "Ideal self" correct the judgment of θ according to π_1 by observing. Under rational expectation,

$$E(\theta/\pi_1) = (1-\tau)\bar{\theta} + t(\pi_1 - \bar{a}_1)$$

In the formula τ is between 0 and 1, because if there isn't prior uncertainty – $\sigma_\theta^2 = 0$, then $\tau = 0$,realistic life need not to correct; on the other hand, if the prior uncertainty is large – $\sigma_\theta^2 \to \infty$, or if there isn't exogenous uncertainty – $\sigma_u^2 = 0$, then $\tau = 1$, realistic environment will change its judgment for θ fully according to π_1 by observing. So in general τ is between 0 and 1.

τ reflects information about π_1: the more large τ is, the more information π_1 contains, hence "actual self" prefer to working in the first stage, therefore it will produce more reputation, so a_1 is an increasing function of τ, this is a good incentive mechanism, "actual self" will work hard.

However, negative perfectionist's "ideal self" usually uses the thinking model of "better and better", thus the requirement of its "Actual self" is dynamic. The strategy adopted by "Ideal self" is signing the explicit contract at every stage, instead of signing long-term contracts across two stages with "Actual self". Because after "Actual self" reached the "standard" of the previous stage industriously, it always want to develop a higher "standard" to the "Actual self" through the development of contract for the next stage as quickly as possible. And it hoped that the "Actual self" can reach the new "standard" through efforts in new stage. Videlicet, the contract for the second stage must be developed after the "Ideal self" observed the π_1, and this is the real reason why the "Ideal self" can't promise the "Actual self" for long-term contracts. This moment, the optimal incentive mechanism design requires that "Actual self" assumes the entire risk and the "Ideal self" receive a fixed "amounts that turned over" at each stage, record it as α_t, then,

$$w_t = \pi_t - \alpha_t \tag{28.1}$$

According to Formula 28.1,

$$w_1 = \pi_1 - \alpha_1;\ w_2 = \pi_2 - \alpha_2$$

At the first stage, the expectation of the individual's internal production capacity is

$$E(\theta) = \bar{\theta}$$

It's the amounts that individuals should turn over to "Ideal self". It is "Ideal self" that who control the individual's thinking, equivalent to that the ownership of a company belongs to the principal, thus

$$\alpha_1 = E(\theta) = \bar{\theta};\ \alpha_2 = E(\theta\,/\pi_1) = (1-\tau)\bar{\theta} + \tau(\pi_1 - \bar{a}_1)$$

Adapting α_1 & α_2, w_1 & w_2 into the "Actual self" utility function, $U = w_1 - c(a_1) + w_2 - c(a_2)$, we have

$$\begin{aligned}
U &= \pi_1 - \alpha_1 - c(a_1) + \pi_2 - \alpha_2 - c(a_2) \\
&= a_1 + \theta + u_1 - \bar{\theta} - c(a_1) + a_2 + \theta + u_2 - (1-\tau)\bar{\theta} - \tau(a_1 + \theta + u_1 - \bar{a}_1) - c(a_2) \\
&= [(1-\tau)a_1 + (1-\tau)\theta + (1-\tau)u_1 - c(a_1)] + a_2 + \theta + u_2 - c(a_2) + (2-\tau)\bar{\theta} + \tau\bar{a}_1
\end{aligned}$$

Based on the condition of the first order

$$\frac{\partial U}{\partial a_1} = 0, \frac{\partial U}{\partial a_2} = 0,$$

we have, $c'(a_1) = 1 - \tau < 1;\ c'(a_2) = 1$.

The effort level of stage 2, a_2, is selected after the contracts of stage 1 & 2 are given. That means, the effort level of "Actual self" in stage 2 is the Pareto first order optimal, yet the effort level in stage 1 is less than Pareto first order optimal level. The reason is that the "Actual self" is the only residual claimants, and "Actual self" selects a_2 in stage 2, made that the marginal product equals the marginal cost. However, the situation is different in stage 1. "Actual self" must consider not only the direct yield brought by increasing effort (the increase of product in stage 1), but also the indirect effect brought by it. With the increasing of a unit of a_1, the expectation of the individual's internal production capacity (θ) which from "Ideal self" will increase τ unit (Because $E(\theta\,/\pi_1) = (1-\tau)\bar{\theta} + \tau(a_1 + \theta + u_1 - \bar{a}_1)$). Thereby, the "amounts that turned over" (α_2) increase τ unit. Thus the net marginal income of a_1 is $1 - \tau$ (<1).

In this theoretical model, the individual's internal production capacity (θ) has large uncertainty, with the increasing of τ, the incentive loss is increasing. In other words, the "Ideal self" has not mobilized the enthusiasm of the work of "Actual self" through putting forward the "better and better" incentive constantly. Conversely, "actual self" takes all kinds of hidden initiatives to reduce their own effort, in order to avoid "Ideal self" putting forward a higher "standard" for itself. We can see from the derivations above, the ratcheting effect played a role because of the thinking model what "Ideal self" consist. Thus, in its own thinking model, "Ideal self" seems as putting forward higher demands to the "Actual self" constantly and expecting that it can take effort to do its best. However, the method what "Ideal self" takes reduces the effort of "Actual self" and makes "Actual self" lose of happiness through its effort, and then presents the behavioral fatigue and mental depression in fact.

28.3 Conclusion

Engineering practice education demands their students have certain qualities, such as being careful with each flaw, pursuit of 100 % correct and never being careless. Because of each engineering problem related to the safety of person's life directly, there is no engineering problem which is unimportant. High standards and strict requirements are the constant compass in engineering education field. However, if the way of thinking becomes extreme in some particular time, places or conditions, the thinking mode of negative perfectionism which has been discussed in this paper may take place. And it would erode the mental health of the individual seriously. As we all know, if the people who take part in the engineering discipline face mental problems or work with unhealthy attitude, it must be the maximal danger for engineering field.

In this paper as an example, there are some university students from engineering disciplines who has a personality trait of negative perfectionism. If these students consist the thinking modes below, for example, measuring the performance of "Actual self" only based on the academic rank and evaluating "Actual self" under rigid rules. Do not understand the sense that "Actual self" can grow a lot in some other sides through its hard working, though it may not arrive the height what "Ideal self" demanded. In this condition, the students are likely to become the people who do not know how to adjust the "standard" based on environmental changes flexibility once go to work. Especially for self-evaluation, at the moment of facing the challenges of life, the individuals should have correct assessments of the situation; develop reasonable advance goals and evaluating standards, and then pursue excellence with positive attitudes. However, the engineering disciplines students who have personality trait of negative perfectionism hold an inflexible and rigid evaluation system, instead of a flexible and dynamic evaluation system. If such individuals coexist in an organization, not only do their own the "Ideal self" and "Actual self" is difficult to reconcile, but also the interpersonal frictions between individuals may easy to occur, and it may cause the disharmony of the organization, decline in the efficiency of the organization directly.

Therefore, in this paper, we consider that, during the education and training of engineering disciplines, we should study the problems such as ability of engineering practice, innovation ability and leadership and so on. What's more, we should standing on psychological perspective, paying more attention to the personality of the engineering discipline students.

In summary, we must explore the ways how to improve content of courses and method of teaching in engineering education, in order that engineering discipline students can develop the excelsior and meticulous thinking habit when they face the engineering practice. At the same time, they should not bring the thinking habit into their ordinary lives. Executing rigid standard when it is needed, and not rigid when "flexible" is needed, adding more harmonious notes in live. Hope that more scholars can pay attention to the study of negative perfectionist personality which may lead to the mental problems of successor of engineering disciplines. I believe that, the study of this issue must be helpful for us to cultivate more and better engineers.

References

Bergman AJ, Nyland JE, Burns LR (2007) Correlates with perfectionism and the utility of a dual process model. Personal Individ Differ 43(2):389–399

Burns DD (1980) The perfectionists' script for self-defeat. Psychol Today 14(6):34–52

Dunkey DM, Zuroff DC, Blankstein KR (2003) Self-critical perfectionism and daily affect: dispositional and situational influences on stress and copying. J Pers Soc Psychol 84(4):234–252

Fang X, Qian M, Zi F (2007) Psychological study of perfectionism. Chin Ment Health J 21(3):208–210 (in Chinese)

Husserl E (1997) Husserl selections, vol 2. Shanghai Sanlian Bookstore, Shanghai, pp 1057–1058 (in Chinese)

Nahoko ITO (2004) The influence of perfectionism with inappropriate motivation on psychological maladjustment. J Jpn Clin Psychol 22(5):542–552 (in Japanese)

Pacht AR (1984) Reflections on perfectionism. Am Psychol 39(4):386–390

Slade PD, Owens RG (1998) A dual process model of perfectionism based on reinforcement theory. Behav Modif 22(3):372–390

Sun Y (2012) Flexibility of ideological and political education and students' psychological well-being. China Social Science Press, Beijing (in Chinese)

Sun Y, Feng C (2012) Discussion on negative perfectionism. Guangming Dly, 2012-6-12(011) (in Chinese)

Weisinger H, Lobsenz N (1981) Nobody's perfect. Warner Books, New York

Xie Y (2004) Research on self-oriented perfectionism tendencies of college students and their relationships with abnormal psychology. J Henan Norm Univ (Philos Soc Sci Ed) 31(4):186–188 (in Chinese)

Yang L, Zhou X, Ze X, Wang S (2004) Discussion about training modes for IE professionals. Ind Eng J 7(5):61–64 (in Chinese)

Zhang W (2003) Game theory and information economics. Shanghai People's Publishing House, Shanghai (in China)

Zhang B, Wang Y, He D, Cai T (2010) Perfectionism and cognitive-behavioral therapy. Med Philos (Humanist Soc Med Ed) 31(4):37–39 (in Chinese)

Zi F, Zhou X (2005) The relationship between students perfectionism and parental rearing behavior. Chin J Health Psychol 13(5):321–323 (in Chinese)

Chapter 29
An Empirical Study on the Formation of Inter-firm eBusiness Capabilities: The Moderating Effect of Formal Governance

Xiao-kuan Shen and Mao-mao Chi

Abstract Our research integrated the Internet-enabled Process View (IEPV) and the IT-izing process of business capability to explore the internal formation mechanism of inter-firm eBusiness capabilities. According to the enterprise survey data, we used SPSS18 and LISREL8.72 to verify the formation model of inter-firm eBusiness capabilities and found that the formation of the inter-firm eBusiness capabilities was the process of inter-firm eBusiness capability needs → inter-firm eBusiness resources → inter-firm eBusiness capabilities. We also revealed the moderating effect of formal governance in the process of inter-firm eBusiness resources → inter-firm eBusiness capabilities. These findings have important theoretical and practical implications when guiding enterprises to establish the inter-firm eBusiness capabilities in multiple organizations environment.

Keywords Formal governance • Formation mechanism • Inter-firm eBusiness capabilities • Moderating effect

29.1 Introduction

With the rapid development of information technology in the 21st, it is hard for a single firm to gain competitive advantage. More and more firms are trying to build strategic relationships with partners to gain the competitive advantage (Kohli and Grover 2008). Establishing an agile electronic value chain becomes the key of firm's success (Steerman 2003). Therefore, how to use IT to establish inter-firm eBusiness capabilities became the focus of attention.

X. Shen (✉) • M. Chi
School of Economics and Management, China University of Geosciences, Wuhan, China
e-mail: busyshen@163.com

E. Qi et al. (eds.), *The 19th International Conference on Industrial Engineering and Engineering Management*, DOI 10.1007/978-3-642-37270-4_29,

Today, many researchers have realized that the emphasis of studies about IT capabilities will transfer from single organization perspective to multiple organization perspective (Grover and Kohli 2008). However, previous study still focused on examining the role of IT capabilities from the utilization of IT resources, and made limited effort to understand the characters of inter-firm eBusiness capabilities. What's more, previous study didn't take into consideration the management factors' role in the formation process of inter-firm eBusiness capabilities.

Breaking through previous study paradigm, this study integrated the Internet-enabled Process View (IEPV) and the IT-izing process (Grover and Kohli 2008) of business capability to explore the formation mechanism of inter-firm eBusiness capabilities. First, we proposed the formation model of inter-firm eBusiness capabilities. Second, according to the enterprise survey data, we used LISREL to do the path analysis of the model and verified the path of inter-firm eBusiness capability needs (IFECN) → inter-firm eBusiness resources (IFER) → inter-firm eBusiness capabilities (IFEC). Third, we used SPSS to analyze the moderating effect of formal governance (FG) in the process of IFER → IFEC.

29.2 Theory and Hypothesis

29.2.1 Internet-Enabled Process View and IT-Izing Process of Business Capability

Internet-enabled Process View (IEPV), proposed by J. Zhao, means recognizing the value innovation activities from the role of internet-enabled process (Zhao et al. 2008). According to IEPV, internet-enabled process is defined as the integration of participants, network information flow and activities. Depending on the network information flow, internet-enabled process is participant-oriented, and it intent to meet people's needs for information interaction. IEPV regards eBusiness activities as the interaction process of network, information flow and participants (Zhao and Zhu 2010). Therefore, needs for the interaction of network information flow promote the generation of internet-enabled process. What's more, IEPV argues that eBusiness capabilities are rooted in the process. EBusiness capabilities are not easily transferred between firms (Soto-Acosta and Merono-Cerdan 2008). In multiple organizations environment, identifying and determining the needs of inter-firm capabilities become the primary consideration.

The concept of IT-izing was proposed by Kohli and Grover (2008). IT-izing process indicates that a firm identifies the needs of capabilities to execute a business imperative firstly and then understand how to use the appropriate information or technology to establish the capabilities (Kohli and Grover 2008). So firms must identify their inter-firm eBusiness capability needs then confirm which resources are required to establish these inter-firm eBusiness capabilities. So far, there isn't models and empirical research.

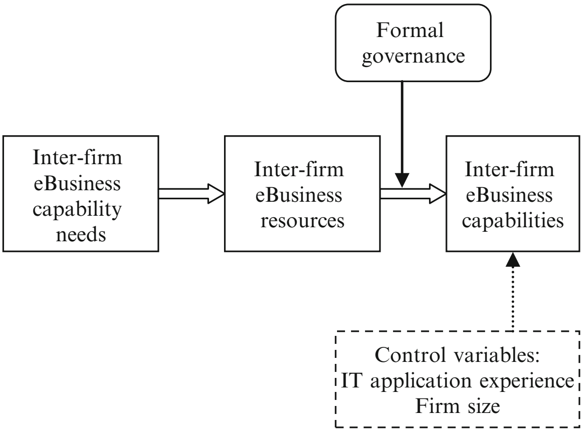


Fig. 29.1 Formation model of inter-firm eBusiness capabilities

Table 29.1 Explanation of variables

Variables	Explanations
IFECN	It is defined on the base of business functions of internet-enabled process (Sarkar et al. 2009; Dahui et al. 2010)
IFER	According to the characters of internet-enabled process, IFER are divided into IT resources and business resources
IFEC	Based on previous studies, we divided the IFEC into the capabilities of sharing information and cooperation process capabilities (Zhao and Zhu 2010)
FG	FG is a kind of coordination activity that uses regulations and contracts to regulate the cooperation activities among firms (Goo et al. 2009; Rai and Tang 2010; Hoetker and Mellewigt 2009)

IFECN inter-firm eBusiness capability needs, *IFER* inter-firm eBusiness resources, *IFEC* inter-firm eBusiness capabilities, *FG* formal governance

This study integrate the IEPV and the IT-izing process of business capability, consider the inter-firm eBusiness capability needs first and then confirm which resources are required to establish these inter-firm eBusiness capabilities. It breaks though previous RBV (resource-based view) study paradigm and develops new ideas for studies about the formation of inter-firm eBusiness capabilities.

29.2.2 B. Research Model and Theoretical Hypotheses

According to the characters of inter-firm eBusiness capabilities, we proposed the research model, and then determined the variables and items based on previous studies. Considering the role of management factors, this study introduced formal governance as moderator. In addition, we added two control variables (IT application experience, firm size). Figure 29.1 shows the research model. Table 29.1 shows the explanation of variables.

Based on Kohli's research, firms should identify their inter-firm eBusiness capability needs then confirm which resources are required to establish these inter-firm eBusiness capabilities. According to the characters of internet-enabled process, the better firms identify their inter-firm eBusiness capability needs, the better they establish inter-firm eBusiness resources (Grover and Kohli 2008). Thus:

Hypothesis 29.1 Inter-firm eBusiness capability needs play a positive role to the establishment of inter-firm eBusiness resources.

After establishing inter-firm eBusiness resources, firms will gain the material basis of constructing inter-firm eBusiness capabilities. Drawing on RBV, inter-firm eBusiness resources are heterogenous and hard to imitate, these characters will help firms to construct inter-firm eBusiness capabilities (Bharadwaj 2000). Thus the hypothesis:

Hypothesis 29.2 The establishment of inter-firm eBusiness resources plays a positive role to the formation of inter-firm eBusiness capabilities.

Formal governance is a kind of eBusiness management and coordination activity that uses contracts and roles to regulate their cooperation (Goo et al. 2009; Rai and Tang 2010; Hoetker and Mellewigt 2009). A large number of practices show that high formal governance level will help firms to construct inter-firm eBusiness capabilities. Thus:

Hypothesis 29.3 Organizations with a higher formal governance level will construct inter-firm eBusiness capabilities more easily.

29.3 Research Methodology

According to the theoretical model, we designed the survey questionnaire for inter-firm eBusiness capabilities. It adopted the common form of Likert 5 points scale table. For example, selecting "1" stands for "very disagree" and selecting "5" stands for "strongly agree". Table 29.2 shows the variables and items. For control variables, we used the length of time of IT application to measure the IT application experience and used the number of regular employees to measure the firm size.

The objects of this investigation are Chinese manufacturing firms. With the help of CECA, we adopted typical sampling method to conduct a questionnaire survey. We sent 600 questionnaires to firms with a higher degree of informatization through E-mails and interviews. All questionnaires were filled in by managers or CIOs. There were 240 questionnaires returned. Finally, we got 224 valid responses that accounted for 93 % of our returned responses. What's more, we tested the bias of managers and IS managers' answers, the result showed that there is no obvious answer deviation ($p > 0.05$).

Table 29.2 Variables and items of the scale table

Variables	Items
IFECN	Planning the strategy of eBusiness (D1), managers' value-conscious of IFECN (D2), making IT resources planning (D3) (Sarkar et al. 2009; Dahui et al. 2010)
IEER	Inter-firm system integration (R1), transformation of internal processes (R2), system integration between departments (R3) and long-term relationship with partners (R4) (Zhao and Zhu 2010)
IFEC	Sharing dynamic information of market with partners (C1), doing online business with partners (C2), online order activities (C3), sharing needs change online (C4), tracking orders (C5) (Zhao and Zhu 2010)
FG	Planning the implementation program of eBusiness cooperation (G1), establishing the cooperation mechanism to encourage partners to join the distribution channels (G2), reaching an agreement with partners to respond to market changes together (G3) (Goo et al. 2009; Rai and Tang 2010; Hoetker and Mellewigt 2009)

IFECN inter-firm eBusiness capability needs, *IFER* inter-firm eBusiness resources, *IFEC* inter-firm eBusiness capabilities, *FG* formal governance

29.4 Data Analysis

29.4.1 Analysis of Validity and Reliability

Before the empirical research, we analyzed variables' validity. First, we used SPSS to do the Exploratory Factor Analysis (EFA) to extract four variables (IFECN, IFER, IFEC, FG). The KMO statistic was 0.951 and the Bartlett test passed the examining at the significance level of 0.001. Therefore, the factor structure was clear, and it met the basic standards of validity preliminarily. Second, we used LISREL to do the Confirmatory Factor Analysis (CFA) to examine the restrained validity and discriminant validity (Fornell and Larcker 1981). We found all items' factor loadings of questionnaire were bigger than 0.72 and significant. So the restrained validity met the study requirement. We analyzed the discriminant validity through the average variance extracted (AVE) values. The results showed that the questionnaire had good discriminant validity (W C W 1998).

We used SPSS to examine the reliability of the questionnaire and found all the Cronbach's α were between 0.84 and 0.93 and all Composite Reliability (CR) values were bigger than 0.84. The analysis above indicated that scale table's reliability achieved a high level. Table 29.3 summarizes the results of analysis of validity and reliability.

29.4.2 Model Verification

In this study, we used modeling method to verify model's paths. In model's path analysis, path coefficients can reflect the influence degree between latent

Table 29.3 Reliability analysis and AVE

	IFECD	IFER	IFEC	FG
Correlation matrix and square roots of AVE				
IFECD	**0.807**			
IFER	0.588	**0.857**		
IFEC	0.715	0.754	**0.857**	
FG	0.616	0.743	0.731	**0.867**
Reliability analysis				
Cronbach's α	0.843	0.915	0.933	0.921
CR	0.848	0.917	0.932	0.923

IFECN inter-firm eBusiness capability needs, *IFER* inter-firm eBusiness resources, *IFEC* inter-firm eBusiness capabilities, *FG* formal governance

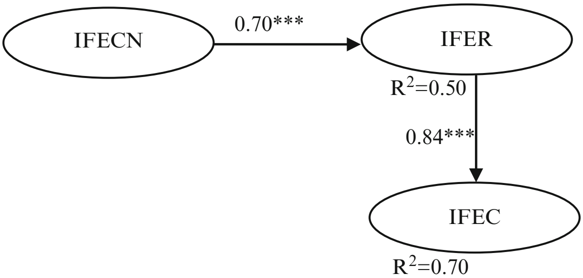


Fig. 29.2 Path diagram of LISREL structural equation. All tests are two-tailed *$p < 0.1$; **$p < 0.05$; ***$p < 0.001$

variables, and R^2 value indicates the influence degree of independent variables on the dependent variable. We used LISREL to analyze the model's path to get path coefficients and R^2 values and found that all path coefficients were significant ($p < 0.001$). Therefore, Hypothesis 29.1 and Hypothesis 29.2 were supported. Figure 29.2 shows the results of path analysis. In addition, we used SPSS to examine the control variables and found that firm size had no effect on the formation of inter-firm eBusiness capabilities. While IT application experience's impact was significant. So it should be controlled in further studies.

Based on the path analysis, we used SPSS to analyze the moderating effect of formal governance. We computed the mean values of four variables, and every mean value was centered. Then we got an interaction item (IFER × FG). Using the OLS linear regression of SPSS, we adopted step-by-step analysis method (Chin et al. 2003) to examine the relationships between dependent variable, independent variable and moderator. Firstly, we put the independent variable into the regression model (Model 1). Secondly, we put the moderator into the regression model (Model 2). Finally, we added the interaction item in the model (Model 3). Table 29.4 summarizes the analysis results.

It could be seen from Table 29.4 that every R^2 value was bigger than 0.56. It indicated that the explanatory variables had a high explanation degree on dependent variable. The fact that all D-W test values were close to 2 indicated that among

Table 29.4 Analysis of the moderating effect

	Model 1 (β)	Model 2 (β)	Model 3 (β)
Constants	0.000	0.000	0.000
IFER	0.754***	0.469***	0.491***
FG		0.383***	0.379***
IFER × FG			0.079*
Range of VIF	1.000	1.064	1.063–2.232
Durbin-Watson	1.936	1.958	1.952
R^2	0.568	0.634	0.640
F test	291.732***	191.071***	130.095***

All tests are two-tailed **p* <0.1; ***p* <0.05; ****p* <0.001

residuals of explanatory variables there was no autocorrelation. VIF values were between 1.0 and 2.3, and it indicated that there was no multicollinearity problem. In summary, regression models had a good overall effect.

Particularly, after we added the interaction item in the model, we found it had a significant effect on inter-firm eBusiness capabilities and R^2 value was changed from 0.634 to 0.640. It indicated that the moderating effect of formal governance was effective. As a result, Hypothesis 29.3 was supported.

29.5 Discussion and Conclusion

We used LESREL to do the path analysis and found that two path coefficients were significant ($p < 0.05$) and their values were 0.70 and 0.84 respectively. All R^2 values were bigger than 0.5 and the Goodness of Fit (GoF) was 0.9. So the model had a high explanation level and a good GoF. This study showed that the driven role of inter-firm eBusiness capability needs was divided into two parts. First, inter-firm eBusiness capability needs drove the establishment of inter-firm eBusiness resources. Second, the construction of inter-firm eBusiness resources promoted the formation of inter-firm eBusiness capabilities. So firms should identify their inter-firm eBusiness capability needs and determine how to use inter-firm eBusiness resources to construct inter-firm eBusiness capabilities. We used SPSS to analyze the moderating effect of formal governance and found the path coefficient of the interaction item was 0.079 ($p < 0.1$). The results showed that firms with higher formal governance level would construct inter-firm eBusiness capabilities more efficient. Therefore, to promote the formation of inter-firm eBusiness capabilities, firms should constitute some regulations and contracts to improve the formal governance level before they cooperate with their partners.

Our research integrated the Internet-enabled Process View (IEPV) and the IT-izing process of business capability and proposed the research model. Based on the enterprise survey data, we verified the model and revealed the internal formation mechanism of inter-firm eBusiness capabilities. We analyzed the relationships among inter-firm eBusiness capability needs, inter-firm eBusiness resources and inter-firm eBusiness capabilities. What's more, we examined the moderating effect of formal governance. In this study, eBusiness capabilities, which used to be studied in single organization environment, were expanded to multiple organizations environment. It provided a new idea for studies about inter-firm eBusiness capabilities.

References

Bharadwaj AS (2000) A resource-based perspective on information technology capability and on firm performance: an empirical investigation. MIS Q 24(1):169–196
Chin WW, Marcolin BL, Newsted PR (2003) A partial least squares latent variable modeling approach for measuring interaction effects: results from a Monte Carlo simulation study and an electronic-mail emotion/adoption study. Inf Syst Res 14(2):189–217
Dahui L, Chau PYK, Lai F (2010) Market orientation, ownership type, and E-Business assimilation: evidence from Chinese firms. Decis Sci 41(1):115–145
Fornell CR, Larcker DF (1981) Structural equation models with unobservable variables and measurement error. J Market Res 18(3):55–67
Goo J, Kishore R, Rao HR et al (2009) The role of service level agreements in relational management of information technology outsourcing: an empirical study. MIS Q 33(1): 119–145
Grover V, Kohli R (2008) Cocreating IT value: new capabilities and metrics for multi-firm environments. MIS quarterly, special issue-call for paper
Hoetker G, Mellewigt T (2009) Choice and performance of governance: matching alliance governance to asset type. Strateg Manag J 30(10):1025–1044
Kohli R, Grover V (2008) Business value of IT: an essay on expanding research directions to keep up with the times. J Assoc Inf Syst 9(1):23–39
Rai A, Tang X (2010) Leveraging IT capabilities and competitive process capabilities for the management of inter organizational relationship portfolios. Inf Syst Res 21(3):516–542
Sarkar M, Aulakh PS, Madhok A (2009) Process capabilities and value generation in alliance portfolios. Organ Sci 20(3):583–600
Soto-Acosta P, Merono-Cerdan AL (2008) Analyzing e-business value creation from a resource-based perspective. Int J Inf Manag 28(1):49–60
Steerman H (2003) A practical look at CPFR: the Sears-Michelin experience. Supply Chain Manag Rev 7(4):46–53
W C W (1998) Issues and opinion on structural equation modeling. MIS Q 22(1):vii–xvi
Zhao J, Zhu Z (2010) eBusiness value creation process model. J Manag Sci China 13(12):47–60 (in Chinese)
Zhao J, Huang WV, Zhu Z (2008) An empirical investigation of E-Business implementation process in China. IEEE Trans Eng Manag 55(1):54–65

Chapter 30
A Driven Model of IT Usage: Determinants and Moderating Effects of Situational Variables

Rong-jie Lv, Li-jun He, Xiao-chun Chen, and Zhao Zhao

Abstract Prior research has provided valuable insight into how and why individual make a decision on the acceptance and use of information technologies (ITs) in the organizations. In practice, however, the diversity of IT and managerial task make these theoretical models not always work. To reinforce understanding about the drive mechanism of users' IT usage, we draw from the representative researches on the technology acceptance model (TAM), and: (i) review and develop an integrated model of determinants of individual level IT usage from both cognitive and emotional perspectives; (ii) discuss the moderating effects of experience, commitment to use, task complexity, network externalities and instrumental on the relationships between determinants and user behavior.

Keywords Cognitive factor • Emotional factor • Experience • Instrumental • IT usage • Network externalities

30.1 Introduction

The presence of information technologies (ITs) in today's organizations has expanded dramatically. From 2000 to 2004, the investment worldwide grew by 5.1 % year-on-year, and the annual growth rate had been increased to 7.7 % in next 4 years (Venkatesh and Bala 2008). Yet, for technologies to improve productivity, they must be accepted and used by employees in organizations. How we can

---

This research is supported by the Natural Science Foundation (G2011202064, G2011202154) and Social Science Foundation (HB12GL050) of Hebei Province.

R. Lv • L. He • X. Chen (✉) • Z. Zhao
School of Economics and Management, Hebei University of Technology, Tianjin, China
e-mail: chenxc2009@126.com

E. Qi et al. (eds.), *The 19th International Conference on Industrial Engineering and Engineering Management*, DOI 10.1007/978-3-642-37270-4_30,

accurately explain user acceptance or usage of new technology is an important and long-standing research question in information systems (DeLone and McLean 1992). Researches in this area have developed serial theoretical models, based on information systems, psychology, as well as social and behavioral sciences. And the dominant paradigm is Davis' technology acceptance model (TAM), which posits that user's behavior can be explained by two beliefs: perceived usefulness and perceived ease of use. Later, an important supplement to this model refers to the theory of planned behavior (TPB), which indicates that human behavior is guided by three kinds of considerations: attitude, subjective norm and perceived behavioral control. Integrated the two theories, Taylor and Todd (1995) suggest the two beliefs – perceived usefulness and perceived ease of use – produce a favorable or unfavorable attitude toward usage. Then in combination, attitude, subjective norm and perception of behavioral control drive to the formation of usage intention.

This integrative model enhances explanatory power of users' IT usage, and has been widely used to predict users' acceptance of IT. Even so, it is still regrettable that the related models just consider cognitive factors driving behavior, little attention is paid to emotional and situational factors. Thus, in order to understanding the driving mechanism of IT usage behavior better, there is a need for a review and synthesis from the perspectives of cognitive factors, emotional factors, as well as situation such as task and technical characteristics.

30.2 Theoretical Models of IT Usage and Literature Review

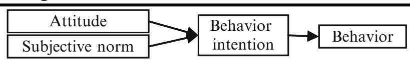
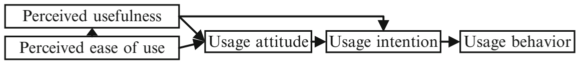
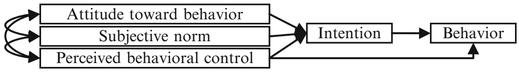
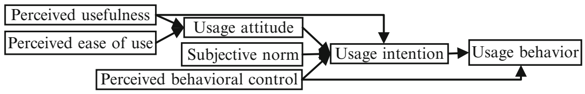
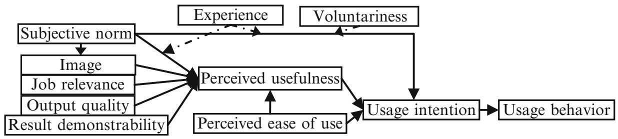
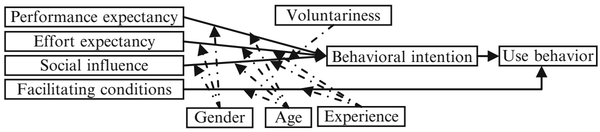
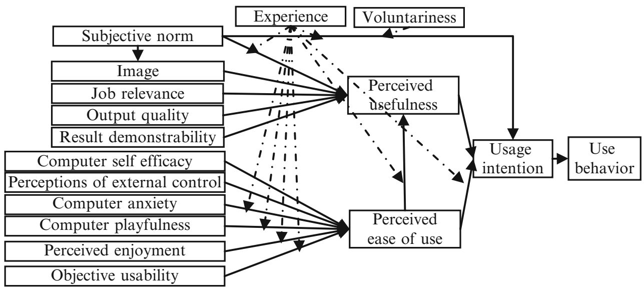
IT adoption research on usage from the perspective individual-level has been an important stream in the management information system area. Vast bodies of researches on IT acceptance and usage exist, with some research are about the evolution of acceptance over time. In this section, we review, comparison and synthesis the theoretical models which define usage intention or usage behavior as the key dependent variable. Table 30.1 provides a brief overview of eight dominating theoretical models.

The same purpose of these models is to understand IT usage as dependent variable. As a predictor, behavioral intention is critical to the behavior of IT usage, and has been well-established in IS (Taylor and Todd 1995; Ajzen 2002). All of them are based on the theoretical constructs of TAM, with extended in the three perspectives depicted in Fig. 30.1.

30.2.1 Other Cognitive Variable

Borrowing from theory of social psychology and behavioral science, researchers define additional several variables that have not been mentioned in TAM to explain the usage of new IT (e. g., perceived behavioral control, subjective norm, facilitating conditions, and self-efficacy).

Table 30.1 Theory model of IT acceptance

Theory	Diagram of model
Theory of reasoned action (Fishbein and Ajzen 1975)	
Technology acceptance model (Davis 1989)	
Theory of planned behavior (Ajzen 1991)	
Combined TAM and TPB (Taylor and Todd 1995)	
Technology acceptance model 2 (Venkatesh and Davis 2000)	
Unified theory of acceptance and use of technology (UTAUT) (Venkatesh et al. 2003)	
Technology acceptance model 3 (Venkatesh and Bala 2008)	

The TAM (Table 30.1) is an adaptation of the theory of Reasoned Action (TRA) proposed by Fishbein and Ajzen, which specifies perceived usefulness and ease of use as determinants of IT usage (Davis et al. 1989). And usage intentions are the sole direct determinant of usage behavior. In other words, the formation of intentions is identified as a necessary precursor to behavior (Fishbein and Ajzen 1975). TAM can be considered a special case of TRA, and the appeal is that it suggests a small number of factors which are specific, easy to understand and can be manipulated through system design and implementation to explain for IT usage jointly. Nevertheless, it excludes the influence of social, as well as personal control factors on behavior, which have been found having significant impacts on IT acceptance (Hartwick and Barki 1994), also emphasized in TPB. Thus Taylor and Todd (1995) suggest the integration of TAM and TPB might be expected to increase understanding of users' behavior.

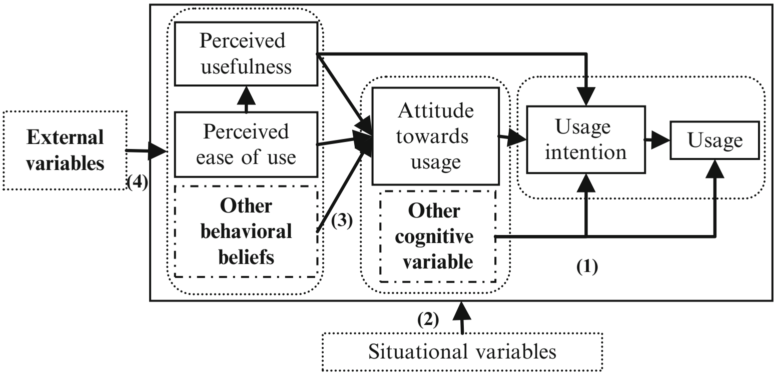


Fig. 30.1 Perspectives of extending towards TAM

The TPB accounts for situations where individuals do not complete control over the behavior in their will. It asserts that behavior does not only depend on behavioral intention but also is directly affected by perceived behavioral control, and that each of the determinants of the behavioral intention (i.e., attitude, subjective norm and perceived behavioral control) (Ajzen 2002), in turn, is respectively determined by underlying attitudinal beliefs, normative beliefs and control beliefs.

With the support of TPB, Taylor and Todd (1995) consider attitude towards IT usage, subjective norm and perceived behavioral control as key factors influencing users' behaviors. Usage attitude reflects individual positive or negative evaluation about IT usage. It is equated with attitudinal belief that performing a behavior will lead to a particular outcome, weighted by an evaluation of the desirability of that outcome. The attitude would be positive if the user believes that using IT will result in improving task performance and this outcome will come true in all probability.

Subjective norm is formed as user's normative belief concerning a particular reference weighted by the motivation to comply with that reference. For instance, a user may believe that others think that he should use the technology, but complying with those people's wishes is relatively unimportant, thus, the subjective norm is weak.

According to TPB, perceived control over the usage behavior reflects the extent to which user perceives the resources and opportunities needed to IT use, or perhaps extent to the internal and external factors that may impede perform using that technology. The two components similar to Trandis's facilitating conditions and Bandura's self-efficacy are included in this concept. Perceived control over IT usage is formed as the sum of control beliefs weighted by the perceived facilitation of the control belief that may facilitate or inhibit performance of using. For example, one may perceive that he does not have capabilities to use IT and those capabilities are critical to determining IT usage.

30.2.2 Situational Factors

A variety of models that discuss antecedents of user behavior have been advanced to explain IT usage at individual-level, of which the TAM and its extended model (Taylor and Todd 1995) are most well-known. Goals of such models are to develop diagnostic tools to predict IT usage and facilitate design changes before user involves into its implementation. Though studies later show these models have reasonable explanatory power generally, tests of the relationships in the models have not produced consistent results. Typical results showed perceived usefulness as direct determinant of usage intention. Moreover, the roles of subjective norm and perceived ease of use determining IT usage are somewhat unclear, and to a large extent mediated by perceived usefulness.

To address searching for the reason of such phenomenon and strengthening the suitability of models, a number of studies (Venkatesh and Bala 2008; Venkatesh and Davis 2000; Venkatesh et al. 2003) have paid attention to identify situational factors that moderate the impacts of antecedents on IT usage (as depicted in Fig. 30.1). Extant literatures on conditional relationship in IT acceptance or usage models are generally from the point of view of the users involved in IT implementation, such as gender, age, prior experience and voluntariness of use. Above all, experience and voluntariness are most mentioned, and will be reviewed as following.

30.2.2.1 Experience

How might the effects of those antecedents change with increasing experience of using a target IT? Existing theories and evidences suggest that the direct effect of subjective norm on usage intentions may weaken with increased experience on IT using (Venkatesh and Bala 2008; Venkatesh and Davis 2000). Hartwick and Barki (1994) explain that when users' knowledge and beliefs about IT are ill-informed, they may rely more on the opinions of significant others as reference to form intentions. Conversely, as they become informed, the direct effect of normative influence would become insignificant. This conclusion is supported by Agarwal and Prasad. Venkatesh and Davis (2000) carry out a longitudinal study to analyze experience in terms of the role of subjective norm caused by the difference. Their conclusion shown as TAM2 in Table 30.1 also emphasizes that both the subjective norm-intention and subjective norm-perceived usefulness relationships are significantly moderated by experience. Additionally, according to recent researches (Venkatesh and Bala 2008; Venkatesh et al. 2003), the experience factors might have influence on the role of perceived ease of use.

30.2.2.2 Voluntariness

Voluntariness is also supposed as a key moderator in IT acceptance process. As already mentioned, the contingency underlying the mixed findings regarding subjective norm are identified by Hartwick and Barki (1994). After controlling experience factor, they find that subjective norm have more significant impact on intention in mandatory situation than in voluntary situation. Venkatesh and Davis (2000) refer to the causal mechanism underlying the effect as compliance: the direct compliance effect of subjective norm on intention produces while individual perceives that others want him/her to carry out a target behavior, either the others are significant or has the ability to reward the behavior/punish non-behavior (French and Raven 1975). To validate the difference between mandatory and voluntary usage situations, Venkatesh and Davis construct TAM2 which supposes voluntariness, the extent to which potential adopters perceive the use decision to be non-mandatory, as another moderating variable.

30.3 Theoretical Framework and Propositions Development

In this section, we propose a theoretical framework and develop propositions. The roles of determinants incorporating cognitive and emotional factors are first discussed, followed by presenting the moderating effects of situational variables related to user, task and technology characteristics.

30.3.1 Roles of Cognitive and Emotional Factors

It has already been mentioned that C-TAM-TPB model (see Table 30.1) involving three key drivers (i.e., attitude, social influence and behavioral control) and two behavior beliefs forming attitude are believed to be more comprehensive to interpret the IT usage behavior so far.

But the premise of TPB is that production of individual behavior is a rational decision making passing from beliefs to beliefs' evaluation, then, forming intentions which finally drive behavior (Venkatesh et al. 2003). Individual behavior is belief-based cognitive and deliberate process. Based on this theory, the pervasive models on user IT acceptance, not excepting C-TAM-TPB, only take cognitive factors into account as well, with no attention on emotional factors.

Actually, individual behaviors have relation to the emotional experience on whether the realization of target behavior meeting needs or not (Fredrickson 1998), and each emotion has its behavioral cues (Lazarus 1991). In practical terms, individual positive emotions would impel him/her to overcome restrictions to show helpful behaviors or form initiative behavioral intention (Fredrickson 1998). Therefore, from the perspective of the theory affective experience (Seo et al. 2010), this research will introduce user's effective commitment to usage as typical emotional factor to explain the motivation (internal driving force) of IT usage. The essence of the effective commitment is emotional element initially used to express employee's positive emotional attachment to the organization, and viewed

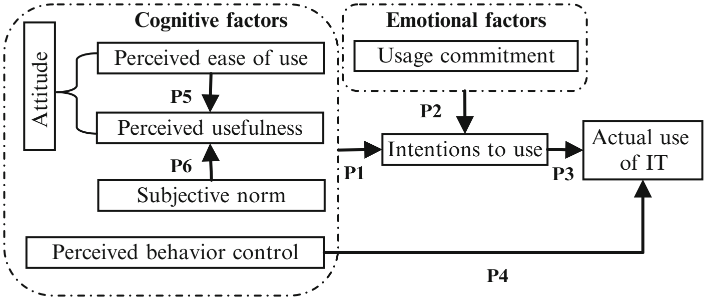


Fig. 30.2 The model of cognitive and emotional driven of IT usage

as individual identifies with the goals and values of the organization, as well as affective experience formed by such identification (Meyer and Allen 1991). Then it is extended to define emotional attachment to a given job or career (Li-rong et al. 2000). In this way, usage commitment presented here indicates user's emotional attachment to IT usage.

Previous research has found individual affective commitment has a positive impact on citizenship behavior or extra-role behavior (Meyer et al. 2006). As an internal driving force making individual drawn into a series of behaviors, affective commitment most likely affects the intentions to a certain behavior (Zhen-jiang and Ji-ping 2007).

All the above suggests our research should been built on the assumption that both cognitive and emotional factors would be determinants of IT usage behavior. So we have a set of propositions (see *P1–P4* in Fig. 30.2):

P1: Cognitive factors including attitude toward usage, subjective norm and perceived behavioral control lead to the formation of intention to IT use.
P2: Commitment to IT usage as typical emotional factor also drives actual use of IT.
P3: Intention to IT use is an immediate antecedent of actual use of IT.
P4: Perceived behavioral control not only affects actual use of IT indirectly through behavioral intention, but also affects it directly.

According to attitude-behavior relationship theories, users may have a large number of behavioral beliefs relate to IT usage before deciding whether to use IT or not. Behavioral (usage) beliefs are the foundation of attitude which refers to user's beliefs that using IT leads to certain outcomes and his or her evaluations the probability of attaining these outcomes. But for a given situation, only dominating usage beliefs can be captured as salient beliefs. That is to say, salient usage beliefs about IT usage build the overall attitudes towards IT use. Using the method of questionnaire, we elicit salient beliefs related to IT usage.

After encoding and classifying the response from 54 users, we obtain the two highest frequency considerations -"whether it can help me improving performance or not" and "how easy it will be to use"-respectively reflect the perceived usefulness and perceived ease of use presented in TAM. We suppose the two considerations form user as salient usage beliefs to build attitude. In order to detect respective effects of the two significant beliefs, we propose using perceived usefulness and perceived ease of use as cognitive factors instead of attitude (see Fig. 30.2). And referring to the suggestion of previous research, we deduce the following two propositions:

P5: Perceived ease of use not only determines intention to use directly, but also affect it through perceived usefulness.

P6: Similarly, as another cognitive factor to determine intention, subjective norm have indirect effect passed by perceived usefulness as well.

30.3.2 Effects of Users, Task and Technology Characteristics

For IT usage in individual-level, user is the subject using technology to complete task, technology is recipient of behavior (usage), and that task is closely related to the function of usage. So attentions should be paid on characteristics of all the three elements when we expect to explain user's behavior in terms of situational factors.

30.3.2.1 User Characteristics—Experience and Commitment

To explain the inconformity of determinants and its effects on IT usage, extant studies have generally emphasized the interaction of user's prior experience, voluntariness, as well as demographic (see Table 30.1). In this research, we continue the discussion of the interactive effect of experience and voluntariness.

Davis (1989) early notices the impact of perceived ease of use on attentions to IT use varies from different experience. He and his colleagues conduct series studies to analyze and verify the interaction mechanism of experience on IT usage (Venkatesh and Bala 2008; Venkatesh et al. 2003). The conclusion is also proven in our interview study as follow: "as IT implementation goes on, users have to accustom themselves gradually to it and accumulate useful experience increasingly, and then not much for them to consider whether technology is easy to use or not".

Beyond that, in the interviews, we are told: "In the beginning, what I could do was listening to advices from experienced others and reference, consult or emulate their practices, since I know little about the new technology and was not sure of my ability to cope with problems during usage". We find that veteran users mainly rely on his own sense of what's usefulness or whether usage or not, rather than looking for affirmation from others. Thus, there is proposition as following (see Fig. 30.3):

P7: Increasing hands-on experience with IT, a user will have more procedural knowledge on how to use technology, while perceived ease of use as well as others' influences (subjective norm) may not be as important determinants in forming behavior intention.

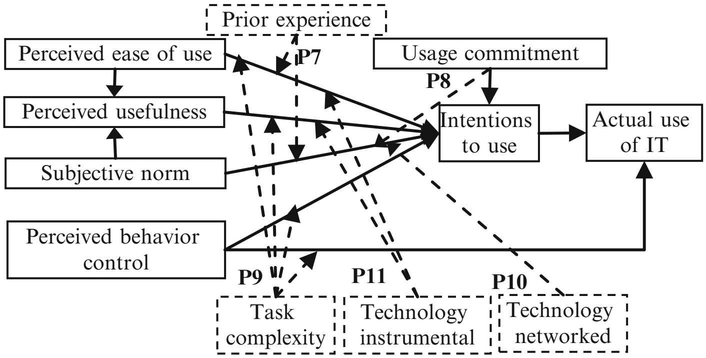


Fig. 30.3 The model of cognitive and emotional driven of IT usage

Then we discuss the effects of user's effective commitment in our research model. In fact, we already know the characteristic of voluntariness has been confirmed to lower the relationship between subjective norm and usage intentions (Venkatesh and Davis 2000). Specifically, in comparison with mandatory, if user volunteers to use IT, he will pay less attention to others. Theoretically, voluntariness itself is thought to be principal factor in the development of individual's commitment (Salancik 1977). Indeed, on a voluntary basis, individual will come to be committed to the behavior before taking action (Wei-peng and Kan 2004). Therefore, we suppose the following proposition:

P8: The positive impact of subjective norm on intentions to use will be weakened by user's effective commitment to IT usage.

30.3.2.2 Task Characteristics—Complexity

Previous research has shown that transfer of factors related to IT usage considered by users from one task situation to another (Xiaowen et al. 2006). Thus, user acceptance model constructed to specific task situation does not apply in all case (Karahanna and Straub 1999). Task complexity, in particular, has been often mentioned as situational variable in researches of task-technology fit (Goodhue and Thompson 1995; Sun 2010).

While facing ambiguity question, users may think more about the usefulness of information related to decisions (Choo 1996). It seems that user focus far more on usefulness of method and tool, and pay less attention on its ease of use. Having compared four different types of IT application, Deng et al. find that in database application, the effect of perceived usefulness is the most significant while the effect of perceived ease of use is the most insignificant. Given this conclusion, Sun (2010) explains that the task of database application is complicated, and it highlights the role of perceived usefulness.

Another result has shown that if task complexity is higher, training based on behavior modeling would be more efficient than on lecturing, because the former could bring more resources and technology convenience (Bolt et al. 2001). Moreover, Bolt finds whatever training method is taken; the self-efficacy would result in bigger drive for IT usage in higher task complexity situation. Taken together, the conditions of resource and convenience, as well as self-efficacy are essential of perceived control (Ajzen 2002). Thus, we present corresponding proposition:

P9: Task complexity strengthens the impacts of perceived usefulness, perceived behavior control on intentions, while lessens the effect of perceived ease of use on intentions.

30.3.2.3 Technology Characteristics—Network Externalities and Instrumental

As far as technology situation factors, the characteristic of network externalities and instrumental are involved in this paper.

Firstly, some studies existing concern the role of perceived user's perception of usefulness, ease of use and social influences on adoption of IT with network externalities. It concludes that user perceptions concerned with network externalities have a positive impact on use of electronic communication systems (Strader et al. 2007). Social norm, in particular, is an important antecedent for the acceptance of technology with network externalities (Dickinger et al. 2008).

ITs can be classified into two categories: networked and non-networked, in terms of the characteristic of network externalities. The former is used to promote communication and cooperation. And to what degree the users accept highly depends on whether others use or not, or information from third-party (Strader et al. 2007). As a result, social influences, such as those significant or experienced colleagues, will play more important roles in networked ITs situations. Thus, the proposition related to network externalities is proposed as follow:

P10: The effect of social norm on intentions to IT use will be strengthened with network externalities.

Secondly, according to Van der and Heijden's idea, ITs can be divided into utilitarian and hedonic. Utilitarian technology aims to provide instrumental value to the user, in contrast to hedonic technology, which aims to provide self-fulfilling value to user. Basing on this, there is a study (Amoako and Salam 2004) confirming that the perceived ease of use has double impetus function than perceived usefulness in the situation of using hedonic technology. After all, instrumentality implies there is an objective external to interaction between user and technology, so the purpose for user to use those ITs is increasing task performance. Naturally, for this purpose, users will think highly of quality, functionality, fitness, and so on, which reflects the usefulness of IT itself. Yet, the value of hedonic IT is a function of the degree to which the user experiences fun when using it. Therefore, the importance of hedonic content, animated images, esthetically appealing visual layouts, as well as ease of use will be emphasized by users. So together, we point out the proposition as:

P11: IT instrumental has a positive influence on perceived usefulness-intentions relationship, yet, has a negative impact on perceived ease to use-intentions relationship.

30.4 Conclusions

ITs are becoming increasing complex and implementation costs are very high. But low acceptance and high underutilization have become a major problem for organizations in term of realizing the benefits of IT implementations. This research aims at improving understanding of the formation of IT acceptance and usage. Prior researches explain such a question from a cognitive behavioral perspective, and propose experience and voluntariness as situational factors moderating the adoption of IT. Our research supports the existing cognitive factors and their general roles on IT acceptance, including perceived ease to use, perceived usefulness, subjective norm and perceived behavioral control.

However, the current models are incapable of answering why some users' behaviors are driven not by planned reason, but by emotions. Therefore, as an important supplement to existing literatures, we analyze and present users' affective commitment to IT usage as emotional factor driving their behaviors. Furthermore, our research pays attentions on characteristic of all of user, task and technology, and discusses the moderating effects of experience, commitment to use, task complexity, network externalities and instrumental on the relationships between determinants and user behavior. And future empirical research to test the presented model and theory propositions would be a reasonable next step.

References

Ajzen I (1991) The theory of planned behavior. Organ Behav Hum Decis Process 50(2):179

Ajzen I (2002) Perceived behavioral control, self-efficacy, locus of control, and the theory of planned behavior. J Appl Soc Psychol 30(4):665–683

Amoako K, Salam A (2004) An extension of the technology acceptance model in an ERP implementation environment. Inf Manag 41(6):731–745

Bolt MA, Killough LN, Koh WC (2001) Testing the interaction effects of task complexity in computer training using the social cognitive model. Decis Sci 32(1):1–20

Choo CW (1996) How organizations use information to construct meaning, create knowledge and make decisions. Int J Inf Manag 16(5):329–340

Davis FD (1989) Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q 13(3):319–340

Davis FD, Bagozzi RP, Warshaw PR (1989) User acceptance of computer technology. Manag Sci 35(8):982–1003

DeLone WH, McLean ER (1992) Information systems success: the quest for the dependent variable. Inf Syst Res 3(1):60–95

Dickinger A, Arami M, Meyer D (2008) The role of perceived enjoyment and social norm in the adoption of technology with network externalities. Eur J Inf Syst 17(1):4–11

Fishbein M, Ajzen I (1975) Belief, attitude, intention, and behavior. Addison-Wesley, New Jersey

Fredrickson BL (1998) What good are positive emotions. Rev Gen Psychol 2(3):300–319

French J, Raven BH (1975) The bases of social power. In: Cartwright D (ed) Studies of social power. Institute for Social Research, Ann Arbor

Goodhue DL, Thompson RL (1995) Task-technology fit and individual performance. MIS Q 19(2):213–236

Hartwick J, Barki H (1994) Explaining the role of user participation in information system use. Manag Sci 40(4):440–465

Hu Wei-peng, Shi Kan (2004) A review of the research about organizational commitment. Adv Psychol Sci 12(1):103–110 (in Chinese)

Karahanna E, Straub DW (1999) The psychological origins of perceived usefulness and ease-of-use. Inf Manag 35(4):237–250

Lazarus RS (1991) Emotion and adaptation. In: Pervin LA (ed) Handbook of personality: theory and research. Oxford University Press, New York, pp 609–637

Long Li-rong, Fang Li-luo, Ling Wen-quan, Li Ye (2000) Theory and measurement of career commitment. J Dev Psychol 8(4):39–45 (in Chinese)

Meyer JP, Allen NJ (1991) A three-component conceptualisation of organizational commitment. Hum Resour Manag Rev 1(1):61–89

Meyer JP, Becker TE, Van Dick R (2006) Social identities and commitments at work: toward an integrative model. J Organ Behav 27(5):665–683

Qi Zhen-jiang, Zhu Ji-ping (2007) The theory of organizational commitment and its research development. J Zhejiang Univ 37(6):90–98 (in Chinese)

Salancik GR (1977) Commitment is too easy. Organ Dyn 6(1):62–80

Seo M-G, Bartunek JM, Barrett LF (2010) The role of affective experience in work motivation: test of a conceptual model. J Organ Behav 31(7):951–968

Strader TJ, Ramaswami SN, Houle PA (2007) Perceived network externalities and communication technology acceptance. Eur J Inf Syst 16:54–65

Sun Yuan (2010) The unified technology acceptance model based on task-technology fit theory. PhD dissertation, Zhejiang University, Hangzhou (in Chinese)

Taylor S, Todd PA (1995) Understanding information technology usage: a test of competing models. Inf Syst Res 6(2):144–176

Venkatesh V, Bala H (2008) Technology acceptance model 3 and a research agenda on interventions. Decis Sci 39(2):273–315

Venkatesh V, Davis FD (2000) A theoretical extension of the technology acceptance model. Manag Sci 46(2):186–204

Venkatesh V, Morris MG, Davis GB, Davis FD (2003) User acceptance of information technology. MIS Q 27(3):425–478

Xiaowen F, Chan S, Brzezinski J, Shuang XU (2006) Moderating effects of task type on wireless technology acceptance. J Manag Inf Syst 22(3):123–157 (in Chinese)

Chapter 31
A Semantic Ontology of Requirement – Product – Process – Resource for Modeling of Product Lifecycle Information

Biao Yu and Han Zhao

Abstract A semantic ontology of requirement-product-process-resource (RPPR) is developed to model the whole product lifecycle information in this study. The ontology is consisting of four parts, i.e., requirement model, product model, process model and resource model, which represents requirement information, product information, process information and resource information of the product life cycle, respectively. The ontology model can offer well-defined semantics for information sharing and reuse between different organizations and software systems. A case study is also presented in the paper to demonstrate the application of the ontology model.

Keywords Process • Product • Product lifecycle • Requirement • Resource • Semantic ontology

31.1 Introduction

The main objective of product lifecycle management (PLM) is to effectively manage the information which is generated or used by the whole product lifecycle processes. It is believed that sharing and reuse of information among different lifecycle processes can improve the efficiency of PLM and thereby to short product lifecycle and reduce product cost. But since the product lifecycle information is managed by different software systems, and each of them has different information structures, the information in different systems cannot be exchanged directly.

B. Yu (✉) • H. Zhao
School of Mechanical and Automotive Engineering, Hefei University of Technology, Hefei 230009, Anhui, China
e-mail: yubiao619@gmail.com; han.zhao@hf.ah.edu.cn

E. Qi et al. (eds.), *The 19th International Conference on Industrial Engineering and Engineering Management*, DOI 10.1007/978-3-642-37270-4_31,

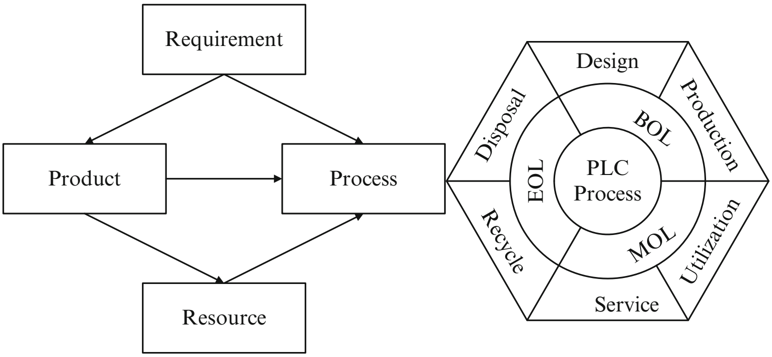


Fig. 31.1 Relationships among requirement, product, process and resource

With the emergence of networked firm and networked economy, the product lifecycle information becomes more and more complex and distribution, management of the lifecycle information becomes more and more difficult.

From the perspective of whole PLM, the product lifecycle information can be categorized into four types: requirement information, product information, process information and resource information. The requirement information is mainly related to customer requirements and engineering requirements. The product and all lifecycle processes are required to satisfy all of the requirements. The product information is related to the design of a product which includes both geometric and non-geometric aspects. The process information is consisting of the whole product lifecycle processes: begin of life (BOL) which includes product design and production process, middle of life (MOL) which includes user utilization and service process, and end of life (EOL) which includes recycle and disposal process. The resource information is the information which is generated or used by the processes and the product. The detailed relationships among them are given by Fig. 31.1.

The product life cycle information modeling has attracted a lot of interests from both academic and industrial during the last decade. Researchers proposed a lot of methods to model the product life cycle information, but most of them just focus on modeling one or two aspects of the following information, i.e. requirement modeling (Jiang and Qu 2006; Lin et al. 1996), product modeling (Sudarsan et al. 2005; Barbau et al. 2012; Zhao and Liu 2008; Steven et al. 2008), process modeling (Imre and László 1997) or resource modeling (Steele et al. 2001). To our knowledge, there is no one literature has covered all the four aspects of the product lifecycle information.

In this study, a semantic ontology of requirement-product-process-resource (RPPR) for product lifecycle information model is developed. This semantic ontology can model the whole product lifecycle information including requirement, product, process and resource, and provide well-defined semantics for information sharing and reuse among different software systems and organizations.

31.2 Ontology Modeling

31.2.1 Ontology

One of most acceptable definitions of ontology is given by Gruber who defined that "ontology is an explicit specification of a conceptualization" (Gruber 1993).

Ontology, as a way to unambiguously represent concepts, terminologies and relations among them in a domain, offers a good mean to model knowledge in a lot of fields especially in the field of information modeling and semantic web. Ontology models can provide some very useful features: to share common understanding of structure of information among human and/or software agents, to enable reuse of domain knowledge, to make domain assumptions explicit, to separate domain knowledge from operational knowledge, to provide formal analysis of terms and based on them, to analysis the domain knowledge (Matsokis and Kiritsis 2010). In this study, web ontology language based on description logic (OWL-DL) (World Wide Web Consortium 2004) is employed to model the product lifecycle information. Since the OWL-DL is based on description logic, the model described by OWL-DL can provide well-defined semantics, thereby avoiding the ambiguity that caused by inconsistencies of terminologies in the product lifecycle information (Yang et al. 2009).

31.2.2 Ontology Development Process

Development of ontology is very complex system engineering. In this study, the methodology proposed by Noy and McGuinness (2001) is applied to develop the product lifecycle information ontology. The detailed process is shown in Fig. 31.2.

31.2.3 Architecture of Product Lifecycle Ontology Model

Four layers ontology model architecture is proposed in the study, i.e., representation layer, meta-ontology layer, domain ontology layer and instance layer. Figure 31.3

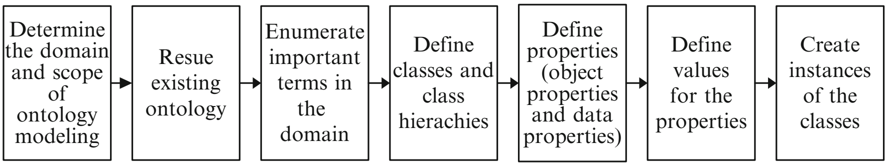


Fig. 31.2 Ontology development process

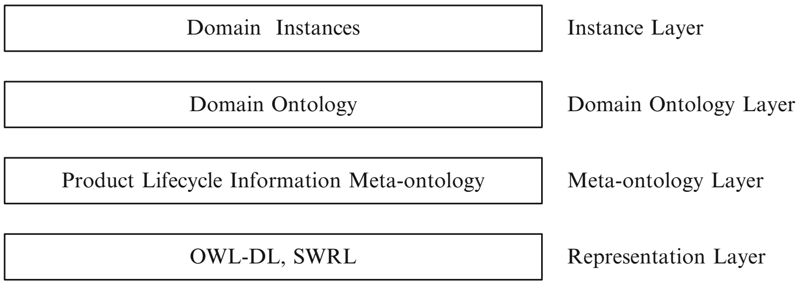


Fig. 31.3 Architecture of ontology model

shows the architecture. In the representation layer, OWL-DL is used as a language to describe the concepts and relations in the other three layers. In the meta-ontology layer, it defines some meta-concepts and meta-relations in the product lifecycle. These concepts and relations are common for various product domains. Domain ontology layer defines some special concepts and relations in a special product domain, e.g., for a direct methanol fuel cell (DMFC) system, in which some concepts and relations needs to defined to describe its product lifecycle information. The instances or individuals of concepts and relations in the domain ontology layer are described in the instance layer.

31.3 Product Lifecycle Meta-Ontology

31.3.1 Overview

The product lifecycle meta-ontology includes some meta-concepts and meta-relations between them. The meta-concepts can be classified into two types: *RPPR Core Object* and *RPPR Utility Object*. The *RPPR Core Object* includes all core concepts which are used to model the whole lifecycle information. These concepts are *Requirement Object*, *Product Object*, *Process Object* and *Resource Object*. The *RPPR Utility Object* is used to model some auxiliary information which is required by the *RPPR Core Object*. The meta-relations between these meta-concepts can be categorized into six types: *relations in Requirement Object*, *relations in Product Object*, *relations in Process Object*, *relations in Resource Object*, *relations between RPPR* and *relations between RPPR Core Object and RPPR Utility Object.* The detailed descriptions of these meta-concepts and meta-relations are given by following sections.

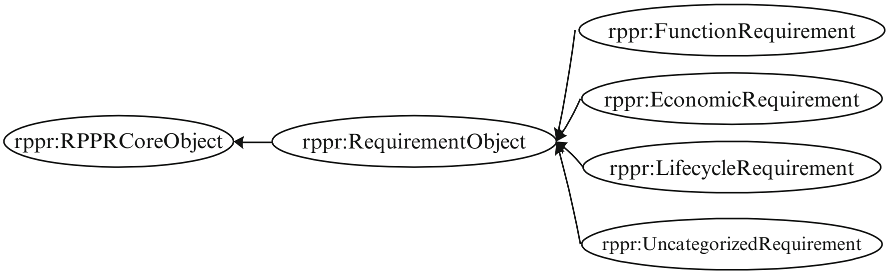


Fig. 31.4 Conceptual hierarchy relationships in the requirement objects

31.3.2 Requirement Object

Requirement Object represents the requirements of customers and engineering on the product and product lifecycle processes. These requirements can be classified into four types: *Function Requirement*, *Lifecycle Requirement*, *Economic Requirement* and *Uncategorized Requirement.*

The *Function Requirement* usually refers to the functions of a product. The *Lifecycle Requirement* refers to the requirements on product lifecycle processes, such as product maintainability, product recycling and disposal, etc. The *Economic Requirement* refers to some economic factors of the product, such as product price, maintenance cost and so on. All other requirements which cannot be classified into the above three types are categorized as *Uncategorized Requirement*. Figure 31.4 shows the conceptual hierarchy relationships in the requirement objects.

31.3.3 Product Object

Product Object represents all design information associated with the product. It should include both geometric and non-geometric information. The geometric information can be modeled by some geometric modeling systems, such as Solidworks, Pro/Engineering, etc. and expressed in some formats such as STEP (Gu and Chan 1995). In this study, the modeling of geometric information is not the focus of the research. Instead, the geometric information of a product is considered as a document resource associated with the product object.

The core concepts in the product object are *Entity*, *Product Specification* and *Property*. An *Entity* is a physical unit for modeling a product design and it can be classified into *Artifact* and *Feature*. An *Artifact* represents a distinct entity in a product. Typical *Artifact* includes *Component* and *Assembly*. *Component*

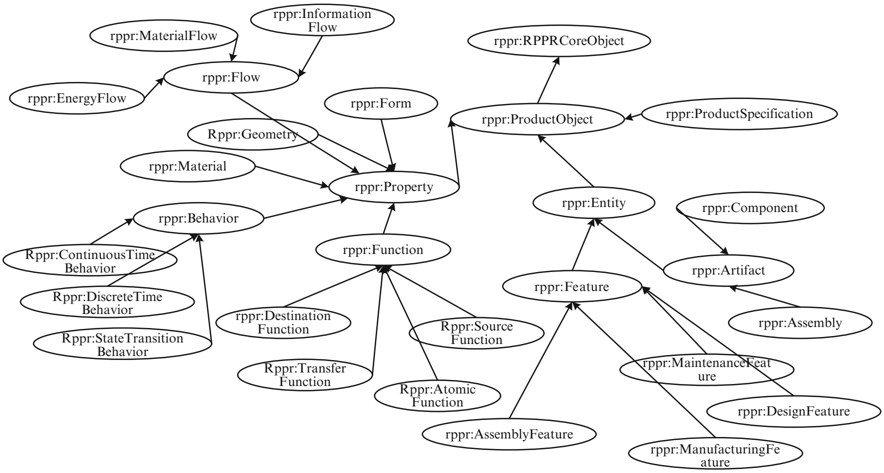


Fig. 31.5 Conceptual hierarchy relationships in the product objects

represents a primitive for modeling a design, and an *Assembly* is composed by some components and sub-assemblies in a product. *Feature* is a portion of the artifact's form that has some specific functions. It usually can be classified into *Design Feature*, *Manufacturing Feature*, *Assembly Feature*, *Maintenance Feature*, etc., according to different lifecycle proposes. *Property* is the inherited characteristic belonging to the entity. Typical properties include *Function*, *Behavior*, *Form*, *Geometry*, *Material* and *Flow*. *Function* is the teleological description of what is supposed to do by an *Artifact*. A *Function* is usually associated with some input flows and output flows. The *Flow* also can be categorized into three types: *Material Flow*, *Energy Flow* and *Information Flow*. According to the input flows and output flows, the *Function* can be classified into four categories: *Atomic Function*, S*ource Function*, *Destination Function* and *Transfer Function*. *Atomic Function* has no input and output flows, *Source Function* only has output flows, *Destination Function* only has input flows and *Transfer Function* has both input and output flows. *Behavior* describes how the artifact's form implements its function. *Behavior* can be categorized as *Continuous Time Behavior*, *Discrete Time Behavior* and *State Transition Behavior*. A *Continuous Time Behavior* describes the changes of the artifact descriptions when the value of time parameter changes continuously. A *Discrete Time Behavior* describes the changes of the artifact descriptions when the time parameter value can only be selected from a set of discrete values. A *State Transition Behavio*r describes a number of states and conditions to change among these states (Xue and Yang 2004). *Form* is the proposed physical solution to the artifact's function. A *Form* is composed by some geometries, features and materials. *Geometry* is the spatial description of a *Form*. *Material* is the description of the internal composition of a *Form*. *Product Specification* represents the design requirements and constraints on the *Artifact*. Figure 31.5 shows the conceptual hierarchy relationships in the product objects.

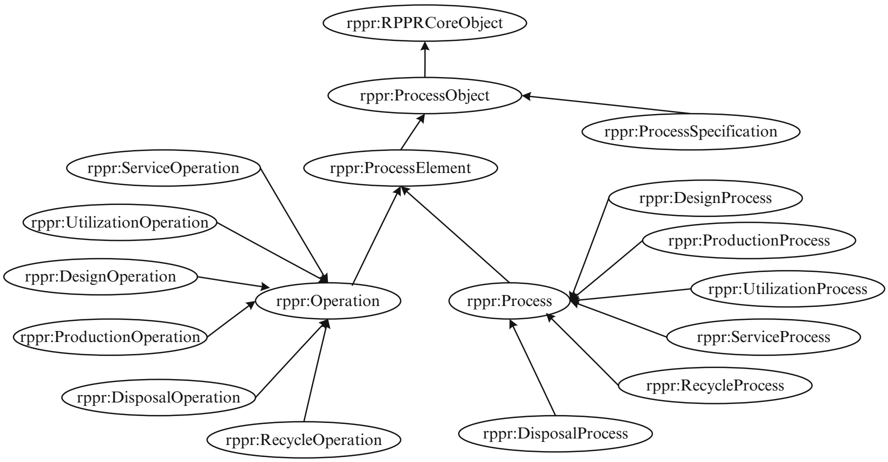


Fig. 31.6 Conceptual hierarchy relationships in the process objects

31.3.4 Process Object

Process Object describes the whole product lifecycle processes which include BOL, MOL and EOL. The core concepts in the process object are *Process Element* and *Process Specification*. The *Process Element* is the description of the structure of a process. The *Process Specification* gives the requirements and constraints on the lifecycle processes. The *Process Element* can be classified into *Operation* and *Process*. An *Operation* is a primitive process element. A *Process* includes some operations and sub-processes. According to lifecycle processes, the *Process* can be classified as *Design Process*, *Production Process*, *Utilization Process*, *Service Process*, *Recycle Process* and *Disposal Process*. Accordingly, the *Operation* also can be categorized as *Design Operation*, *Production Operation*, *Utilization Operation*, *Service Operation*, *Recycle Operation* and *Disposal Operation*. Figure 31.6 shows the conceptual hierarchy relationships in the process objects.

31.3.5 Resource Object

Resource Object represents all the resource information that required or generated by the product objects and process objects. Typically, the resources in an organization can be classified as *Organization*, *Workplace*, *Document*, *Material Resource* and *Actor*. *Organization* describes the organization structure in a company. It is categorized as *Company* and *Department*. The *Workplace* represents the locations where the work is conducted in an *Organization*. It is classified as *Office*, *Factory*, *Workshop*, *Production Line*, *Workgroup*, *Workstation* and *Workcell*.

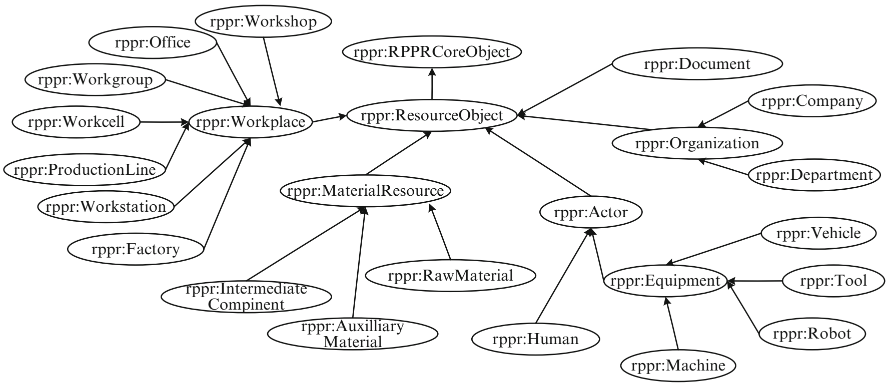


Fig. 31.7 Conceptual hierarchy relationships in the resource objects

Material Resource is the materials to support the lifecycle processes such as manufacturing process. It is classified as *Raw Material*, *Intermediate Component* and *Auxiliary Material*. *Document* represents the document resources that are generated or required by the lifecycle processes. *Actor* is operator which is required by some processes. It is categorized as *Human* and *Equipment*. The *Equipment* also can be classified as *Machine*, *Robot*, *Tool*, *and Vehicle*. Figure 31.7 shows the conceptual hierarchy relationships in the resource objects.

31.3.6 Utility Object

There are two kinds of utility objects: *Description* and *Constraint*. *Description* object provides some descriptions such as *object ID*, *object name* and other information to the *RPPR Core Object*. Each *RPPR Core Object* should associate with a *Description* object. *Constraint* describes the constraints on the core objects and relations.

31.3.7 Relation

In the proposed ontology model, the relations among the concepts are very complicated. We categorize them into six types: *relations in Requirement Object*, *relations in Product Object*, *relations in Process Object*, *relations in Resource Object*, *relations between RPPR* and *relations between RPPR Core Object and RPPR Utility Object*. Table 31.1 gives the main meta-relations in the meta-ontology model.

Table 31.1 Main meta-relations in the meta-ontology

	Relation		
Category	Relation name	Domain	Range
Relation in requirement object	hasSubRequirement	Requirement	Requirement
Relation in product object	hasFeature	Artifact	Feature
	hasFunction	Aftifact	Function
	hasForm	Artifact	Form
	hasBehavior	Artifact	Behavior
	hasGeometry	Form	Geometry
	hasMaterial	Form	Material
	hasInputFlow	Function	Flow
	hasOutputFlow	Function	Flow
	isSpecifiedByProduct Specification	Artifact	Product Specification
Relation in process object	hasOperation	Process	Operation
	isSpecifiedByProcess Specification	ProcessElement	Process Specification
Relation in resource object	hasWorkplace	Organization	Workplace
	requireDocument		Document
	requireMaterialResource		MaterialResource
	requireActor		Actor
Relation between RPPR	productRequireResource	ProductObject	ResourceObject
	processRequireResource	ProcessObject	ResourceObject
	productSpecification SatisfyTo	ProductObject	Product Specification
	processSpecification SatisfyTo	ProcessObject	Process Specification
Relation between RPPR core object and RPPR utility object	hasDescription	RPPRCoreObject	Description
	hasConstraint	RPPRCoreObject	Constraint

31.4 Case Study

A direct methanol fuel cell (DMFC) system is used as an example to demonstrate the application of the RPPR meta-ontology. Figure 31.8 shows part of the product structure of the DMFC system.

The part of manufacturing process of component BPP_A in the DMFC system is given by Fig. 31.9.

Based on the developed product lifecycle meta-ontology, the DMFC domain ontology is developed. For example, the stack, a sub-assembly of DMFC system is defined as follows:

<owl:Class rdf:ID = “Stack”>
<rdfs:subClassOf rdf:resource = “http://www.owl-ontologies.com/rppr_metaontology.owl#Assembly”/>
<rdfs:subClassOf>

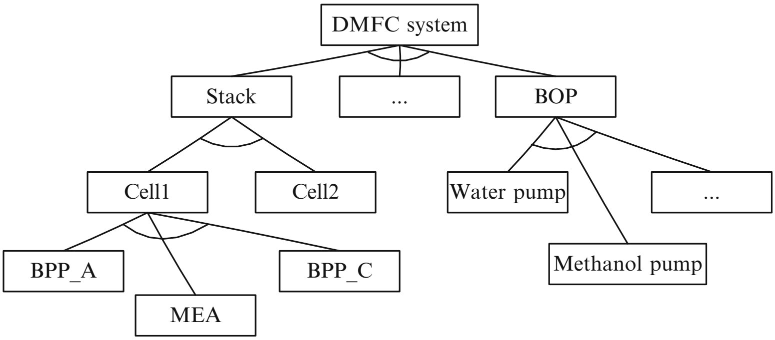


Fig. 31.8 DMFC system product structure

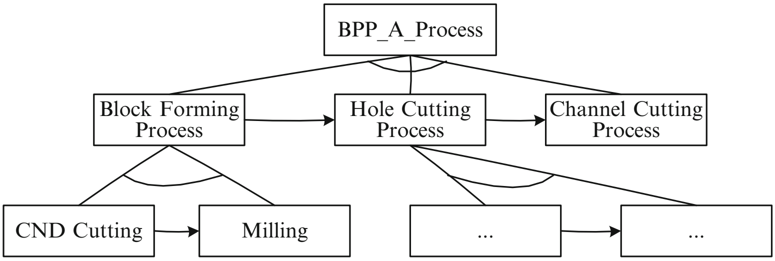


Fig. 31.9 Manufacturing process of BPP_A

```
    <owl:Restriction>
 <owl:onProperty>
        <owl:TransitiveProperty rdf:ID = "hasCell"/>
      </owl:onProperty>
       <owl:minCardinality rdf:datatype = "http://www.w3.org/2001/XMLSchema
          #int"
      >1</owl:minCardinality>
    </owl:Restriction>
  </rdfs:subClassOf>
  <rdfs:subClassOf>
    <owl:Restriction>
      <owl:minCardinality rdf:datatype = "http://www.w3.org/2001/XMLSchema
          #int"
      >0</owl:minCardinality>
      <owl:onProperty>
        <owl:ObjectProperty rdf:ID = "stackOf"/>
      </owl:onProperty>
    </owl:Restriction>
  </rdfs:subClassOf>
 </owl:Class>
```

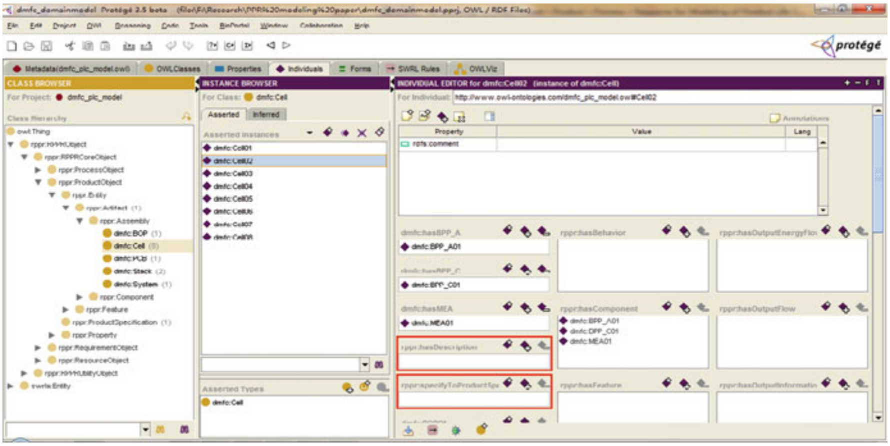


Fig. 31.10 Snapshot of the RPPR ontology modeling system

Then an instance of this stack is represented as:

<dmfc:Stack rdf:ID = "Stack01"/>
The CNC-cutting manufacturing operation is also can be defined as:
<owl:Class rdf:ID = "CNCCutting">
<rdfs:subClassOf rdf:resource = "#Cutting"/>
</owl:Class>

An instance of this CNC-cutting is given by:

<dmfc:CNCCutting df:ID = "BlockCNCCutting"/>

A system based on Protégé is also implemented. Figure 31.10 shows a snapshot of the system.

31.5 Conclusion

A semantic ontology of requirement-product-process-resource (RPPR) for product lifecycle is developed in this study. This ontology model can provide well-defined semantics for modeling all aspects of product lifecycle information, i.e., requirement, product, process and resource, thereby to achieve information sharing and reuse. A case study is also conducted to demonstrate the application of the model.

Acknowledgment The authors would like to acknowledge all reviewers for this paper.

References

Barbau R, Krima S, Rachuri S, Narayanan A, Fiorentini X, Foufou S, Sriram RD (2012) OntoSTEP: enriching product model data using ontologies. Comput Aided Des 44(6):575–590

Gruber T (1993) A translation approach to portable ontologies. Knowl Acquis 5(2):199–220

Gu P, Chan K (1995) Product modeling using STEP. Comput Aided Des 27(3):163–179

Imre JR, László H (1997) Modeling of manufacturing processes using a Petri-net representation. Eng Appl Artif Intel 10(3):243–255

Jiang L, Qu FZ (2006) Study on construction and application of product requirement design ontology. In: International technology and innovation conference, ITIC 2006, Hangzhou, pp 96–101

Lin JX, Fox MS, Bilgic T (1996) A requirement ontology for engineering design. Concurr Eng 4(3):279–291

Matsokis A, Kiritsis D (2010) An ontology-based approach for product lifecycle management. Comput Ind 61(8):787–797

Noy NF, McGuinness DL (2001) Ontology development 101: a first guide to creating your first ontology. Stanford knowledge system laboratory technical report KSL-01-05 and Stanford medical informatics technical report SMI-2001-0880

Steele J, Son YJ, Wysk RA (2001) Resource modeling for the integration of the manufacturing enterprise. J Manuf Syst 19(6):407–427

Steven JF, Foufou S, Bock C, Sriram RD (2008) CPM2: a core model for product data. J Comput Inf Sci Eng 8(1):1–14

Sudarsan R, Fenves SJ, Sriram RD, Wang F (2005) A product information modeling framework for product lifecycle management. Comput Aided Des 37(13):1399–1411

World Wide Web Consortium (2004) OWL web ontology language overview. http://www.w3.org/TR/2004/REC-owl-features-20040210/

Xue D, Yang H (2004) A concurrent engineering-oriented design database representation model. Comput Aided Des 36(10):947–965

Yang D, Miao R, Wu HW, Zhou YT (2009) Product configuration knowledge modeling using ontology web language. Expert Syst Appl 36(3):4399–4411

Zhao W, Liu JK (2008) OWL/SWRL representation methodology for EXPRESS-driven product information model: part I. Implementation methodology. Comput Ind 59(6):580–589

Chapter 32
Research on Turning the Toll Mode of the Periphery Freeway into Congestion Charging

Jian Tong, Yuan-hua Jia, and Zhong-ming Guo

Abstract The traffic flow of periphery area between freeway and urban expressway needs scientific and accurate control. We should establish reasonable policies and measures to control the traffic; guide travel by economic measures, and reduce traffic congestion. The adaptability of existing congestion charging and traffic growth is different. This article summarizes the periphery freeway congestion mechanism by supply and demand theory, resource scarcity theory. Based on the comparison between the charging models at home and abroad, it builds the bi-level programming model for periphery freeway congestion, turns the traditional toll mode into congestion charging, and takes freeway G2 as an example. Practice proves that the methods are feasible and efficient and it could provide the foundation for alleviating traffic congestion of periphery area (Fund Project: 863 Program, Project number: 2007AA11Z213).

Keywords Periphery area • Freeway • Congestion charging • Elastic demand • Stochastic user • Equilibrium • Bi-level programming

32.1 Introduction

The traffic jam of periphery road network in Beijing, Shanghai and Guangzhou becomes more and more serious at present. We have lack of management and control about periphery road network. At present, the basic method is using predetermined main line traffic state threshold to decide whether the ramp opening or closing. The considering elements are simple.

J. Tong (✉) • Y. Jia • Z. Guo
School of Traffic and Transportation, Beijing Jiao-tong University, Beijing, China
e-mail: 11120989@bjtu.edu.cn

E. Qi et al. (eds.), *The 19th International Conference on Industrial Engineering and Engineering Management*, DOI 10.1007/978-3-642-37270-4_32,

Vickrey (1969) offers a charging model which bases on the queue theory about bottleneck congestion. User can travel on the road connecting the living and working areas. The freeway has a bottleneck entrance, whose capacity is limited. It divides user's cost into three parts that contain link travel time cost, plan delay cost and road use cost. Although the bottleneck model can analyze the increase and decrease of queue and the behavior of user's time choice, it is different from the actual traffic because of the stern constraints. Hau (1992) uses social welfare effects analysis of the static congestion charging. He thinks that the benefit of government is the largest. So we must pay more attention to the income redistribution. Cohen (1987) groups the travelers according to their characteristics, and analyzes the income effect of costs on different users qualitatively. Liu Weiming (Xie et al. 2012) makes the research on the optimal method of road network charging rate, the charging road network and charging system in his "Road Pricing System Optimization Model Calculating method". Lu Zhengfeng uses the route choice behavior theory of the road user as the basis theory (Rong Chaohe 2011) of the freeway rate optimization. We solve the problem by opening or closing the ramp, information dissemination, while traffic management demand especially congestion charging is less considered.

32.2 Supply and Demand Theory and Resource Scarcity Theory

32.2.1 Supply and Demand Theory

Supply and demand are the two basic aspects of transportation market. Demand is the reason of supply while supply is the foundation of demand. Thus, transportation market is the mechanism that both demand side and supply side well cooperates and regulates price and quantity (Xu and Wei 2012). Then the mechanism adjusts itself to bring the regular movement between supply and demand, and appear the balance of price and quantity. As time goes on, supply varies its own rule as well as demand. Therefore the current balance between price and quantity is broken and a new one moves forward. In road transportation, travelers represent demand side just like road infrastructures and services which are supplied by supply side, government and management.

32.2.2 Resource Scarcity Theory

In this paper, scarcity is insufficiency that products are relative to meeting human needs. Scarcity is closely related to demand and surely it does not mean the absolute quantity of products or resource but is related together and has certain conditions.

For instants, scarcity in freeway resource indicates people cannot possess, use or construct road resource without limit. As a result, scarcity causes basic features of existence and activity about the freeway resource value, is the existence of freeway competition and the price mechanism of freeway. If the freeway is lack of control measures, the competition normally will become unhealthy, such as the continued spread of traffic congestion (Li Dongmei 2004). Necessarily, freeway resource scarcity requires price mechanism to control the number of users by charging to control the number and type of users in freeway. In addition, time property of freeway resource scarcity is also its dynamic feature. For actually, rush hour with large traffic flow has scarcity, however, period that is free from congestion at midnight is not.

32.3 Congestion Charging Model for Freeway in Periphery Area

32.3.1 Introduction of Model

Bi-level programming which considers two levels problem is Multi-objective decision-making problem. Each level represents different interests, and has their own objective function and constraint condition. The calculation process of bi-level programming is to obtain a set of optimal solutions from lower constraint condition and objective function, put it into the upper objective, feed back to the lower level to intervene and adjust the solution of function, then put it into the upper level again, circulate until finding the optimal value, this is the upper level control lower method to control and dispose decision-making variable (Salemi et al. 2011; Louie Nan and McDonald 1998; Wei and Chen 2009; Haijun 1994; Leurent 1993).

32.3.2 The Lower Model

In the existing traffic assignment model, we mostly pay attention to deterministic UE and fixed demand SUE problem, but little attention to elastic demand SUE problem. There is a certain deviation in travelling environment, personal attributes and estimate of traffic situation and road network status for users. The deviation is a random variable. So the stochastic user equilibrium of elastic demand is more suitable to descript the travel behavior. Traffic demand is not fixed in any OD because users could adjust the travel plans on the network congestion degree and charging strategy. The elastic demand SUE reappear traffic demand elasticity and the randomicity of path choice. This study describes the lower model as the elastic demand SUE problem.

1. The condition of route choice. Define the $C_k^w = c_k^w + \varepsilon_k^w, \varepsilon_k^w$ are mutual independence, obey GumbeL distribution with parameter θ. The Objective Function based on SUE state of Logit is:

$$f_k^w = p_k^w q^w = \frac{e^{-\theta c_k^w}}{\sum_{k \in kw} e^{-\theta c_k^w}} q^{\mathrm{w}}, \forall_k \in k_w, w \in W \tag{32.1}$$

2. The elastic demand function. q^{w} is travel demand in OD (r, s), S^{w} is expectation travel resistance in OD (r, s). If q^{w} is the strongly monotonous decreasing function of S^{w} with upper limit; q^{w} is travel demand function in OD (r, s), $D^{w-1}(q^w)$ is the inverse function of demand function $D^w(S^w)$, where π_w is the positive parameter of elastic demand function. Its value is 0.01.

   The stochastic user equilibrium condition of private car is:

$$f_{1k}^w = p_{1k}^w q_1^w = \frac{e^{-\theta_1 c_{1k}^w}}{\sum_{k \in k_w} e^{-\theta_1 c_{1k}^w}} q_1^w \tag{32.2}$$

   The inverse function of private car on the road k is:

$$D_1^{w-1}\left(q_1^w\right) = -\frac{1}{\theta_1} ln \sum_k e^{-\theta_1 c_{1k}^w} \tag{32.3}$$

   The stochastic user equilibrium condition of taxi is:

$$\mathrm{f}_{2\mathrm{k}}^{\mathrm{w}} = \mathrm{p}_{2\mathrm{k}}^{\mathrm{w}} \mathrm{q}_2^{\mathrm{w}} = \frac{\mathrm{e}^{-\theta_2 \mathrm{c}_{2\mathrm{k}}^{\mathrm{w}}}}{\sum_{\mathrm{k} \in \mathrm{k_w}} \mathrm{e}^{-\theta_2 \mathrm{c}_{2\mathrm{k}}^{\mathrm{w}}}} \mathrm{q}_2^{\mathrm{w}} \tag{32.4}$$

   The inverse function of taxi on the road k is:

$$D_2^{w-1}\left(q_2^w\right) = -\frac{1}{\theta_2} ln \sum_k e^{-\theta_2 c_{2k}^w} \tag{32.5}$$

User's travel resistance function is contacted with the flow of taxi and private car on the road α. Their influence is symmetrical, which is shown as follows:

$$\frac{\partial c_{1a}\left(x_{1a}, x_{2a}\right)}{\partial x_{2a}} = \frac{\partial c_{2a}\left(x_{1a}, x_{2a}\right)}{\partial x_{1a}} \tag{32.6}$$

In a word, the minimum value model is:

$$\begin{aligned} & minZ\left(x, f_i, q_i, q\right) \\ & = \sum_i \left[\begin{array}{c} \frac{1}{\theta_i} \left(\sum_w \sum_k f_{ik}^w \left(ln\, f_{ik}^w - 1\right) - \sum_w q_i^w \left(ln\, q_i^w - 1\right) \right) \\ - \sum_w \int_0^{q_i^w} D_i^{w-1}(w) dw \end{array} \right] \\ & \quad + G^m(\cdot) \end{aligned} \tag{32.7}$$

$$s.t. \quad \sum_{k} f_{ik}^{w} = q_i^{w}, \forall i, k \tag{32.8}$$

$$\begin{cases} x_{ia} = \sum_{w} \sum_{k} f_{ik}^{w} \delta_{a,k}^{w}, \forall a, i \\ f_{ik}^{w} \geq 0, x_{ia} \geq 0, \forall i, k, a \end{cases} \tag{32.9}$$

Where the expression of $G^m(\cdot)$ is different based on the value of m, when $m = 2$, the expression is:

$$\begin{aligned} G^2(\cdot) = \frac{1}{2} \sum_{a} \int_0^{x_{1a}} & c_{1a}(w, x_{2a})\, dw \\ + \frac{1}{2} \sum_{a} \int_0^{x_{1a}} & c_{1a}(w, 0)\, dw \\ + \frac{1}{2} \sum_{a} \int_0^{x_{2a}} & c_{2a}(x_{1a}, w)\, dw \\ + \frac{1}{2} \sum_{a} \int_0^{x_{2a}} & c_{2a}(w, 0)\, dw \end{aligned}$$

32.3.3 The Upper Model

The objective functions which are made by upper decision-makers will be different according to the different starting points. We should consider the manager financial goals on charging road and the user response to charging strategy. The upper object will be the maximum customer surplus in the whole system when road flow is less than traffic capacity. The difference between TUB (Total User Benefit) (Total User Benefit) and TSC (Total Social cost) is customer surplus CS which is shown as follows:

$$\mathrm{CS} = \mathrm{TUB} - \mathrm{TSC}$$

The elastic demand is:

$$TUB = \sum_{w} \int_0^{q^w(u)} D^{w-1}(x) dx$$

$$TSC = \sum_{a} C_a(x_a(u),\ u)\, x_a(u)$$

So the upper model is:

$$
\begin{aligned}
minF\,(u, x(u),\, q(u)) &= -CS \\
&= \sum_a C_a\,(x_a(u),\, u)\, x_a(u) \\
&- \sum_w \int_0^{q^w(u)} D^{w-1}(x)dx \\
s.t.\ x_a(u) &\le v_a \quad \forall a \in A \\
u_a &\ge 0 \quad \forall a \in A
\end{aligned}
$$

32.3.4 Algorithm Solving

Bi-level programming problem is NP, there is no polynomial solution because of its non-convexity, and the solution is usually not a global optimal solution but a local optimal solution. This paper adopts directly search method of step accelerate method and penalty function method, uses penalty function to change upper model into unconstrained extremum problem. Algorithm is shown as follows:

Step 0: Set the initial penalty factor $\gamma_i = 10$, let $i = 1$;

Step 1: Chose the initial charging plan: $u = u^0$, $\chi(u)$ and $q(u)$ are given from solving the lower model, put them into the upper model to calculate the $G\,(u, \gamma)$, choose initial step δ and acceleration factor σ, let $\beta = 1, j = 1, k = 1$ Table 32.1;

Step 2: detection search. if $j > |A|$, then go to Step 3, let $u^* = u^0 + \beta\delta e_1$, solve the lower problem. Calculate the upper objective function $G\,(u^*, \gamma) < G\,(u, \gamma)$, let $u = u^*$, $G\,(u, \gamma) = G\,(u^*, \gamma)\quad j = j + 1, \quad \beta = 1$, repeat step 2, otherwise let $\beta = -1$, if the answer is still $G\,(u^*, \gamma) > G\,(u, \gamma)$, let $u^* = u^0, j = j + 1$, repeat step 2, namely:

$$
u^* = \begin{cases} u^0 + \beta\delta e_1, & G\left(u^0 + \beta\delta e_1,\ \gamma\right) < G\,(u,\ \gamma) \\ u^0 - \beta\delta e_1, G\left(u^0 - \beta\delta e_1,\ \gamma\right) < G\,(u,\ \gamma) < G\left(u^0 + \beta\delta e_1,\ \gamma\right) \\ u^0, \quad G\,(u,\ \gamma) < min\left[G\left(u^0 - \beta\delta e_1,\ \gamma\right), G\left(u^0 + \beta\delta e_1,\ \gamma\right)\right] \end{cases}
$$

Step 3: pattern search if $G\,(u^*, \gamma) > G\,(u, \gamma)$, $\delta > \varepsilon$, let $u^{k+1} = \max\left[0, u^k + \sigma\left(u^{k+1} - u^k\right)\right]$, $j = 1$, $k = k + 1$, go to Step 2, if $G\,(u^*, \gamma) > G\,(u, \gamma)$, $\delta > \varepsilon$, go to Step 4, otherwise let $\delta = 0.5\,\delta, j = 1$, $u = u^k$, go to Step 2;

Step 4: if $\chi_a > v_a, \forall_a \in A$, then increase the penalty factor, which is $\gamma_{i+1} = \lambda\gamma_i$, $(\lambda > 1)$, let $i = i + 1$, $u^0 = u^k$, go to step 1, otherwise end the evaluation when u^k is output as the optimal solution.

Table 32.1 Parameters of travel cost function

Road	s-5	5-4	4-3	s-1	1-2	2-3	s-6	6-7	7-3	1-5	2-4	6-5	7-4
l_a (kilometer)	1.2	3.5	3.2	1.2	3.5	3.2	1.2	3.5	3.2	0.345	0.394	0.607	0.402
t_0 (second)	36	105	96	72	210	192	72	210	192	31.05	35.46	54.63	36.18
Y_a (yuan)	0.984	2.870	2.624	1.968	5.740	5.248	1.968	5.740	5.248	0.283	0.323	0.498	0.330
T_a (yuan)	2.4	7	6.4	4.8	14	12.8	4.8	14	12.8	0.69	0.788	1.214	0.804

32.4 Case Study

32.4.1 *The Periphery Freeway Congestion Charging Rate Constitution Method*

In this paper, the study takes the G2 freeway morning rush hour in urban direction as an example, uses periphery freeway and road network make up of adjacent road to distribute the traffic. It takes the private cars and taxis from urban expressway into freeway during morning rush hour as the research objects, verifies the charging model above.

1. Basic requirements:
   Charge sections are periphery freeway only, charging objects are taxis and private cars ramp into or existing on the periphery freeway.
   Suppose that the length and width of entrance ramp have no effect on traffic, the weaving section has no effect on impedance function;
   Suppose that the ramp beginning and end of entrance ramp can be connected to ordinary road.
2. Periphery road network is shown in Fig. 32.1, the meaning of symbols are:
   S: The charging starting point of city direction;
   5: The G2 freeway Dayangfang Bridge;
   4: The G2 freeway ShibaLidian Bridge;
   3: The G2 freeway Fenzhongsi Bridge;
   1 and 6, 2 and 7: The ramp entrances of the Fourth Ring and the Fifth Ring;
   S-5-4-3: The G2 freeway charging sections;
   1-5 and 6-5, 2-4 and 7-4: The ramps of G2 freeway which connect to the Fourth Ring and the Fifth Ring;
   S-1 and S-6: The roads between the charging starting point S and the Fifth Ring entrances1 and 6. (Length is equal to S-5);
   1-2 and 6-7: The roads between the Fourth Ring entrances and the Fifth Ring entrances. (Length is equal to S-4);
   2-3 and 7-3: The roads between the Fourth Ring entrances and the Fenzhongsi Bridge. (Length is equal to 4-3).

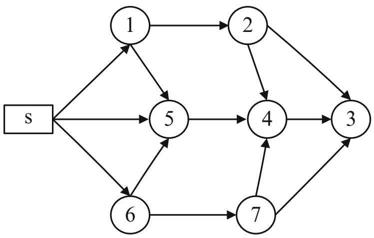


Fig. 32.1 The network simplified structure

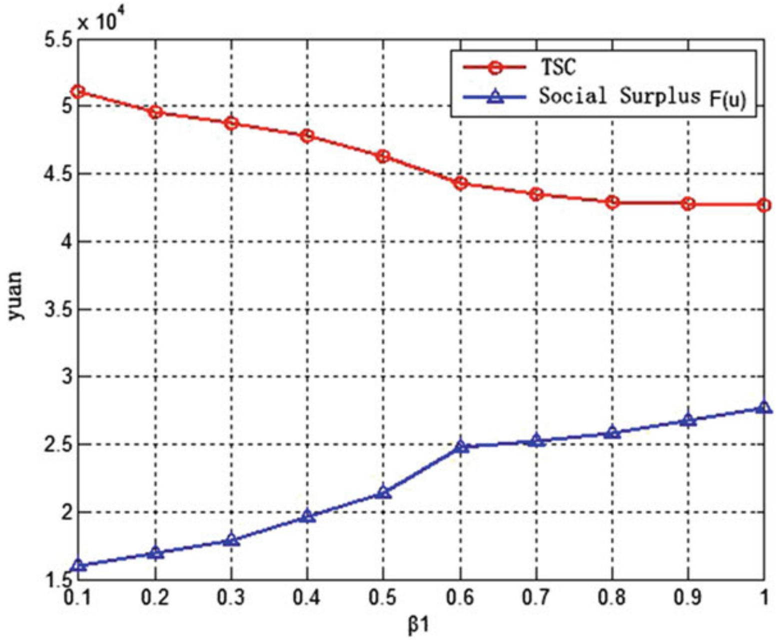


Fig. 32.2 Social surplus $F(u)$ and the TSC change ($\theta_1 = \theta_2 = 0.1$)

3. Basic data are as follows:
   In 2011, Beijing's GDP is 1,600 billion yuan, according to the resident population, the whole city per capita GDP is 80394 yuan, Nml is 9.2 yuan/h (0.003 yuan/s). The unit distance expenses of private car is 0.82 yuan/km. Taxi fare is 2 yuan/km. The private car and taxi each has an average of 1.5 passengers. The Third Ring, the Fourth Ring, the Fifth Ring and the
   G2 freeway, rush hour traffic flow of each are: 4253 pcu/h, 5126 pcu/h, 4532 pcu/h, 4246 pcu/h. Parameters of the time impedance function:

$$\alpha_{11} = 0.513; \quad \beta_{11} = 1.643 \quad \alpha_{22} = 0.5425;$$
$$\beta_{22} = 1.611; \quad \beta_{21} = \beta_{12} = 1; \quad \alpha_{21} = 0.301;$$
$$\alpha_{12} = 0.373.$$

32.4.2 Result Analyze

To research the rate problem above, this paper uses MATLAB programming, obtains the results which are shown in Figs. 32.2, 32.3, and 32.4:

β_1, β_2 are the government decision parameters. The smaller value expresses government pays more attention to the social surplus CS. From Figs. 32.2 and 32.3, the travel demand of car users are more and more, social total cost TSC is lower, taxi

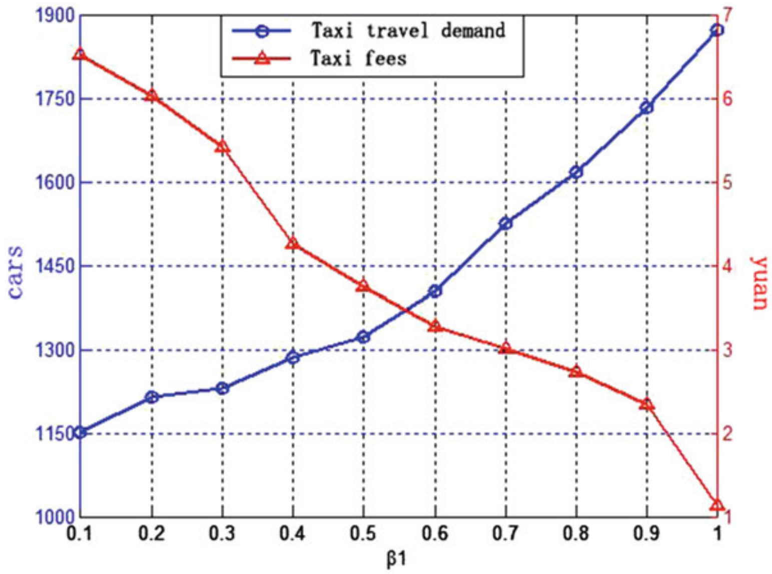


Fig. 32.3 Taxi travel demand and taxi fees change ($\theta_1 = \theta_2 = 0.1$)

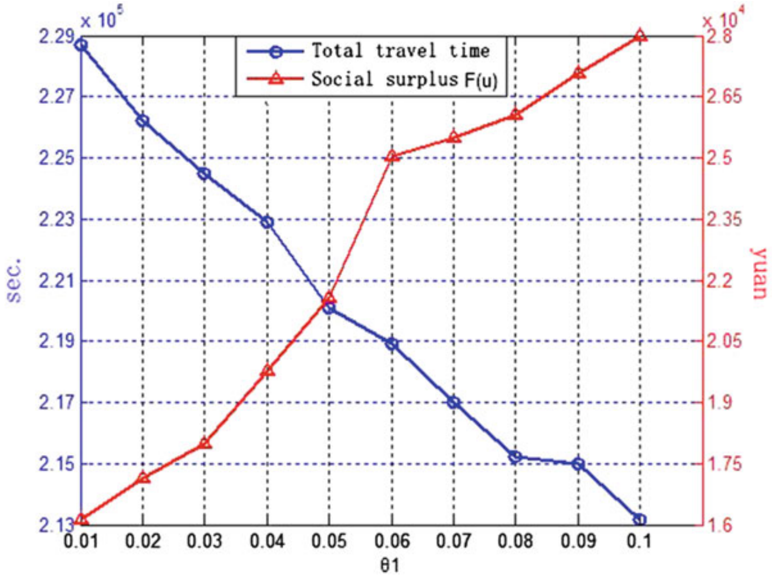


Fig. 32.4 The social total travel time and social surplus change ($\beta_1 = \beta_2 = 0.5$, $\theta_2 = 0.1$)

congestion charging rate is reducing. Because the government pays attention to the social surplus produced by taxi users, which causes the travel demand of taxi users increasing, also the vehicles in the road network are increasing. At the same time, travel expenses of the private car users increases as impedance increases, the total road network operation impedance increases. But the increase of taxi users makes the social surplus bigger. Therefore, government can guide people chose the way of travel through controlling the value of β_2. Reducing β_2 can guide private car users' choice public transportation, give up the line to carpool with others, it reduces the road network traffic pressure and eases traffic congestion.

T Parameter θ reflects the situation about traveler understanding trip travel time. The larger θ is, the deeper users apperceive impedance. The study considers that the taxi users are more sensitive than the private car users about road network, $\theta_2 > \theta_1$, So the article lets $\theta_2 = 0.1$. From the Fig. 32.4, we know that when $\beta_1 = \beta_2 = 0.5$, θ_1 is larger, which means users are more sensitive to road network. With θ_1 becomes larger, the total travel time of road network becomes small and social surplus grows fast. The private users choose the smaller impedance roads or other modes of transportation. Since private car users choose the small impedance road or other traffic pattern. The reduction of road network total time proves that effectiveness of congestion charging has an effect on reducing traffic jam.

32.5 Research Conclusions

Periphery area undertakes the conversion task about urban traffic flow. But freeway is the main transport corridor of periphery area. As a proven traffic demand management measure, we pay more attention to it. Domestic and foreign scholars in this field do a lot of researches and achieve great success. With the No Parking Electronic Toll System introduced, it provides guarantee to freeway congestion charging measure. There are four conclusions in this study:

1. According to the situation whether charge or jam to define the economic attributes of periphery freeway as personal effects, concrete analysis of freeway traffic, find out the reason and expression of periphery freeway congestion.
2. Giving qualitative analyze about the definition of periphery area, clearing the transport network environment of research object. Through the analysis of G2 freeway traffic flow characteristics, this paper gets the freeway traffic congestion periods and roads.
3. According to the problems of freeway in operation in our country, this study combines the fact of freeway congestion, puts forward the transition from toll mode to congestion charging on the existing freeway, demonstrates its necessity, and determines the purpose, principle and object of charge.
4. It builds the bi-level programming model for periphery freeway charging considering two kinds of users—private cars and taxis, takes Freeway G2 as an example, validates the efficiency of model, and further provides the necessity and feasibility of turning the toll mode of the periphery freeway into congestion charging.

References

Cohen Y (1987) Commuter welfare under peak-period congestion tolls: who gains and who loses. Int J Transp Econ 14(1):239–266
Haijun H (1994) Bidge tolls making under traffic equilibrium assignment. Inf Control 23(5):279–284
Hau TD (1992) Economic fundamentals of road pricing: a diagrammatic analysis. Transport Division, Infrastructure and Urban Development Department, The World Bank, Washington, DC
Leurent F (1993) Cost versus time equilibrium over a network. Eur J Oper Res 71:205–221
Li Dongmei (2004) Research on the congestion pricing of freeway. Southeast University, China
Louie Nan, McDonald JF (1998) Efficient congestion tolls in the presence of unpriced congestion: a peak and off-peak simulation model. J Urban Econ, Elsevier, 44(4):352–366
Rong Chaohe (2011) The value of transport-logistics time and its applications in time-spatial analysis. Econ Res J 12(8):66–101
Salemi S, Selamat A, Selamat MH (2011) Analysis of service quality and user satisfaction improvement in public transportation system. Int J Digit Content Technol Appl 5(10):95–104
Vickrey WS (1969) Congestion theory and transportation investment. Am Econ Rev 34(2):66–101
Wei Wang, Chen Xuewu (2009) Simulation research on chaos characteristics of a class of macroscopic traffic models. In: Proceedings of the 2009 international conference on engineering computation, 2009, Hong Kong, pp 150–153
Xie F, Jia Y, Jia R (2012) Algorithm for minimum cost maximum flow in transportation network. J Converg Inf Technol 7(7):165–173
Xu W, Wei Z (2012) Determining the optimal structure of the traffic and transportation network. Adv Inf Sci Serv Sci 4(12):355–361

Chapter 33
The Research of Information-Based Market Manipulation

Xiao-long Liu and Wen-xiu Hu

Abstract According to the theory of behavior finance, the model of information-based manipulation in Chinese stock market is studied in this paper. Meanwhile, the trading strategies of different traders are analyzed by using restriction of financing and security loan, on the basis of which the equilibrium outcomes of stock manipulation in different periods are obtained. Furthermore, the influences of margin deposit ratio and the rationality degree of semi-rational noise traders on stock market are discussed. Finally, some policy advices are given in the light of model analysis.

Keywords Margin deposit ratio • Market manipulation • Rationality degree • Semi-rational noise trader

33.1 Introduction

China's securities market, as an emerging market, faces serious problems of stock price manipulation for many reasons: relatively short time since the establishment, imperfection and immaturity in various aspects. Allen and Gale (1992) have divided market manipulation into information-based manipulation, action-based manipulation and trade-based manipulation in the classic literature. However, many studies have shown that profit of the latter two manipulations is unlikely, and market

The work is supported by the Ph.D. Programs Foundation of Ministry of Education of China (20096118110010) and the Specialized Fund for Key Subjects of Universities in Shaanxi Province (107-00X902).

X. Liu (✉) • W. Hu
School of Business Administration, Xi'an University of Technology, Xi'an 710054, China
e-mail: lxlwolf@126.com

E. Qi et al. (eds.), *The 19th International Conference on Industrial Engineering and Engineering Management*, DOI 10.1007/978-3-642-37270-4_33,

manipulation based on the information, particularly the inside information, is more common in the stock market (Foster and Viswanathan 1994).

Some scholars of the information-based market manipulation conducted extensive theoretical and empirical researches. Kyle (1985) built a theoretical model and studied optimal trading strategies of investors who have inside information in the case of continuous auction, which provide a theoretical basis for the subsequent researches. Back (1992) extended Kyle's model, and made the equilibrium pricing rules on the basis of wider distribution of the assets value in the market. Chakraborty and Yilmaz (2004), based on Kyle's research framework, established the market order correction mode under the conditions of discrete-time, analyzed the balance of the different periods during the information-based market manipulation, and drew a conclusion that informed traders will manipulate the market in the case that private information will take longer time to be revealed. Jiang et al. (2005) did an empirical study of how information-based manipulation influences stock market development, and found that most of the manipulated stock will be ignored by investors over a long subsequent time, thus suggesting that the stock market manipulation is against development. Pan et al. (2003) studied market manipulation of traders who do not fully informed for the quotation-driven market, obtaining the conclusions that informed traders can reverse the manipulation under certain conditions. Zhang and Qiu (2005) used evolutionary game theory combining with perspective of behavioral finance to explore how the market manipulators control the situation by using irrational feedback behavior of retail investor.

From the works above, we find that although information-based market manipulation have been studied by many scholars, there are few studies on information-based manipulation for specific conditions of Chinese market, especially those which take restriction of financing and security loan and investor composition into consideration. Therefore, the aim of this paper is to establish a theoretical model of the general process of market manipulation on the ground of type of Chinese stock market investors and way of financing and securities loan.

33.2 Model Description

33.2.1 Process of Market Manipulation

In Chinese stock market, the manipulator takes many means to manipulate stock by using insider information. For generality, this paper will establish a model of market manipulation based on the general process of market manipulation of the operator informed of insider information. The general process of market manipulation includes the following stages:

1. *The accumulation stage*
   At this stage, the chips are scattered, so no investors can manipulate the market price. The operator informed of the inside information mostly takes a gradual

and slow step to absorb the stock to establish positions for ensuring their ability of manipulation of stock price. For simplification, the model of this paper does not describe the phase significantly, but ensure the ability to maneuver through the establishment of certain restrictions on the stock price manipulators.

2. *The pulling up stage*
   This stage is described in the first period of the model. At the stage, manipulators already have the ability to manipulate stock prices, and pull up the share price high from the cost to reduce the profit pressure and improve profit margins. At this stage, they tend to strictly prevent the leakage of information to prevent other investors (especially rational traders represented by institutional investors such as funds etc.) from awareness. Just as what the model showed, margin deposit limits, rational degree of semi-rational noise traders and information closer to truth and so on, have different effects on the current stock price manipulation.
3. *The shipping stage*
   The second period of the model is a simplification of this stage. Through the manipulation of the previous stage, manipulators have created larger profit margins. The paper emphasizes how manipulators make profits over the shipping of the semi-rational noise traders at this stage, and perceives "free ridings" trading of the passive rational traders.
4. *The recovery stage*
   This stage focuses on the recovering process to the normal price range of the manipulated stock after the shipping back of stock manipulators. But to ensure the integrity of the model, this phase is described in the third period of the model though it does not belong to manipulation phase.

33.2.2 Assumptions

Suppose there are two assets in the market: cash and stock. Cash is no net gain and the stock's net supply is 0. There is the total of four periods in the market as $t = 0, 1, 2, 3, 4$. In the third period, the two stock-holding sides liquidate in accordance with the fundamental stock value $P_0 + \Psi$ and pay the dividends θ. Among them, P_0 is the stock price for the period 0, which can fully reflect the fundamental value of current share. Ψ is the fundamental change of the share value at the beginning of the first period, the information relevant to which is issued in the first period, and we take the possible values of Ψ for $\psi, 0, -\psi$. The fundamental value of shares has no change in other periods. Bonus θ is subjected to normal distribution $N(0, \sigma_\theta^2)$, and all the market participants can learned the value of θ only in period 3.

33.2.2.1 Investor Types and Trading Strategies

According to the general investment features of Chinese stock market investors, this article assumes that there are three categories of investors on the market with similar transactions among the same investors. They are respectively as follows:

(a) *Informed traders*
Informed traders (denoted by i) are rational traders and can be well informed of the accurate signals of Ψ in the period $t = 1$, whose number scale is μ. Such investors are risk-averse investors, and risk aversion coefficient is Υ. The purpose of their transactions is to meet largest expected utility of liquidation in the third period. Assume that the utility function is a negative exponential function $U(W) = e^{-\Upsilon W}$ of CARA (Constant Absolute Risk Aversion), where W stands for the informed traders' resource endowments. Informed traders' maximum requirements on the expected utility in fact meet the mean-variance expression. Therefore, to ensure the maximum expected utility in period 3, informed traders' demand selected in period 2 is given by

$$\begin{aligned} D_{i2} &= \frac{E(P_3) - E(P_2)}{2\Upsilon\sigma_\theta^2} = \frac{P_0 + \Psi - E(P_2)}{2\Upsilon\sigma_\theta^2} \\ &= \alpha[P_0 + \Psi - E(P_2)], \end{aligned} \tag{33.1}$$

where $\alpha = \frac{1}{2\Upsilon\sigma_\theta^2}$ is the function slope of rational investors' demand. $E(P_2)$ and $E(P_3)$ are expectations of share price in period 2 and 3 respectively. In the process of actual transactions of securities market, such informed traders are mostly traders informed of insider information, for example, the broker, funds and so on.

(b) *Passive rational investors*
The number scale of passive rational investors (denoted by r) is $1-\mu$. In period 1, they do not know the existence of the signal related to Ψ, and only trade with passive investment strategy, that is, compared to their recognized value of the stock, such investors buy in low price and sell in high price. The demand in period t is subjected to the balance between its current stock price P_t and the value of the stock $E(P_t|I_t)$ determined by the current information set I_t. The setting is similar to the assumptions of the positive feedback trading model and the model related to market manipulation established in De Long et al. (1990) and Aggarwal and Wu (2006) respectively. Both passive rational investors and informed traders are rational traders, so let's make their demand function slopes same, which is pointed out in Hirshleifer et al. (1994). The demand of such investors in period t is

$$D_{rt} = \alpha[E(P_t|I_t) - P_t]. \tag{33.2}$$

In the process of actual transactions of securities market, such investors are mostly liquidity demanders or non-informed institutional investors, for example, the fund, QFII and brokerage which have not received inside information. Shi (2001) showed that the trading behavior of the institutional investors in China represented by investment funds were basically rational.

(c) *Semi-rational noise traders*
Semi-rational noise traders (denoted by f) are the combinations of noise investors with positive feedback trading characteristics and rational investors of value investment characteristics, whose number scale is 1 and demand function in period t is

$$D_{ft} = (1-k)\beta(P_{t-1} - P_{t-2}) - k\alpha[E(P_t \mid I_t) - P_t], \tag{33.3}$$

where $\beta(P_{t-1} - P_{t-2})$ is the demand function of noise investors with positive feedback trading characteristics, β stands for the slope of the demand curve. The biggest characteristic of positive feedback trading is that investors just simply extrapolate the trend of the stock price in the past for trading but do not consider such new information as that of current trading. That is, only the pre-market price changes affect the demand for the current period. In the actual transaction process, positive feedback traders are those who buy when the price increases and sell when the price decreases. $\alpha[E(P_t \mid I_t) - P_t]$ is the demand function of rational traders based on the investment value strategy, which has the same set as Eq. (33.2). k $(0 \leq k \leq 1)$ stands for the ration level of semi-rational noise traders. The greater k denotes the trading behavior of semi-rational noise investors is more rational. Thus, semi-rational noise traders' demand function combined positive feedback noise traders' demand function with rational value investors' demand function. The basis of its establishment is, when the securities market was first established, medium and small retail investors often have positive feedback trading characteristics. However, as the stock market improved, their trading strategy presents a rational trend. Xie (2004) confirmed this view by studying China's stock market. Therefore, we think the investment of medium and small retail investors have features of positive feedback and rational investment. The setting can better reflect the characteristics of medium and small retail investors in Chinese stock market, which is different from previous simple setting of noise traders. To ensure the informed traders' ability of price control, this paper makes $\alpha > \beta$ on the basis of relevant theoretical assumptions of De Long et al. (1990). In addition, the number setting of the three kinds of investors is only simplified, and has no significant effect on the analysis results.

33.2.2.2 Margin Deposit Setting

Assume that investors can participate in the margin of market. For the margin limitation, the investor must pay the security deposit whose quantity is determined by the stock price at the end of the last period and the applied number of transactions in the current period. Suppose margin ratio is e in period t. Because of different initial resource endowments, some investors failed to submit sufficient margin with their small initial resource endowments and are forced to reduce transaction needs. Obviously, the informed traders and passive rational traders are mostly institutional investors, with larger initial resource endowment and fewer margin deposit constraints; while semi-rational noise traders, mostly medium and small retail investors with less initial resource endowment face more margin constraints. For simplicity, this article assumes that margin constraints have influence only on semi-rational noise investors.

33.2.3 Transaction Process

At the beginning of the first period, when informed traders receipt the signal ε ($\varepsilon \in \{\psi, -\psi\}$), they have two manipulation methods: (a) pulling up stock price by buying a large number of shares in the first period, selling the stock at a higher price in the second period and liquidating shares in the third period; (b) pushing down stock price by selling a large number of shares in the first period, buying stock at lower price in the second period, and liquidating shares in the third period. The present paper discusses the market situation under the first trading strategy.

1. *Period* 0
   This period is reference to provide basis for the subsequent transactions. There is no information disclosed and transactions occurred.
2. *Period* 1
   (a) *Informed traders*
   In this period, informed traders i receive signal ψ about the Ψ. It makes informed traders be better to master the expected price in the liquidating period, consequently ensure that the informed traders can gain greater expected utility than the investors without insider information for market manipulation. The paper proposes "signal closer to truth" included noise information, which combine with that private information is imperfect obtained by informed traders in actual market. Its characteristics lie in that compared to the real information, the signal closer to truth doesn't emerge excessive deviation and jumping although which contains noise information. It is guarantee of obtaining information superiority for the informed traders. This means that when informed traders gain $\varepsilon = \psi$, the signal closer to truth meets the following expressions:

$$\begin{aligned}
&\Pr\ ob[\Psi = -\psi \mid \varepsilon = \psi] = 0, \\
&\Pr\ ob[\Psi = \psi \mid \varepsilon = \psi] = \lambda, \\
&\Pr\ ob[\Psi = 0 \mid \varepsilon = \psi] = 1 - \lambda, \\
&\Pr\ ob[\Psi = -\psi \mid \varepsilon = -\psi] = \lambda, \\
&\Pr\ ob[\Psi = 0 \mid \varepsilon = -\psi] = 1 - \lambda.
\end{aligned} \tag{33.4}$$

We let $\lambda\psi < P_0$ for considering that the expected value of bad news information generally does not exceed the basic value of stock in actual transaction process. This setting doesn't have any important effect on the conclusion just for the simplicity of analysis.

In order to ensuring the expected utility of the third period, the informed traders i determine the demand D_{i1} to pull up the market price according to the receiving signal ε in this period. We only deal with the situation of the positive effect $\varepsilon = \psi$ on the basic value. It is similar when $\varepsilon = -\psi$.

(b) *Passive rational investors*

The passive rational investors only trade according to the current price and the present information's value in the first period because the rational traders neither understand private information nor judge whether private information emerges. The investors' demand in period 1 is as follows:

$$D_{r1} = -\alpha(P_1 - P_0). \tag{33.5}$$

(c) *Semi-rational noise traders*

The paper sets that the semi-rational noise traders do not trade in this period for considering that the medium and small retail investors tend to trade when stock prices have comparative rising or falling, thus probability of the trading is small in this time when current price changes are small. The present paper sets that the semi-rational noise traders have no transactions, and this setting will not have important impact on the key conclusions.

3. *Period* 2

In practice, informed traders can find out true information by the further study of inside information and market condition. Luo (2005) established theoretical model and then concluded that inside trader can modify private information through the market situation which is proved by the American stock market data. Passive rational traders can also be based on the further research of the transaction as well as information to understand the real information, which is then theoretically confirmed by Damodaran and Liu (1993). Therefore, this article assumes that in the second period informed traders and passive rational traders have known the true information and semi-rational noise traders have also obtained partial information, acquired extent of which is determined by the degree of rationality. At this point, the traders of different types have the following trading requirements.

(a) *Informed traders*
Informed traders have pulled up stock price in the first period. The trade purpose of current period lies in attracting the trade of semi-rational noise traders by pulling up share price and making profit by shipping to ensure the maximum expected utility in the third period. Equality (33.1) gives the current number of transactions D_{i2}.

(b) *Passive rational traders*
Passive rational informed traders have achieved precise information and also found the case of market manipulation of informed traders. Now, on the basis of value investment strategy, they make profit by pushing up the stock profits of informed traders. According to Eq. (33.2), the demand of such investors is illustrated as follows.

$$D_{r2} = \alpha(P_0 + \Psi - P_2). \tag{33.6}$$

(c) *Semi-rational noise traders*
The great raising of stock prices in the first period attracts the trade of semi-rational noise traders, which provides a basic prerequisite for the profits of informed investor. Due to the margin deposit limits, the demands of such investors are given in the following equation from (33.3) and (33.4):

$$D_{f2} = \frac{1}{e}\left[(1-k)(P_1 - P_0) - k\alpha(P_0 + \Psi - P_2)\right]. \tag{33.7}$$

4. *Period* 3
In this period, inside information Ψ becomes public information of market and listed company also announces dividend information. The investors liquidate the holding stock in accordance with $P_3 = P_0 + \Psi + \theta$ in this period.

33.2.4 Model Solution

Based on the above assumptions and the various trading strategies, we can obtain the price and quantity of market trading in each period, and then get the results below.

Conclusion 33.1 under the same condition of margin deposit limits, the optimal price of informed traders in the first period is described by

$$P_1 = P_0 + \frac{\lambda\psi}{\dfrac{\lambda(1-\lambda)\psi^2(1-\mu)}{\sigma_\theta^2\mu} - \dfrac{\beta(1-k)}{\alpha(k+e)} + 1}. \tag{33.8}$$

33.3 Model Analysis

33.3.1 Margin Deposit Ratio Analysis

When there is the deposit, we seek the partial derivative of P_1 about e and get:

$$\frac{\partial P_1}{\partial e} = \frac{-\lambda\psi}{\left[\frac{\lambda(1-\lambda)\psi^2(1-\mu)}{\sigma_\theta^2\mu} - \frac{\beta(1-k)}{\alpha(k+e)} + 1\right]^2} \times \frac{\beta(1-k)}{\alpha(k+e)^2} < 0. \tag{33.9}$$

Conclusion 33.2 under the condition of the margin deposit limits, proportional change of deposit has important influence on the price of the first period for informed trader.

The expression (33.9) means that the proportion decrease of margin will cause price increase of the first period. This is because the lower margin ratio will increase the demand of semi-rational investor, which causes the issue of price rise.

But considering the actual situation of the securities market, it is uncommon that the lower margin ratio e leads to the big differences of price P_1 and P_0. The reason in fact is that raising price of informed traders will face three practical difficulties: (1) The excessive increase and abnormal change of price will cause the attention and investigation of securities supervision department, (2) Actual investor usually has "S" type of risk aversion characteristic. When the stock price is abnormally high, whether rational or semi-rational investors are reluctant to approach trading; (3) In order to reduce the risk of market, supervision department set higher margin ratio. This means that in the practice securities market, the higher margin deposit proportion are more likely to increase the difficulty of market manipulation of informed traders, so that market price is not too deviate from the real market value, which plays an important role in effective protection of minority investors.

33.3.2 Rationality Degree Analysis of Semi-rational Noise Traders

For $k > 0$, there is incomplete rational trading behavior for semi-rational noise traders. For the partial derivative of P_1 about k, we get

$$\frac{\partial P_1}{\partial k} = \frac{-\lambda\psi}{\left[\frac{\lambda(1-\lambda)\psi^2(1-\mu)}{\sigma_\theta^2\mu} - \frac{\beta(1-k)}{\alpha(k+e)} + 1\right]^2} \times \frac{\beta(1+e)}{\alpha(k+e)^2} < 0. \tag{33.10}$$

Conclusion 33.3 rational level of semi-rational noise traders has effect on the price of the first period for informed trader.

The economic significance of (33.10) is very clear that the increasing k will result in the decreasing of P_1. According to the Eq. (33.3), it is obvious that the increase of k will reduce the demand of semi-rational noise traders in the second period, which reduces the expected utility of informed traders. If informed traders increase P_1 at this time, then this will result in two results: (1) It does not affect the stock trading volume of semi-rational noise traders determined by positive feedback noise trading characteristics; (2) It reduces the number of stock transactions which decided by rational trading characteristics. Combined with the actual situation and the model results, we believe that a higher rational level of semi-rational noise traders contributes to inhibition of market manipulation. This means that we can enhance the level of rational investment for small and medium-sized retail investors to prevent market manipulation through the education of investment and increase of information disclosure and supervision.

33.4 Conclusions and Suggestions

The paper creates the market manipulation model which conforms to Chinese market informed investor's types and traders' margin deposit limits. We analyze the trading strategies of informed traders, passive rational investors and semi-rational noise traders represented by manipulator, fund investors and medium and small retail traders respectively through introducing margin deposit system, and obtain the equilibrium results for market manipulation in every period. We further analyze the influences of the margin deposit limits and the rational degree of semi-rational noise traders on the equilibrium outcomes. So we draw the following conclusions:

1. By analyzing the effects of margin deposit limits on the market price, we find that margin deposit plays an important role in inhibiting market manipulation behavior. Improving the proportion of e can increase the market manipulation difficulty to the informed traders through setting margin deposit limits for the small and medium-sized;
2. The authorities should enlarge information disclosure extent, prevent insider information leakage and increase margin ratio. On the one hand, it helps the semi-rational noise traders to promote their rational degree. On the other hand, it can increase the facing information risk of informed traders to reduce insider information leakage from the listed company and increase the manipulation's difficulty of informed traders, which are both helpful for preventing the market manipulation.

References

Aggarwal R, Wu G (2006) Stock market manipulation. J Bus 79(4):1915–1953
Allen F, Gale D (1992) Stock price manipulation. Rev Financ Stud 5(3):131–171
Back K (1992) Insider trading in continuous time. Rev Financ Stud 5:387–410
Chakraborty A, Yilmaz B (2004) Informed manipulation. J Econ Theor 114:132–152
Damodaran A, Liu C (1993) Insider trading as a signal of private information. Rev Financ Stud 6(1):79–119
De Long J, Shleifer A, Summers L, Waldmann RJ (1990) Positive feedback investment strategies and destabilizing rational speculation. J Financ 45(2):379–395
Foster D, Viswanathan S (1994) Strategic trading with asymmetrically informed traders and long-lived information. J Financ Quant Anal 29:499–518
Hirshleifer D, Subrahmanyam A, Titman S (1994) Security analysis and trading patterns when some investors receive information before others. J Financ 49(5):1665–1698
Jiang G, Mahoney P, Mei J (2005) Market manipulation: a comprehensive study of stock pools. J Financ Econ 77:147–170
Kyle A (1985) Continuous auctions and insider trading. Econometrica 53(6):1315–1365
Luo Y (2005) Do insiders learn from outsiders evidence from mergers and acquisitions. J Financ 60(4):1951–1982 (in Chinese)
Pan D, Zou Y, Liu HL, Wu CF (2003) Strategic trading when some investors receive partly information. Syst Eng Theor Pract 23(10):22–29 (in Chinese)
Shi DH (2001) The trade practices and market impact of securities investment funds. World Econ 10:26–31 (in Chinese)
Xie HY (2004) An analysis on the movements of treasury yield in China. Stud Int Financ 7:65–72 (in Chinese)
Zhang YP, Qiu PG (2005) Evolutionary game analysis of market manipulation process. Syst Eng Theor Pract 25(3):214–222 (in Chinese)

Chapter 34
Research on Contractor's Claims Caused by Extension of Time

Gong-jie Ao, Xiao-li Zhang, and Xiao-dong Zhang

Abstract In the process of construction, project schedule delay often occurs due to various reasons. So it's especially important for contractors to make reasonable use of claims to compensate for project loss and improve economic performance. Base on that, the paper stood in the perspective of the contractors, dividing the responsibility of the two parties in the claims caused by extension of time, further identified factors that induced extension of time and determined the key factors by SPSS. Then the paper analyzed the calculation methods of claim for extension of time and cost in terms of the key factors, finally confirmed the specific claim content and calculation for contractors once extension of time appeared.

Keywords Contractor's claims • Cost claims • Calculation methods • Extension of time

34.1 Introduction

Along with the competition in the construction industry becoming more and more serious, the construction companies faced a severe situation in economic performance and their opportunity for gaining profits shrank day by day. In 2010, China Railway Construction Co. Ltd announced their huge losses of 4.153 billion

G. Ao (✉)
School of Management, University Du Québec A Chicoutimi (UQAC), Chicoutimi, Canada
e-mail: ao_gongjie@126.com

X. Zhang
Management School, Tianjin University of Technology, Tianjin, China
e-mail: utopiaxiaoli@126.com; utopiaxiaoli@163.com

X. Zhang
China Petroleum Pipeline Bureau Tianjin Design Institute, Tianjin, China
e-mail: luckxiaodong@163.com

E. Qi et al. (eds.), *The 19th International Conference on Industrial Engineering and Engineering Management*, DOI 10.1007/978-3-642-37270-4_34,

Yuan in Saudi Arabia's light rail project; In 2011, China Overseas Engineering Group Co. Ltd of CREC will also face the same losses about 2.545 billion Yuan in Poland project, which all proved the severe situation.

While in the progress of construction, it often appeared that the extension of time caused by sorts of causes. So the contractor who shall survive in such environment should pay more attention to the corresponding claim from the owners and protect their legitimate interests in order to make up for engineering loss and improve the economic benefit (Gou and Li 2011).

Base on that, the discussion of the influence factors of claim risk caused by extension of time and making clear of the contractors' claim calculation methods become the key problem to be solved, which provides the necessary guidance for successful claim to the contractor.

34.2 The Principle of the Responsibility in Extension of Time

34.2.1 The Principle of the Responsibility in Single Extension of Time

In general, according to the different undertakers, the single extension of time can be classified by Non-excusable Delays (NE)—Reason for contractor; Excusable and Compensable Delay (E&C)—Reason for owner; Excusable but not Compensable delays (ENC)—External reason. In accordance with relevant literature (Li Jianshe and Lv Shengpu 2005), the principle of the responsibility in single extension of time can be summarized in Table 34.1.

Table 34.1 Classification of the single extension of time

Non or excusable	Reason for extension	Undertaker	Principle
Excusable	1. Design change 2. Construction conditions change 3. Owner 4. Engineer	Owner	Compensation for time and cost
	1. Severe weather 2. Natural disasters 3. Social factors	External	Compensation for time
Non-excusable	1. Equipment, material 2. Undeserved plan and lack of experience	Contractor	Non-compensation for time or cost

34.2.2 The Principle of the Responsibility in Joint Extension of Time

Actually, in the process of construction, the extension of time was often caused by two or more reasons such as the contractor, the owners or objective reason (Lv Shengpu et al. 2004). However, it's difficult to divide the responsibility of the interference events when joint extension of time happened. Nowadays there were four kinds of principles in such extension: The initial event principle, Contractor-Against principle, Correlation method principle; Responsibility-Share principle (Huang Yao and Wu Shiming 2011a, b; Zhu Jianguo et al. 2012).

34.3 Factors of Contractors' Claims Caused by Extension of Time

34.3.1 Analysis of Risk for Contractor

The whole process of construction projects may encounter many risk, which generally divided into the owner risk and the contractor risk also include some other risk.

Risk of contractor means all the risk in the process of the construction projects except the owner's risk (Zhang Shuibo and He Boshen 2003). In different stages, the contractor undertakes the different risk. So the contractor should pay more attention to the whole process in the construction in order to ensure the project to reach the acceptance standard. The risk for contractor is concluded in project management risk, material management risk, cost management risk, contract ability risk and so on.

The contractor may encounter many other risks that include design risk, contract risk, people risk and so on (Wang Kuan 2005).

34.3.2 Analysis of the Key Risk Factors Due to Extension of Time

34.3.2.1 Factors of Extension of Time

In the process of project implementation, there are many factors that influence the construction period (Yao Xiaogang and Ling Chuanrong 2001). To effectively control the time limit, we must analysis and predict the factors that influence the time. In terms of factors that influence the construction of the project period, they can be divided for man, technology, materials, equipment and components, machinery, capital, hydrology, geology and weather, environmental, social, and other unpredictable factors, the factor of people for the most. According to the related literature, the main factors that influence extension of time focused on the relevant units (the construction unit, design units and other stakeholders), change of the condition, technical errors, poor construction management and the emergence of the accident (Chenghu 2003).

34.3.2.2 The Key Factors of Extension of Time

We should choose the factors of extension of time the contractors concerned about from the above five kinds of risks. Based on this, the paper classifies 31 kinds of the reason while each factor divides into five levels according to the influence degree in project. That's very not important, not important, doesn't matter, general important, very important. The standard of grading is 1, 2, 3, 4, 5.

Based on this, the paper extracts several random construction units, consulting unit and issues 450 questionnaires, which contains 100 questionnaires of the contractor, engineers, and developers. According to the importance role in the questionnaire of qualification and unit in the project, the paper screened 100 questionnaires and received 78 valid questionnaires. Then put Likert scale method into use to score for the factors influence the delay and combined with SPSS to make the quantitative calculation, finally analyzed by the Mean Score.

$$MS = \Sigma\,(fi\ S)\,/N$$

MS: Mean Score of the factors

fi: Numbers of delay factors in degree

i: Influence degree (i = 1, 2, 3, 4, 5)

S: Weight

N: The total numbers of the factors

34.3.2.3 The Results of Key Factor of the Extension of Time

The paper made use of Mean Score method to calculate the formula while analyzed the reason of time delay and sorted the results, eventually got the most four factors that didn't belong to contractor but led to time delay. They are Interim payment not-in-time, Design drawing change, Supply of material equipment default, Force majeure.

34.4 Calculation Methods of Claims Caused by Extension of Time

34.4.1 Calculation Method of Time Claims Caused by Extension of Time

According to the calculation methods of time claim, scholars had put forward calculation methods from different angles. Li Qiming and Zhu Shuying came up with network analysis and contrast analysis method from the angle of the interference of events, They thought network analysis is a scientific and reasonable calculation method and the period claim value could be calculated through the difference between before and after of interference events (Li Qiming et al. 2001). Li Zhengzui presented the most scientific time calculation method for Critical Path Method (CPM) (Li Zhengzui 2007). Chen Yongqiang and Zhang Shuibo classified into whole influence method, plan and actual schedule contrast method, But-for method, plan influence method, progress focus tracking method, time influence method, the ratio analysis method from the angle of difficulty of claims (Chen Yongqiang and Zhang Shuibo 2008). Based on this, this paper analyzes the calculation method of commonly used time claim, as shown in Table 34.2.

34.4.2 Calculation Methods of Cost Claims Caused by Extension of Time

In terms of the extension of time caused by owner, the contractor put forward for cost claims while time claims.

Calculation Methods of cost claims included actual cost method (sub item method), total cost method, the fixed cost method, reasonable value method and

Table 34.2 Calculation methods of time claims

Designation	Advantages	Disadvantages
Network analysis method	A scientific and reasonable calculation method	Apply in large, complex project, the manual calculation is so difficult that need computer to complete.
Ratio analysis method	Simple, convenient, easy to understand and accept	Not science, reasonable, sometimes do not conform to the actual project
Direct method	Simple and easy to calculate	Need a detailed record of the construction, the engineer should confirm, otherwise easy to cause the incident

Table 34.3 Calculation methods of cost claims

Designation	Characteristics	Basic ideas
The actual cost method	1. Rational and scientific	1. Analysis each affected claim events project, no omission
	2. Convenience for claim evaluation, negotiation	2. Calculate each cost project and compare with the cost value in contract sum, get claim value
	3. Easy to accept Difficult, more complex than total cost method	3. Collect the claim value and get the claim value of the total projects
The total cost method	Most simple 1. Reasonable, reflect the real situation	Put the fixed price contract into cost plus fee contract, the claim value is based on contractor's cost plus management fees and profit
	2. The nature of the contract dispute shall not apply to other calculation	The contractor make their own internal records and documents the actual costs and the difference with contract is the claim value

trial found method, etc. The most common method is the actual cost method and the total cost method (Zhu Jianguo and Shen Jie 2008). Based on this, the paper made analysis of the advantages and disadvantages for the two kinds of methods, as shown in Table 34.3.

Based on the analysis of the two kinds of methods, the paper chooses the actual cost method to calculate claim value of time delay.

34.5 Calculation Methods of Extension of Time Caused by the Key Four Reasons

34.5.1 Claims of Interim Payment Not-in-Time

The regulations of international engineering contract conditions are the time of payment. In terms of interim payment, the owner should pay in 28 days after the engineer submitted the signature of the monthly statement to the owner (Qiao and Shi 2006).

In face of the interim payment not-in-time, the contractor has two ways. One is to mat and continue to engineering construction. They have the right to require the owner to delay time and pay them interest rate. The other is to suspend construction and the owner should assume the period and economic loss.

1. Claims process of interim payment not-in-time, as shown in Fig. 34.1.
2. Cost calculation

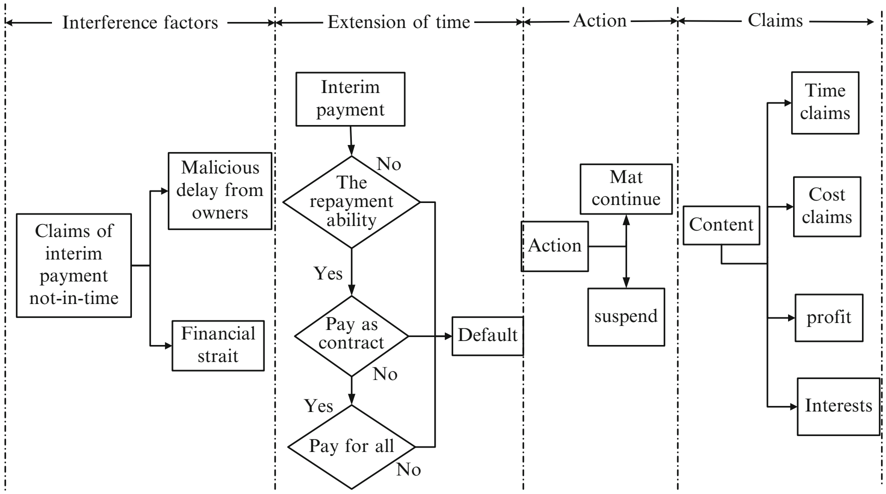


Fig. 34.1 Formation of payment delays

(a) Labor fee = (actual days − days of finished project − other days − loss of labor caused by contractor's risk) × labor unit price;
(b) Materials fee = (current price − basic price) × materials amount
(c) Machine fee = stagnation amount × machinery unit price;
(d) Site management fee = (site management fee in the contract ÷ contract time) × delay time × (finished workload ÷ contract work);
(e) Headquarters management fee = day of management fee × management fee shared in dispute contract ÷ dispute contract execution days;
Claim value of management fee = day of management fee rate × delay time of contract
(f) Interest: I = (C2 − C1) [(1 + P) n–1];

I – Interest of claim for contractor;
C1 – Project amount of contractor from owner in actual period;
C2 – Project amount of contractor that owner should pay for in terms of contract (1 month)
P – Month rate of bank (as the two parties agreed in the contract)
n – Management fee of the drag payment

34.5.2 Claims of Design Drawing Change

Engineering design is the root of the project quality. Before construction, design or drawings should be delivered to the owner for approval, but in many cases, the contractor does not have the ability to review and make the decision by the explanation of designers. If the mistakes happed sometimes may cause damage to engineering. Although the designer should bear legal responsibility for the mistake, it can't mean contractor hadn't suffered from the loss, especially the plan hard to carry, which led to engineering delay that caused heavy loss to the contractor.

The cost calculation of design drawing change is followed as:

1. Extra work fee = labor unit price × work coefficient
2. Labor fee of extra work = labor unit price, day unit price or agreed unit price in contract
3. Materials fee claims = materials consumption + material unit price;
4. Materials fee = (current price − basic price) × materials amount
5. The increased material transportation, purchase, storage costs = the actual cost − quotation cost.
6. Idle machinery fee = machinery depreciation fee × idle time
7. The increased rental machinery fee = the actual of rental machinery × continued time
8. Mechanical operation efficiency lower fees = actual cost of mechanical operation − the planed bid price
9. Site management fee = site management fee rate × basic fee; Headquarters management fee = headquarters management fee rate × (basic fee + Site management fee)

In the project of increased quantities and additional engineering, offers already include profit and the value of claim included profits.

34.5.3 Claims of Supply of Material Equipment Default

Some owners are responsible for purchasing materials by themselves for saving engineering spending or taking care of their relationship. In this case, the business of materials often focus on their interests instead of considering the owner's need, which caused project stopping work, or due to material not to rework required.

The Cost calculation of supply of material equipment default is followed as:

1. Labor cost = (actual days − days of finished project − other days − loss of labor caused by contractor's risk) × labor unit price;
2. Materials cost = (current price − basic price) × materials amount
3. The increased material transportation, purchase, storage costs = the actual cost − quotation cost.
4. Idle machinery fee = machinery depreciation fee × idle time
5. Site management fee = (site management fee in the contract ÷ contract time) × delay time × (finished workload ÷ contract work);
6. Headquarters management fee = day of management fee × management fee shared in dispute contract ÷ dispute contract execution days;

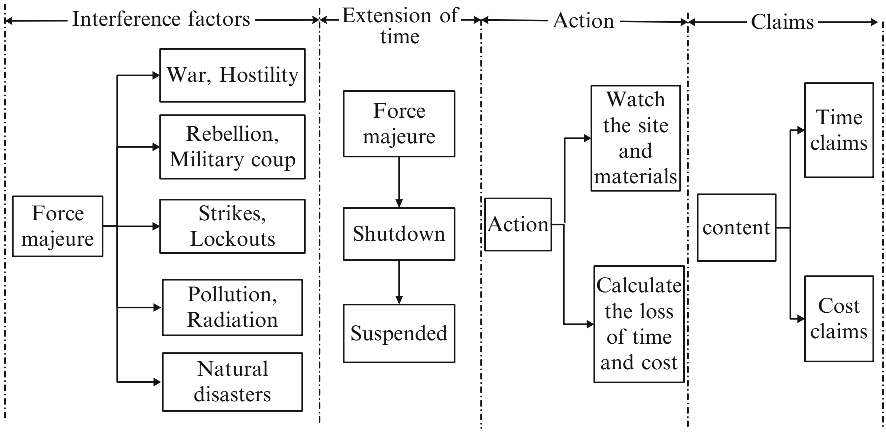


Fig. 34.2 Formation of force majeure delays

34.5.4 Claims of Force Majeure

1. The claims process of force majeure is shown in Fig. 34.2.
2. Cost calculation

In the insurgency, riots, military coup, the contractor can claim for the reason of commotion, noisy, chaos, strike and pollution, radiation, explosions in the delay.

1. Personnel shutdown fee = personnel unit price × days to cease;
2. Materials fee = increased material price + material breach fee
3. Mechanical shutdown fee = mechanical unit price × days to cease;
4. Site management fee = (site management fee in the contract ÷ contract time) × delay time × (finished workload ÷ contract work);
5. Headquarters management fee = day of management fee × management fee shared in dispute contract ÷ dispute contract execution days.

34.6 Conclusion

The paper is established in the contractors how to make use of claims to protect their interests when the extension of time happened in the process of construction, then analyzed the principle of the responsibility in extension of time. Based on this, explicated the risk of the contractor should bear, further identified factors that induced extension of time and determined the four key factors by SPSS, they are interim payment not-in-time, design drawing change, supply of material equipment default, force majeure that caused time delay. Finally the paper uses the calculation methods of time and cost claims for the four reasons, which provide the strong support for the contractor to claim.

References

Chen Yongqiang, Zhang Shuibo (2008) International project claims. China Architecture and Building Press, Beijing

Chenghu (2003) Contract management and claim for construction project. Architecture and Building Press, Beijing

Gou Lina, Li Jianfeng (2011) Analysis on claim opportunity and research on claim for extension of time of contractor under FIDIC contract. Chang'an University

Huang Yao, Wu Shiming (2011a) Discussion of calculation method of time claim in international engineering project. Proj Manag Technol 9:47–50

Huang Yao, Wu Shiming (2011b) Research on the calculation of delay claims in international engineering project. Proj Manag Tech 9(6):47–50

Li Jianshe, Lv Shengpu (2005) Claim method and examples of civil engineering. China Communications Press, Beijing

Li Qiming, Zhu Shuying, Huang Wenjie (2001) Construction project contract and claims management. Science Press, Beijing

Li Zhengzui (2007) Discussion of contractors' claims under FIDIC. Des Hydroelect Power Stn 2:99–104

Lv Shengpu, Li Jianshe, Yang Feixue (2004) Calculation methods of claim under many events disturbance. Constr Technol 12:16–18

Qiao Baihe, Shi Yongming (2006) Research on the claims management and delay claims. Chang'an University

Wang Kuan (2005) Building project risk theory and case studies. University of International Business and Economics, Beijing

Yao Xiaogang, Ling Chuanrong (2001) Complication study on compensate time limit for project under double factors affection. J Tongji Univ 3:362–365

Zhang Shuibo, He Boshen (2003) Allocation and sharing of construction risks between contract parties. J Tianjin Univ 7:257–261

Zhu Jianguo, Shen Jie (2008) Study on the principle of calculation of delay claims for project under multi-factor affection. Constr Econ 8:90–92

Zhu Jianguo, Xu Wei, Shen Jie, Cheng Hu (2012) Calculation on construction delay responsibility sharing with multi-event interference. J Southeast Univ 1:188–193

Chapter 35
The Study on Sewage Treatment BOT Project Concessionaire Selection Criteria: Best Practice from China

Wei-hua Yang, Ya-nan Yang, Feng-hai Zhang, and Xiu-shan Wang

Abstract Concessionaire selection method has drawn much attention worldwide recently. The government of Dalian has developed well-structured concessionaire selection criteria in China. This paper analyzes and summarizes experience from this concessionaire selection practice to improve procurement efficiency of regions lacking such experience or methods.

Keywords BOT project • Best practices • Concessionaire selection • Sewage treatment

35.1 Introduction

BOT (build-operate-transfer) has been used worldwide in the infrastructure and public utility project delivery. Government needs to select concessionaire who finances, designs, and builds the project and then operates it for a specified concession period. The concessionaire is usually a private sponsor with whom government develops long-term partnerships and shares risks (Zhang et al. 2002). As a core member of BOT stakeholders, concessionaire assumes more responsibilities and shares more risks than does a traditional contractor of public project (Liou and

This research is supported by the National Natural Science Foundation of China (Grant No. 71002076).

W. Yang (✉) • Y. Yang • F. Zhang • X. Wang
School of Management, Dalian Polytechnic University, Dalian, People's Republic of China
e-mail: Jasonyangwh@sina.com

E. Qi et al. (eds.), *The 19th International Conference on Industrial Engineering and Engineering Management*, DOI 10.1007/978-3-642-37270-4_35,

Huang 2008). Therefore, it is typically important for the government to select appropriate concessionaire, which is the key to steadily operate BOT projects, and is the assurance for citizens to obtain high quality public products and services.

How to select concessionaire has attracted attention from academia, industry and government. For example, Tiong described the importance and characteristics of six critical success factors for project sponsors to win BOT contracts (Tiong et al. 1992). Tiong argued that an attractive financial package, high level of equity and risk allocation were critical under the conditions when the project was technically certain and the project financing were uncertain (Tiong 1995a, b, c). Tiong suggested that the promoter must fully understand the government's needs and concerns and be able to address them through the right package of the winning elements (Tiong and Alum 1997). Zhang identified 92 concessionaire selection criteria through questionnaire survey in Hong Kong and UK (Zhang 2004, 2005). Qiao also reported survey findings concerning critical success factors for tendering the BOT projects in China (Qiao et al. 2002). Ahadzi studied three key attributes of private sector and ranked 24 selection criteria (Ahadzi and Bowles 2004). Torta and Khadaroo also summarized selection criteria, such as technical ability, concession period, contract arrangement and so on, through case study (Torta 2005; Khadaroo 2008). Kumaraswamy designed a unified framework for Public–Private Partnerships team selection decision-making (Kumaraswamy and Anvuur 2008). Dias proposed a multiattribute evaluation model for assessing the capability of a private-sector company's becoming a promoter (Dias and Ioannou 1996).

Although research findings mentioned above are helpful for the government to select BOT concessionaire during a project tender procedure, the selecting criteria and evaluation methods suggested by researchers cannot be applied directly into specific BOT project practice, such as sewage treatment BOT concessionaire selection decision.

Nowadays, it is common that BOT delivery method is applied actively in the sewage treatment industry of China, in order to utilize resources from private sectors to save public expenditure and improve infrastructure efficiency. According to statistics by Ministry of housing and Urban–rural Development of China, 31.8 % of 1,004 sewage treatment projects take the form of BOT contracts until 2008 in China. Dalian, an eastern coastal city, is the earliest to explore BOT delivery method in the sewage treatment industry in China. Fujiazhuang sewage treatment plant of Dalian is the first sewage treatment BOT project in China, which came into operation in 2001. Up to now, there are more than ten sewage treatment BOT projects under operation in Dalian (shown as in Table 35.1), and sewage treatment rate of the entire city has reached at 95 %, and will be 98 % in 2015.

In the future, all sewage treatment projects in Dalian district will be delivered as BOT. Dalian government has acquired good experience in selecting appropriate concessionaire efficiently and effectively. So it is necessary to benchmark the best practices of Dalian in order to optimize the result of public procurement and to minimize the risk of selecting incompetent concessionaire.

Table 35.1 Sewage treatment BOT projects of Dalian

Sewage treatment projects	Year of construction	Design capacity (10,000 t/day)	Investment (million yuan RMB)
Laohutan	2003	8	103.00
Quanshui	2005	3.5	37.64
Lingshui River	2005	6	66.09
Chunliu River phase II	2006	12	115.10
Malan River phase II	2006	8	81.91
Xiajia River	2006	3	41.31
Siergou	2007	12	169.00
Xiaoping Island	2008	2	23.00
Yingchengzi	2008	1	20.00
Hutan New District	2009	3	48.00

35.2 Concessionaire Selection Procedure

35.2.1 Concessionaire Selection Principles Applied by Dalian

BOT is a specific form of Public-Private Partnerships (PPPs) which is characterized by collaboration and cooperation between public and private sectors. Dalian government realized that concessionaires should be regarded as partners rather than contractors assuming most project risks, and concessionaire selection decision should be made based on the comprehensive evaluation taking account of all important attributions of private sectors. Based on the above mentioned reasons, Dalian government designed selection criteria and decision rules according to three principles.

1. Government decisions aim at selecting appropriate long-term partners, realizing the balance of interests between public and private sectors through close cooperation in the whole lifecycle of BOT projects.
2. Concessionaire candidates should be evaluated comprehensively and have main characteristics of good reputation, strong financing ability, reasonable sewage treatment unit price proposal, similar project experiences.
3. Sewage treatment unit price should be an important criterion to evaluate concessionaire, and components of the price should be elaborated in private sectors' bidding proposals.

35.2.2 Selection Procedure Adopted by Dalian

Usually, the concessionaire selection procedure must be formed on the basis of laws and regulations in a host country. Dalian government developed a three-stage selection procedure under the framework of relevant legislation in China, including

"Invitation and Submission of Bids Law of The People's Republic of China", "Contract Law of the People's Republic of China", "Administrative Measures for the Franchise Operation of the Municipal Public Utilities of the People's Republic of China", etc. Dalian government mainly chose open competitive tendering method in the three-stage selection procedure. The first stage is prequalification evaluation through which government can select tenders that meet all the pre-established criteria, such as qualification, performance and reputation, while reject tenders who fail to satisfy any of these criteria. The second and third stage will focus on technical ability and business performance evaluation of tenders respectively. Without passing the preceding stage, a tender has no chance to enter into a succeeding stage.

35.3 Concessionaire Selection Criteria

35.3.1 Concessionaire Selection Criteria Designed by Dalian

Dalian government established a set of criteria evaluating the concessionaire's technical program and financial ability.

According to the three principles of selecting concessionaire mentioned above, evaluation criteria of technical ability can be divided into three dimensions, namely, construction scheme, operating program, transfer plan, shown as in Table 35.2. Likewise, criteria of business performance have three dimensions in Table 35.3. The above-mentioned six dimensions of criteria are further expanded and defined as specific items. Evaluation criteria of technical ability cover all main aspects in the whole lifecycle of BOT projects, while criteria of business performance focus on three different attributions of concessionaires. Such a criterion framework can avoid overlap and similarity between criteria, and embody a comprehensive evaluation principle.

The six dimensions are assigned different weights according to their relative importance as determined by Dalian government. Bid price is the most important factor when selecting concessionaire, and assigned weight of 50 %. The second important criterion is construction scheme assigned weight of 12 %.

Operating program, finance scheme and performance of similar projects are of equal importance, all assigned weights of 10 %.

A concessionaire candidate's final evaluation score is the sum of scores of technical ability and business performance (denoted as S_1 and S_2 respectively), and denoted as S_i (i means the *ith* concessionaire). The score of technical ability is a weighted score, which means that score of every dimension equals to multiplication between total scores of specific criteria belonging to the dimension and corresponding weights, shown as in row 10 of Table 35.2. The *ith* concessionaire candidate's final score can be calculated as equation below.

$$S_i = S_1 + S_2 \tag{35.1}$$

Table 35.2 Evaluation criteria of technical ability

Evaluation dimensions	Evaluation criteria	Maximum scores
Construction scheme (weight: 12 %)	(1) Treatment technology and process	30
	(2) Main equipments	15
	(3) General layout of treatment plant	15
	(4) Construction organization plan	10
	(5) Measures to ensure project quality	10
	(6) Construction progress plan	5
	(7) Environmental impact analysis and protection	10
	(8) Safety and civilization construction	5
	(9) Total score of construction scheme	100
	(10) Weighted score = (9) × 12 %	12
Operating program (weight: 10 %)	(11) Project human resource management	10
	(12) Operation management plan	20
	(13) Equipment management and maintenance	40
	(14) Risk control and safety management plan	30
	(15) Total score of operating program	100
	(16) Weighted score = (15) × 10 %	10
Transfer plan (weight: 8 %)	(17) Technical state security commitments	60
	(18) Security commitments after transfer	40
	(19) Total score of transfer plan	100
	(20) Weighted score = (19) × 8 %	8
Total score	(21) $S_I = (10) + (16) + (20)$	30

Table 35.3 Evaluation criteria of business performance

Evaluation dimensions	Evaluation criteria	Maximum scores
I. Bid price (weight: 50 %)	Evaluation of proposed sewage treatment unit price (evaluation score denoted as T_i)	50
II. Finance scheme (weight: 10 %)	Sources of funds for project financing; capital structure; financing ability; bank support	10
III. Performance of similar projects (weight: 10 %)	Scale, amount, sewage treatment standards of similar projects; construction period, quality and contract performance of similar projects	10
Total score	$S_2 = (\mathrm{I}) + (\mathrm{II}) + (\mathrm{III})$	70

The tender evaluation committee calculates the average score of every concessionaire candidate by deleting the highest and lowest scores based on group decision. The committee will recommend the first three tenders to be potential concessionaires which have rights to negotiate with government about the details of concession contracts.

Table 35.4 Components of sewage treatment unit price

Item	No.	Cost/ton sewage treated
Variable costs	1	Pharmacy cost
	2	Power consumption cost
	3	Water consumption cost
	4	Sludge landfill cost
Fixed costs	5	Amortization charges
	6	Depreciation
	7	Wage and welfare
	8	Maintenance and repair cost
	9	Management fee
	10	Interest
Total cost	11	Total cost (1–10)
		Including: operation cost
Profit and treatment price per ton claimed by tenders	12	Profit
	13	Corporate income tax
	14	Business tax
	15	Other taxes
	16	Treatment price per ton (P_i)

35.3.2 Evaluation of Proposed Sewage Treatment Unit Price

Dalian government requires private sector tenders to explicate sewage treatment unit pricing method according to the price component template illustrated in Table 35.4, which is provided by government in bidding guidance books. The tenders must calculate the price per ton sewage treated by concessionaire through accomplishing the Table 35.4, and put the table in a sealed envelope without any information about themselves. The Table 35.4 will be analyzed and evaluated by a group of financial specialists independently.

The specialists determine a reasonable amount of every component of sewage treatment unit price by comparison of all tenders' bidding price proposals through checking the information of row 1–15 in Table 35.4. Then a unit price baseline (denoted as B) can be formed by summing all reasonable amounts of the pricing components determined by the specialists.

The score of bidding price can be calculated according to the deviation between sewage treatment unit price (denoted as P_i) proposed by the *ith* tender and the unit price baseline (B). The deviation (denoted as K_i) can be calculated as

$$K_i = (P_i - B) \times 100\% / B \tag{35.2}$$

The evaluation score of sewage treatment unit price proposed by the *ith* tender can be measured as (35.3).

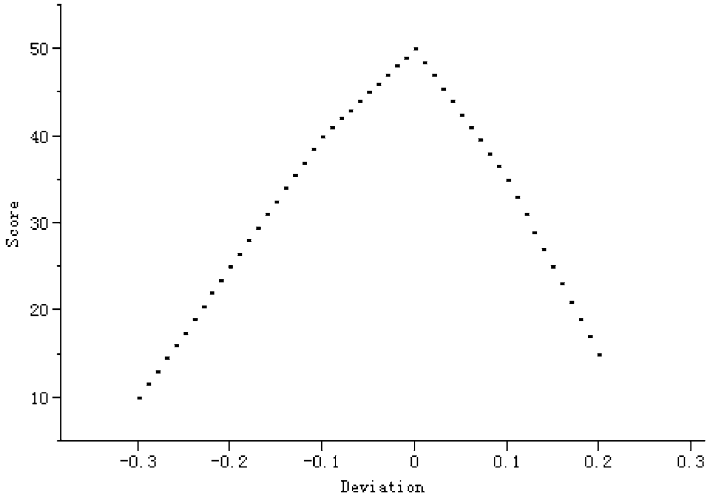


Fig. 35.1 Relation between score and sewage treatment price deviation

$$T_i = \begin{cases} 50 - K_i \times 100 \times 1.5, \ 0 \le K_i \le 10\% \\ 35 - (K_i - 10\%) \times 100 \times 2, \ K_i > 10\% \\ 50 - |K_i| \times 100 \times 1, -10\% \le K_i < 0 \\ 40 - (|K_i| - 10\%) \times 100 \times 1.5, \ K_i < -10\% \end{cases} \tag{35.3}$$

Figure 35.1 depicts the relation between the score of sewage treatment unit price and deviation which is derived from (35.2) and (35.3). The curve in Fig. 35.1 indicates that when P_i is greater than B, the score decreases more quickly than when P_i is less than B. The reason for such scoring system is to encourage tenders to propose reasonable sewage treatment unit price which is mostly close to the unit price baseline.

35.4 Discussion and Conclusion

The concessionaire selection criteria adopted by Dalian government result in improving efficiency of sewage treatment BOT projects. The time of inviting public bidding has reduced from 1 month to 2 weeks, and the time of concession contract negotiation has shortened from 6 months to 1 month. Dalian government has formulated guidance for sewage treatment BOT projects, consisting of model documents of sewage treatment BOT project bidding, model concession contract, model agreements for sewage treatment unit price, etc. The concessionaire selection criteria introduced above are integrated in the guidance documents. All tenders must follow the guidance and present information under the same evaluation framework when bidding on sewage treatment BOT projects in Dalian.

Concessionaire selection criteria should cover main characteristics of BOT project programs and attributions of concessionaires. These criteria can also be applied to other BOT projects, for example, garbage incineration power generation BOT projects, when adjusted accordingly. Concession price, e.g. sewage treatment unit price, is always the most important selection criterion when evaluating concessionaire in China, especially for BOT projects that need subsidy from government. The weight allocated between criteria of technical ability and business performance is 30 and 70 % respectively, which is very common in bidding practice in China. It is feasible that the weights of selection criteria change according to the needs of practice. However, it is suggested that the relative importance of criteria be kept as much as possible.

References

Ahadzi M, Bowles G (2004) Public-private partnerships and contract negotiations: an empirical study. Constr Manag Econ 22(9):967–978

Dias A, Ioannou PG (1996) Company and project evaluation model for privately promoted infrastructure projects. J Constr Eng Manag 122(1):71–82

Khadaroo I (2008) The actual evaluation of school PFI bids for value for money in the UK public sector. Criti Perspect Account 19(8):1321–1345

Kumaraswamy MM, Anvuur AM (2008) Selecting sustainable teams for PPP projects. Build Environ 43(6):999–1009

Liou FM, Huang CP (2008) Automated approach to negotiations of BOT contracts with the consideration of project risk. J Constr Eng Manag 134(1):18–24

Qiao L, Wang SQ, Tiong RLK, Chan TS (2002) Critical success factors for tendering BOT infrastructure projects in China. J Struct Proj Financ 8(1):1–13

Tiong RLK (1995a) Risks and guarantees in BOT tender. J Constr Eng Manag 121(2):183–188

Tiong RLK (1995b) Impact of financial package versus technical solution in a BOT tender. J Constr Eng Manag 121(3):304–311

Tiong RLK (1995c) Competitive advantage of equity in BOT tender. J Constr Eng Manag 121(3):282–289

Tiong RLK, Alum J (1997) Distinctive winning elements in BOT tender. Eng Constr Archit Manag 4(2):83–94

Tiong RLK, YEO K, Mccarthy SC (1992) Critical success factors in winning BOT contracts. J Constr Eng Manag 118(2):217–228

Torta F (2005) The concession through a bid – the new Brescia–Milan highway: a case study. Res Transp Econ 15(1):135–143

Zhang XQ (2004) Concessionaire selection: methods and criteria. J Constr Eng Manag 30(2):235–244

Zhang XQ (2005) Criteria for selecting the private-sector partner in public–private partnerships. J Constr Eng Manag 131(6):631–644

Zhang XQ, Kumaraswamy MM, Zheng W, Palaneeswaran E (2002) Concessionaire selection for build-operate-transfer tunnel projects in Hong Kong. J Constr Eng Manag 128(2):155–163

Chapter 36
Study of Optimal Contract Under Asymmetric Information

Ping Li and Hong-zhi He

Abstract The opportunism and information asymmetry between the two parties of a contract alteration give rise to the phenomenon of hold up and underinvestment, as well as the moral hazard problem of hiding information from the other party. This article built up the optimal contract model under asymmetric information by using the principle-agent theory and studied the above problem. The article proved that, under the conditions that the contractor hides information, the optimal contract would decrease the special purposed investment levels, prolong the time limit of the engineering project, and decrease the owners' utility.

Keywords Contract alteration • Hidden information • Information asymmetry • Project management

36.1 Introduction

Large engineering projects have characteristics such as long time limits, huge investment and complex technologies and involve various uncertain factors such as natural conditions, design alteration and the revision and abolition of laws and regulations. Engineering contract is a typical incomplete contract. In the standard project contract samples both home and abroad, conducting negotiation and contract alteration under uncertain natural condition types is permitted (China National Association of Engineering Consultants/International Federation of Consulting Engineers 2003; Ministry of Construction of The People's Republic of China and Commerce 1999).

P. Li (✉) • H. He
Department of Engineering Management, Luoyang Institute of Science and Technology, Luoyang, China
e-mail: lip05@163.com

E. Qi et al. (eds.), *The 19th International Conference on Industrial Engineering and Engineering Management*, DOI 10.1007/978-3-642-37270-4_36,

During the negotiation of contract alteration, the position of the contractor is changed due to the magnitude of special purposed investments in the early stage. Usually the more investment the poorer the position of the contractor is in. Under such circumstances, the conduct of hold up is produced if the owner possesses the contractor's production surplus by taking advantage of the better position in negotiation and makes the contractor under-invest (Hart and Moore 1994; Kultti and Takalo 2002). As for the change of natural conditions, the contractor usually has information advantage. By using information advantage, the contractor may hide the disadvantageous change of natural conditions and possess the owner's production surplus under the drive of profit. Moral hazard risk is thus produced. The design of optimal contract that overcomes the act of hold up and moral hazard becomes an important research topic.

Information asymmetry between the two parties of a contract and opportunistic behavior are the final reasons for the act of hold up and moral hazard. Moral hazard includes types of hiding actions and hiding information. The moral hazard that arises from the alteration processes of engineering project contracts belongs to the realm of hiding information. The standard framework of the moral hazard model that hides information means that, information is symmetric when signing the contract, but nature chooses the status after signing the contract. The agent observes the choice of nature and then chooses his action. The principal can observe the agent's action but cannot observe the choice of nature. Therefore the problem for the principal is to design an incentive contract that induces the agent to choose the most favorable action for the principal given the conditions of natural choice (Zhang 1996). Over the past decade a great deal of research has been done both home and abroad concerning the problems of choosing logistics service provider under asymmetric information (Li and Gao 2010; Gong and Peng 2009), reducing the information- asymmetry effect between investor and manager through corporate governance, corporate supervision and the professionalism of auditing (McLaughlin and Safieddine 2008; Cormier et al. 2010; Almutairi et al. 2009), the role of the labor unions in negotiations under asymmetric information (Hilary 2006) and moral hazard (Dewatripont and Tirole 1994; Martimort 1996; Agrawal 2002), but the research of the moral hazard of hiding information in engineering projects is scarce. Under the standard framework of the moral hazard of hiding information, this paper studies the problems of underinvestment of the contractor for engineering projects and the characteristics and efficiency of optimal contracts.

36.2 Description of the Problem

According to the principle of risk sharing, risks that can be correctively evaluated should be assumed by the party that can best control risk; risks that cannot be evaluated or controlled should be assumed by the party that has the ability to assume or the party that can get insurance (Kreps 1990). In engineering projects, the owner is usually the party that assumes risks that cannot be evaluated or predicted. During contract implementation, once such unpredictable and uncontrollable risk happens, the contractor is permitted to claim refunds and the owner gives the contractor refunds through lifting up contract price or prolonging the project time limit. This paper calls these risk factors that cannot be controlled by the contractor but are assumed by the owner as construction condition ε. Without loss of generality, this paper assumes that there are only two types of status, the good condition ε^G and the bad condition ε^B, where $\varepsilon^B > \varepsilon^G,\ \varepsilon \in \{\varepsilon^B, \varepsilon^G\}$. Suppose the owner and the contractor are uncertain about the value of construction condition ε before signing the contract, and the probability of the advantageous construction condition ε^G is a common knowledge between the two parties. After signing the contract, nature selects the construction condition. The contractor can get the exact value of the construction condition ε from environment, and determines his most favorable special purposed investment level i. However, the exact value of the construction condition ε and the special purposed investment level i cannot be observed by the owner, or cannot be verified even when observed by the owner. The owner can only determine the contract price p through the optimal project time limit $\bar{t}$ of the contractor's decision.

By the mix of direct fees and indirect fees, the contractor's cost will be increased whether the project time limit is too long or too short. Thus there exists optimal time limit satisfying $\mathrm{C}'(\bar{\mathrm{t}}) = 0$, where the optimal project time limit $\bar{t}$ is a function of the contractor's special purposed investment i and the construction condition ε, which can be expressed as $\bar{t}(i, \varepsilon)$, $\frac{\partial \bar{t}(i,\varepsilon)}{\partial i} < 0$, $\frac{\partial \bar{t}(i,\varepsilon)}{\partial \varepsilon} > 0$. Suppose the optimal time limit is a linear function, then the compression of optimal project time limit due to the contractor's special purposed investment is $\delta t(i, \varepsilon) = t_0 - \bar{t}(i, \varepsilon) = ki - \varepsilon$, where t_0 can be understood as the expected optimal project time limit when the special purposed investment is 0, the construction condition ε is the negative effect of the contractor's special purposed investment, k is a coefficient larger than 0.

Suppose the owner is risk-neutral and the contractor is risk-averse. Then the owner's utility is $B(\delta t, p) = V(\delta t) - p = V(ki - \varepsilon) - p$, where V is the evaluation of the owner toward the optimal project time limit compression, $V' > 0,\ V'' < 0$;

the contractor's utility is $U(p, i) = u(p) - v(i)$, where u is the positive effect due to the contractor obtains payment, $u' > 0,\ u'' < 0$; v is the negative effect due to the contractor's special purposed investment, $v' > 0,\ v'' > 0$.

36.3 The Optimal Contract Under Symmetric Information Conditions

Under symmetric information conditions, the owner and contractor cannot predict the hardness of the implementation or the uncertain events that possibly arise from the implementation before signing the contract. After signing the contract, the contractor knows the true value of the construction condition, whether it be advantageous or disadvantageous, and informs the owner the true construction

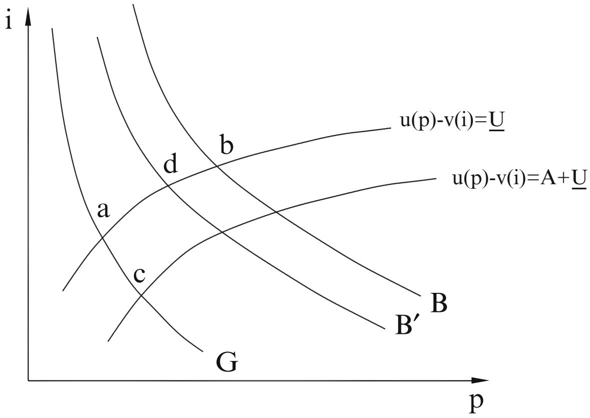


Fig. 36.1 Optimal contract under asymmetric information

condition. According to the observed construction condition, the contractor selects the special purposed investment level. The owner can observe and verify the special purposed investment conducted by the contractor. The contractor's contract items include the contractor's special purposed investment I and the project's contract price p. the optimal contract {p, i} designed by the owner is to maximize personal utility guaranteeing the contractor's conserve utility, satisfying

$$\underset{p,i}{Max} \quad V(ki - \varepsilon) - p$$
$$u(p) - v(i) \geq \underline{U}$$

Where U is the contractor's conserve utility, i.e. the opportunity cost; $\varepsilon \in \{\varepsilon^B, \varepsilon^G\}$. Given initial assumptions, the optimal contract $\{p^{G^*}, i^{G^*}\}$ when the construction condition is advantageous that is concave in {p, i} satisfies the following conditions:

$$u(p^{G^*}) - v(i^{G^*}) = \underline{U} \tag{36.1}$$

$$kV'(ki^{G^*} - \varepsilon^{G^*}) = \frac{v'(i^{G^*})}{u'(p^{G^*})}, \tag{36.2}$$

Where (36.1) is the participation condition and (36.2) is the efficiency condition as shown by the curve G in Fig. 36.1. Similarly, the optimal contract $\{p^{B^*}, i^{B^*}\}$ when the construction condition is disadvantageous satisfies the following conditions:

$$u(p^{B^*}) - v(i^{B^*}) = \underline{U} \tag{36.3}$$

$$kV'(ki^{B^*} - \varepsilon^{B^*}) = \frac{v'(i^{B^*})}{u'(p^{B^*})}. \tag{36.4}$$

Because $V'' < 0$, $u'' < 0$, $v'' > 0$ and $\varepsilon^B > \varepsilon^G$, (36.2) is under (36.4) in (p, i) space, as shown by Curve B in Fig. 36.1. a, b are the optimal contracts respectively under two construction conditions given initial assumptions, and $i^{G*} < i^{B*}$, $p^{G*} < p^{B*}$.

The contractor's special purposed investment level i can be observed and verified by the owner under symmetric information conditions and the construction condition ε is the common information shared by the two parties and thus there is no moral hazard of hiding information. The contractor's special purposed investment level i determines the owner's payment $p = u^{-1}[\underline{U} + v(i)]$, which is irrelevant from the construction condition and the owner would assume all risks that arise from the uncertainty of the construction condition. Therefore the owner can devise a pure contract $\{p^{G*}, i^{G*}\}$ before signing the contract. After signing the contract, if the contractor observes that the construction condition is bad ε^B, then he conducts an i^{B*} investment and claims compensation to the owner. The owner alters the contract $p^{G*} \rightarrow p^{B*}$ and compensates the contractor according to the change of the construction condition and the special purposed investment, which can realize an optimal contract in the Pareto sense.

36.4 Optimal Contract Under Asymmetric Information

Under asymmetric information, the contractor's special purposed investment and the construction condition cannot be observed or verified. The owner can only conduct payment according to the practical optimal project time limit. If the owner devises a pure contract $\{p^{G*}, \bar{t}^{G*}\}$ before signing the contract, if the construction condition is bad after signing the contract, the contractor must increase his investment and claim compensation to the owner. But because the investment and the construction condition cannot be observed or verified, the owner cannot agree to contract alteration. Thus the contractor would assume risks that arise from the uncertainty of the construction condition that are supposed to be assumed by the owner, producing underinvestment or hold up phenomena. If the owner devises a pure contract $\{p^{B*}, \bar{t}^{B*}\}$, if the construction condition is good after signing the contract, the contractor would hide the true construction condition by taking the information advantage and damages the owner's benefit and thus produces moral hazard of hiding information. If the owner offers contract menu $\{p^{G*}, \bar{t}^{G*}\}$, $\{p^{B*}, \bar{t}^{B*}\}$, the contractor would prefer $\{p^{B*}, \bar{t}^{B*}\}$ because the expected utility is no less than the conserve utility by selecting the latter no matter what the construction condition is. Therefore the optimal contract under symmetric information is no longer optimal under asymmetric information.

According to principle (Kreps 1990), the menu of the optimal contract must be able to motivate the contractor to reveal the true information about the construction condition after signing the contract. Therefore the format of the optimal contract must be self-selected and the contract menu must make the utility of the contractor that reveals the true information about the construction condition after signing the contract larger than the utility that deceives the owner. The owner's problem is to maximize his expected profit under the above constraint conditions. The contract menu $\{p^G, \bar{t}^G\}$ $\{p^B, \bar{t}^B\}$ thus must satisfy: $\underset{\{p^G, i^G\},\{p^B, i^B\}}{Max}\ q[V(ki^G - \varepsilon^G) - p^G] + (1-q)[V(ki^B - \varepsilon^B) - p^B]$

$$u(p^G) - v(i^G) \geq \underline{U} \tag{36.5}$$

$$u(p^B) - v(i^B) \geq \underline{U} \tag{36.6}$$

$$u(p^G) - v(i^G) \geq u(p^B) - v[i^B - (\varepsilon^B - \varepsilon^G)/k] \tag{36.7}$$

$$u(p^B) - v(i^B) \geq u(p^G) - v[i^G - (\varepsilon^G - \varepsilon^B)/k] \tag{36.8}$$

The participation constraints of the above programming are afterwards constraints (36.5), (36.6). Whether good or bad the construction condition is, the contractor always has interest to accept the contract. (36.7), (36.8), the incentive compatibility constraints, are also called self-selection constraints, motivating the contractor to choose $\{p^G, \bar{t}^G\}$ when the construction condition is good and $\{p^G, \bar{t}^G\}$ when the construction condition is bad, or else his utility is reduced. In the above equations, $i^B - (\varepsilon^B - \varepsilon^G)/k$ is the special purposed investment level of the contractor when the optimal project time limit is $\bar{t}^B$ under good construction condition; $i^G - (\varepsilon^G - \varepsilon^B)/k$ can be similarly understood. From (36.7), $u(p^G) - v(i^G) \geq u(p^B) - v(i^B) + \{v(i^B) - v[i^B - (\varepsilon^B - \varepsilon^G)/k]\} \geq u(p^B) - v(i^B) \geq \underline{U}$. Thus (36.6) implies (36.5) and (36.5) can be eliminated. Therefore the contractor's utility when the construction condition is good is no less than his utility when the construction condition is bad under asymmetric information conditions. Therefore low-efficiency construction condition that is relevant to the unique participation constraint must be paid attention to.

Theorem 36.1 *the optimal contract menu* $\{p^{G^{**}}, \bar{t}^{G^{**}}\}$, $\{p^{B^{**}}, \bar{t}^{B^{**}}\}$ *under asymmetric information is defined by the following equations*:

1. $u(p^{B^{**}}) - v(i^{B^{**}}) = \underline{U}$
2. $u(p^{G^{**}}) - v(i^{G^{**}}) = \underline{U} + A$
3. $kV'(ki^{G^{**}} - \varepsilon^G) = \frac{v'(i^{G^{**}})}{u'(p^{G^{**}})}$
4. $kV'(ki^{B^{**}} - \varepsilon^B) = \frac{v'(i^{B^{**}})}{u'(p^{B^{**}})} + \mu\frac{v'(i^{B^{**}}) - v'[i^{B^{**}} - (\varepsilon^B - \varepsilon^G)/k]}{1-q}$,

where $A = v(i^{B^{**}}) - v[i^{B^{**}} - (\varepsilon^B - \varepsilon^G)/k] > 0$, $\bar{t}^{G^{**}} = t_0 - (ki^{G^{**}} - \varepsilon^G)$, $\bar{t}^{B^{**}} = t_0 - (ki^{B^{**}} - \varepsilon^B)$.

Proof Let λ, μ, ν denote the Lagrange coefficients of (36.6), (36.7) and (36.8) respectively. The first order conditions for i^G, i^B, p^G, p^B are:

$$qkV'(ki^G - \varepsilon^G) - \mu v'(i^G) + \gamma v'[i^G - (\varepsilon^G - \varepsilon^B)/k] = 0 \tag{36.9}$$

$$(1-q)kV'(ki^B - \varepsilon^B) - \lambda v'(i^B) + \mu v'[i^B - (\varepsilon^B - \varepsilon^G)/k] - \gamma v'(i^B) = 0 \tag{36.10}$$

$$-q + \mu u'(p^G) - \gamma u'(p^G) = 0 \tag{36.11}$$

$$-(1-q) + \lambda u'(p^B) - \mu u'(p^B) + \gamma u'(p^B) = 0. \tag{36.12}$$

From (36.11), (36.12), $\lambda = \frac{1-q}{u'(p^B)} + \frac{q}{u'(p^G)} > 0$. Equation (36.1) thus holds.
Again from (36.11), $\mu - \gamma = \frac{q}{u'(p^G)} > 0$. Then $\mu > \gamma$.
Next we prove that $\mu > \gamma$ and $\gamma = 0$.
From (36.7), (36.8),

$$v[i^G - (\varepsilon^G - \varepsilon^B)/k] - v(i^G) \geq v(i^B) - v[i^B - (\varepsilon^B - \varepsilon^G)/k]. \tag{36.13}$$

Assume $f(i) = v[i - (\varepsilon^G - \varepsilon^B)/k] - v(i)$, then $f'(i) > 0$.
From (36.13),

$$f(i^G) \geq f[i^B - (\varepsilon^B - \varepsilon^G)/k] \tag{36.14}$$

If $\mu > \gamma$ and $\gamma \neq 0$, then (36.13), (36.14) hold and $i^G = i^B - (\varepsilon^B - \varepsilon^G)/k$, $u(p^G) = u(p^B)$, $p^G = p^B$ is true.

From (36.9), (36.10), $\lambda = \frac{1}{u'(p)} = \frac{kV'(ki-\varepsilon)}{v'(i^B)}$. Then:

$$\mu = \frac{q}{u'(p)} + \gamma = q\lambda + \gamma \tag{36.15}$$

From (36.8),

$$\mu = \frac{qkV'(ki-\varepsilon)}{v'(i^G)} + \gamma \frac{v'[i^G - (\varepsilon^G - \varepsilon^B)/k]}{v'(i^G)} = q\lambda \frac{v'(i^B)}{v'(i^G)} + \gamma \frac{v'[i^G - (\varepsilon^G - \varepsilon^B)/k]}{v'(i^G)} \tag{36.16}$$

If (36.15), (36.16) are identity equations, then $i^G = i^B$, $\varepsilon^G = \varepsilon^B$ must hold, contradicting the initial assumption.

Therefore it must hold that $\mu > \gamma$ and $\gamma = 0$ for the optimal contract satisfying Kuhn-Tucker conditions. Thus (36.7) holds, $u(p^G) - v(i^G = \{u(p^B) - v(i^B)\} + \{v(i^B) - v[i^B - (\varepsilon^B - \varepsilon^G)/k]\} = \underline{U} + A$,

Where $A = v(i^B) - v[i^B - (\varepsilon^B - \varepsilon^G)/k]\} > 0$. Equation (36.2) holds.

From this, the utility has no difference whether the contractor chooses the project time limit as $\bar{t}^{G**}$ or $\bar{t}^{B**}$ under bad construction condition; the utility when the contractor chooses the project time limit as $\bar{t}^{G**}$ is strictly superior to $\bar{t}^{B**}$ under good construction condition. The utility is larger than the conserve utility. The owner should pay the information rent.

From (36.9), (36.11), $kV'(ki^G - \varepsilon^G) = \frac{v'(i^G)}{u'(p^G)}$.

From (36.10), (36.12), $kV'(ki^B - \varepsilon^B) = \frac{v'(i^B)}{u'(p^B)} + \mu \frac{v'(i^B) - v'[i^B - (\varepsilon^B - \varepsilon^G)/k]}{1-q}$.

Then Eqs. (36.3) and (36.4) hold. Thus we have: The optimal contract can realize Pareto optimum when the construction condition is good; the optimal contract has low efficiency when the construction condition is bad, as shown by curve B' in Fig. 36.1. The investment level cannot meet the payment requirement of the owner. The two points c, d represent the optimal contracts under two kinds of construction conditions.

From Theorem 36.1 we can draw conclusion that the optimal contract under asymmetric information has the following characteristics:

1. The contractor's utility under good construction condition is no less than the utility under bad construction condition. He obtains utility that is larger than the conserve utility level because of his personal information; He obtains merely the conserve utility when the construction condition is bad.
2. The contractor's utility has no difference whether he chooses the project time limit as $\bar{t}^{G^{**}}$ or $\bar{t}^{B^{**}}$ when the construction condition is bad and the incentive constraint is invalid; the contractor's utility when he chooses project time limit as $\bar{t}^{G^{**}}$ is strictly better than when he chooses project time limit as $\bar{t}^{B^{**}}$ under good construction condition and the incentive constraint is effective.
3. The optimal contract can realize Pareto optimum when the construction condition is good. However, the owner must pay the contractor information rent, as represented by Point c in Fig. 36.1. The efficiency is irrelevant to the contract price when the contractor is risk-neutral.
   $p^{G^{**}} = p^{G^*}$, $i^{G^{**}} = i^{G^*}$, $\bar{t}^{G^{**}} = \bar{t}^{G^*}$; the efficiency is relevant to the contract price when the contractor is risk-averse. $p^{G^{**}} > p^{G^*}$, $i^{G^{**}} < i^{G^*}$, $\bar{t}^{G^{**}} > \bar{t}^{G^*}$, as shown in Fig. 36.1.
4. Contract twist and low efficiency occur under bad construction condition and the contractor's investment level cannot meet the owner's payment requirement. Contract twist shifts the efficiency curve down in p, i space, as shown by Curve B' in Fig. 36.1. $p^{B^{**}} < p^{B^*}$, $i^{B^{**}} < i^{B^*}$, $\bar{t}^{B^{**}} > \bar{t}^{B^*}$.

36.5 Conclusion

For incomplete contract in the class of contracts for large engineering project, the two parties' flexibility such as the owner's single-sided design alteration right and the contractor's right of relief is increased, risks can be effectively avoided and shared, the two parties are motivated to increase investment, transaction cost is reduced and efficiency is promoted through the alteration items in a contract. However, because of the opportunism and information asymmetry in contract alteration between the two parties of a contract, acts of hold up, underinvestment and the problem of moral hazard of information hiding are produced, lowering down efficiency and increasing internal risks.

In the light of the above problems, this paper built the optimal contract model under asymmetric information using the principal-agent theory and investigated the characteristics of the optimal contract under the moral hazard of hiding information. The study demonstrates that the phenomenon of hold up and moral hazard of hiding information in the contract alteration of the engineering project can be effectively overcome through the optimal contract menu designed by the owner and the contractor's self-selected contract alteration. However, under asymmetric information, because the payment of the information rent and the lowering down of efficiency, the owner's utility is reduced, the contractor's special purposed investment level is decreased and the optimal project time limit is prolonged. The conclusion offers theoretical reference for solving the phenomenon of hold up and the problem of moral hazard of hiding information in contract alteration of engineering projects.

References

Agrawal P (2002) Double moral hazard, monitoring, and the nature of contracts. J Econ 75(1):33–61

Almutairi AR, Dunn KA, Skantz T (2009) Auditor tenure, auditor specialization, and information asymmetry. Manag Audit J 24(7):600–623

China National Association of Engineering Consultants/International Federation of Consulting Engineers (2003) The FIDIC contracts guide. Machinery Industry Press, Beijing, pp 250–255

Cormier D, Ledoux M, Magnan M, Aerts W (2010) Corporate governance and information asymmetry between managers and investors. Corp Gov 10(5):574–589

Dewatripont M, Tirole J (1994) A theory of debt and equity: diversity of securities and manager-share holder congruence. Q J Econ 109:1027–1054

Gong Z, Peng Q (2009) Analysis on cost control model for outsourcing logistics for enterprises under information asymmetry. In: Proceedings of the international conference on transportation engineering 2009, Chengdu, pp 3106–3111

Hart O, Moore J (1994) Default and renegotiation: a dynamic model of human capital. Q J Econ 109:841–879

Hilary G (2006) Organized labor and information asymmetry in the financial markets. Rev Account Stud 11:525–548

Kreps D (1990) A course in microeconomic theory. Harvester, Wheatsheaf, New York, pp 230–281

Kultti K, Takalo T (2002) Hold-ups and asymmetric information in a technology transfer: the micronas case. J Technol Transf 27:233–243

Li D, Gao L (2010) Mechanism design of TPLs based on information asymmetry. In: Proceedings of 10th international conference Chinese transportation professionals, ICCTP'10, Beijing, 2010, pp 4016–4023

Martimort D (1996) Exclusive dealing common agency and multiprincipal incentive theory. Rand J Econ 27:1–31

McLaughlin R, Safieddine A (2008) Regulation and information asymmetry: evidence from the performance of industrial and utility firms issuing seasoned equity in the USA. J Financ Regul Compliance 16(1):59–76

Ministry of Construction of The People's Republic of China, The State Administration for Industry and Commerce (1999) Construction contracts (GF-1999-0201), pp 12–16

Zhang W (1996) Game theory and information economics. The People's Publication House of Shanghai, Shanghai, pp 397–441

Chapter 37
Intelligent Control via Power-Line Carrier for Illumination and Air Condition in Buildings

Rong-Ching Wu, Ke Zhu, En-Chih Chang, and Jia-Chu Lee

Abstract This paper establishes the monitoring system of advanced metering infrastructure by power line carrier. The communication between these modules is accomplished with power lines. The master module consists of a signal transform circuit, a LCD circuit and a keyboard circuit. The master module can measure the real-time states, including the voltage, the current, the real power and the reactive power, etc.,. The control modules can be divided into types, one is air conditioning control and another is illumination control. Each control module consists of a signal transform circuit, an AVR microprocessor, a keyboard decode circuit, a LCD circuit and a power line carrier circuit. It has the functions of local and remote monitors. The air conditioning control module is installed an infrared rays control circuit for the air condition device, and the illumination control module is added a power control circuit for the illumination load.

Keywords Advanced metering • Infrastructure • Load management • Power line carrier

R.-C. Wu • E.-C. Chang
Department of Electrical Engineering, I-Shou University, Kaohsiung, Taiwan

K. Zhu (✉)
Department of Electrical Engineering and Renewable Energy, Three Gorges University, Yichang, China
e-mail: zhu_zhu_ke@qq.com

J.-C. Lee
Department of Electrical Engineering, National Sun Yat-Sen University, Kaohsiung, Taiwan

E. Qi et al. (eds.), *The 19th International Conference on Industrial Engineering and Engineering Management*, DOI 10.1007/978-3-642-37270-4_37,

37.1 Introduction

In recent years, as social economy grows and people's living standard improves, requirements of commercial electricity and residential electricity have been increasing substantially. To meet electricity demands of users, it is necessary for electric power enterprises to develop power supply with new power plant constructions or extensions and PTD (power transmission and distribution) equipments. However, because of the limited petroleum energy, its price fluctuates greatly; In Chinese Taiwan, due to people's environmental awareness and resistance, exploitation of energy resources and electric power construction are confronted with great difficulties in that it is the unshirkable responsibility for everyone to save energy. The improper electric power utilization will lead to increases of power consumption that exceeds contract capacity and electric charge. Efficient use of energy should be regarded as right energy-saving concept.

According to Economic Ministry statistics, Taiwan's total power consumption of residential commercial buildings accounts for 34 %. Power using in illumination and air condition of Taiwan's old apartment is about 52 %. Therefore, to achieve energy saving goals, we can improve the residential and commercial building equipment at the beginning. Although the government has controlled the power use of the new building, in Taiwan there are still more than 94 % old buildings which did not consider the energy-saving. This paper aims to succeed in that the advanced metering infrastructure optimizes load operation with power line carrier and achieves the load management. Two subsystems of the hardware are composed of a maser module and two control modules. The master module consists of a signal transform circuit, an AVR microprocessor, a keyboard decode circuit, a LCD circuit, and a power line carrier circuit, which has functions of local and remote monitoring (Wu et al. 2006). Control modules are classified into two categories- air conditioning control module and illumination control module. Each control module consists of a signal transform circuit, an AVR microprocessor, a keyboard circuit, a LCD circuit, and a power line carrier circuit, which has the functions of manual and automatic load control. An infrared control circuit is installed in air conditioning control module for controlling air conditioner, and a power control circuit for controlling illumination device is added to the illumination control module.

These modules realize communications via power lines (Shwehdi 1996). The distributed management is adopted to allocate information resources to every module. The distributed form is also applied to the information system development and data store. Its main advantage is that it can satisfy the information needs of each module and meet the demand that each module controls information resource. With convenience for use and maintenance, it has the biggest difference from centralization that if a module is down, other modules can still keep running and the overall system will not come to a standstill. The research result of this paper is estimated to be smoothly applied in the occasion that the load management is required to reduce electric charge (Wu et al. 2008).

37.2 Methodology

37.2.1 Load Management Plan

This paper makes use of distributed control model to serve the purposes of humanity and load control. Depending on highly automatic operation, it will take proper actions according to real-time and accurate data to reduce labor cost and electricity expenditure, and improve power efficiency. This system includes a master module and multiple control modules, show as Fig. 37.1. The master module monitors the users' practical electric power utilization, and transmits signal to control modules. The control modules will adopt certain control technologies to switch off the electricity or adjust the equipments' power consumption in accordance with load characteristics. Considering that users know little about the service condition of electric power, this paper will develop a system that is suitable for ordinary people to operate and there will be a LCD screen for observing the service condition of electric power. Therefore, users can know their electric power quality or change the schedule of electric power utilization accordingly.

The master module, a digital meter, is used to measure the parameters including the effective values of voltage and current, total power, total reactive power, etc., (Mortensen and Johnson 1988). The module has external communication interface, which can transfer data via power line carrier, thus, wherever the control modules of the building are, once the LAN is connected, any load can be remotely monitored and controlled.

Control modules are open, which receive service condition of electric power and properly controls the equipments. The working conditions of equipments are considered as judgments of control. Control modules can be divided into air conditioning control and illumination control on the basis of load characteristic. This paper controls large loads in family including air conditioners and lighting equipments. Air conditioner adopts infrared rays to exercise air conditioning control, while lighting equipments are for illumination control. With rapid development of technology, load control will be more compact and reliable; whose hardware will complete a monitoring system in cooperation with automatic meter reading for communication technology is becoming more developed.

37.2.2 Hardware System

See Fig. 37.1 for hardware framework. The system of this plan includes three subsystems, namely master module, air conditioning control module and illumination control module. The master module can measure the real-time states, including the voltage, the current, the real power, and the reactive power, etc. (Phadke et al. 1983), which will be revealed on LCD screen. The measured data can be transmitted to control modules for further handling via power line, and the control modules

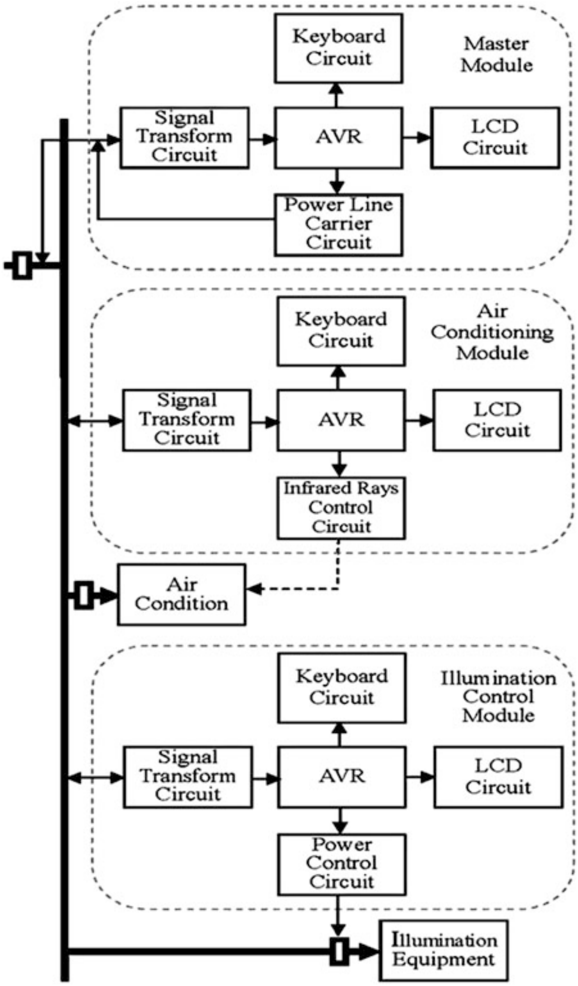


Fig. 37.1 Hardware framework

can control loads based on the received data. The master module consists of a signal transform circuit, an AVR microprocessor, a keyboard decode circuit, a LCD circuit, and a power line carrier circuit. It has the functions of local and remote monitoring. The signal transform circuit can transform the actual values of voltage and current into the available voltage level in line with analogical/logical conversion. AVR microprocessor can take samples of continuous signals, process commands of keyboard and computer, output the result to LCD and command the controller's actions. The keyboard circuit commands AVR to display different parameters. The LCD circuit can reveal the measured wave and measurements and so on. The aim of power line carrier circuit is modulating the service condition of electric power to the power line carrier signal, which will be transmitted to control modules for controlling purpose.

Air conditioning control module consists of a signal transform circuit, an AVR microprocessor, a keyboard circuit, a LCD circuit; infrared rays control circuit and a power line carrier circuit. It has functions of manual and automatic control load. AVR microprocessor can process the commands of keyboard, generate control commands according to master module, output results to LCD and d rive infrared rays control circuit. The keyboard circuit can command AVR to reveal different parameters, or instruct AVR to take action to infrared rays control circuit. The LCD circuit can display operation states. Infrared rays control circuit may control switch and temperature regulation of load in accordance with commands. The aim of power line carrier circuit is to decode the carrier signal of power line and transform it into the data that is available for controlling of control modules. This paper takes air conditioner as the controlled entity of air conditioning control modules.

The contents of illumination control module are similar to those of air conditioning control module and only power control circuit, which controls switch of load according to commands, replaces infrared rays control circuit. This paper considers lighting equipments as the controlled entity of illumination control module.

The communication between master module and control modules is established with power line. The data of master module is output via the built-in RS-232 interface; the module will transmit the signal to power line or receive commands from power line via power line carrier circuit. The system will work in the coverage of power line. It utilizes transfer ability of power line carrier signal and network function to construct monitoring system of load management.

37.2.2.1 AVR Microprocessor

This plan takes AVR chip microcontroller as STK128 development board (ATmega128). See Fig. 37.2. The AVR single chip, developed by ATMEL Company, is the high-speed 8-bit single chip with enhanced CPU (Reduced Instruction Set CPU, RISC) with built-in flash. AVR single chip is widely applied in instruments and meters, communication equipments, home appliances and so on. It mainly features AVR 8-bit microcontroller of high performance and low power consumption, and advanced RISC framework. When 16 MHz is adopted, it will achieve 16 MIPS (million instructions per second), and system of 128 K byte can conduct flash programming, also optimal external memory space of 64 K bytes can realize programming of system. It has JTAG interface (compatible with standard IEEE 1149.1). Via JTAG interface, programmings of Flash, EEPROM, fuse point and lock point are implemented.

It has two independent prescalers and 8-bit timer/counter with function of comparator, two 16-bit timers/counters with functions of prescaler, comparing and capturing, real-time timer with independent prescaler, eight routes and 10-bit ADC, eight single channels, seven difference channels, two programmable serial USART and so on. I/O and encapsulation have 53 programmable I/O, working voltage of 4.5–5.5 V, and frequency range of 0–16 MHz. In conclusion, ATmega128 is an 8-bit microprocessor with low power consumption based on AVR RISC

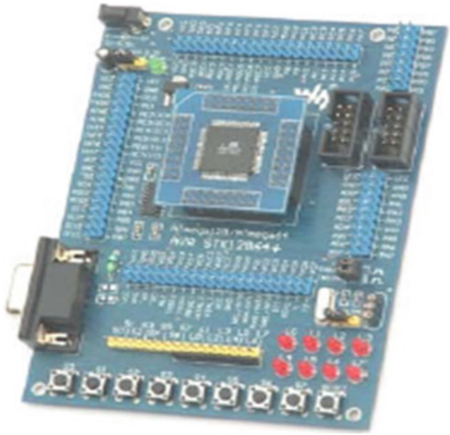


Fig. 37.2 ATmega128

framework. Because of the advanced instruction set and monocyclic instruction time, the throughput of ATmega128 reaches 1 MIPS/MHz, which can alleviate the contradiction between power consumption and processing rate.

37.2.2.2 Communication Format

The communication format of system is To + From + Command + Data.

The concrete instructions are in Table 37.1. Take the communication of master module and illumination control module as an example. If a master module inquiries illumination control module about operation state, the master module will send out CAA00000. The string from the AVR UART end to PLC is in series, via the power line. UART is set to 9600, N, 8, 1. Each module of the PLC reads in the communication format (8 bytes). The first byte is to be recognized whether it belongs to it. If it is, process; if not, skip. If the identification code belongs to it, UART will go into AVR, reading instruction for action. In this case illumination control module will identify whether the instruction is from master module and then sends the operation state back to master module. Suppose all loads are the OFF at first, then the return data is ACQ00000. If master module commands lamp 1 ON, the master module will send CAC11000. Illumination control module would follow the mentioned method to accept the instruction, and sends operating condition state of ACQ10000 (Wu and Soon 2007).

37.2.2.3 Power Line Carrier Circuit

This circuit adopts data transmission module of power line carrier (CZ-03) made by Caizhi Technology. CZ-03 is a power line carrier module of low cost and high performance. It integrates all carrier transceiver circuits that it can form a complete

Table 37.1 Communication format

Format	Length	Instruction	
To	1 word	Master module	A
		AC control module	B
		Lighting module	C
From	1 word	Master module	A
		AC control module	B
		Lighting module	C
Command	1 word	Inquiry	A
		Reply	Q
		AC control module and lighting module	C
		Temperature up	U
		Temperature down	D
Data	5 words	Inquiry	00000
		AC control module reply	Running 10000
			Stop 00000
		Lighting control module reply	D1D2D3D4 D1 = 1 indicates lamp 1 on, D1 = 0 indicates lamp 1 off, and so on
		AC control module operation (on/off)	10000
		Lighting control module	i1000 Di on i = 1, 2, 3, 4
			i0000 Di off i = 1, 2, 3, 4

power carrier application system when the assorted coupling coil is connected and the high-voltage capacity is isolated, and it transmits signal by distribution network of end pressure on electrical cable. After the signal is modulated, it will be sent by power line. The module can achieve low-speed bidirectional data transmission function under the baud rate 300 bps. In a certain circumstance, the module with anti-collision built in allows multiple carrier equipments to transmit and receive data without mutual influence. The required communication function will be carried out under mains voltage of 110/220 V and frequency range of 50 60Hz. And carrier module's protocols between other applications don't need to be set especially because it can work independently. Power line carrier module is the entity and hardware communication equipment. When multiple power line carrier modules on power line transmission network are needed to be connected, the ID of every data equipment should be added to transmittal data packet to determine if other packet data should be received. In other words, when the data is transmitted, the transmission protocol is supposed to be defined according to communication data, such as setting the equipment ID, data confirmation and time-out, and so on. This module is widely applied in many application areas, such as equipments of system monitoring and controlling and other data communication. All can consider power line as transmission medium, such as home automation, lighting control, HVAC control, low-speed data network, automatic meter reading signal and data display, solar water heater, fire alarm and anti-theft warning and so on.

37.3 Discussion

The Complete system is shown as Fig. 37.3. This section describes communication monitoring, basic function and connection function.

After starting, the modules will detect the network connection and display connection states, contributing to users' understanding of connection states. Once the connection is done, remote control can start. In spite of connection failure, all modules will keep working independently and the whole system won't come to a standstill (Gao et al. 2008).

Master module is applied to single-phase 110 V/220 V power system. Measurement parameters are as follows: Basic parameters: AC voltage 0 220 (V), AC current 0 3 (A), phase angle $-\pi$ π (rad), Effective power 0 660 (W), reactive power (VAR). LCD backlit display (Papadopoulos and Papagiannis 2008) LCD display updates every 2 s, which can also fix signal analysis. It also shows the voltage and current waveforms, in order to identify the relationship between the two. Text Display: 3 and 1/2 LCD display, with functions, units, decimal point. It also shows four kinds of electrical parameter data. Master module connects with a computer or control module to provide power parameters analysis via power line, thus composing a smart grid. Master module can remotely control air conditioning module and illumination control module. Master module can also run automatic control according to electricity situation.

The air conditioning module has the operations of ON/OFF, modifying temperature setting and so on. Control signal: infrared Control, self-learning function,

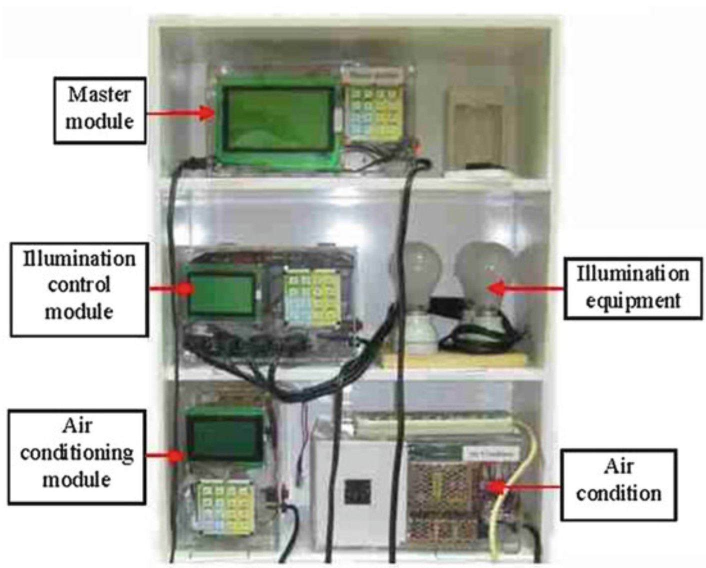


Fig. 37.3 Practical system

suitable for different electrical equipments, or ON/OFF contact control. The module has backlit screen with back-lit display function (Wang and Li 2010). Character display: 3 and 1/2LCD displaying shows the details of function, unit and radix point and four power parameters are displayed at the same time. As for air conditioner, it adopts thermoelectric cooling chip to make backlit display to imitate air conditioner operation. The load with rated parameters of 12VDC, 3ADC, 36 W, can reach 22 °C in actual operation which meets the demand of imitation. The controller of air conditioner adopts normal air conditioner control circuit with LCD wireless remote control and the microcomputer automatic control of temperature and wind speed can improve its practicability. The local control function of air conditioning control module is consistent with the function of air conditioner controller. Besides, the air conditioning control module can be connected to master module with power line to achieve function of remote control or automatic control.

The illumination control module controls its load by ON/OFF contact. The module has LCD backlit screen with backlit display function to display the operation states. Character display: 3 and 1/2LCD displaying shows the details of function, unit and radix point and four power parameters are displayed at the same time. The bulbs are the controlled targets suitable single phase 110 V/60 Hz power system. The rated power for each bulb is 60 W. The illumination control module controls local load switch. Connecting to master module via power line, it can also achieve the function of remote control and automatic control.

Connecting the modules of the system, remote control or automatic load management can be conducted to optimize load utilization. The connected control modules can transfer operation states to master module, and the master module can transfer commands to the control modules as well to achieve remote control. The users can change the load management model according to requirements (Andreadou et al. 2010).

37.4 Conclusion

This system has one master module and two control modules. The master module can detect local voltage, current, active power, reactive power, apparent power and so on. The computed result will be displayed on the LCD screen to upgrade the AVR application with respect to digital meter. The digital meter contains signal transform circuit, microprocessor, keyboard circuit, LCD circuit, power control circuit, power line carrier circuit and so on. Control module can automatically and manually control load operation states. It can set demand control for the system according to the setting of power consumption. In the condition of minimum operating environment influence, a series of control arrangement should be made for electric equipments to improve power efficiency, reduce electric charge such extra cost that comes because the power utilized exceeds that of contract capacity and the floating electric charge, and enhance the reliability of power supply in the summer peak particular. The control module includes microprocessor, keyboard circuit, LCD circuit, infrared control circuit, power control circuit, power line carrier circuit and so on.

This system has following features:

(a) The communication between these modules is accomplished with power lines. It doesn't need any additional communication wire or alteration for the circuit structure of loads
(b) The distributed management is used in this system. If a certain module is down, the else parts can still run their works and avoid the whole system coming to a stop.
(c) The control modules not only can manually control loads, but also can automatically control objects according to the load flow.
(d) A friendly man–machine interface will be developed in this system, the concept and application of the load management will be more popularly for common users.
(e) This system is expected to optimize the load operation and achieve the aim of load management.

In brief, this paper completes load management system with microprocessor and power line carrier. It is estimated to achieve the purposes of saving power, reducing electric charge, prolonging service life of equipments, improving electric power quality and so on.

Acknowledgements The authors thank the projects of NSC 99-2628-E-214 -001 which support the research of this paper.

References

Andreadou N, Member IEEE, Pavlidou F-N (2010) Modeling the noise on the OFDM power-line communications system. IEEE Trans Power Deliv 25(1):150–157
Gao Q, Yu JY, Chong PHJ (2008) Solutions for the "silent node" problem in an automatic meter reading system using power-line communications. IEEE Trans Power Deliv 23(1):150–156
Mortensen AN, Johnson GL (1988) A power system digital harmonic analyzer. IEEE Trans Instrum Meas 37(4):537–540
Papadopoulos TA, Papagiannis GK (2008) Low-voltage distribution line performance evaluation for PLC signal transmission. IEEE Trans Power Deliv 23(4):1903–1910
Phadke AG, Thorp JS, Adamiak MG (1983) A new measurement technique for tracking voltage phasors, local system frequency, and rate of change of frequency. IEEE Trans Power Appar Syst PAS-102(5):1025–1038
Shwehdi M-H (1996) A power line data communication interface using spread spectrum technology in home automation. IEEE Trans Power Deliv 11(3):1232–1237
Wang Y, Li W (2010) Reliability analysis of wide-area measurement system. IEEE Trans Power Deliv 25(3):1483–1491
Wu QL, Soon YT (2007) A distributed circuit model for power-line communications. IEEE Trans Power Deliv 22(3):1440–1445

Wu RC, Chiang CT, Tsai JI (2006) The real-time optimal spectrum analysis system based on personal computer. In: IASTED international conference on applied simulation and modeling, Rhode, pp 296–301

Wu RC, Kang JC, Sung CL, Chiang CT (2008) Realization of load management by power line carrier and lab VIEW. In: IEEE international conference on sustainable energy technologies, Singapore, pp 600–604

Chapter 38
Design and Implementation of a Remote Equipment Monitoring and Management System Based on IOT

Naren Gaowa, Ying Liu, and Mu-yan Li

Abstract Equipment monitoring and management based on IOT technology improves the efficiency of equipment monitoring and management, makes equipment management intelligent. A remote equipment monitoring and management system based on IOT is introduced in this paper. It is used in the lab of Beijing Jiaotong University. This system consists of three parts. The first is a perception node module, the second is a wireless multi-hop data transmission module and the last is a center management server. The design of those three parts is introduced in the paper. The real performance of the implemented monitoring and management system shows that the system works effectively and efficiently. This system can be expanded to suitable management systems in various trades and industries.

Keywords Equipment • IOT • Monitoring and management • Multi-hop networks • RFID

38.1 Introduction

Due to considerations of costs, technology and other factors, traditional barcode is popular in the general approach of equipment monitoring and management for fixed assets. In the process of equipment monitoring and management, the administrator first inputs the equipment information to the computer management systems, and then fixes barcode on equipment. In order to monitor the performance of equipment, the administrator has to go to the lab and check the equipments on regular time schedule. By this manual-based management approach, the inspection of assets

N. Gaowa (✉) • Y. Liu
School of Electronic and Information Engineering, Beijing Jiaotong University, Beijing, China
e-mail: narengaowa1989@gmail.com

M. Li
School of Computer Science and Engineering, Beihang University, Beijing, China

E. Qi et al. (eds.), *The 19th International Conference on Industrial Engineering and Engineering Management*, DOI 10.1007/978-3-642-37270-4_38,

costs a lot of time every year. And during the interval of inspection, equipment file will not be updated timely if the equipment is missing or has been taken away for personal business. So it is difficult to manage equipments (Xingbiao Mao 2010).

IOT (Internet of Things) aims to achieve automatic recognition of goods through the RFID (radio frequency identification), infrared sensors, etc. and achieve information sharing by connecting to the Internet. Ultimately, it makes management intelligent and constitutes a more intelligent system of production and living (Puchun Yu 2010). It is a very complex, various-formed and systemic technology. IOT connects the separated physical world with information space and represents the trend of development of future network. The physical world will form a huge net by IOT. The integration of IntelliSense technology, network technology and software technology in IOT is widely used in transportation control, parking management, warehouse management and other automated fields (Xinsheng Wu et al. 2012).

Equipment monitoring and management based on IOT can combine equipment management with IOT and ensure the optimal use of resources. It can solve the complex problems of equipment management. The most fundamental purpose of equipment management based on IOT is to achieve real-time equipment monitoring and improve the speed and accuracy of the equipment asset management and improve the efficiency of management.

A design of remote equipment monitoring and management system based on IOT technology for university lab is introduced in this paper. It is applied in the lab of Beijing Jiaotong University. The design combines RFID, wireless multi-hop transmission networks, database and Internet technology. An application system based on IOT, including a perception node module, a wireless multi-hop data transmission module and a center management server, is implemented by authors. Regarded flexibility and general-purpose as the main design ideas, the system can also be expanded to suitable management systems in various trades and industries. The system will improve management efficiency and reduce the labor intensity of the administrator. In this paper, the first section introduces the concept, main applications, current status of IOT, problems of equipment management and illustrates the significance of equipment management in IOT. The framework of overall design of this remote equipment monitoring and management system based on IOT is introduced in the second section. The third section concentrates on the hardware design of the perception node module. Section 38.4 mainly introduces the wireless multi-hop transmission routing strategy of the data transmission module. The software architecture of the center management server is presented in Sect. 38.5. Finally, the remote equipment monitoring and management system is concluded.

38.2 Framework of the System

The remote equipment monitoring and management system consists of perception node module, wireless multi-hop data transmission module and center management server three parts, achieves the goal of university laboratory equipment management in IOT approach.

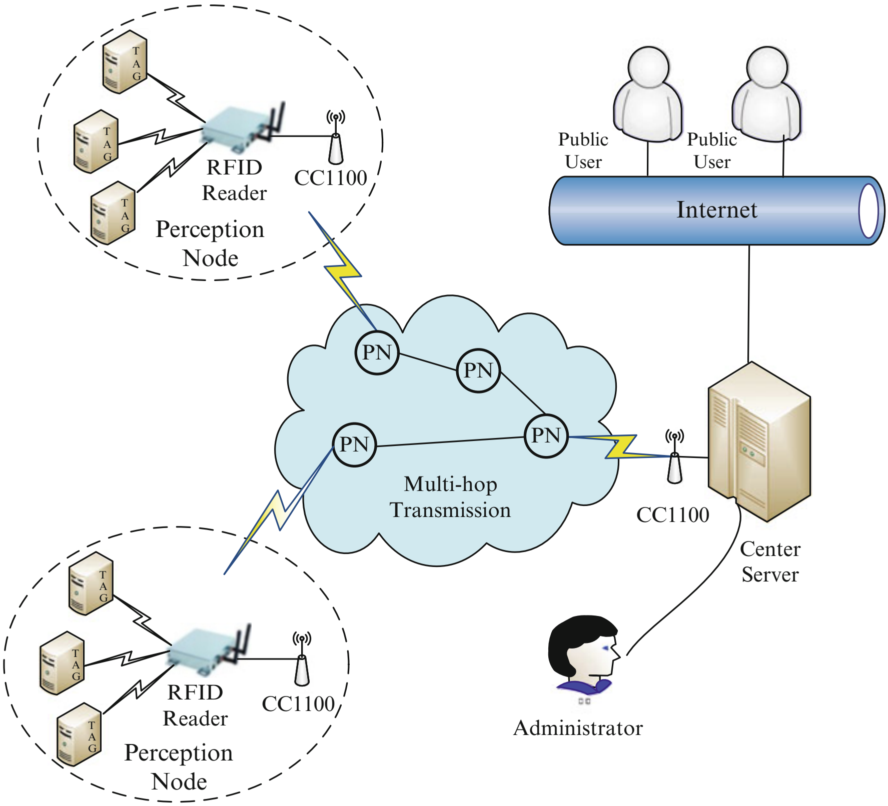


Fig. 38.1 Framework of the remote equipment monitoring and management system

Public users can browse laboratory equipment information at any time in any place with Internet access and the administrator can achieve real-time monitoring easily. The system framework is shown in Fig. 38.1. At perception node module, the electronic tag is attached to the equipment needed to be monitored. The RFID reader is responsible for reading the information in electronic tag. In order to extend the distance of data transmission and expand the coverage of the system, wireless multi-hop transmission is employed. The CC1100 wireless long-distance data transceiver based on CC1100 chips is responsible for the data transmission. The center management server collects all information data of the perception node for managing and monitoring equipments, a database based on the information data is established on the server for easier management. The center management server is connected to the Internet for remote sharing equipment information and usage data.

38.3 Hardware Design of Perception Node Module

The perception node module consists of five parts, power unit, RFID unit, wireless data transmission unit, RAM unit and controller unit. The hardware design of the perception node module is shown in Fig. 38.2. The controller is responsible for the entire perception node scheduling; the RFID unit is responsible for reading the electronic tag information and data collection. The information data collected by the RFID reader is temporarily stored in RAM while the wireless data transmission unit is responsible for sending out the electronic tag information in RAM storage and receiving the information from other perception nodes.

From the current perspective of the development of the IOT, RFID (Radio Frequency Identification Radio Frequency) tags technology is one of the very important technical indicators. RFID is a noncontact automatic identification technology. Because of its noncontact, identification without human intervention, rapid identification of high speed moving objects, fast and convenient to operate and easily automated features, RFID is now widely used in production, logistics, retail and other industries. RFID applications system includes RFID radio frequency identification systems and industry application systems. RFID identification system collects information of the reader through various means (RS232, the TCP/IP, middleware, front-end, protocol converters, etc.), then transmits the information to the industry application systems (Wenbing Fan et al. 2007; Hongyue Dai et al. 2007). The RFID readers and electronic tags used in this system are from Beijing Dingchuang Hengda Intelligent Technology Co., Ltd. The working frequency of initiative Omni-directional reader DC0301A is 2.45 GHz. The effective reading distance is 80 m or more. The RFID readers can identify multiple tags simultaneously. It communicates with the controller via serial communication.

In order to extend the distance of data transmission and expand the coverage of the system, the information data transmission uses wireless multi-hop transmission. By wireless multi-hop transmission, the perception node can join and exit monitoring system flexibly and easily. The increased difficulty of space wiring and construction bring by the wired transmission can be avoided. And the Ethernet IP resources occupied by data transmission on various monitoring points will be saved. As the system is planned to be used in a university laboratory environment, after comparing the performance of a variety of wireless data transceiver and actual inspection, the CC1100 wireless transceiver is finally chosen for data transmission

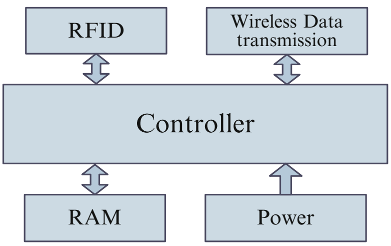


Fig. 38.2 Hardware design of perception node module

(Lijun Li et al. 2007). The CC1100 wireless transceiver is based on TI's high-performance CC1100 wireless communications chips. CC1100 is a single-chip UHF transceiver for low-power, low-cost, wireless applications. Its RF transceiver integrates a highly configurable modem. The modem supports different kinds of modulation formats, and its data transmission rate can reach up to 500 kbps. The transmission distance of CC1100 transceiver is up to 200–300 m in open space and up to 600–800 m after signal amplification. With high reliability, this transceiver can be widely used in the field of various occasions of short distance wireless communications.

The task of the controller in perception node module is to control RFID reader reading and collecting information in electronic tags based on the command received from center management server and store the collected information in RAM, then inform the wireless data transmission unit to send the data. The working mechanism is simple, and does not need a very high requirement of execution speed and storage space. A simple microcontroller can complete the above task well. To this end, the simple microcontroller STC12C5A60S2 control chip is taken as the perception node controller in this system. STC12C5A60S2 is a low-power high-speed general MCU with a 24.5 MHz oscillator, and a programmable flash memory of 32 KB. The flash memory has a security option, which can avoid its data be modified and erased maliciously. It has dual serial ports and with excellent performance.

38.4 Wireless Multi-hop Transmission Routing Strategy

Due to unpredictable nature of wireless environment, link breakage, disorder packets and other issues may occur during the data transmission process. In order to improve the link reliability, a wireless multi-hop transmission strategy is specifically designed for this system to ensure the reliability of the data transmission based on the ad hoc network routing strategy of AODV and the DSDV (Perkins and Royer 1999; Perkins and Bhagwat 1994). Set the perception node connected to center data management server as center perception node and other perception nodes as sub-perception nodes. Center perception node requests to obtain data of subperception nodes in the entire wireless network, and subperception nodes response to center perception nodes passively. Center perception node asks for data from all sub-perception nodes periodically or requests data from a specific node irregularly. Define data transmission from the center perception node to subperception nodes as the downlink; in contrast, the data transmission from subperception nodes to center perception node is uplink. Each perception node maintains a routing table and each route table entry contain the following two fields.

– Destination
– Next hop the next hop to destination

The data packets in transmission follow a unified format. The format of the data packets is illustrated as follow, and contains the following fields:

<header, next_hop, sequence_num, pre_hop, destination, data>

- Header set AA as packet header
- next_hop the next hop to destination
- sequence_num unique identifier for each packet that center perception node sends at downlink, set to 0 at uplink and incremented at downlink
- pre_hop the previous hop where the data packet comes from
- destination the final destination of packets
- data the equipment information collected by the sub perception nodes, set to 0 at downlink

The following two subsections describe how to generate data requests and the process after receiving data requests, respectively.

38.4.1 Generating Data Requests

The center perception node sends a data request packet when it requires data of a subperception node with no route to. Set the next_hop field as broadcast address and broadcast the data request packet.

38.4.2 Receiving Data Requests

When a subperception node receives a data request packet, the process is shown in Fig. 38.3. When a sub perception node receives a data request packet, the node first checks whether the next-hop is a broadcast address or not. If the next-hop is a broadcast address, the node will only respond to the packets with an undeceived sequence_num, and drop the packet if the sequence_num had been received. A perception node records the pre_hop of the packet from which it received the first copy of the data request as the next hop to center perception node in routing table. And then determine whether the node itself is the destination node or not, if so, that means the broadcast data request packet find a destination, the node will record the pre-hop and send data to pre-hop. Otherwise, the node just plays a forwarding node role, replaces the pre-hop of the data request packet by its own address and rebroadcasts the packet. If the next_hop in the packet equals to the perception node address, indicating that the node is the next hop in this route path.

The packet is for uplink or downlink is determined by whether the sequence_num value is 0. For uplink packet, send this packet according to the next hop in routing table and replace the pre_hop by itself. Rebroadcast this packet if its routing table has not been set up and add a route entry for the pre_hop, record the pre_hop as the next hop to itself. For downlink data request packet, record the pre_hop of the

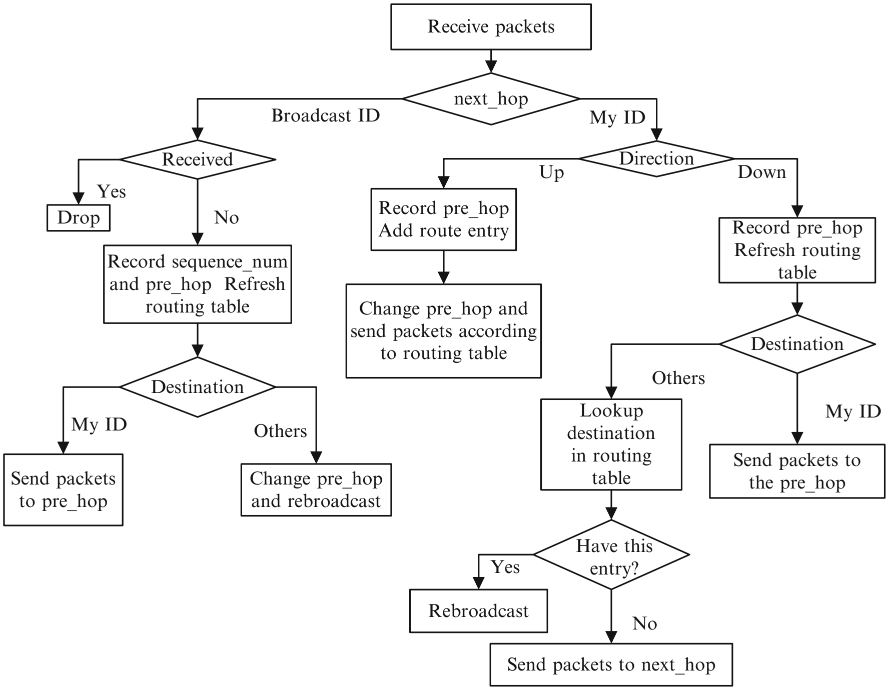


Fig. 38.3 The process after receiving data requests

packet from which it received the first copy of the data request as the next hop to center perception node in routing table. And then determine whether the destination address equals itself, if so, record the pre_hop and send data to pre_hop, otherwise, unicast the data request packet to next hop if this node possesses a current route to the destination, otherwise, rebroadcast this data request packet.

38.5 Software Architecture of Center Management Server

The software, developed by Microsoft visual C++, MySQL service and Eclipse, is a comprehensive management system that integrates the data collection, information processing and sharing. Software architecture of the Center Management Server consists of three parts: MySQL database (Xuhui Lan et al. 2004), real-time monitoring MFC module and JSP.NET module (Feng et al. 2006). Software architecture of the center Management Server is shown in Fig. 38.4.

Center Management Server stores equipment information collected by all perception nodes, to facilitate the equipment monitoring and management, an equipment information database is established on the management server. Structure of the

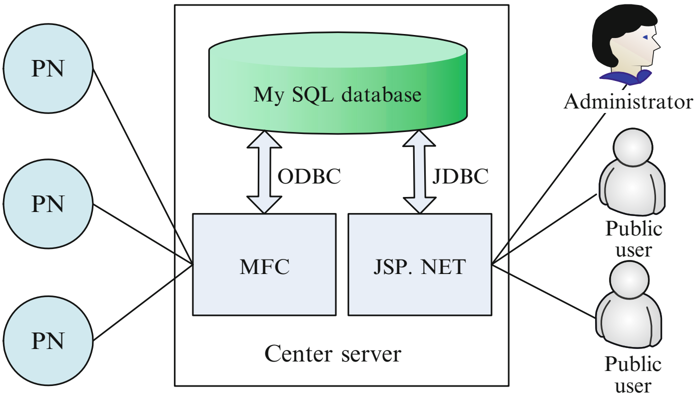
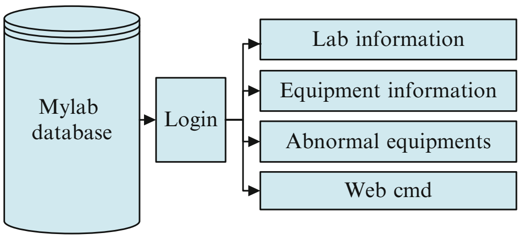


Fig. 38.4 Software architecture of the center Management Server

Fig. 38.5 Structure of the equipment information database



equipment information database is shown in Fig. 38.5. The pair < lab information, equipment information > uniquely identifies equipment. Abnormal equipments entry represents one or more equipment is missing or out of its location. Web cmd entry is particular used for administrator's remote monitoring command. Distribute RFID readers and the corresponding address according to the size and number of equipment of the laboratory. Equipment information stored in the database makes equipment locating and information searching easily.

MFC real-time monitoring module communicates with wireless data transceiver via the serial port (Ying Wang and Huajun Sun 2011; Jiabin 2010). It is responsible for generating data request packet, gathering information data collected by perception node and data processing. MFC real-time monitoring module interacts with MySQL database through the specific ODBC driver (Gong Cheng and Dianfu Yang 2002; Miao Mai and Ciyong Luo 2003). The MFC real-time monitoring module can collect data from all subperception nodes periodically or from a specific node irregularly. Its software interface is shown in Fig. 38.6.

In order to facilitate remote sharing of equipment information and related equipment usage information, JSP.NET module designs html pages for client browser and connects to MySQL database through the specific JDBC driver (Cailan Zhou et al. 2006). Client browser has access to database, forming the B/S mode

Fig. 38.6 Software interface of MFC real-time monitoring module

(Yajuan Wang et al. 2011) by JSP.NET module. Through a browser, public users can get the equipment names, type, and location information based on remote information sharing, at the same time, the administrator can send commands to do real-time equipment monitoring remotely.

38.6 Conclusion

A demonstration system is established within the Beijing Jiaotong University laboratory for remote monitoring and management. With a computer accesses to the Internet, administrator can scan the equipment information and usage information which making daily management become more convenient and easier. The perception nodes join the network in wireless multi-hop transmission way, without rewiring or moving the placement of equipments during the process of the system established.

After a period of actual operation, CC1100 wireless multi-hop transmission network can be steady operated without intervention, at the same time, the daily usage information of the equipment is normal. The usage information of equipment records is stored in the center management server database, and report generated directly through the web browser.

Remote equipment monitoring and management system based on IOT not only facilitates equipment management, but also provides a convenient platform to improve the utilization of equipments. With the rapid development of IOT, monitoring and management based on IOT will be widely applied in the future time.

Acknowledgment This paper is supported by National Natural Science Foundation of China (Grant No. 61172130) and RCS2009K008 from Beijing Jiaotong University.

References

Cailan Zhou, Lin Sun, Sufen Li (2006) Web database connection technology based on JSP. Comput Technol Dev 16(4):209–211 (in Chinese)

Feng Bao, Dongyun Lu, Xianfeng He (2006) JSP.NET technology and its architecture design. In: Proceedings of the episode of the scientific database and IT, Shanghai, China, pp 230–235 (in Chinese)

Gong Cheng, Dianfu Yang (2002) Comparison and selection about visiting technology of several database in VC. Appl Res Comput 19(2):82–84 (in Chinese)

Hongyue Dai, Xiaozheng Lai, Hailong Zhu, Shengli Lai (2007) Application of FPGA in real-time digital signal processing for UHF RFID system. Electr Meas Instrum 44(12):58–61 (in Chinese)

Jiabin Li (2010) The research of communication and management of meter reading system based on WSN. Ph.D. dissertation, Beijing University of Posts and Telecommunications (in Chinese)

Lijun Li, Daihua Wang, Jing Zu (2007) Design of wireless data transmission system based on CC1100. Foreign Electron Meas Technol 26(12):42–45 (in Chinese)

Miao Mai, Ciyong Luo (2003) The exploitation of database application by MFC ODBC Technique in VC. Inf Technol 27(3):75–76 (in Chinese)

Perkins CE, Bhagwat P (1994) Highly dynamic destination-sequenced distance-vector routing (DSDV) for mobile computers. In: Proceedings of the SIGCOMM '94 conference on communications architectures, protocols and applications, ACM New York, NY, USA, pp 234–244

Perkins CE, Royer EM (1999) Ad-hoc on demand distance vector routing. In: Proceedings of the WMCSA '99. Second IEEE workshop on mobile computing systems and applications. IEEE Computer Society, Los Alamitos, pp 90–100

Puchun Yu (2010) On the internet of things and internet of things technology development in China. Silicon Val 13:18 (in Chinese)

Wenbing Fan, Xiaoguang Cao, Yan Chen (2007) Design of RFID communication component based on middleware technology. Zhengzhou Univ (Eng Sci) 28(4):47–50 (in Chinese)

Xingbiao Mao (2010) Construction of colleges and universities equipment fixed asset management in IOT based on RFID technology. Educ Teach Forum 30:99–100 (in Chinese)

Xinsheng Wu, Xiaohong Mao, Keqin Li (2012) Design of colleges and universities asset management system based on IOT technology. Internet Things Technol 2(2):68–69 (in Chinese)

Xuhui Lan, Jiajun Xiong, Gang Deng (2004) Development of application program based on MySQL. Comput Eng Des 25(3):442–443 (in Chinese)

Yajuan Wang, Yi Zhao, Tao Wei (2011) Development and application of dynamic production facilities management system based on B/S mode. Silicon Val 17:106–107 (in Chinese)

Ying Wang, Huajun Sun (2011) Design and implementation of serial communication based on VC + +. Mod Electron Technol 34(12):19–20 (in Chinese)

Chapter 39
An Integrated Planning and Execution Model for Supporting Emergency Response

Pan Tang and Jian Zhou

Abstract This paper proposes an integrated planning and execution model for supporting emergency command and controlling process during emergency response. Firstly, a common incident action plan is designed for representing all the actions of responders and relations among them. Additionally, the planning process and execution process based on the incident action plan are introduced.

Keywords Emergency response • Incident action plans • Integrated planning and execution

39.1 Introduction

Extreme events are of large scale, and can have unpredictable outcomes and serious consequence. The nature of crisis situations today can extend far beyond a local area and requires actions and resources of many different organizations from the governmental and private sectors. Therefore, emergency response for extreme events involves response efforts from various functional teams and multiple jurisdictions (Smith and Dowell 2000). The crux of the coordination problem of emergency response is that exact actions and responsibilities of the individuals cannot be

The work was supported by the Fundamental Research Funds for the Central Universities of China (projects No. 21612301), the third phase of 211 major project of Guangdong province "Emergency management theory and practice", and National Science Foundation Grant of China (projects No. 90924301).

P. Tang (✉) • J. Zhou
School of Public Administration/School of Emergency Management, Research Center of Emergency Management, Jinan University, Guangzhou, China
e-mail: ptanghust2010@gmail.com; zhoujianjnu@126.com

E. Qi et al. (eds.), *The 19th International Conference on Industrial Engineering and Engineering Management*, DOI 10.1007/978-3-642-37270-4_39,

pre-defined (Murray and Michael 2004). Therefore, emergency managers develop unified incident action plans to coordinate the large response groups (Comfort 2004). However, managing long-term missions with limited resource requires deliberation mechanisms which are often computationally expensive, whereas evolving in a violate environment requires reactive behaviors for adapting to the dynamic response requirements. Additionally, the emergency response process is full of uncertainty. The execution process of the actions may not be the same as that when planning. It is not possible to create a planning model and a domain knowledge database that allows somebody to predict everything. In fact, emergency planning can also be direct toward improving and facilitating coordination by feedback of the execution process. The coordination by feedback assumes internal control by those in the front lines of a disaster situation.

Our research interest is to provide decision support for emergency command and controlling process to coordinating actions of different responders during emergency response. The objective of this paper is to propose an integrated planning and execution model (Lemai and Ingrand 2003) for supporting emergency command and controlling of extreme events response process. It takes into account the planning process and plan execution process and interactions between them for supporting the emergency command and controlling process. One critical work is to design a common incident action plan, to represent all the emergency response system's activities, from high-level planning to low-level behaviors and represent how all these activities are related to each other, along with their execution history. Moreover, the plan model is not designed for planning, but for a rich representation of the system's situation during execution. Additionally, how the planning process and execution process interacts should be defined based on the common action plan.

The remainder of this paper is list as following: Sect. 39.2 describes the process of developing and deploying incident action plans during emergency response. Then, an integrated planning and execution model for supporting it is proposed in Sect. 39.3. Section 39.4 defines the knowledge formalisms for representing incident action plans which take into account the characteristics of emergency response and execution context. Then, the concrete planning process and execution process is introduced in Sect. 39.5. Finally, the conclusions and future work are given in Sect. 39.6.

39.2 Description of Emergency Response Process

The emergency response involves multiple function disciplines (e.g., fire, law enforcement, medical services) and multiple levels of jurisdictions. For reducing the complexity of emergency response management, the process is divided to several periods, which are called operational periods (OP for short), as shown in Fig. 39.1. An operational period is the time which is scheduled for executing a given set of operation actions, as specified in the incident action plans (Annelli 2006). Operational periods can be of various lengths, although usually they last 12–24 h. The

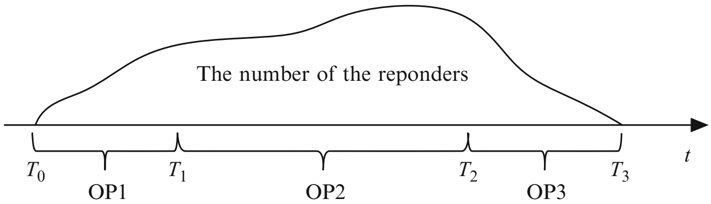


Fig. 39.1 Different operation periods in emergency response

emergency situation is dynamic and complex. Moreover, the number of participating responders changes in different operation periods during the entire process.

During each operation period, the broad spectrum of activities and organizations utilize both governmental and nongovernmental resources to plan for, respond to, and recover from extreme events. For coordinating the activities of multiple responders in different geographic location, the emergency command team (ECT) in emergency operation center commands and controls responders which achieve specific tasks during emergency response process (Wybo and Kowalski 1998). A formal description of command and controlling process for ECT is list as following:

Step 1. Develop and maintain the situation awareness. The first phase includes gathering, recording, analyzing and displaying incident, environment and resource information, which are also called emergency situation. The situation elements include magnitude, complexity and development trend of incidents, environment and available resource, weather conditions and so on. The situation awareness for ECT is the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning and the projection of their status in the near future (Endsley 1995).

Step 2. Establish a set of incident objectives and strategies. This step involves identifying the incident objectives by experience in specific domain based on current emergency situation. A set of incident objectives and strategies can be identified for this step. To be effective, they should be specific, measurable, assignable, reasonable and time-related (Hereth 2005). Additionally, evaluation criteria that represent preferences of ECT and constraints stating the limit and performance boundaries are also needed to be present in this step, such as preserving safety of human life, minimizing response costs, minimizing time involvement and so on.

Step 3. Develop incident action plans. After implementing the first two steps, the decision problem is identified, which states the responders must achieve incident objectives in the current emergency situation. Subsequently, step 3 is executed to find a set of feasible incident action plans for solving emergency response problem. The cognitive process can be regarded as a sequence of retrieval and refinement stages (Mendonca and Wallace 2007). In the first stage, emergency managers search a set of plan fragments named emergency operation plans or standard operation procedures (U.S. Department of Homeland Security 2012), which might be taken in the given planning state to achieve specific incident goals. Always, they define a set of steps with temporal constraints and can be viewed as fundamental building blocks of incident action plans. Then, in the second stage, they evaluate the alternatives and select an optimum or feasible one that actually will be taken next based on the evacuation criteria.

Step 4. Issue operation orders and supervise the execution process. This step involves translating the actions in incident action plans into operation orders, issue each operation order to specific responder. Then, the responders receive the operation orders and try to achieve the tasks describing by them. Meanwhile, the ECT monitors the execution process and contingencies in the response environment. Because the environment is dynamic, at every execution step the situation is observed and the responders try to execute the next step at the observed state. If the execution process is not the same as the developed incident action plans, the ECT modified the plans in the new situation and go on executing it.

According to the description of the emergency response process, the essence of emergency response commanding process is to develop and deploy incident action plans real-time. AI planning technologies will provide effective decision aid for supporting step 3. However, there is little research on how to support emergency managers to execute step 4 and the interaction between step 3 and step 4.

39.3 An Integrated Planning and Execution Model for Supporting Emergency Response

According to the emergency response process, the key steps are to develop and deploy incident action plans, supervise the execution process, and modify it according new situations. In this section, an integrated planning and execution model for supporting the emergency response process is proposed. Additionally, an integrated planning and executing cycle is introduced for describing the operation process of this model.

The integrated planning and executing cycle is shown in Fig. 39.2. Obviously, the planning process and execution process is interleaved. At time point t_0, the planning process is invoked at the beginning of each operation period. At time point t_1, the emergency managers decide to act, the planner terminates the planning process and sends the most optimal search node to plan manager module for execution. During the process of time point t_0 and t_1, the planning process constructs a search nodes space for planning the incident action plan to achieve the incident objectives identified at the time of t_0. When the incident action plan should be modified at time point t_2, the execution process is stopped and a new planning process to develop a new incident action plan is triggered. Additionally, the iterative planning and executing procedure proceed during the entire emergency responding process.

During the process, the planning module and execution module operate independently and interact continuously for achieving the emergency response tasks. The integrated planning and execution model is given in Fig. 39.3. At beginning

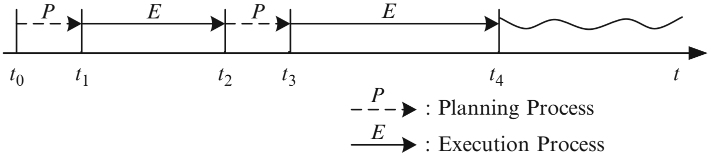
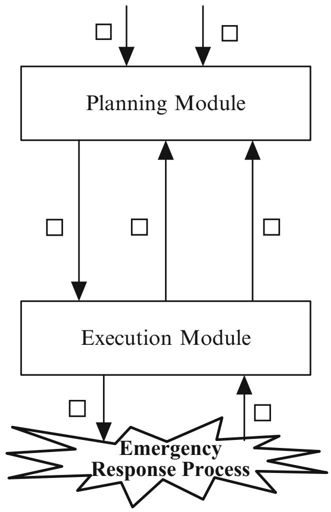


Fig. 39.2 Integrated planning and execution cycle

Fig. 39.3 An integrated planning and execution model



of each operation period, the initial emergency situation ① and the initial incident objectives ② which are identified based on it are input to the planning module. Then, the incident action plan ③ is generated by planning module and is input to execution module. The action in incident action plan is transformed to operation orders ④, which are executed by responders during the emergency response process. Meanwhile, the execution information ⑤ during emergency response process is fed back to execution module for evaluating whether the current incident action plan is feasible and effective. Otherwise, the current emergency situation ⑥ and new identified incident objectives ⑦ are input back to the planning module for generating new incident action plan in the new emergency situation.

39.4 Knowledge Formalism for Representing Incident Action Plans

An incident action plan represents all the responders will do in the future in a specific time during emergency response. Obviously, an incident action plans is interface of planning process and execution process in the above model. An incident action plan is a form of contract between the responders participating in the emergency response process. All the responding groups agree on the common plan and should promise to achieve specific tasks. However, the emergency response environment is dynamic and uncertain. Therefore, it is hard to develop a perfect and specific incident action plans for coordinating actions of multiple responders. In this section, a temporal action taking into account the execution context is proposed. Additionally, event, operator model are proposed for representing incident action plans and its execution context. Finally, the concrete knowledge formalism of the incident action plans is introduced.

39.4.1 Event Model

Definition 39.1 An event $event = (p\ flag\ \ tp)$ represents the state transition in a specific time point. p is an predicate instance representing an effect and tp is time point, when the event is scheduled to occur. The $flag_k$ is a Boolean value. If $flag_k$ has the Boolean value True, add the logical atom a_k to the current state, after the event is triggered. Otherwise, delete it from current state.

Events can be either contingent if they occur only because of external conditions or controllable if the emergency teams can make them occur. The time point representing the time the event occurs includes: the time when the execution effects of the operators or the state transition caused by the environment. An event is emitted or it occurs if the situation it represents has been met. In the following, the event structure of incident action plan describes the execution flow during the plan execution.

39.4.2 Action Model

An action is a basic unit of work performed by the specific elementary emergency response team. Firstly, the action is executed in an interval. Moreover, due to the uncertainty of the emergency response environment, the duration of the action can be specified before the action is executed. Secondly, there are complex processes behind the elementary actions. The effects can occurs in any time during execution. For example: the emergency actions executed by transportation teams include multiple steps by the members of them. Therefore, it is necessary to model the characteristics of these processes, such as variable action durations, conditions before specific time points and for an invariant time, and effects taking place at any time points during execution.

To get a more flexible representation for elementary emergency operations, we make the operators to be temporally annotated, including instantaneous preconditions and invariant preconditions. The instantaneous preconditions are those that must be satisfied before a specific time point. The invariant preconditions are logical atoms which are required to be hold for an interval. Additionally, an action composes of a set of events describing the change of the world induced by the task, which represent milestones in during the task execution.

Based on the operator model in PDDL2.1 (Fox and Long 2003), our representation of operators is defined as following:

Definition 39.2 Action is of the form (: action *head*, *instantPres*, *invarientCons*, $Del_i, Add_i, Del_d, Add_d$):

1. $head = (name, x_1, \cdots, x_n)$ is a primitive task, consisting of the operator's name and a list of parameters for describing it;
2. $instantPres = \{instantPre_i = (logicalExp_i @ t_i)\}$ is a set of instantaneous preconditions; $instantPre_i$ is a instantaneous precondition indicating that the logical expression $logicalExp_i$ must be satisfied by the state before the time t_i;
3. $invariantCons = \{invariantCon_i = (p_i - (st, et))\}$ is a set of invariant conditions, $invariantCon_i$ is a invariant condition representing that the predicate p_i cannot be deleted during the time interval $(startTime + st, startTime + et)$ during the planning process. (*startTime* is the time when the operator is executed).
4. $Del_i = \{p_i\}$ and $Add_i = \{p_i\}$ are a set of predicates representing the delete list and add list for the instantaneous effects. They give the negative and positive effects when apply this operator, respectively.
5. $Del_d = \{e_i = (p_{\mathrm{i}} - t_i)\}$ and $Add_d = \{e_i = (p_{\mathrm{i}} + t_i)\}$ (t_i is less than the duration of the action) define delete list of event and add list of event for representing the execution effects, consisting of a collection of events used to update the event queue. The events Del_d and Add_d is scheduled to be triggered in the event's specific time delta, which can be start time, end time or any time point during the execution.

39.4.3 Incident Action Plan Model

When the planning process and execution process interacts, a common action plan model is needed. Designing an effective incident action plan model for supporting planning and execution process is a key problem in our research. Due to the emergency environment is dynamic, making a perfect plan is nearly impossible. The plan generation for developing incident action plans manipulates simplified models of the environments. Moreover, it is only partially known in limed time during the emergency response process. Therefore, the state estimation of the emergency response process is also based on imperfect models and as such is subject to errors.

In most of HTN planning paradigms, the action plan only represents partial orders between tasks. In our incident action plan model, we assume that the actions are executed in a duration, which is not known. In addition, the exact time constraints between tasks must be represented for adapting to the characteristics of this application domain.

According to the requirement, we add two time points representing the start time and end time of each action and represent the time constraints between them by a Simple Temporal Network $STN = (X, D, C)$ (Dechter et al. 1991). $X = \{TR, t_1, t_2, \cdots, t_{2n}, t_{2n+1}\}$, t_{2i} and t_{2i+1} represent the start time and end time of an action; D is the domain of time variable in, and is always the domain $[0, +\infty]$; $C = \{c_{ij}\}$ is set of time constraints over on the time variables in X and the time constraint between time points tp_i and tp_j have the standard form: $a_{ij} \leq tp_j - tp_i \leq b_{ij}$. Based on the analyzing, the incident action plan is defined as following:

Definition 39.3 The incident action plan has the form *incidentActionPlan* $=$ (*actionSet*, *STN*), *actionSet* $= (a_1 \cdots a_m)$ represents a set of actions; *STN* is a simple time net defining all the time constraints over the start time and end time of these task.

The output of the planning process should be transformed to the above knowledge formalism. Then, execution of the above model supports the emergency managers to supervise the emergency response process.

39.5 The Planning and Execution Process

39.5.1 Planning the Incident Action Plan

It is necessary to represent a partial view of the future responding actions in emergency management. Therefore, it is very important to develop a command incident action plan. AI planning model is similar to cognitive process of developing incident action plans, which is shown in step 3 of emergency response model. Several HTN planners are designed for providing decision support to incident action plans developing process (Munoz-Avila et al. 1999). In this paper, a temporal HTN planner SIADEX (Castillo et al. 2006) for developing incident action plans is adopted for generating incident action plan. The input to the planner is the initial state, which describes the emergency situation at the beginning of the emergency response process identified by execution step1 in Sect. 39.2, and incident objectives identified by execution step 2. The output of the action plan can be transformed to the incident action plan model as Definition 39.3.

39.5.2 Temporal Execution Control

Execution control of action plans is a very active research domain which can support emergency managers to supervise the emergency response process. The incident action plan is a typical joint action plan, which is executed by multiple groups in different geographical locations. As the incident action plan is a temporal action plan, how to control the temporal network is the key of execution process.

The temporal execution controlling of STN is to decide the execution of actions and maps the time points in the incident action plans to their real execution time. On one side, the temporal executive decides the execution of tasks which is to determine the start time and end time of it. On the other side, when the time point related to the contingent events is met, it emits.

During the execution process, the temporal executive determines the next time point which should be executed and its current execution time. The action in incident action plan is fully instantiated just before starting its execution. The description of the execution process for an action is list as following:

1. Firstly, the action is transformed into operation order. Then, it is sent to specific emergency team for execution. Then, the current time is assigned to the time point representing the start time of this action. Then, a time constraint representing this real time for the start time point is added to the STN of the current incident action plan.
2. During the executing process, a report is sent back by the responders to emergency command team each time when the operation order is completed. Meanwhile, the current time is mapped into the end time point of the action. In fact, the real execution time assigned to the end time point of the task is the time when the report was received. Then, a time constraint representing that is added to the STN of the current incident action plan.
3. Additionally, the report which describes the completion of an action contains the information: the completion status (success or failed), and the actual exertion effects in case of failure. When the completion status is failed, the current incident action plan is abandoned and the new state is input to the planning process for developing new ones. Otherwise, the execution of the current incident action plan is going on.

To achieve efficient management of the temporal constraints generated during execution, it is necessary to check consistency of the STN associated to the current executing incident action plan, which records all the temporal constraints of actions in the current incident action plan. Once the time constraint representing the real execution time is added to STN, the algorithm PC-2 (Dechter et al. 1991) is invoked to check whether the STN associating to the current incident action plan is consistency. This is an incremental time constraint propagation method useful for planning problem, where the time constraints are posted increasingly as the problem is being solved. $O(|STN.X|^3)$. During the plan execution, when the STN is not consistent, that is to say some task violates a temporal constraint which is specified in the plan; the current action cannot be executed. Moreover, the incident action plan is abandoned.

39.6 Conclusion and Future Work

The temporal execution of a plan can lead to various needs for re-planning, including: (1) uncontrollable and controllable time points are time out; (2) new incident objectives are identified; (3) the tasks are failed to be executed. The future work is to develop the re-planning mechanism for adapting these characteristics. Additionally, the planning process and execution process are interleaved in this paper, which is not the real situation during emergency response. In fact, planning for emergency response is a continuous process. Therefore, another future work is to design effective methods for integrating planning process and execution process.

Acknowledgment Pan Tang author thanks Doctor Ming-lei Li for discussing the characteristics of the emergency command and controlling, which give me the inspire for designing the integrated planning and executing model.

References

Annelli JF (2006) The national incident management system: a multi-agency approach to emergency response in the United States of America. Rev Sci Tech 25(1):223–231

Castillo L, Fdez-Olivares J, García-Pérez O, Palao F (2006) Temporal enhancements of an HTN planner. Lect Note Comput Sci 4177:429–438

Comfort LK (2004) Coordination in complex systems: increasing efficiency in disaster mitigation and response. Int J Emerg Manag 2(1):62–80

Dechter R, Meiri I, Pearl J (1991) Temporal constraint networks. Artif Intell 49:61–95

Endsley MR (1995) Toward a theory of situation awareness in dynamic systems. Hum Factors 37(1):32–64

Fox M, Long D (2003) PDDL 2.1: an extension to PDDL for expressing temporal planning domains. J Artif Intell Res 20:60–124

Hereth LL (2005) Setting objectives in a unified command: the "cost" of leadership. In: Proceeding of 2005 international oil spill conference, 2005, Washington, DC, pp 7663–7671

Lemai S, Ingrand F (2003) Interleaving temporal planning and execution: IxTeT-eXeC. In: Proceedings of the ICAPS workshop on plan execution, Trento

Mendonca D, Wallace WA (2007) A cognitive model of improvisation in emergency management. IEEE Trans Syst Man Cybern Part A (Syst Hum) 37(4):547–561

Munoz-Avila H, Aha DW, Breslow L, Nau D (1999) HICAP: an interactive case-based planning architecture and its application to noncombatant evacuation operations. In: Proceedings of the sixteenth national conference on artificial intelligence, AAAI-99, Orlando, pp 870–875

Murray T, Michael C (2004) The design of a dynamic emergency response management information system. J Inf Technol Appl 5(4):1–36

Smith W, Dowell J (2000) A case study of coordinative decision-making in disaster management. Ergonomics 43(8):1153–1166

U.S. Department of Homeland Security (2012) National incident management system. http://www.icisf.org/articles/NIMS.pdf

Wybo JL, Kowalski KM (1998) Command centers and emergency management support. Saf Sci 30:131–138

Chapter 40
Structural Effects on Consumer Purchase Intention Towards Slimming Underwear

Wen Xiong, Hang Zhang, Shuang-jia Ren, and Jian-ping Wang

Abstract The purpose of this paper was to identify and get insight into the structural effects of the main factors that affected consumer purchasing intention towards slimming underwear by Structural Equation Modeling, based on the Theory of Planned Behavior. This study provided useful and practical suggestions for the seamless underwear industry, as the information provided in this study was essential for the development of marketing strategies. In spite of many studies that researchers had made of apparel consumer behavior, little was applied to the research of the underwear market. Therefore, the value of this paper lay in its presentation of a structural equation model for factors that affect consumer purchasing intentions towards slimming underwear, with the originality of taking both perceived benefit and perceived risk into consideration, whereas only one of which was explored in past studies of consumer behavior.

Keywords Perceived benefit • Perceived risk • Purchase intention • Slimming underwear • Structural equation modeling

40.1 Introduction

As women put more emphasis on good looks, the demands for solutions that can help create and maintain attractive body figures are growing increasingly strong. Slimming underwear is now considered a booming market where underwear no longer serves as a mere functional equipment of supporting, but an effective tool of shaping and slimming. As a result, brands from home and abroad are trying hard to gain market share in this valuable market. The competition is fiercer than ever.

W. Xiong • H. Zhang • S. Ren • J. Wang (✉)
Fashion Design Institute, Donghua University, Shanghai, China
e-mail: wangjp23@yahoo.cn

E. Qi et al. (eds.), *The 19th International Conference on Industrial Engineering and Engineering Management*, DOI 10.1007/978-3-642-37270-4_40,

However, with more choice than ever before, power has shifted from suppliers to consumers (Pitt et al. 2002). Thus, it becomes vital that marketers need not only improve their product quality, but also develop a comprehensive understanding of consumers. Hence, research on factors that influence consumer purchase intention of slimming underwear can make the difference between success and failure. The research subject was to identify and explore the structural effects that the factors (drawn from highly relevant literature) had on customers purchase intention and also to what extent the influences were, through developing an investigation model.

40.2 Methodology

40.2.1 Research Structure

A research model was established on the theoretical foundation of the Theory of Planned Behavior (TPB) as well as other highly relevant literature. As some theories and researches pointed out, potential variables connected with each other under causal relationships. The formation of potential variables and existence of causal relationships were used to create a complete model and the model fit was verified by observing the variables to explain the causal relationships (Reisinger and Turner 1999). The following hypotheses were proposed accordingly, as shown in Table 40.1.

40.2.2 Questionnaire and Data Collection

Subjects of this study were targeted on women who were over age 18, living in Shanghai. The questionnaires were handed out directly to consumers in five major

Table 40.1 Hypotheses

Item	Hypothesis
H1	There is relationship between perceived risk and perceived benefit.
H2	There is relationship between perceived risk and product knowledge.
H3	There is relationship between perceived risk and subjective norm.
H4	There is relationship between perceived benefit and product knowledge.
H5	There is relationship between perceived benefit and subjective norm.
H6	There is relationship between product knowledge and subjective norm.
H7	Perceived risk has a negative effect on the attitude.
H8	Perceived benefit has a positive effect on the attitude.
H9	Product knowledge has a positive effect on the attitude.
H10	Subjective Norm has a positive effect on the attitude.
H11	Perceived risk has a negative effect on the purchase intention.
H12	Subjective Norm has a positive effect on the purchase intention.

metropolitan areas. And the respondents were provided with information regarding the research subject and objective face-to-face by the author and other five trained fashion design and engineering students, so that the rate of effective questionnaires was insured. 320 questionnaires were distributed and 292 effective questionnaires were returned, the return rate being 91.25 %. The result conformed to the value (the number of returned samples should be at least five times the number of measuring variables) suggested by given that there are 6 variables 36 questions in this study.

40.3 Data Analysis

The quantitative data were processed using survey analysis software SPSS 17.0 (Statistical Package for the Social Sciences) and SEM (Structural Equation Modeling) program Amos 17.0.

40.3.1 Reliability and Validity Analysis (Refer to Table 40.2)

To identify the internal consistency of each attribute, the Cronbach's alpha coefficient was evaluated. Suggested that alpha values between 0.70 and 0.90 were sufficient. The Corrected Item-Total Correlation (CITC) was also tested to evaluate how each individual item acted with respect to the entire scale. Items with weak correlation (<0.20) were revised, excluded or maintained on the scale and the decision whether to exclude an item or not was based on a joint analysis of the CITC and the scale's adjusted alpha if the item was deleted (Di lorio 2005). The result of the calculation indicated that no item was to be deleted (CITC of all items >0.3), and the Cronbach's alpha values of all variables were higher than 0.7, indicating an excellent level of reliability, therefore, further statistical analysis was appropriate for using these measuring scales.

Subsequently, exploratory factor analyses were carried out with the suggestion of good construct validity under the condition that the factors extracted could explain more than 50 % of the total variance and factor loadings were over 0.5. To verify the samples' adequacy of the application of factor analysis, the Bartlett's test and the KMO (Kaiser–Meyer–Olkin) measurement were preceded. The index of KMO ranges from 0 to 1, where values of 0.90–1.00 indicate a remarkably high correlation, values of 0.80–0.90 a good level of correlation, values of 0.70–0.80 a medium level of correlation, values of 0.60–0.70 a reasonable level of correlation and values below 0.50 indicate unacceptable correlation (Hair et al. 2009). Based on Bartlett's test and the KMO value ($KMO = 0.870$, $p < 0.001$), the variables and data in this study were found to be well suited for factor analysis.

In this study, the items of each variable were factor-analyzed under the principal component analysis (PCA) method and the number of factors was extracted by

Table 40.2 Reliability and validity analysis

Variable and Items	Corrected item-total correlation	Cronbach's alpha if item deleted	Cronbach's alpha	Factor loading	Variances (%)
Perceived risk			0.815		44.136
PR1: SU products purchased may be counterfeit products	0.486	0.801		0.609	
PR2: There's an overuse of commercial ads for SU	0.571	0.788		0.684	
PR3: Ill comfortableness of SU may cause health problems	0.600	0.784		0.742	
PR4: Bad design of SU may cause health problems	0.637	0.780		0.771	
PR5: Vendors are unable to solve health problems caused by SU	0.533	0.793		0.677	
PR6: Slimming effects of SU are not good	0.526	0.794		0.645	
PR7: After service of SU stores is bad	0.457	0.804		0.580	
PR8: SU is not worth of the money	0.458	0.804		0.580	
Perceived benefit			0.777		48.664
PB1: SU can adjust body shape	0.542	0.740		0.704	
PB2: SU is a necessity for maintaining perfect figure	0.475	0.756		0.651	
PB3: SU can be used for quick slimming in case of emergency	0.471	0.757		0.641	
PB4: SU can improve self-image	0.690	0.702		0.838	
PB5: SU can help gain confidence	0.655	0.710		0.804	
PB6: SU can satisfy curiosity	0.342	0.790		0.490	
Product knowledge			0.866		71.404
PK1: I'm familiar with the brand of SU	0.723	0.826		0.853	
PK2: I'm familiar with the price of SU	0.738	0.820		0.863	
PK3: I'm familiar with the function of SU	0.672	0.847		0.811	
PK4: I'm familiar with the style of SU	0.731	0.822		0.852	

(continued)

Table 40.2 (continued)

Variable and Items	Corrected item-total correlation	Cronbach's alpha if item deleted	Cronbach's alpha	Factor loading	Variances (%)
Subject norm			0.769		68.479
SN1: My reference group considers SU a good choice	0.621	0.669		0.840	
SN2: People who are important to me agree my swearing SU	0.625	0.664		0.843	
SN3: My families and friends are wearing SU	0.563	0.734		0.798	
Attitude			0.798		71.771
A1: It's a wise choice to buy SU	0.690	0.680		0.872	
A2: I'm very interested in wearing SU	0.628	0.753		0.834	
A3: I hold a preferable evaluation toward SU	0.625	0.744		0.835	
Purchase intention			0.903		72.329
PI1: I'm willing to pay more attention to SU	0.748	0.584		0.844	
PI2: I'll wear SU very often	0.754	0.592		0.848	

Table 40.3 Reliability and validity analysis after modification

Variable	KMO	Bartlett's test	Cronbach's alpha	Variance (%)
Perceived risk	0.779	0.000	0.815	54.968
Perceived benefit	0.785	0.000	0.790	50.201
Product knowledge	0.751	0.000	0.866	71.404
Subject norm	0.793	0.000	0.769	68.479
Attitude	0.704	0.000	0.798	71.771
Purchase intention	0.879	0.000	0.903	72.329

Kaiser' criterion, in which only factors with eigenvalues >1 were considered adequate (Cahill et al. 2009; Ford et al. 1986; Hair et al. 2009; Hayton et al. 2004).

The result of factor analysis (Table 40.2) showed the factor loadings of each item (all above 0.5) and the total variance explained by the extracted factors (perceive risk = 44.136 %, perceived benefit = 48.664 %, product knowledge = 71.404 %, subjective norm = 68.479 %, attitude = 71.771 %, purchase intention = 72.329 %). Thus, it suggested that the items with the lowest loadings should be removed, until the variance explained by extracted factors was over 50 %. The modified results were shown in Table 40.3, where it was observed that both reliability and validity of all the measuring scales now reached the recommended level, with some yielding statistically significant results, verifying that those items eliminated had no significance and the scales have equally better results without these items, further indicating that the measurement scales can thus be applied to SEM testing.

40.3.2 Hypothesis Testing

The proposed model was analyzed by AMOS 17.0, using maximum likelihood estimation. The result of the chi-squared ($\chi 2$) test was 584.052($p < 0.001$), with 286 degrees of freedom, indicating that it was not a perfectly fitted model. However, more indices needed to be examined to further evaluate the fit level of the hypotheses. The root mean square error of approximation (RMSEA) was 0.060, which exceeded the acceptable range of 0.05. The goodness of fit index (GFI = 0.863) and the normed fit index (NFI = 0.858) were found to be below the recommended critical limit of 0.90. While the comparative fit index (CFI = 0.921 > 0.90) and the adjusted goodness of fit index (AGFI = 0.832 > 0.8) yielded significant result. Therefore, it could be inferred that the proposed model did not fit the data satisfactorily, which necessitated modifications of the proposed model.

Model modification often involves model building and model trimming, namely to add new paths or remove some of them to make the model more refined and recognizable. Two modification indexes were provided by AMOS, which were modification indexes (MI) and critical ratio (CR). Refinements were made step by step by inspecting MI results, which indicated that by adding paths between the residual errors with the highest MIs, $\chi 2$ would drop significantly. Moreover, paths with low coefficients but contributing to a great degree of MI were deleted. Subsequent modifications were undertaken by referring to CR results. The modified results were shown in Fig. 40.1.

It was shown that both consumers' attitude (Y13 = 0.43, $p < 0.01$) and subjective norm (Y12 = 0.59, $p < 0.01$) had positive effects on purchase intention towards slimming underwear. It could be concluded that consumers' willingness to buy slimming underwear were greatly influenced by the opinions of their families and peers, and could be reinforced by ways of forming preferable attitude. Perceived risk (Y7 = −0.24, $p < 0.01$), perceived benefit (Y8 = 0.3, $p < 0.01$) and subjective norm (Y10 = 0.68, $p < 0.01$) were found to have significant influences on attitude, while the influence that product knowledge had on attitude was denied. It indicated that consumers not only had high perceptions of risk and benefit of slimming underwear, but also paid a lot of attention to the opinions, beliefs, attitudes, and ways of acting of their reference groups. In addition, subjective norm was also closely connected with perceived benefit (Y5 = 0.36, $p < 0.01$) and product knowledge (Y6 = 0.43, $p < 0.01$), further indicating that subjective norm was one of the major factors directly and indirectly influence purchase intentions of consumers. There was a significant relationship between "perceived risk" and "perceived benefit" (Y1 = 0.29, $p < 0.01$), as some studies pointed out, consumers' purchasing

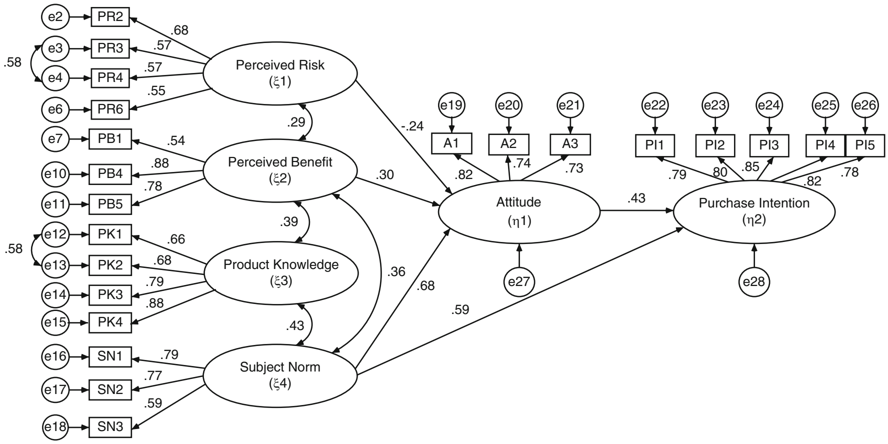


Fig. 40.1 Results of the revised structural equation model. $\chi 2 = 307.727$, $df = 198$, $P = 0.000$, GFI = 0.967, AGFI = 0.889, RMSR = 0.044, NFI = 0.914, CFI = 0.913

decisions were made based on the overall value perceived deriving from purchasing or using a specific product, and value was identified to be closely associated with perceived risk and benefit. Perceived benefit and product knowledge ($Y4 = 0.39$, $p < 0.01$) was also closely correlated, suggesting that by proper communication strategies, marketers could achieve substantial market share by means of enhancing consumer awareness of the slimming underwear.

40.4 Conclusion

First, this research focused and highlighted the structural effects of the factors that influence consumer attitude and purchase intention towards slimming underwear. Through the constructing and revising of the research model, a holistic view that offered detailed information of path coefficients allowed a better understanding of how factors affected purchase intention.

Second, this study offered theoretical enrichment to the limited research on consumer behavior in the slimming underwear industry. Despite numerous studies on consumer behavior of other kinds of apparels, studies of slimming underwear were scare, not to mention exploring the full relationships among attitudinal variables and their structural impacts on purchase intention.

Third, this study provided useful and practical suggestions for the seamless underwear industry, as the kind of information provided in this study was essential for slimming underwear companies to develop marketing strategies. Our findings regarding the impact of the attitude on purchase intention indicated that a positive attitude should result in a strong willingness of purchasing slimming underwear.

In addition, the buying intentions of surveyed consumers were quite strong, more than half of the subjects responded that they would like to purchase slimming underwear in the future. Consumers' preferable attitudes towards slimming underwear could be encouraged by increasing consumers' perception of either basic or elevated benefits, or reducing health, financial or psychological risks of slimming underwear, through effective and credible communication and promotion strategies. Subjective norm acted as a direct cause of consumers' attitudes towards slimming underwear, and stimulated higher purchase intention. However, the product knowledge gained by consumers was found to be not comprehensive enough to help consumers make their decisions.

All this indicated that consumers had yet to build firm beliefs about slimming underwear, as neither knowledge was sufficient enough, nor were relevant messages effectively conveyed to consumers. As a result consumers mainly drew references from their families and peers instead of getting information directly from the market. Hence, it is critical that effective promotion and communication strategies to be designed and stressed to convey sufficient and meaningful product information to consumers.

References

Cahill JM, Freedland-Graves JH, Shah BS, Lu HX, Klohe-Lehman DM (2009) Development and validation of the eating stimulus index in low-income, minority women in early postpartum. J Am Diet Assoc 109(9):1593–1598

Di lorio CK (2005) Measurement in health behavior methods for research and education, 1st edn. Jossey-Bass, San Francisco

Ford JK, Mac Callum RC, Tait M (1986) The application of exploratory factor analysis in applied psychology. A critical review and analysis. Pers Psychol 39:291–314

Hair JF, Black W, Babin BJ, Anderson RE, Tatham RL (2009) Análise multivariada de dados. Bookman, Porto Alegre

Hayton JC, Allen DG, Scarpello V (2004) Factor decision in exploratory factor analysis. A tutorial on parallel analysis. Organ Res Method 7(2):191–205

Pitt LF, Berthon PR, Watson RT, Zinkhan GM (2002) The internet and the birth of real consumer power. Bus Horiz 45(4):7–14

Reisinger Y, Turner L (1999) A cultural analysis of Japanese tourists: challenges for tourism marketers. Eur J Market 33(11/12):1203–1227

Chapter 41
Ant Colony Algorithm-Based Mixed-Model Assembly Line

Balancing Problem Modeling with Precedence Conflicts for Models

Yong-feng Xiong and Wen-sheng Yang

Abstract The existence of common tasks among models is a major feature of the mixed-model assembly line, traditional research were always done without precedence conflicts, so that the problem is simplified. However, the different production processes for different products often lead to conflicts in the actual production. The problem considered in this paper is how some common tasks can be duplicated to improve a mixed-model assembly line considering the precedence conflicts among common tasks for different products. Model was subjected to the constraint of precedence, assignment and cycle time; the objective function combined the efficiency with the smoothing index which can effectively distinguish the same number of workstations solutions. In addition, an ant colony algorithm with hybrid search mechanism is designed. Finally, the proposed mathematical models are illustrated and validated by means of a numerical illustration.

Keywords Ant colony algorithm • Duplicable tasks • Mixed-model assembly line balancing • Precedence relationships conflict

---

Foundation item: This work is supported by National High-Tech. R&D Program of China under Grant NO. 2011AA040603, Natural Science Foundation of China under Grant NO. 70871060 and the Fundamental Research Funds for the Central Universities of China NO. NUST 2011XQTR10.

Y. Xiong (✉) • W. Yang
School of Economics and Management, Nanjing University of Science & Technology, Nanjing, China
e-mail: xiongyongfeng14610@163.com

E. Qi et al. (eds.), *The 19th International Conference on Industrial Engineering and Engineering Management*, DOI 10.1007/978-3-642-37270-4_41,

41.1 Introduction

Since the assembly line balancing problem is formally proposed in 1954, a lot of researches have been done on the problem (Baykasoglu 2006; Song Huaming and Han Yuqi 2002, 2003). At first, their research is mainly focused on the single-model assembly line balancing problem (SALBP). With the changes in market conditions at home and abroad, as well as deepening of the study, a growing number of scholars and production practitioners have found that there is still a wide gap between the SALBP and the actual production, it cannot fully reflect the actual situation, eventually leading to the limited application of the obtained research results in reality. Thus more and more scholars focused their study on the mixed-model assembly line balancing problem (MALBP) which is more similar to the actual production. Essentially, MALBP is to assign tasks of P models into several successive workstations by satisfying many constraints like cycle time, precedence and assignment constraints so that single or multiperformance measures such as minimizing the number of workstations can be optimized. Merengo et al. (1999) analyzed some typical problems of manual, mixed-model assembly lines and presented new balancing and production sequencing methodologies to minimizing the rate of incomplete jobs and reducing WIP (Merengo et al. 1999). Noorul Haq et al. (2005) dealt with mixed-model assembly line balancing for N models, and a classical genetic algorithm was applied to minimize the number of workstations (Noorul Haq et al. 2005). Özcan and Toklu (2009) presented a new mathematical model for the two-sided mixed-model assembly line balancing problem to minimize the number of mated-stations and stations simultaneously (Özcan and Toklu 2009).

Compared with the literatures on SALB, the literature on MALB is relatively small. However, we can conclude two characteristics of the traditional research on the deterministic MALBP from the existing studies:

1. The basic assumptions of most research are almost the same.

As the MALBP is rather complicated, most scholars always have some assumptions of their own research in order to simplify and clear the problems. Some basic assumptions of these studies are roughly the same:

(a) Task time of each models are known constants.
(b) Precedence diagrams for each model are known.
(c) No WIP buffer is allowed between stations.
(d) Parallel stations are not allowed.
(e) Common tasks exist between models, which must be assigned to the same stations.

2. Most of the research methods are basically the same.

The products assembled on the same line are varied from the one basic model, so there are several common tasks among the products, and these tasks almost have similar precedence relationships. Thus, the similarity between the precedence relationships of different products has been utilized. And the traditional research on the deterministic MALBP is completed by adjusting tasks' completion times (Roberts and Villa 1970) or by combining the individual precedence diagrams of different products (Boysen et al. 2009).

On the other hand, ALB problem is a typical NP-hard problem, its complexity exponentially grows with the increasing number of tasks, and there are no accurate polynomial time algorithms for the optimal solution. Methods on solving the ALBP can be divided into optimization methods, heuristic methods and artificial intelligence. Optimization methods can establish a rational mathematical model and get the optimal solution, but the computing time is too long; Heuristic algorithms can overcome its weakness by using computer, but the solutions are the feasible solutions not the optimal ones; In recent years, artificial intelligence algorithms (genetic algorithms, simulated annealing, artificial neural network) have been developed rapidly, and performanced well in solving ALB problems.

Ant colony algorithm (ACO) proposed in 1991, was originally applied to solve the traveling salesman problem (TSP). ACO has so many advantages (such as self-organization, parallel, positive feedback, etc.) for solving complex combinatorial optimization problem. It's widely used to solve the problems such as production scheduling, logistics and so on. In recent years, it has also been applied to solve ALBP: Bautista and Pereira (2002) proposed an ant colony algorithm based on Dorigo's ant colony system for SALBP (Bautista and Pereira 2002). Then in 2007, they applied the algorithm to solve the ALBP with time and space constraints (Bautista and Pereira 2007). With the excellent performance in solving the SALBP, ACO has gradually been applied to the MALBP. McMullen etc. in 2003 proposed four heuristic algorithms based on ACO technology to address the MALBP with stochastic task time and parallel tasks (McMullen and Tarasewich 2003). Vilarinho etc. in 2006 proposed an ACO for MALBP Type-1: given cycle time C, to minimize the number of workstations and balance the load of each workstation simultaneously (Vilarinho and Simaria 2006).

Admittedly, the studies on the MALBP have made remarkable achievements. However most of the research has been done under the assumption that the precedence relationships of the common tasks for different products are not conflicting, so that the problem can be transformed into the SALB problem. However, in this paper, the traditional assumption is relaxed, the existence of conflict is allowed and common tasks are divided into duplicable tasks and nonduplicable tasks, the problem is called as "mixed-model assembly line balancing with precedence conflicts and duplicable common tasks (MALB-PCDC)". As far as the authors know, the conflict precedence relationships have been mentioned in a few literature but have not been formulated to date (Bukchin and Rabinowitch 2006), and some other scholars have taken into account the duplicable tasks but the conflict precedence relationships is ignored (Fokkert 1997). Here a mathematical model is proposed on the previous studies to address the problem. In addition, an ant colony algorithm based on the problem is designed with the consideration of avoiding the poor and inadequate results, and the numerical example shows its feasibility.

41.2 Description and Modeling of the MALB-PCDC Problem

41.2.1 Description of the MALB-PCDC Problem

What's the difference between the MALB-PCDC problem and the traditional MALBP? Here, we define MALB-PCDC problem as follow: Assume precedence conflicts exist among common tasks of different products, and research on how to improve the performance measures such as the line efficiency by making some common tasks duplicable. Obviously, the biggest difference between the two is that MALB-PCDC relaxed a most common constraint of the MALBP research: there are no precedence conflicts among common tasks for different products, and must be assigned to the same workstation. However, in practice, this may not always be possible due to the "conflicting precedence" caused by design differences. Thus the method of transform the MALBP into the SALBP does not always work. Conflicts in precedence relationships are an important barrier for a synchronized flow in mixed-model assembly lines and should be considered on its own in balancing.

41.2.2 Model for the MALB-PCDC Problem

The proposed model here was based on the assumptions as follows:

1. Products assembled on the assembly line are quite similar;
2. Task times are known and deterministic.
3. Common tasks among the products exist, and their completion times may not be the same.
4. Conflicting precedence relationships may exist among some of the common tasks of different products.
5. Some common tasks can be duplicated, and they can be assigned to different workstations. However, the other tasks can only be assigned to only one workstation.
6. Workload of each workstation cannot exceed the predetermined cycle time.

The following notations are used in the proposed model:

P: Products set assembled one the line
Tp: Tasks set of product p
N: Total number of tasks
n: Total number of the different tasks of all the products
TD: Common tasks set that can be duplicated
TND: Common tasks set that cannot be duplicated
J: Workstations set
C_T: Cycle time

TNP_i: Total number of products that require task i
PR_p: Precedence relationships set of product p
$(g, h) \in PR_p$: g is an immediate predecessor of h
t_{ip}: task time of task i of product p
c_i: Duplication cost of task i
AC: Available capital for task duplication

$$x_{pij} = \begin{cases} 1 \; \textit{if task i of product p is assigned to workstation j} \\ 0 \qquad \textit{else} \end{cases}$$

$$\mu_j = \begin{cases} 1 \; \textit{if workstation j is open} \\ 0 \qquad \textit{else} \end{cases}$$

$$y_{ij} = \begin{cases} 1 \; \textit{if common task } i \in TND \textit{ is assigned to workstation j} \\ 0 \qquad \textit{else} \end{cases}$$

$$z_{ij} = \begin{cases} 1 \; \textit{if common task } i \in TD \textit{ is assigned to workstation j} \\ 0 \qquad \textit{else} \end{cases}$$

Then the proposed model is as follow:

$$Min \sum_{j \in J} \mu_j \tag{41.1}$$

$$S.T. \; \sum_{j \in J} x_{pij} = 1 \quad \forall p \in P \;\; \forall i \in T_P \tag{41.2}$$

$$\sum_{a=1}^{n} a x_{pgj} - \sum_{b=1}^{n} b x_{phj} \le 0 \;\; \forall p \in P \;\; \forall (g, h) \in PR_p \tag{41.3}$$

$$\sum_{j \in J} t_{ip} x_{pij} \le C_T \quad \forall p \in P \;\; \forall j \in J \tag{41.4}$$

$$\sum_{p \in P} \sum_{i \in T_p} x_{pij} - N\mu_j \le 0 \quad \forall j \in J \tag{41.5}$$

$$\sum_{p \in P} x_{pij} - TNP_i \, y_{ij} = 0 \quad \forall i \in TND \;\; \forall j \in J \tag{41.6}$$

$$\sum_{p \in P} x_{pij} - TNP_i \, z_{ij} \le 0 \quad \forall i \in TD \;\; \forall j \in J \tag{41.7}$$

$$\sum_{i \in TD} \sum_{j \in J} c_i z_{ij} - \sum_{i \in TD} c_i \le AC \tag{41.8}$$

Obviously, Eq. (41.1) is the objective function, that's minimizing the number of workstations, and it's equal to maximizing the line efficiency as Eq. (41.11). In order to distinguish the same solution, smooth coefficient is taken into account as Eq. (41.12). That's:

$$m = \sum_{j \in J} \mu_j \tag{41.9}$$

$$T_{jp} = \sum_{j \in J} t_{ip} x_{pij} \quad \forall p \in P \, \forall j \in J \tag{41.10}$$

$$LE_p = \frac{\sum_{i \in T_p} t_{ip}}{m^* c} \quad \forall p \in P \tag{41.11}$$

$$SI_p = \sqrt{\frac{\sum_{j=1}^{m} (\max(T_{j\,p}) - T_{j\,p})^2}{m}} \quad \forall p \in P \quad \forall j \in J \tag{41.12}$$

Then we can transform Eq. (41.1) into the new objective function from Eqs. (41.9), (41.10), (41.11), and (41.12):

$$Max\, f(m) = Max \sum_{p \in P} (\omega LE_p - SI_p) \quad \forall p \in P \quad \omega \textit{ is user-defined and } \omega > 1 \tag{41.13}$$

Equations (41.2), (41.3), and (41.4) are the common constraints and followed by assignment, precedence and cycle time constraint. Assignment constraint ensures that each task should be assigned to at least and at most one workstation. Precedence constrain is that all precedence relationships among tasks should be satisfied. Cycle time constraint assures that the workload of a workstation should not exceed the cycle time.

In addition, Eq. (41.5) determines whether workstation j is utilized. If any task is assigned to workstation j, then $\mu_j = 1$; Otherwise, it will be 0. Eq. (41.6) is added to assure that the nonduplicable common tasks are assigned to the same workstation. On the contrary, Eq. (41.7) is to determine whether the duplicable task i is assigned to workstation j. $\sum z_{ij}$ represents the total number of different workstations which common task i is assigned to. Eq. (41.8) ensures that the task duplication cost does not exceed the available capital of the firms.

41.3 ACO Algorithms for MALB-PCDC Problem

ACO algorithms are population-based procedures inspired on the behavior of real ant colonies which imitates ants' foraging activity. While researching, the problem was abstracted to the node-model; different nodes correspond to different objects.

When it's applied to the ALB problem, the tasks assembled on the line turn into the nodes that ants need to travel while foraging, connections between the tasks and workstations are considered as the edges that ants travel through, then the process of selecting and assigning tasks to each workstation is treated as the process that the ant colony move along the edges under the influence of the pheromone intensity and heuristic information.

While designing the ACO algorithms for MALB-PCDC problem, there are two key steps including the construction of feasible balancing solution and the definition of the pheromone update rule.

41.3.1 Constructing Feasible Balancing Solutions

Before we describe the construction process of the feasible balancing solutions, the following two concepts are defined:

1. no-assigned task: task which is not assigned to any workstation;
2. assignable task: no-assigned tasks which satisfy the precedence constrain and cycle time constraint.

Here we get a feasible balancing solution by adding the assignable tasks to the corresponding workstations one by one until all the tasks are assigned. Then the constructing steps are following:

1. Open the first workstation;
2. Select tasks satisfying the precedence and cycle time constraint from the no-assigned tasks set to generate the assignable tasks set; if the no-assigned task set is empty, then go to step (5);
3. Select one task from the assignable tasks set to the current workstation according to task selection strategy. If the workstation is not full, then go to step (2); otherwise, go to step (4);
4. Open a new workstation then goes to step (2);
5. End.

After one iteration, m feasible solutions are available after the m ants finish their travel. And there is always a relatively optimal one among the m solutions.

The performance of ACO algorithm is influenced by the selected pheromones cumulative manner and heuristic rules. Here the pheromone trail τ_{pij} on the edge is defined as the expectations of assigning task i of product p to the workstation j, and the heuristic information of task i is taken as its position weight, that's:

$$\eta_{i\,p} = pw_{ip} = t_{ip} + \sum_{k \in F_i} t_{kp} \quad \forall i \in T_p \;\; \forall p \in P \tag{41.14}$$

Where F_{ip} is the owing tasks set of task i of product p.

General ACO algorithm always falls into the local optimal and has low explore ability. To overcome its weaknesses, a random selection is added to the general search mechanism, that's to say, tasks selection strategy is on the basis of the following hybrid search mechanism:

$$i = \begin{cases} I_1 \quad \underset{i \in N_j}{argmax}((\tau_{pij})^{\alpha}(\eta_{i\,p})^{\beta}) \quad 0 \le r \le r_1 \quad exploitation \\ I_2 \quad p_{ij} = \frac{(\tau_{pij})^{\alpha}(\eta_{i\,p})^{\beta}}{\sum\limits_{s \in N_j}(\tau_{psj})^{\alpha}(\eta_{s\,p})^{\beta}} \quad r_1 \le r \le r_2 \quad exploration \\ I_3 \; random \;\; selection \quad i \in N_j \quad r_2 \le r \le 1 \;\; random \;\; selection \end{cases} \tag{41.15}$$

Let r be a random number between 0 and 1, r_1 and r_2 are user-defined parameters; N_j is the assignable tasks set of the current workstation for the ant, and S is the assignable task; α β are parameters that determine the relative importance of pheromone intensity versus heuristic information.

According to the rules above, the tasks selected and assigned by the ant m are added to the $tabu_m$ list, when all the tasks have been added to the list, the ants will complete a tour, and a solution is available for the problem.

41.3.2 Pheromone Update Strategy

Pheromone update strategy of this article is a combination of local update and global update which is also commonly used in the ant colony system. The local pheromone update can reduce the impact of the selected assignment on the later ants, and then they can have stronger exploration ability on the edges that have not traveled yet. When it comes to the constructing process of the feasible balancing solutions, it's executed as follows: while the ant select task *i* to assign to workstation *j*, pheromone of the edge *ij* updates as Eq. (41.16):

$$\tau_{\mathrm{pij}_{(n+1)}} = (1 - \rho_1)\tau_{\mathrm{pij}_{(n)}} + \rho_1\tau_0 \quad \rho_1 \in [0, 1] \tag{41.16}$$

Where ρ_1 is the local evaporation coefficient, τ_0 is the initial pheromone level.

In order to enhance the search purpose, only the global optimal solution will increase the pheromone after the iteration and update the global pheromone as Eq. (41.17):

$$\begin{gathered} \tau_{\mathrm{pij}_{(t+1)}} = (1 - \rho_2)\tau_{\mathrm{pij}_{(t)}} + \rho_2\Delta\tau^{gb}{}_{pij\,(t+1)} \quad \rho_2 \in [0, 1] \\ \Delta\tau^{gb}{}_{pij\,(t+1)} = \begin{cases} f(m) \; (i,j) \; is \; belong \; to \; the \; current \; optimal \; solution \\ 0 \qquad\qquad else \end{cases} \end{gathered} \tag{41.17}$$

Where ρ_2 is the global evaporation coefficient, $f(m)$ is the value of the current optimal solution.

41.3.3 Steps of the ACO Algorithm

In summary, the algorithm can be summarized as the following six steps:

Step 1: Access to relevant data, such as task time matrix, precedence relationship matrix, common tasks set and the duplicable tasks set;
Step 2: Initialize the related variables, such as the pheromone matrix, the tabu list and so on;
Step 3: Open the first workstation, and select the first assignable tasks for each ant;
Step 4: To see if the workstation is full. If it's full, then open a new workstation; otherwise, go on selecting the next tasks. while selecting the next task, it's necessary to see if it's belong to the non-duplicable tasks set; If it's yes, then we should see if the same tasks of the other products are assigned to the workstation (if it's yes, the task should be assigned to the workstation, or another selection should be done); Otherwise, it's assigned to the workstation. The local pheromone updates after every selection, and generate a new assignable tasks set; Repeat until all the tasks are assigned.
Step 5: Record the optimal solution of the iteration and update the global pheromone after the iteration. Then compare it with the global optimal solution, if it's better than the global solution, and replace it.
Step 6: To see if the termination condition is satisfied: yes, then output the optimal solution and exit the algorithm; otherwise, clear the tabu list, and go to step 3.

The specific flowchart of the algorithm is shown in Fig. 41.1:

41.4 Numerical Illustration

As for research on the MALB-PCDC problem is at a preliminary stage now, there is still no standard test-data. In this section, a two-product problem is generated as the example and solved to illustrate the proposed model and the algorithm. While designing the example, the characteristics of the problem are taken into account. By solving the example, the following two points can be explained: first, it can clearly show how to solve the MALB-PCDC problem by duplicating some of the common tasks; second, the proposed mathematical model is well illustrated.

41.4.1 Description of the Example

In a firm, two similar products (product 1and product 2) are assembled on the same line, and their precedence diagrams are shown in Fig. 41.2. In Fig. 41.2, numbers in the circle indicate the task number, numbers a/b outside the circle denote the task time and duplication cost. From the precedence diagrams, we can know: (1) tasks 1,

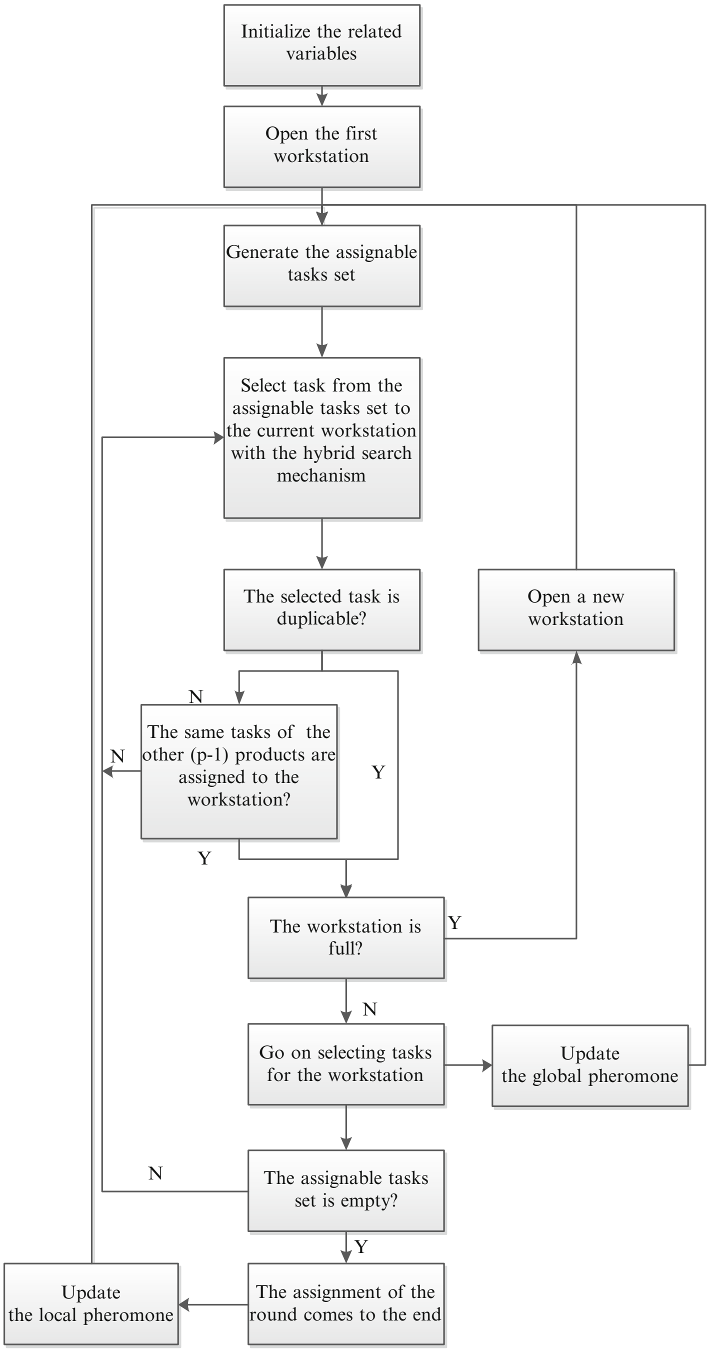


Fig. 41.1 The flowchart of the ACO algorithm

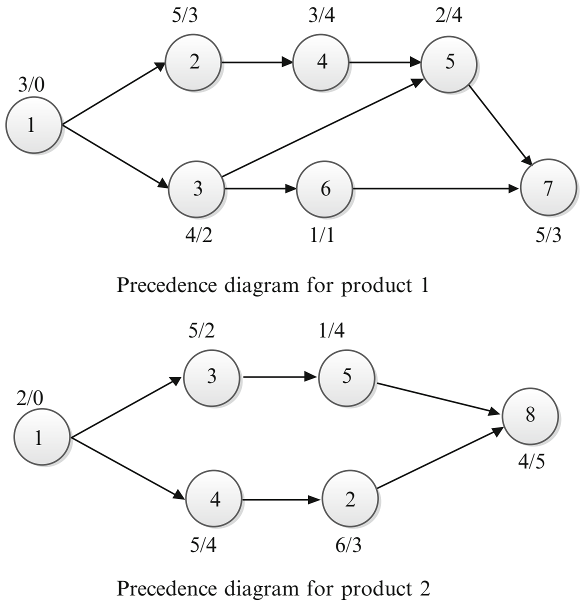


Fig. 41.2 Precedence diagrams

2, 3, 4, and 5 are common to both products. We supposed that tasks 1, 2, 4, and 5 can be duplicated, while task 3 cannot because of its specific technical operations; (2) conflicts are existed between their precedence: in product 1, task 2 is the immediate predecessor of task 4, while in product 2 is the opposite. That's a typical MALB-PCDC problem.

41.4.2 Computational Results and Analysis

The procedure was coded in MATLAB and run on a 2.6-GHz Pentium IV computer. The following values of the numeric parameters are used: the number of ants $N_{ant} = 15$, maximum iterations $Q_{\max} = 50$, and initial pheromone level $\tau_0 = 1$, pheromone intensity $\alpha = 1$, the relative importance heuristic information $\beta = 2$, local evaporation coefficient $\rho_1 = 0.2$, global evaporation coefficient $\rho_2 = 0.2$, task selection $r_1 = 0.7$ $r_2 = 0.9$,the maximum number of workstations allowed open $J_{\max} = 5$,available capital that can be invested for task duplication $AC = 15$ $\omega = 2.5$.We can know that when the $C_T = 7, 8, 9, 10$, the minimum number of the workstations are 4, 4, 4 and 3, the specific assignment shows in Table 41.1 by running the procedure.

From the computational results, we can know:

Firstly, the results show how the tasks are assigned, and which of the common tasks are duplicated to balance the line. For example, given $C_T = 7$, the minimum number of workstation is 4, 11 units funds is spent on the tasks 2, 4, 5 duplication to obtain the optimal configuration of the assembly line.

Table 41.1 Computational results

C_T		S_1	S_2	S_3	S_4	TD	AC	f(m)
7	**Product 1**	1, 3	2	4, 5, 6	7	2, 4, 5	11	
	Workload	**7**	**5**	**6**	**5**			1.11
	Product 2	1, 3	4	2	5, 8			
	Workload	**7**	**5**	**6**	**5**			
8	**Product 1**	1, 2	3, 4, 6	5, 7		2, 4, 5	11	2.12
	Workload	**8**	**8**	**7**				
	Product 2	1, 4	3	2	5,8			
	Workload	**7**	**5**	**6**	**5**			
9	**Product 1**	1, 2	3, 4, 6	5, 7		2, 4, 5	11	
	Workload	**8**	**8**	**7**				1.65
	Product 2	1, 4	3	2	5, 8			
	Workload	**7**	**5**	**6**	**5**			
10	**Product 1**	1, 3, 6	2,4	5, 7		2, 5	7	
	Workload	**8**	**8**	**7**				0.37
	Product 2	1, 3	4, 5	2, 8				
	Workload	**7**	**6**	**10**				

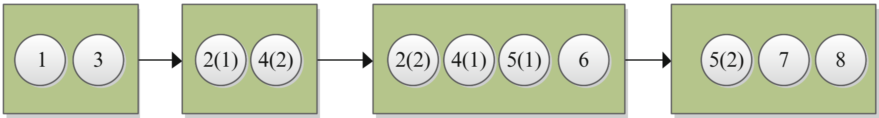


Fig. 41.3 Layout of the line given $C_T = 7$

Table 41.2 Solution contrast while $C_T = 7$

		S_1	S_2	S_3	S_4	LE	SI	f(m)
Optimal solution	**product 1**	1, 3	2	4, 5, 6	7	0.82	1.5	1.11
	workload	**7**	**5**	**6**	**5**			
	product 2	1, 3	4	2	5, 8	0.82	1.5	
	workload	**7**	**5**	**6**	**5**			
Non-optimal solution	**product 1**	1, 3	2	4, 5, 6	7	0.82	1.5	0.80
	workload	**7**	**5**	**6**	**5**			
	product 2	1, 3	4	2, 5	8	0.82	1.8	
	workload	**7**	**5**	**7**	**4**			

Secondly, the layout of the assembly line is available from the results. For example, while $C_T = 7$, the layout of the line is shown in Fig. 41.3. Each box means a workstation, circle in the box indicates task done in the workstation, number in the circle $T(p)$ represents task T of product p, for example, 2(1) represents that task 2 of product 1 is assembled in workstation 2, and the others are similar.

Finally, we can easily know that the model proposed can distinguish the same solution by comparing the results with the single object, see Table 41.2. Let takes the occasion of given $C_T = 7$ for example. There are many solutions if the object is to minimize the number of workstations, two of them is in Table 41.2. However, the model proposed can get the optimal solution from the same solution. From the contrast, the assembly line efficiencies are the same for the same solution, while the smoothing coefficients are different for the different assignments.

41.5 Conclusions

Traditional researches on the deterministic MALB problem do not consider the precedence conflict among different products, thus it can be transformed into SALB problem by combining the precedence diagrams, while it is not always the truth. Though the conflicting precedence relationships have been mentioned in a few literature but have not been formulated to date, and some other scholars have taken into account the duplicable tasks in their study but the conflicting precedence relationships is ignored. In this paper, duplicable tasks and conflicts are considered, and model was proposed. Model distributions under the traditional precedence relationships, assignment and cycle time constraints, and the assignment constraints of the duplicable and nonduplicable tasks are specified; besides the duplication capital constraint is added. The objective function combined the efficiency with the difference of the smoothing index which can effectively distinguish the same number of workstations solutions. In addition, an ACO algorithm is designed according to the model. To overcome falling into the local optimal and low explore ability, the hybrid search mechanism with exploitation, exploration and random selection is proposed, and the pheromone update strategy is a combination of local update and global update. Numerical illustration shows the feasibility of the model and algorithm.

Furthermore, deterministic MALB-PCDC problem is discussed here, and it can be extended to the more complexity and realistic problem such as stochastic problem. And the synergy of multiobjective optimization is also among the future work. Minimizing the duplication cost can be one of the objectives, so that decision makers can select the most appropriate assignment according to personal preferences.

Acknowledgment The authors of this paper gratefully acknowledge the financial support of National High-Tech. R&D Program of China under Grant NO. 2011AA040603, Natural Science Foundation of China under Grant NO. 70871060 and the Fundamental Research Funds for the Central Universities of China NO. NUST 2011XQTR10.

References

Bautista J, Pereira J (2002) Ant algorithms for assembly line balancing. Lect Notes Comput Sci 2463:65–75

Bautista J, Pereira J (2007) Ant algorithms for a time and space constrained assembly line balancing problem. Eur J Oper Res 177:2016–2032

Baykasoglu A (2006) Multi-rule multi-objective simulated annealing algorithm for straight and U type assembly line balancing problems. J Intell Manuf 17(2):217–232

Boysen N, Fliedner M, Scholl A (2009) Assembly line balancing: joint precedence graphs under high product variety. IIE Trans 41(3):183–193

Bukchin Y, Rabinowitch I (2006) A branch-and-bound based solution approach for the mixed-model assembly line-balancing problem for minimizing stations and task duplication costs. Eur J Oper Res 174(1):492–508

Fokkert JIvZ-d (1997) The mixed and multi model line balancing problem a comparison. Eur J Oper Res 100:399–412

McMullen PR, Tarasewich P (2003) Using ant techniques to solve the assembly line balancing problem. IIE Trans 35:605–617

Merengo C, Nava F, Pozzetti A (1999) Balancing and sequencing manual mixed-model assembly lines. Int J Prod Res 37(12):2835–2860

Noorul Haq A, Rengarajan K, Jayaprakash J (2005) A hybrid genetic algorithm approach to mixed-model assembly line balancing. Int J Adv Manuf Technol 28(3–4):337–341

Özcan U, Toklu B (2009) Balancing of mixed-model two-sided assembly lines. Comput Ind Eng 57(1):217–227

Roberts SD, Villa CD (1970) On a multiproduct assembly line-balancing problem. AIIE Trans 2(4):361–374

Song Huaming, Han Yuqi (2002) Genetic algorithms-based U-shaped assembly line balancing. J Syst Eng 17(5):424–429 (Chinese)

Song Huaming, Han Yuqi (2003) Combined U-shaped assembly line balancing. China Mech Eng 17(5):37–41, Chinese

Vilarinho PM, Simaria AS (2006) ANTBAL: an ant colony optimization algorithm for balancing mixed-model assembly lines with parallel workstations. Int J Prod Res 44(2):291–303

Chapter 42
Institutional Investors and Corporate Short-Term Debt Financing

Fang-fei Ding, Chan Gu, and Jin-hua Chen

Abstract China's listed companies have short-term debt financing preference. But institutional shareholders have the ability to influence the company's financing decisions. Through the analysis of 2009–2011 data, we found that the proportion of institutional investors holding has a significant negative correlation with the level of short-term debt. The results show that institutional investors effectively participate in corporate finance. Institutional shareholders inhibit short-term debt financing preference. The positive action is conducive to the development of capital markets and the improvement of governance structure of listed companies.

Keywords Corporate governance • Institutional investors • Short-term debt finance

42.1 Introduction

With the continuous development of the global economy, institutional investors have become an important force in the capital market. It played a positive role in stabilizing the market and protecting the interests of investors. Also, institutional investors influence the corporate governance. The financing decision is an important part of corporate governance. Studies have shown that in debt financing, due to the constraints of the short-term debt is relatively low, the listed companies had short-term debt financing preferences which has led to excessive liquidity risk of the listed companies.

F. Ding • C. Gu (✉) • J. Chen
Department of Accounting, University of Hunan, Changsha, China
e-mail: guchan68@sina.com

E. Qi et al. (eds.), *The 19th International Conference on Industrial Engineering and Engineering Management*, DOI 10.1007/978-3-642-37270-4_42,

As the supernormal development of institutional investors, it may have a certain impact on financing. Institutional investors can participate in the discussion of financing options, and vote on financing decisions, and submit proposed reporting and program of the EGM. These pathways will produce certain effect to financing. The structure of corporate finance relate to funding sources and operational efficiency, which playing a crucial role in the survival and development of the enterprise. But a short-term debt financing preference makes the financing of the structure imbalance. Institutional investors involved in corporate governance, whether dose it inhibit the short-term debt financing preference and balance the financial structure has a new theoretical and practical significance.

Many scholars have studied the relationship of institutional investors and corporate debt governance. Firth (1995) found that there is a positive relationship between the shareholding of institutional investors and asset-liability ratio. Also, institutional investors tend to increase long-term debt financing and improve the company's financial leverage ratio. Deming Zeng et al. (2006) considered that institutional investors with its strong financial strength to equity share holders of listed companies, changing the governance structure of listed companies (Deming Zeng et al. 2006). Institutional investors can actively participate in corporate governance. Dongliang Luo (2010) analyzed the stake of institutional investors and corporate debt level. He found that institutional investors are actively involved in corporate governance which affects the long-term debt levels.

Review of the literature, we know that the existence and development of institutional investors had a definite impact in corporate finance. Previous studies focus on the relationship between institutional investors holding and long-term debt levels, while neglecting short-term debt financing. Institutional investors play an important role in corporate governance. Therefore depending on the debt maturity; we study the affection of institutional investors on short-term debt financing. This enriches the existing institutional investors to participate in corporate governance theory research. Also, this is our contribution.

42.2 Theoretical Predictions

As the investors institutionalized, there exists Double Principal-Agent relationship (Friend and Lang 1988). On one hand, institutional investors hold List Company's share, forming shareholders and management Principal-Agent relationship. Institutional investors have the strength to participate in corporate governance due to its abundant capital and professional competence. On the other hand, institutional investors gather retail investor's funds, forming individual investors- institutional investors Principal-Agent relationship. Whether institutional investors invest company can have a better performance reflects its professional ability. So, it is the internal pressure of institutional investors to enhance their holding company's level of performance and affect the stock price gains. Thus, institutional investors have the motility and pressure to participate in corporate governance. Institutional investors take direct and indirect measures to impact financing decisions.

Institutional investors participate directly in financing governance mechanisms, mainly through the general meeting of shareholders, board of supervisors and board of supervisors. First, the general meeting of shareholders is the company's highest authority. Institutional investors have a larger share (Gillian and Starks 2000). Financing options and the use of funds report submitted to the general meeting of shareholders, institutional investors can ask questions and comment it. Institutional investors hold a higher proportion of shares which has more voting rights. Thus, it proposed financing program in general meeting have a higher probability to accept. Second, the board is the executive body of the shareholders' meeting (Fama and Jensen 1983). Financing options determined by the Board. Then board reported the financing program to shareholders' meeting to vote. Institutional investors through the election of members of the Board directly involved in the company's financial decision-making of the major options, which select the appropriate mode of financing. Again, the Board of Supervisors is the supervisory bodies. When the choice of financing options will maximize shareholder value, the Board of Supervisors mechanism will play a role.

Institutional investors indirectly involved in the mechanism by external capital markets. To a certain extent, the company's share price reflects the business performance and management of operating capacity, which can produce the governance effect. Operating performance and financing decisions are inseparable, which the financing affect a direct impact on corporate performance. If the means of financing for management take unreasonable, making the decline of corporate performance, thereby affecting the stock price, institutional investors will be replacing the company's management or sell stocks. These actions can threat the operator's control power, forcing the operators fulfill their duties and improve the use of capital efficiency. Therefore, institutional investors also take corporate control mechanisms to force operators to choose appropriate financing programs which can improve the company's performance.

In China, owing to the constraints of long-term debt financing and the under developed of bond market, relatively short-term debt financing is easier to obtain, which resulting a preference for short-term debt. This preference lead to liquidity risk and higher refinancing costs. Also it makes the corporate finance structure imbalance, which has a negative impact to companies. The existence and development of institutional investors through direct and indirect mechanisms to influence the financing decision and produce governance effect. Institutional investors in favor of balanced financing structure through inhibit short-term debt financing preferences. Based on the above analysis, we propose the following hypothesis:

H1: Institutional investors tend to reduce the company's level of long-term debt.

42.3 Data and Research Methodology

42.3.1 Sample Description

We elect 2009, 2010, 2011 data as the object of study. We selected these companies listed on the Shanghai Stock Exchange and Shenzhen Stock Exchange which the company's shares were held by institutional investors in 3 years. These data have non-ST and non-PT companies. There are 1,520 samples in 2009, 1,879 samples in 2010, 2,149 samples in 2011. First of all, because of the particularity of the finance and insurance companies, we removed the data. In 3 years, we excluded 112 firms belong to the financial industry. Secondly, we excluded 15 firms which asset-liability ratio was greater than 1. Finally, we got the last observed value of 5,421. The stake of institutional investor's data came from the WIND database, other data from CSMAR database.

42.3.2 Measurement of Major Variables

1. *Institutional investor's shareholding ratio*
Ownership of institutional investors is the listing Corporation's equity share of the quantization. Shareholding ratio determines the general meeting of shareholders voting rights. So, we choose the institutional investor's shareholding ratio as explanatory variables. This variable is abbreviated as INSTIT.

2. *Current debt ratio*
Current debt ratio is a measure of a short-term debt level indicator. Listed companies prefer to borrow more short-term debt financing, which has led to excessive liquidity risk. We choose the current liability ratio as explanatory variable. This variable is abbreviated as CDR.

3. *Operating income of the logarithm*
Due to higher revenue levels of listed companies, we choose the logarithmic of revenue to replace. Xiaoping Tan (2008) recommended that it is easy for listed companies to rely on short-term debt financing. We choose the Operating income of the logarithm as control variable. This variable is abbreviated as LNSALES.

4. *Liquidity ratio*
Current assets are an important guarantee for the repayment of current liabilities (Grossman and Hart 1980). The level of liquid assets restricted the ability to repay current liabilities. So we choose the liquidity ratio as control variable. This variable is abbreviated as LR.

5. *EBIT ratio of assets*
This is the indicator of the profitability of the business. According to a study by Myers and Majluf (1984), short-term debt financing period are mostly less than 1 year, and the constraints are relatively small which can meet the financing needs of enterprises in the short term. This variable is abbreviated as ROI.

6. *Operating profit ratio*
This is corporate performance indicators. It was used to explain the efficient use of funds in corporate finance (Cruchley and Jensen 1999). So we choose this indicator to measure the short-term debt on the business performance. This variable is abbreviated as ROC.

42.3.3 Models

According to the research needs, we create the following linear regression model 1 to verify the hypothesis 1.

$$\begin{aligned} CDR_{i,t} &= \alpha + \gamma_1 INSTIT_{i,t} + \gamma_2 LNSALES_{i,t} \\ &\quad + \gamma_3 LR_{i,t} + \gamma_4 ROI_{i,t} + \gamma_5 ROC_{i,t} \end{aligned} \tag{42.1}$$

42.4 Empirical Results

42.4.1 Descriptive Statistics

Table 42.1 is the short-term debt model descriptive statistics. We can see that the listing corporation current liabilities ratio average reached 68.34%. The maximum reached 1, indicate that listing corporation preference short-term debt financing. The proportion of institutional ownership mean reached 37.45%, which has less standard deviation. This phenomenon views that institutional investors have become a force in the capital market that cannot be ignored.

42.4.2 Correlation Test of Variables

Form the correlation Table 42.2 we can see that from 2009 to 2011 the current liability ratio and the ownership of institutional investors exists a significant negative correlation. The operating income ratio and liquidity ratio also have negative correlation with current liability ratio. ROI has positive correlation with the dependent variable but ROC is not. Form the correlation coefficient table we can

Table 42.1 The descriptive statistics of variables

Variables	Min	Max	Mean	Standard deviation
INSTIT	0.00	0.982	37.455	0.233
CDR	0.015	1.000	0.683	0.247
LNSALES	12.80	28.549	21.17	1.483
LR	0.019	1.000	0.743	0.226
ROI	−0.422	2.646	0.069	0.064
ROC	22.22	28.57	0.1170	0.789
N	5,421	5,421	5,421	5,421

Table 42.2 The result of correlation test

	CDR	INSTIT	LR	LNSALES	ROI	ROC
CDR	1.000					
INSTIT	-0.13^{***}	1.000				
LR	0.12^{***}	-0.14^{***}	1.000			
LNSALES	-0.11^{***}	0.36^{***}	-0.15^{***}	1.000		
ROI	0.07^{***}	0.12^{***}	0.05^{***}	0.08^{***}	1.000	
ROC	−0.001	0.01	−0.01	-0.07^{***}	0.12^{***}	1.000

Notes: ***. **. * present significant at the 1 %. 5 %. 10 % level significant is unilateral test results

Table 42.3 The result of the regression

	Non-standardized coefficient		Standardized coefficient		
	B	Standard error	Beta	T value	Sig.
(Constant)	0.885	0.052		16.967	.000
INSTIT	−0.103	0.015	−0.097	−6.727	.000
LR	0.100	0.015	0.092	6.765	.000
LNSALES	−0.012	0.002	−0.073	−5.041	.000
ROI	0.319	0.052	0.083	6.084	.000
ROC	−0.007	0.005	−0.022	−1.638	.102

see that the relationship between the independent variables are less than 80 %. Thus there is no significant collinear problem in this model which can enter the regression equation.

42.4.3 Regression Analysis

We can see from Table 42.3, the institutional investor's shareholding has a significant negative correlation with ratio of short-term debt in the 1 % level. Multiple linear regression results show that institutional investors effectively inhibited the short-term debt financing preferences. Greater the proportion of institutional investor's holdings exists lower short-term corporate debt level. The result indicates that institutional investors are actively engaging in corporate governance and balancing financing structure. While operating income of the company is negatively relate to short-term debt levels. On the one hand, the small companies prefer short-term debt financing; on the other hand, better performance companies have low short-term debt. The ratio of current assets and short-term debt level showing a significant positive correlation, indicating that the higher the level of short-term debt, the more liquid assets required to mortgage. Based on liquidity risk, the more debt required more protection for short-term. ROI has a positive correlation with the current ratio. Good performance companies have higher external debt demand on short-term debt. The operating margin does not pass the test which the relationship between the levels of short-term debt is not significant.

Regression tests on the model, indicating that the presences of institutional investors inhibit short-term debt financing preferences. This positive governance practices can balance the financing structure of listed companies which is conducive to the healthy development of listed companies and the stability of the securities market.

42.5 Conclusion

Financing options determine the effect of corporate governance contract. The short-term debt financing preference in China's listed companies made the financing structure imbalance. We analyzed that institutional investors have the ability and incentive to participate in the financing governance and affect the choice of financing methods. Then, we tested the 3 years of data of listed companies. The results show that the ownership of institutional investors has negatively relationship with short-term debt level. This shows institutional investors effectively reduce preferences short-term debt financing preferences which balance the financing structure of listed companies. Also, it shows that institutional investors actively participate in the financing governance which is conducive to the health of listed company's developments. This study provides new evidence for institutional investors to participate in corporate governance.

The balance of the capital structure is an important guarantee for the healthy development of listed companies. Institutional investors participate in corporate governance but we don't investigate the efficiency of the use of debt funds. We leave the puzzling problems for further study.

References

Cruchley C, Jensen M (1999) Agency problems and the simultaneity of financial decision making: the role of institutional ownership. Int Rev Financ Anal 8(2):177–197

Deming Zeng, Yuan Liu, Chong Wang (2006) China's institutional investors participating in the company management the necessity and obstacle analysis. J Hunan Univ (Soc Sci Ed) 20(5):76–78 (in Chinese)

Dongliang Luo (2010) Empirical analysis on institutional investor's propelling enlisted companies 'debt financing. Acad Exch 9(12):96–99 (in Chinese)

Fama E, Jensen M (1983) Separation of ownership and control. J Law Econ 26(2):301–315

Firth M (1995) The impact of institutional stockholders and managerial interests on the capital structure of firms. Manage Decis Econ 16(2):167–175

Friend I, Lang L (1988) An empirical test of the impact of managerial self-interest on corporate capital structure. J Financ 43(2):271–281

Gillian S, Starks L (2000) Corporate governance proposals and shareholder activism: the role of institutional investor. J Financ Econ 57(2):275–305

Grossman S, Hart O (1980) Takeover bids, the free rider problem and the theory of the corporation. Bell J Econ 11(1):42–64

Myers C, Majluf N (1984) Corporate financing and investment decisions firms have information that investors do not have. J Financ Econ 13(2):187–221

Xiaoping Tan (2008) Institutional analysis of preference for short-term debt financing of listed companies in China's. J Beijing Technol Busi Univ 23(2):78–84 (in Chinese)

Chapter 43
Multi-product Supply Planning for Combat Units in Battlefield Environment

Ji Ren, Xiao-lei Zheng, and Yue-jin Tan

Abstract The multi-product supply planning problem is investigated in the battlefield environment. The practical quantity of the products consumed by the combat units is stochastic, while the supplying process is also uncertain because of the random loss caused by attacks from the enemy. A nonlinear programming model is proposed to optimize the problem with both uncertain demand and supply consideration, and a solution algorithm based on Lagrangian relaxation is developed to obtain the optimal solution. Randomly generated examples involving 10, 100 and 1,000 commodities respectively are solved by the proposed algorithm. The computational performance of the algorithm is analyzed, which shows that the proposed algorithm can obtain optimal solutions for all examples with different sizes in short time.

Keywords Battlefield • Lagrangian method • Uncertain demand • Supply planning

43.1 Introduction

Current logistic support in battlefield environment emphasizes to supply the right fighting support commodities to the proper combat units, at right time, and in right quantity (Focused Logistic 2010; Tripp et al. 2006). There are many uncertainties in the battlefield, which make supply planning for the combat units very difficult.

J. Ren (✉) • Y. Tan
School of Information System and Management, National University of Defense Technology, Changsha, Hunan, China
e-mail: renji201203@163.com

X. Zheng
Department of Logistics, Lagistic Command College, Beijing, China

E. Qi et al. (eds.), *The 19th International Conference on Industrial Engineering and Engineering Management*, DOI 10.1007/978-3-642-37270-4_43,

The demand for the combat unit is highly uncertain. The quantity of commodities consumed during the combat process depends on the size of the battle, the armed equipments, tenacity for the enemy, the battlefield environment and so on. Many of these elements cannot be exactly anticipated. Therefore, an approximate estimate can only be obtained through analyzing the information about the battlefield situation.

There are many uncertainties in the supplying process. In battlefield, the logistic support system is always one of the most important targets attacked. In recent decades, as the application and generalization of information technologies, more and more reconnaissance equipments with high precision and long-distance attacking weapons are developed and equipped in army. Therefore, in current wars, the military logistic systems in battlefield suffer more frequent and severe attack from the enemy, which cause much loss when the commodities are transported from backside base to the front combat units. Besides, unexpected disasters, such as storms, can also cause some loss of commodities. All these uncertain events, including the attacks from the enemy and the disasters from the environment, result in many unexpected loss of commodities, which cannot be precisely calculated.

All the uncertainties have been observed in recent wars, such as the Golf War, the Iraq War. Their influence on logistic support systems has received increasing attention from commanders and researchers on military logistics. It is widely realized that both over and insufficient supply for commodities required by combat units has negative influence on the completion of missions. When the supplying quantity is over the practical requirement, it decreases the combat units' mobility and stealth. When the supplying quantity is lower than the practical demand, the combat unit can not sufficiently utilize its power.

Many new logistic theories are developed to deal with the logistic support problems in battlefield environment, like Focused Logistic (2010), Sense and Response Logistic (Tripp et al. 2006). But there is still lack of mathematical models for the supply planning problem that considers both the uncertain demand and supply in battlefield environment.

Motivated by the above challenges, the multi-commodity supply planning problem in battlefield environment is investigated, which consider the uncertainties in the supply and demand. A nonlinear programming model is developed to minimize the overall costs of the supplying commodities and the negative influence from the uncertainties, subjected to one critical capacity constraint. A solution algorithm based on Lagrangian relaxation is developed to calculate the optimal solution.

The paper is organized as follows. Section 43.2 presents the literature review, and the model formulation is presented in Sect. 43.3. The solution algorithm is developed in Sect. 43.4, and computational performance is analyzed in Sect. 43.5. The paper is concluded in Sect. 43.6.

43.2 Literature Review

The multi-product supply planning problem in battlefield environment considers the uncertainties from demand and supply, and it is related with the multi-item newsboy problem and the supply planning in battlefield. Therefore, the literatures related with these two aspects are reviewed here.

Hadley and Whitin (1963) first studied the multi-item newsboy problem with budget constraint and developed a solution approach. For related research before 1999 we refer the reader to (Khouja 1999), which presented a comprehensive review on NP.

As multi-product newsvendor problem is close to practical business, it is an active research topic in operation research area in recent decades. Many different formulations for the problem is developed, such as minimum regret formulations (Vairaktarakis 2000), quadratic programming (Abdel-Malek and Areeratchakul 2007), approximating programming (Niederhoff 2007), risk decision model (Zhou et al. 2008), mixed integer programming (Shi and Zhang 2010). And also, different solution approaches are developed, like dynamic programming procedures (Moon and Silver 2000), heuristics algorithm (Erlebacher 2000), generic iterative algorithm (Abdel-Malek et al. 2004), Lagrangian-based method (Abdel-Malek and Montanari 2005a, b), and hybrid intelligent algorithm (Shao and Ji 2006). Multi-product newsvendor problem is also be extended to the closed loop supply chain situation, where the uncertainties from used product return and demand are considered (Shi et al. 2010, 2011).

Logistic support is an important research topic in military operation research area. Gue (2003) considers the inventory problem in a sea-based logistic system, and develops a multi-period model by integrating the facility location and material flow decisions to support a given operational plan. Barahona et al. investigate the transportation and inventory allocation problem for logistics support in the network centric battlefield (Barahona et al. 2007). Toyoglu et al. (2011) study an ammunition distribution problem in the battlefield and develop a three-layer facility location and transportation routing model that can help distribute multiple commodities to the combat units, subject to time window constraints. Seen from current literatures in business management and military operation research, there are fewer models that are developed for the supply planning problem in battlefield environment with the uncertain demand and supply consideration.

43.3 Model Formulation

43.3.1 Notation

We first introduce some notations used to formulate the problem.

Indices:

$i = 1, \ldots, I$: index of commodities

Parameters:

$c_i =$ the cost for one unit of product i
$h_i =$ the consuming quantity of the transportation capability for delivering each product i
$HC =$ the maximum capacity of transportation capability
$g_i =$ the punishing cost for each product i when the received quantity is lower than the actual demand
$s_i =$ the punishing cost of each product i when the received quantity is more than the actual demand
$\alpha_i =$ the parameter that is used to balance the weight between the military value and the economic value of product i
$u_i =$ stochastic demand for product i
$f_i\ (\cdot) =$ pdf for the distribution of u_i
$F_i\ (\cdot) =$ cdf for the distribution of u_i
$x_i =$ proportion of product i that can safely arrive at the combat unit (a random variable)
$r_i\ (\cdot) =$ pdf for the distribution of x_i
$R_i\ (\cdot) =$ cdf for the distribution of x_i
$\bar{x}_i =$ the mean of x_i
$\sigma_i =$ the standard deviation of x_i.

Variables:

$Q_i =$ the allocation quantity of product i.

43.3.2 Model

The multi-product supply planning problem with one capacity constraint can be described as follows:

Min

$$\Pi = \sum_{i=1}^{I} \alpha_i c_i Q_i + \sum_{i=1}^{I} (1-\alpha_i) \begin{bmatrix} s_i \int_0^1 \int_0^{x_i Q_i} (x_i Q_i - u_i)\, f_i\,(u_i)\, r_i\,(x_i) du_i dx_i \\ +g_i \int_0^1 \int_{x_i Q_i}^{\infty} (u_i - x_i Q_i)\, f_i\,(u_i)\, r_i(x) du_i dx_i \end{bmatrix} \tag{43.1}$$

subject to

$$\sum_{i=1}^{I} h_i Q_i \le HC, \tag{43.2}$$

$$Q_i \geq 0, \forall i. \tag{43.3}$$

In function (43.1), the first term illuminates the proportion of total cost for all the products, while the second term is the proportion of loss caused by uncertain demand. Constraint (43.2) is the capacity constraint. Constraints (43.3) are nonnegative constraints.

43.4 Solution Approach

The solution algorithm consists of three main parts. First, constraint (43.2) is relaxed to obtain the Lagrangian dual formulation for the problem, which is divided into a number of single item subproblems. Then a bisection algorithm is developed to solve the dual problem.

43.4.1 Lagrangian Relaxation

Capacity constraint (43.2) is relaxed, and the Lagrangian relaxation problem can be formulated as.

Min

$$\begin{aligned} L\Pi = {} & \sum_{i=1}^{I} \alpha_i c_i Q_i - \lambda \left(HC - \sum_{i=1}^{I} h_i Q_i \right) \\ & + \sum_{i=1}^{I} (1-\alpha_i) \begin{bmatrix} s_i \int_0^1 \int_0^{x_i Q_i} (x_i Q_i - u_i) \, f_i\,(u_i)\, r_i\,(x_i) d u_i d x_i \\ + g_i \int_0^1 \int_{x_i Q_i}^{\infty} (u_i - x_i Q_i) \, f_i\,(u_i)\, r_i(x) d u_i d x_i \end{bmatrix} \end{aligned} \tag{43.4}$$

subject to (43.3),

where λ is the Lagrange multiplier for capacity constraint (43.2).

The Lagrangian relaxation problem can be divided into I single item subproblems.

Subproblem $L\Pi_i$:

Min

$$\begin{aligned} L\Pi_i = {} & (\alpha_i c_i + \lambda h_i)\, Q_i \\ & + (1-\alpha_i) \begin{bmatrix} s_i \int_0^1 \int_0^{x_i Q_i} (x_i Q_i - u_i) \, f_i\,(u_i)\, r_i\,(x_i) d u_i d x_i \\ + g_i \int_0^1 \int_{x_i Q_i}^{\infty} (u_i - x_i Q_i) \, f_i\,(u_i)\, r_i(x) d u_i d x_i \end{bmatrix} \end{aligned} \tag{43.5}$$

Substituting (43.5) into (43.4), the relaxed model can be rewritten as follows:
Max

$$L\Pi = \sum_{i=1}^{I} L\Pi_i + \lambda HC$$

subject to (43.3).

The relaxed problem can be solved by solving subproblems $L\Pi_i$.

In order to make the problem more tractable and obtain an optimal supply policy, we assume that the demand for commodity i is uniformly distributed with parameters $[a_i, b_i]$. Then a critical property can be obtained for function $L\Pi_i$, which is stated in Proposition 43.1.

Proposition 43.1 *Function $L\Pi_i$ is concave in the order quantities Q_i when the demand is uniformly distributed with parameters* $[a_i, b_i]$, *and the optimal solution is*

$$Q_i^* = \frac{\bar{x}_i\,(1-\alpha_i)\,(a_i s_i + b_i g_i) - (\alpha_i c_i + \lambda h_i)\,(b_i - a_i)}{(1-\alpha_i)\,(s_i + g_i)\left({\sigma_i}^2 + \bar{x}_i^2\right)}.$$

Proof of Proposition 43.1 When demand of commodity i is uniformly distributed with parameters $[a_i, b_i]$, the objective function of the subproblem can be reformulated as follows.

$$L\Pi_i\,(Q_i) = (\alpha_i c_i + \lambda h_i)\,Q_i + \frac{1-\alpha_i}{2\,(b_i - a_i)}\begin{bmatrix}(s_i + v_i)\left({\sigma_i}^2 + \bar{x}_i^2\right)Q_i{}^2 \\ -\,2\bar{x}_i\,(a_i s_i + b_i v_i)\,Q_i + s_i {a_i}^2 + v_i {b_i}^2\end{bmatrix}.$$

Then

$$\frac{d\Pi_i\,(Q_i)}{dQ_i} = \alpha_i c_i + \lambda h_i + \frac{1-\alpha_i}{b_i - a_i}\left[(s_i + v_i)\left({\sigma_i}^2 + \bar{x}_i^2\right)Q_i - 2\bar{x}_i\,(a_i s_i + b_i v_i)\right],$$

and $\frac{d^2\Pi_i(Q_i)}{d{Q_i}^2} = (1-\alpha_i)\,\frac{(s_i+v_i)\left({\sigma_i}^2+\bar{x}_i^2\right)}{b_i - a_i} > 0$.

Therefore, $L\Pi_i\,(Q_i)$ is concave.

Through solving $d\Pi_i\,(Q_i)\,/dQ_i = 0$, we can obtain that $Q_i^* = \frac{\bar{x}_i(1-\alpha_i)(a_i s_i + b_i g_i) - (\alpha_i c_i + \lambda h_i)(b_i - a_i)}{(1-\alpha_i)(s_i+g_i)\left({\sigma_i}^2+\bar{x}_i^2\right)}$.

43.4.2 Solving the Lagrangian Dual Problem

Bisection search is widely used to calculate the roots for nonlinear equations. In fact, solving the Lagrangian relaxation problem is to search the optimal multiplier λ which can make $B_{error} = 0$ where $B_{error} = \sum_{i=1}^{I} h_i Q_i^{L*} - HC$. Thus, a bisection iteration algorithm is developed by the basic idea of bisection method, which is presented as follows.

Step 0: Set $\lambda_1 = 0$ and $\lambda_2 = \lambda_{\max}$ ($\lambda_{\max}$ is the upper bound of λ).
Step 1: Let $\lambda = (\lambda_1 + \lambda_2)/2$, solve all the subproblems $L\Pi_i$ for $i = 1, \ldots, I$ by Proposition 43.1, and get their optimal solutions Q_i^*.
Step 2: Calculate $B_{error} = \sum_{i=1}^{I} h_i Q_i^{L*} - HC$,
Step 3: If $abs\,(B_{error}) \leq \delta_1$ or $abs\,(\lambda_1 - \lambda_2) \leq \delta_2$, STOP, else GOTO Step 4
Step 4: If $B_{error} < 0$, set $\lambda_1 = \lambda$, else set $\lambda_2 = \lambda$. GOTO Step 1.

Here δ_1 and δ_2 are set by the decision maker according the practice requirement. In our case, we set $\delta_1 = 1$ and $\delta_2 = 0.001$.

43.5 Computational Performance

Some performance analysis is presented in this section, which illustrates the effect of the solution algorithm developed in Sect. 43.4. The algorithm is programmed with MATLAB. The computational work for all examples is implemented on IBM T420 with 2GB of RAM.

In order to test the effect of the solution algorithm, we use the bisection algorithm to solve stochastic numerical examples. Sixty examples are stochastically generated and solved: 20 examples in small scale which include supply planning for 10 products, 20 examples in middle scale involving 100 products, and 20 examples in large scale that include 1,000 commodities. The demand of all products is assumed to be uniformly distributed, while the proportion of commodities that can safely arrive at the combat unit is following normal distribution.

All the parameters are stochastically generated in predetermined ranges which are set based on the study from data of history wars and practical war games. Table 43.1 summarizes the computational performance. As seen in Table 43.1, the average computational time for all examples can be treated as linearly increasing with the number of commodities. Even for examples in large scale, whose computational time is the longest, the proposed algorithm can solve the problem in 233 s.

Table 43.1 Computational performance of the solution algorithm

	Computational time (s)		
Problem size	Max	Min	Average
10	2.908	1.347	1.845
100	29.846	13.107	19.407
1,000	232.745	151.485	185.586

43.6 Conclusion

In the paper, we study the multi-product supply planning problem in battlefield environment with the considerations of uncertain demand and supply. A nonlinear programming model is developed to optimize the problem with one critical capacity constraint. A Lagrangian based algorithm is presented to solve the problem. Randomly produced instances that involve supply planning for 10, 100 and 1,000 commodities respectively are solved by the algorithm. Computational performance analysis shows that the proposed algorithm can obtain optimal solutions for all examples in short time.

References

Abdel-Malek LL, Areeratchakul N (2007) A quadratic programming approach to the multi-product newsvendor problem with side constraints. Eur J Oper Res 176(3):1607–1619

Abdel-Malek L, Montanari R (2005a) An analysis of the multi-product newsboy problem with a budget constraint. Int J Prod Econ 97:296–307

Abdel-Malek L, Montanari R (2005b) On the multi-product newsboy problem with two constraints. Comput Oper Res 32:2095–2116

Abdel-Malek L, Montanari R, Morales LC (2004) Exact, approximate, and generic iterative models for the multi-product newsboy problem with budget constraint. Int J Prod Econ 91(2):189–198

Barahona F, Chowdhary P, Ettl M (2007) Inventory allocation and transportation scheduling for logistics of network centric military operations. IBM J Res Dev 51(3/4):391–408

Erlebacher SJ (2000) Optimal and heuristics solutions for the multi-item newsvendor problem with a single capacity constraint. Prod Oper Manag 9(3):303–318

Focused Logistics (2010) A joint logistics roadmap, joint vision. www.acq.osd.mil/log/lmr/programs/focuslog.pdf. Last obtained in March 2012

Gue KR (2003) A dynamic distribution model for combat logistics. Comput Oper Res 30:367–381

Hadley G, Whitin TM (1963) Analysis of inventory systems. Prentice-Hall, Englewood Cliffs

Khouja M (1999) The single-period (new-vendor) problem: literature review and suggestions for future research. OMEGA Int J Manag Sci 27:537–553

Moon I, Silver EA (2000) The multi-item newsvendor problem with a budget constraint and fixed ordering costs. J Oper Res Soc 51(5):602–608

Niederhoff JA (2007) Using separable programming to solve the multi-product multiple ex-ante constraint newsvendor problem and extensions. Eur J Oper Res 176(2):941–955

Shao Z, Ji X (2006) Fuzzy multi-product constraint newsboy problem. Appl Math Comput 180(1):7–15

Shi J, Zhang G (2010) Multi-product budget-constrained acquisition and pricing with uncertain demand and supplier quantity discounts. Int J Prod Econ 128:322–331

Shi J, Zhang G, Sha J, Amin SH (2010) Coordinating production and recycling decisions with stochastic demand and return. J Syst Sci Syst Eng 19(4):385–407
Shi J, Zhang G, Sha J (2011) Optimal production planning for a multi-product closed loop system with uncertain demand and return. Comput Oper Res 38:641–650
Toyoglu H, Karasan OE, Kara BY (2011) Distribution network design on the battlefield. Nav Res Log 58:188–209
Tripp RS, Amouzegar MA, McGarvey RG et al (2006) Sense and respond logistics: integrating prediction, responsiveness, and control capabilities. Rand Corporation, Santa Monica
Vairaktarakis GL (2000) Robust multi-item newsboy models with a budget constraint. Int J Prod Econ 66(3):213–226
Zhou Y-J, Chen X-H, Wang Z-R (2008) Optimal ordering quantities for multi-products with stochastic demand: return-CVaR model. Int J Prod Econ 112(2):782–795

Chapter 44
An Optimal Layout Design for Storage Yard of Container Terminal

Jian-xun Tang and Li-xin Tang

Abstract This problem is to optimize the layout design of blocks in new container yard. The yard cranes have restrictions on the width of blocks, and the layout of blocks and berths also have important impact on the driving route of yard trucks. Making use of bays, the lengths of blocks are discretized. An integer programming model is proposed for this problem, whose objective is to maximize the efficiency of trucks and the number of bays. In this model, there is also a lower bound on the minimum number of bays. Based on the characteristics of this model, a Filter-and-Fan algorithm is developed to solve this problem. The experimental results show the effectiveness of algorithm.

Keywords Berth position • Container yard layout • Filter-and-Fan • Integer programming model • Yard truck transport

44.1 Introduction

The container and its supporting transport system have been widely used in the global logistics system, now the container terminal has gradually become an important part of large commercial ports. With the development of economic globalization and world integration, the container throughput increases year by year. So many ports are to establish or extend more container terminals.

This research is partly supported by State Key Program of National Natural Science Foundation of China (71032004), the Fundamental Research Funds for the Central Universities (N090104002, N100704002) and by the 111 Project under Grant B08015.

J. Tang (✉) • L. Tang
Liaoning Key Laboratory of Manufacturing System and Logistics, The Logistics Institute, Northeastern University, Shenyang, China
e-mail: jxtang@tli.neu.edu.cn; qhjytlx@mail.neu.edu.cn

E. Qi et al. (eds.), *The 19th International Conference on Industrial Engineering and Engineering Management*, DOI 10.1007/978-3-642-37270-4_44,

In order to increase competitiveness, ports are intent on improving their operation efficiency. Container yard is the most complex part of container terminal, the level of yard operating efficiency is a measure of port competitive strength. Reasonable arrangements for the yard operation can enhance the speed of terminal handling, and improve the utilization of container yard. Therefore, the majority of studies for container yard focused on this. Zhang et al. studied a storage space allocation problem in the storage yards of terminals. The problem is related to all the resources in terminal operations, including quay cranes, yard cranes, storage space, and yard trucks (Zhang et al. 2003). Sciomachen and Tanfani studied a container loading stowage plan problem to reduce the load time and improve the work efficiency of terminal equipment (Sciomachen and Tanfani 2007). Petering et al. studied a yard scheduling problem to reduce the number of stack shuffling and improve the utilization of yard cranes, and developed a real-time yard crane scheduling system (Petering et al. 2009).

Kang et al. proposed a method based on a simulated annealing search to derive a good stacking strategy for containers with uncertain weight information (Kang et al. 2006). Linn et al. studied a yard crane scheduling problem and gave an optimal configuration of the cranes (Linn et al. 2003). Li et al. studied a problem of truck scheduling and storage allocation. An improved tabu search heuristic with diverse neighborhoods was developed to obtain the solutions (Li and Tang 2009). Zhao et al. studied a problem of the simultaneous quay crane and truck scheduling at a container terminal, and proposed an improved PSO algorithm (Zhao and Tang 2011).

The above studies are all the related scheduling researches based on the existing container yard. There are not many researches about the design for new container yard. Container yard planning design has an important impact for the operating efficiency of ports. How to optimize the design of container yard has some theoretical value and practical significance. Kim et al. researched the yard layout optimization, storage space optimization and block area optimization respectively (Kim and Kim 2002; Kim et al. 2008; Lee and Kim 2010). Petering studied the length and width of blocks, crane configuration and yard layout for the overall performance of container terminals (Petering and Murty 2009; Petering 2009).

In the above studies, blocks in the yard are the same size, but in the actual investigation found that the size of blocks within a yard is not necessarily equal. Taking into account the functional division of yard's blocks and the different numbers of different function areas needing's bays, it is practical significance that blocks are designed for different lengths. This paper analyzes the effect that the berths distribution and yard layout on the routing of yard trucks, and gives the optimal design for blocks to maximize the truck transport efficiency and the number of bays.

The paper is organized as follows: Sect. 44.2 gives the description of container yard layout design problem, analyzes the effect that the layout of blocks on the routing of yard trucks. Section 44.3 gives a mathematical model for the problem. Section 44.4 gives a Filter-and-Fan algorithm to solve the problem. The results are analyzed in Sect. 44.5. Section 44.6 summarizes this article.

44.2 Problem Description

A container terminal is usually composed of berths and yards. Berths are provided for container vessels to dock. Yards are for container stacking. A yard is divided into rectangular regions called blocks. And blocks consist of bays. There are equipped with yard truck (YT) and yard crane (YC) in yards. YTs are used to transport the containers between container vessels and yards. YCs are used to transfer containers between YTs and stacks in blocks. A zone is a sequence of blocks that together form a single lane for YC movement. Figure 44.1 shows a kind of general layout of a yard.

The width of a block is limited by the span of YC. In this study, in order to let YCs do cross gantry movement between zones in the existing yards and the new yard, the width of new yard must be same to that of existing yards.

In the previous studies, the lengths of blocks are all equal in a yard. But it is found in actual research that the lengths of blocks are not necessarily the same in a yard, that is , the number of bays in blocks does not necessarily equal. Figure 44.2 shows three different layouts of blocks. As can be seen from the figure, for a yard with a given size, the number of blocks is fewer, the lengths of some blocks are longer, and then the number of bays in the blocks is more. In this paper the lengths of blocks are not necessarily equal.

Because YTs have to move around blocks, the length and the layout of blocks not only affect the number of bays, but affect the driving route of YTs, thereby the transport efficiency of YTs. Figure 44.3 shows the impact of blocks' length to the YTs' driving routes. In the figures "X" indicates the starting and ending point of driving routes, the arrows mean the direction of travels. It can be seen from the figures that when the block is smaller, the driving route is shorter, and the transport efficiency is relatively high, but the driving lane takes more yard area, and the number of bays will be reduced. Thus, there is need to decide the length and layout of blocks to maximize the transport efficiency of YTs and the number of bays.

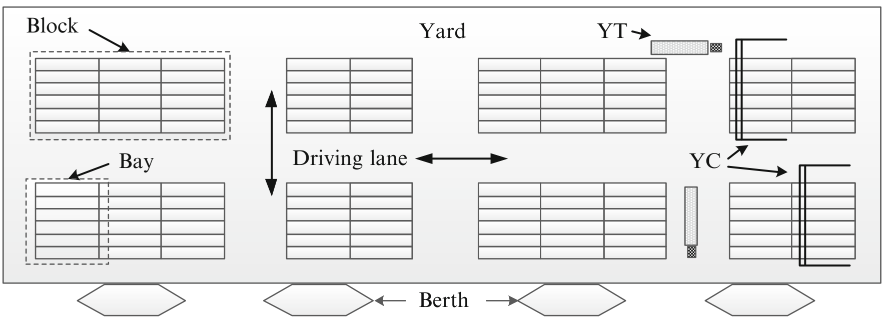


Fig. 44.1 Illustration of a kind of yard layout

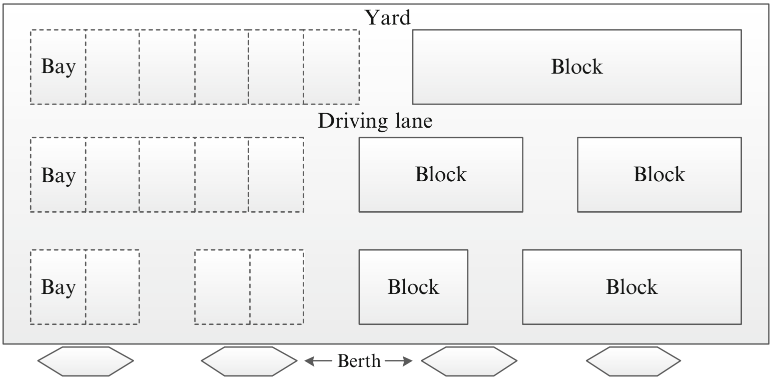


Fig. 44.2 Illustration of different blocks layout

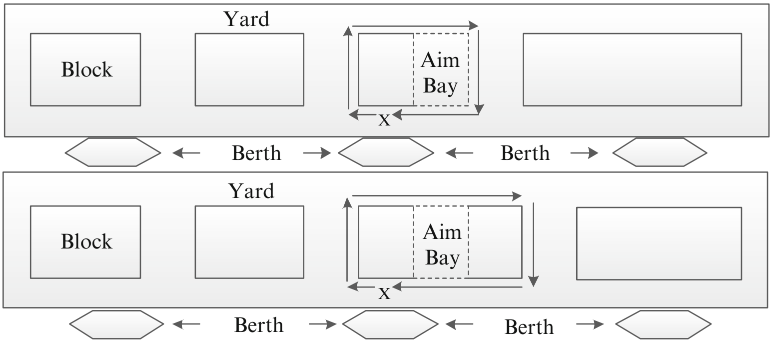


Fig. 44.3 Illustrations of impact to driving route

As shown in Fig. 44.4, for a given block layout design, from a same berth to the different blocks which positions are vertical to berth, the difference of driving route is the multiples of blocks' width, and unrelated to the layout design.

44.3 Problem Model

For a given yard, according the width of blocks the yard can be divided into several zones. Here the layout design of blocks is decided in the zone which is closest to berths. Without loss of generality, each zone is divided into n bits according the size of bay, and the bits are denoted as set $I = \{1, \ldots, n\}$from left to right, indexed by i.

Here berths' positions are known. For a given set of m berths, it is noted as $Q = \{1, \ldots, m\}$, indexed by q. Assume that the center position of q berth corresponds to d_q bit. These are shown in Fig. 44.5.

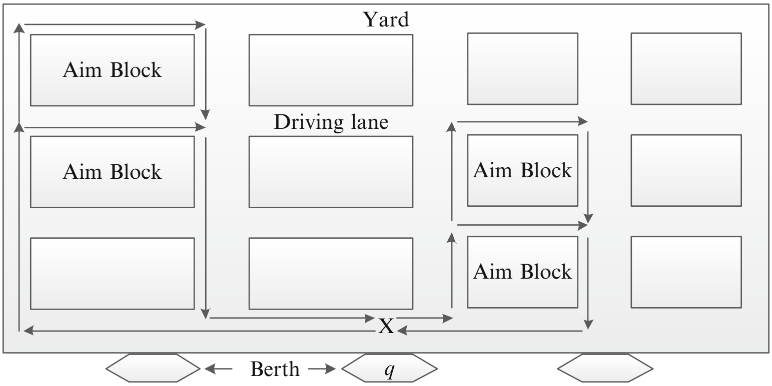


Fig. 44.4 Difference of travels by a yard truck with different aim blocks

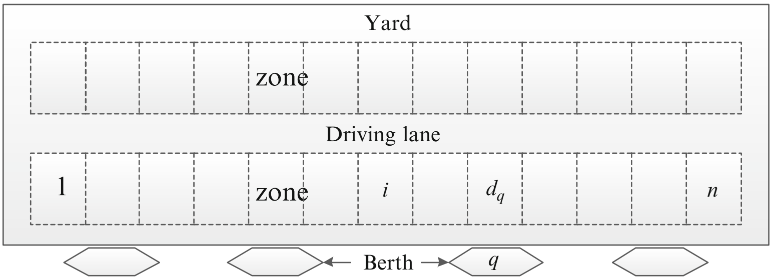


Fig. 44.5 Illustration of bits divided

A binary variable x_{ij} is introduced to describe the state of each bit. For bit i, $x_{ij} = 1$ means that the bit i is assigned to block j to stockpile container. $\sum x_{ij} = 0$(for all j) means that the bit i are not assigned to any one of blocks, and used as driving lane. A block is composed of the continuous bits which state is 1. There is driving lane between blocks in a zone. Assume that each driving lane only occupy one bit. When the states of all bits are determined, the layout of blocks is got in a zone. It is assumed that there are at most n blocks in a zone. Variable $Y_j = 1$ means that block j is used.

In a zone, the driving routes from a berth to all blocks are not same for different layout designs. For a given layout design, we can calculate the driving distance that YT starts from a berth to all blocks. In order to solve the problem easily, it is assumed that the center of each berth is the starting and ending point of driving routes for YT.

In order to express the driving distance, the following decision variables are introduced. The variable L_j means the index of the bit that is the left bound of block j. The variable R_j means the index of the bit that is the right bound of block j. Constant W_L means the width of block and W_B means the width of bay. Figure 44.6 gives an illustration.

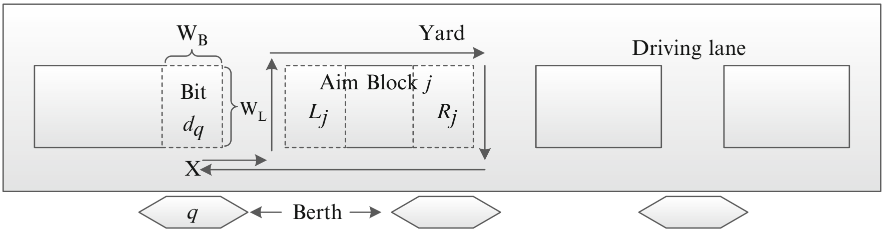


Fig. 44.6 Illustration of driving distance by YT

Then the driving distance from berth q to block j can be expressed as:

$$D_{qj} = 2W_L + W_B \left(\left|d_q - R_j\right| + 1 + \left|d_q - L_j\right| + 1 + R_j - L_j\right)$$

It is assumed that there is the same chance of loading and unloading container for YT to every block. For a given layout design, firstly calculate the driving distance of YT from a berth to each block, then get the product of distance and the number of bays in each block. The sum of products for all berths is used to evaluate the transport efficiency of YTs in the given layout design. In this way the model of the problem can be formulated as the following:

$$Min \sum_{q=1}^{m} \sum_{j=1}^{n} D_{qj} \left(\sum_{i=1}^{n} x_{ij} \right) + C \sum_{i=1}^{n} \sum_{j=1}^{n} b x_{ij} \tag{44.1}$$

$$s.t. \quad \sum_{i=1}^{n} \sum_{j=1}^{n} b x_{ij} \geq B \tag{44.2}$$

$$\sum_{j=1}^{n} x_{ij} + \sum_{j=1}^{n} x_{i+1,j} \geq 1, \forall i = 1, \ldots, n \tag{44.3}$$

$$x_{i-1,j} \geq x_{i,j}, \forall i = 1, \ldots, n; j = 1, \ldots, n \tag{44.4}$$

$$R_j = \max\{i x_{ij} | i = 1, \ldots, n\}, \forall j = 1, \ldots, n \tag{44.5}$$

$$L_j = R_j - \sum_{i=1}^{n} x_{ij}, \forall j = 1, \ldots, n \tag{44.6}$$

$$L_j - R_{j-1} \geq 1 + (Y_j - 1)M, \forall j = 2, \ldots, n \tag{44.7}$$

$$L_j - R_{j-1} \leq 1 + (1 - Y_j)M, \forall j = 2, \ldots, n \tag{44.8}$$

$$Y_j \geq (x_{ij} - 1)M + 1, \forall i = 1, \ldots, n; j = 1, \ldots, n \tag{44.9}$$

$$Y_{j-1} \geq Y_j, \forall j = 2, \ldots, n \tag{44.10}$$

$$0 \leq \sum_{j=1}^{n} x_{ij} \leq 1, \forall i = 1, \ldots, n \tag{44.11}$$

$$x_{ij} \in \{0, 1\}, \forall i = 1, \ldots, n; j = 1, \ldots, n \tag{44.12}$$

$$Y_j \in \{0, 1\}, \forall j = 1, \ldots, n \tag{44.13}$$

Here C is a penalty coefficient, M is a maximal positive number, b is the maximum number of containers stacking in a bay, B is a lower bound on the minimum number of containers stacking in a zone. In the objective function (44.1) the first section means the transport efficiency of YTs, and the second is the capacity of container stock. Constraints (44.2) ensure that the capacity of container stock of a layout design can meet the demand. Constraints (44.3) ensure that there have no two consecutive blocks assigned to driving lane. Constraints (44.4) ensure that the bits assigned to the same block must be sequential allocation. Constraints (44.5) and (44.6) define the right and left boundary to block j respectively. Constraints (44.7) and (44.8) define the difference between right border of block located in the left side and left border of block adjacently located in right side is 1. Constraints (44.9) define whether a block is used. Formulas (44.10) are the continuous constraints of blocks. Constrains (44.11) ensure that bit i is assigned to exact one block, if it has not been used as driving lane. Constraints (44.12) and (44.13) are the integer constraints for decision variables.

44.4 Problem Solving

Taking into account that other decision variables can be expressed by x_{ij}, and the value range of x_{ij} is $\{0, 1\}$, the best solution can be searched by changing the status of each bit, that is to change the value of x_{ij}. Filter-and-Fan (F&F) algorithm is more suitable to solve the problem, which neighborhood movements are used to change the status of bits.

44.4.1 Filter-and-Fan Algorithm

F&F is a heuristic algorithm to create a dynamic adaptive neighborhood for local search. This method was first proposed by Glover who used it to improve the scatter search (Glover 1998). Rego extended the algorithm to solve the combinatorial optimization problems (Rego and Glover 2010). Conceptually, the algorithm is composed of a filtering mechanism and a fan list of candidates. There are two search strategies, one is local search to find local optimum, the other is filter and fan search to skip from local optimum. When a strategy finds a new local optimum, the algorithm will switch to another search strategy, and keep the alternating until not improve the current best solution. To find a better solution than the current solutions, or to get a given search tree layer, the algorithm will stop, and return a new solution. This solution will be a starting point for the next search.

44.4.2 Initial Solution

By the constraints (44.2) it can be obtained that the minimum number of bits is $\lceil B/b \rceil$, which status are 1. It means that there are not more than $n - \lceil B/b \rceil$ bits that are arraigned to be driving lane. Here a simple strategy is taken to get an initial solution by a heuristic. The procedure of this heuristic can be described as follows.

Step 1: set the status of all bits to be 1.
Step 2: divide the zone into $n - \lceil B/b \rceil + 1$ blocks among which the difference of the number of bits is 0 or 1.
Step 3: select $n - \lceil B/b \rceil$ blocks from the left in turn, and set the status of the lase bit to be 0 in each block.
Step 4: an initial solution is got.

44.4.3 Neighborhood Structure

It is very necessary to introduce the standard terminology before the description of algorithm. In the local search, the neighborhood structure can be divided into simple neighborhood and compound one, depending on that the neighborhood solution is achieved through a separate submove or multiple ones. In a composite neighborhood, if more than one neighborhood is considered, then it is called a mixed neighborhood. Further, if the number of composite neighborhood‘s mobility is changing in the process of the algorithm, the algorithm can be called a variable depth neighborhood. F&F algorithm framework provides a means to construct such a variable depth design (Ranjbar 2008).

The F&F algorithm starts as a standard descent local search method by performing moves as long as they improve the best current solution. For the F&F, the initial solution is taken as the root node of search tree. Then on the root node μ_1 moves are conducted to create the first level of the F&F neighborhood tree. The next levels can be created as follows. The μ_1 solutions obtained by the μ_1 moves are denoted as μ_1 nodes. For each node, μ_2 moves are conducted to generate a total of $\mu = \mu_1 \times \mu_2$

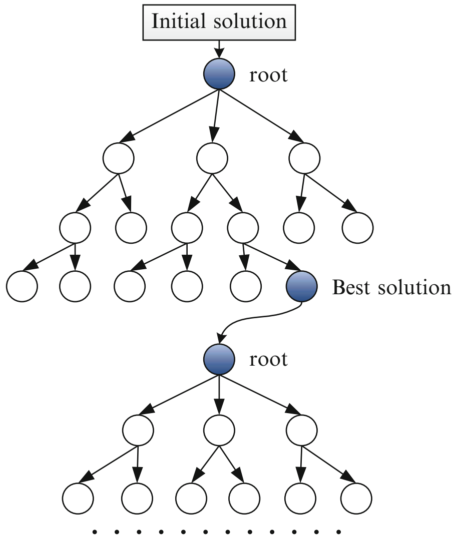


Fig. 44.7 Example of F&F procedure

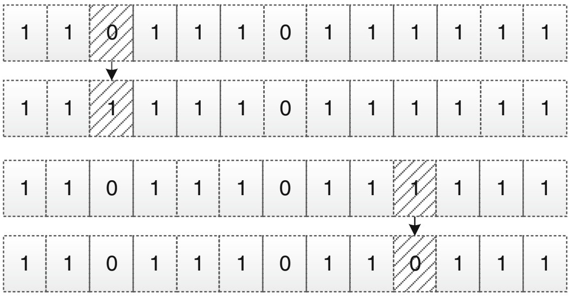


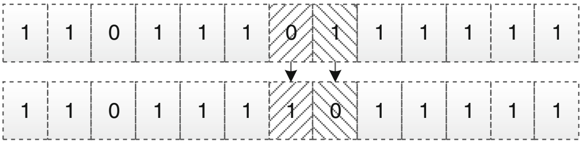
Fig. 44.8 Illustration of switch move

trial solutions. Then, from these μ solutions, μ_1 best ones are selected as the nodes for the second level of the F&F neighborhood tree. The third level can be obtained by repeating the above procedure. Here variables μ_1 and μ_2 are input parameters and can be changed in different problems. If an improved trial solution is found, or a given search tree level is got, the method stops branching and switches back to the descent phase starting with this new improved solution as the root node. Figure 44.7 gives an example of the flow chart of the F&F with $\mu_1 = 3$, $\mu_2 = 2$ and level $= 3$.

Here two neighborhoods are proposed based on simple moves. One is the switch neighborhood that is to switch the status of a bit from 1 to 0 or from 0 to 1. Through the switch move, the number of blocks can be increased or reduced in a zone. The other is the swap neighborhood that is to swap the bit whose status is 0 with the adjacent bit whose status is 1. The swap move can change the number of bays in two adjacent blocks, which is to change the position of the driving lane.

An example of switch neighborhood is given in Fig. 44.8.

Fig. 44.9 Illustration of swap move



In the left there is a bit which status is changed from 0 to 1, which means that a driving channel is cannel and the number of blocks reduces from three to two. In the right there is a bit which status is changed from 1 to 0, which means that a driving channel is added and the number of blocks increases from three to four. Figure 44.9 gives the example of swap neighborhood. There are two bits which statuses are 0 and 1, and have been swapped. By the swap the driving channel changes the position. The number of blocks is unchanging, but the numbers of bays in adjacent blocks are changed.

44.5 Computational Experiments

To test the performance of proposed F&F algorithm, the computational experiments are carried out on representative problem instances which are randomly generated on the extendibility of actual planning data of one famous port in the reasonable bound. The algorithm is implemented in C++, and all the tests are run on a personal computer with Pentium-IV 2.8 GHz CPU and 2 GB memory.

The number of bits is used to measure the size of yards. The combinations of the number of bits and the number of berths are {(100, 1) (300, 1) (300, 2), (500, 2), (500, 3) (700, 3), (700, 5), (1,000, 3), (1,000, 5), (2,000, 5), (2,000, 8), (2,000, 10)}. The positions of berth centers are randomly generated and there are not less than 100 bits between adjacent ones. For each scale of the problem, according to the position of berth center 20 instances are generated randomly. So there are 240 instances in total.

In F&F the setting of μ_1 and μ_2 affects the size of search space and the rate of decline. Through experiment, we prefer to adopt the following parameter setting: $\mu_1 = 10$ and $\mu_2 = 5$. A level limit is given for the neighborhood tree. If the limit is reached before getting improved solution, then the F&F will terminate and return the best solution it has found. In the experiment the level limit is set to 5.

The computational results are given in Table 44.1. In the table we give the average time and the maximum time, the average improvement margin compared with the initial solution.

Based on the results given in Table 44.1., the following observations can be made.

1. The average time of all test instances is not over 10 s. The maximum time is 10 s or so. These mean that the proposed F&F algorithm can provide solutions for the problem in a very short time.

Table 44.1 Computational results

	F&F algorithm		
Instance	Average time (s)	Maximum time (s)	Average improvement (%)
100×1	0.413	0.790	13.17
300×1	0.408	0.647	13.55
300×2	0.625	0.947	14.06
500×2	0.833	1.018	13.97
500×3	1.028	1.561	14.36
700×3	1.244	1.590	15.44
700×6	2.337	2.820	16.37
$1{,}000 \times 5$	3.725	4.135	19.21
$1{,}000 \times 8$	4.388	4.926	22.64
$2{,}000 \times 5$	6.325	6.918	25.47
$2{,}000 \times 8$	7.941	8.638	26.55
$2{,}000 \times 10$	9.021	10.334	28.33

2. The average improvements show that the algorithm can obviously improve the objective function within a reasonable computation time, so the F&F algorithm can solve the problem effectively.
3. As the size of problem increases, it is more difficult to solve, and the layers of the search increase, so that the computing time increases.

By the simulation experiments we found that the number and the location of berths have a direct impact on the layout of blocks.

44.6 Conclusion

In this paper we study the layout design of container yard where the width of blocks has been limited by the existing yard cranes. We discretize the lengths of blocks by using of bays, and formulate the effects of block's layout and berths on the driving routes of yard trucks, give an integer programming model for the problem with the objective of maximizing truck transport efficiency and the number of bays. Based on the characteristics of this model, a Filter-and-Fan algorithm is proposed to solve this problem. The experimental results show the effectiveness of algorithm.

References

Glover F (1998) A template for scatter search and path relinking. In: Hao J, Lutton E, Ronald E, Schoenauer M, Snyers D (eds) Artificial evolution, lecture notes in computer science, vol 1363. Springer, Berlin/Heidelberg, pp 1–51

Kang J, Ryu KR, Kim KH (2006) Deriving stacking strategies for export containers with uncertain weight information. J Intell Manuf 17(4):399–410

Kim KH, Kim HB (2002) The optimal sizing of the storage space and handling facilities for import containers. Transp Res Part B Methodol 36(9):821–835

Kim KH, Park YM, Jin MJ (2008) An optimal layout of container yards. OR Spectr 30(4):675–695

Lee BK, Kim KH (2010) Optimizing the block size in container yards. Transp Res Part E Logist Transp Rev 46(1):120–135

Li K, Tang L (2009) A tabu search algorithm for the integrated truck scheduling and storage allocation in quay. 2009 second international symposium on computational intelligence and design, Changsha, Hunan, pp 275–278

Linn R, Liu J, Wan Y, Zhang C, Murty KG (2003) Rubber tired gantry crane deployment for container yard operation. Comput Ind Eng 45(3):429–442

Petering MEH (2009) Effect of block width and storage yard layout on marine container terminal performance. Transp Res Part E Logist Transp Rev 45(4):591–610

Petering MEH, Murty KG (2009) Effect of block length and yard crane deployment systems on overall performance at a seaport container transshipment terminal. Comput Oper Res 36(5):1711–1725

Petering MEH, Wu Y, Li W, Goh M, Souza RD (2009) Development and simulation analysis of real-time yard crane control systems for seaport container transshipment terminals. OR Spectr 31(4):801–835

Ranjbar M (2008) Solving the resource-constrained project scheduling problem using filter-and-fan approach. Appl Math Comput 201(1–2):313–318

Rego C, Glover F (2010) Ejection chain and filter-and-fan methods in combinatorial optimization. Ann Oper Res 175(1):77–105

Sciomachen A, Tanfani E (2007) A 3D-BPP approach for optimizing stowage plans and terminal productivity. Eur J Oper Res 183(3):1433–1446

Zhang C, Liu J, Wan Y, Murty KG, Linn RJ (2003) Storage space allocation in container terminals. Transp Res Part B Methodol 37(10):883–903

Zhao J, Tang L (2011) The simultaneous quay crane and truck scheduling problem in container terminals. 2011 fourth international conference on intelligent computation technology and automation, Shenzhen, Guangdong, pp 279–282

Chapter 45
Analysis of Environmental Governance for Construction Project Based on a Dynamic Non-cooperative Game

Jiao-jie Han and Rui Zhao

Abstract The model of extended producer's responsibility (EPR) has been widely used for construction projects supervision. This paper mainly focuses on environmental governance based upon EPR model, from the perspective of game theory application. Moreover, a dynamic non-cooperative game theoretical model is built aimed at maximizing both economic and environmental benefits, in order to find out the main factors that may affect the producer's effort level. This paper also provides theoretical support for the environmental management of construction project as well as aids to improve the exiting supervisory mechanism of environmental protection.

Keywords Construction project • Dynamic non-cooperative game • Environmental governance • Extended producer responsibility

45.1 Introduction

With the social development and progress, more and more construction projects are launched within various industrial or municipal fields. However, any construction project cannot but give rise to the impact of environment from its inception until final completion, for instance the noise, dust etc. In particular, construction is a major source of waste materials, which accounts for 15 % of the total waste (Mcdonald and Smithers 1998). Moreover, the proportion is even higher in the developing

J. Han (✉)
School of Management and Economic, Southwest Jiaotong University, Chengdu, China
e-mail: hanjiaojie@gmail.com

R. Zhao
Faculty of Geosciences and Environmental Engineering, Southwest Jiaotong University, Chengdu, China
e-mail: Rui.Zhao@2009.hull.ac.uk

E. Qi et al. (eds.), *The 19th International Conference on Industrial Engineering and Engineering Management*, DOI 10.1007/978-3-642-37270-4_45,

countries, about 30–40 % in China (Ye and Li 2007). Although environmental protection and management has been put much more and more emphasis in the context of sustainable development, the economic profit of the project is still given more deliberation by the administrator. It is commonly considered by senior project manager that the project cost and schedule are much more important than the environmental management during project construction cycle (Shen and Tam 2002). Thus, how to enhance performance attached to the environmental management of engineering construction project is in need of solution, so as to achieve both economic and environmental profit. Here, the authors propose an approach using game theory to study the incentive and the restraining mechanism for environmental management of engineering construction project, thus to provide strategic options in order for developing new model of environmental governance.

Current studies on environmental management of engineering construction project can be mainly divided into two areas: how to implement the environmental governance and how to manage the construction waste in the project construction. For the former issue, Shen and Tam (2002) took the engineering project construction in Hong Kong as a case to study the difficulties of how to promote environmental management. They found the major factors to impede its implementation were additional administrative cost, employees or engineers who were lack of sufficient environmental education, lack of the support from the owner and other construction parties involved. Relevant research can be found by Ofori et al. (2002) who discussed the executive issue of environmental governance for construction project in the context of ISO14000. Used case studies in Malaysia, to analyze the factors which the project contractors are willing to pay in order to improve the environmental management, as well as factors given rise to the change of strategic behavior and attitude (Begum et al. 2007, 2009). For the construction waste management, Teo and Loosemore (2001) applied the psychological theory based upon a questionnaire survey to discussing how to minimize the construction waste, and then to help project managers improve attitudes towards waste. Later selected different evaluative criteria to enhance the waste sustainability, especially to use construction and demolition waste as a case example (Chuang and Lo 2003). These indicators are environmental desirability, economic optimization, social acceptability and equity, administrative diligence, respectively. Similarly, Kartam et al. (2004) studied the recycling efforts aimed at minimizing construction and demolition waste in Kuwait, and moreover discussed the feasibility to build the relevant recycling facility according to the economically efficient and environmental analysis. In addition, the incentive mechanism design for construction waste disposal has drawn an increasing attention of research scholars recently. For instance, Tam and Tam (2008) developed a Stepwise Incentive System (SIS) to strengthen the awareness of managers on controlling waste generation (Tam and Tam 2008). According to their local case study, the rate of waste reduction is demonstrated as 23 %. However, SIS lacks applicability of the engineering projects that have been signed by fixed price contract by the reason that rational owners are not willing to offer any additional rewards to the construction side.

From the above analysis, the authors indicate that total economic cost is one of the most important factors which deeply affect the implementation of environmental management during project construction. Although studies on environmental compensation and supervision are still progressing, incentive system is still an effective measure to help managers raise the level of environmental management. Therefore, this paper presents a dynamic game model based upon the EPR (Extended Producers' Responsibility) mode of engineering project construction with a fixed price contract, to analyze the major factors involved in the stakeholders' behaviors on environmental protection, as well as discuss the incentive and supervision mechanism for environmental decision-making on project construction.

45.2 Game Theoretical Modeling

45.2.1 Game Action Based on EPR Mode

Due to particular requirements of engineering project construction, *e.g.* safety, reliability *etc.*, the lifetime of construction products should be much longer than general manufactured products. Consequently, the construction side as a producer cannot be responsible for the whole lifecycle of project construction especially for the post-consumer stage, but Government, project owner or a third party will take over the duty for further environmental management. Thus, the key difference between the EPR implementation for manufactured products and construction products can be reflected as producer's responsibility in the specific stage of product life cycle. For engineering projects, we suggest in this study construction organization should take the duty of environmental protection only in the process of construction.

45.2.2 Fundamental Assumption

There are three players in this game, government, project owner, and construction organization, separately. Assume that they are all perfect rationality. The project profits are presumed to be divided into two parts: economic profits fixed by R_1, and environmental impact fixed by R_2. In this study, we assume the contracts are fixed price contract, defined as Y_1, Y_2 and Y_3 for construction organization, project supervisor and to others, respectively, satisfy $Y = Y_1 + Y_2 + Y_3$.

Except for production efforts, we assume, the project also requires the construction organization paid efforts on environmental management and protection, which is no less than a_0. Denote $a \in [0, 1]$ to be the effort of environmental behavior paid by construction origination. Set $p_a \in [0, 1]$ to be the probability that construction organization chooses $a = a_0$. So the probability that $a \neq a_0$ is $1 - p_a$. But if

$a < a_0$, then it will take the risk of being punished. If it is being found by owner, the owner will punish him ϕ_p value of money. Otherwise, if being found by both owner and government, the loss will be ϕ_g and $\phi_g > 0$. Meanwhile, the reputation of project owner will also be damaged thus to generate a loss η. η is related to the environmental impact, $\eta = \gamma R_2$ $(0 \le \gamma \le 1)$, where γ is the loss factor of owner's reputation.

In addition, the supervision behavior is costly for both project owner and construction organization. Assume the supervision efficiency of project owner on environmental protection is p_p $(0 \le p_p \le 1)$, the cost of supervision is $M(p_p)$, satisfying $M(0) = 0$, $M(1) = \propto$, which indicates it is impossible to monitor the construction organization's behavior completely to environmental management. In the real process of project construction, owner usually supervises the construction organization by means of project supervisor. Thus, the supervision cost can be equal to the contract value that project owner has paid to supervision organization $M(p_p) = Y_2$. Similarly, the supervision efficiency of Government is defined by p_g $(0 \le p_g \le 1)$, and the cost of supervision is $M(p_g)$, satisfying $M(0) = 0$, $M(1) = \propto$.

45.2.3 Dynamic Game Model

This game model should not only maximize the payoffs of all the involved stakeholders, but also satisfy with the incentive compatibility (IC) constraint and participation constraint (IR). Let project owner's reservation utility be V_0 and construction organizations be W_0. The game model shown as in following expression is built under the asymmetric condition.

$$
\begin{aligned}
& \max\ U = R - M\left(p_g\right) \\
s.t.\ & \max\ V = R_1 - p_g\left(1 - p_p\right)\left(1 - p_g\right)\eta - Y \quad (IC) \\
& \max W = (1 - p_a)\, W_1 + p_a W_2\, (IC) \\
& R_1 - p_g\left(1 - p_p\right)(1 - p_a)\,\eta - Y \ge V_0 \quad (IR) \\
& (1 - p_a)\, W_1 + p_a W_2 \ge W_0 \quad (IR)
\end{aligned}
$$

Decision-making on environmental protection of project construction based upon EPR mode can be understood as a dynamic game model. The sequence of game action is as follows: Firstly, Government sets up the environmental standard a_0 and the standard of economic sanction ϕ_g. Project owner starts to invite tenders, as well as construction parties bid on the price Y_1, which should satisfy their participant constraints. Then, project owner selects the successful bidder in terms of maximizing self-interest, and predetermines the supervisory degree p_p and punishment ϕ_p on the environmental management of project. Finally, construction party will choose the work effort as a to start construction, in order to maximize self-benefit.

45.3 Solution of Game Model

45.3.1 Equilibrium of Project Owner and Construction Organization

Assume a_0, ϕ_g and p_g are given, and the project owner satisfies his participation constraints, then the game model in the pervious section can be transformed as follows:

$$\begin{aligned} & \max\ V = R_1 - p_g(1-p_p)(1-p_a)\eta - Y \\ st\quad & \max\ W = (1-p_a)W_1 + p_a W_2 \qquad (IC) \\ & (1-p_a)W_1 + p_a W_2 \geq W_0 \qquad (IR) \end{aligned} \tag{45.1}$$

According to the Kuhn-Tucker condition, the in-equation of participation constraint (IR) can be transformed into an equation by the reason that project owner signs the fixed price contract with construction organization. Thus, the project owner will not undertake additional expense, $(1-p_a)W_1 + p_a W_2 = W_0$. For construction organizations, their optimal environmental effort satisfied that $\partial W/\partial a'^* = 0$. Thus, their optimal environmental efforts' level is $a'^* = 0$. From the project owner's perspective, the environmental effort that maximizes their utility a^* can be solved by $\partial V/\partial a^* = 0$, that is:

$$a^* = p_g \gamma k(1-p_p)(1-p_a)/c \tag{45.2}$$

In order to maximize the social benefit and minimize the side effect of construction, the environmental efforts from construction organization cannot be deemed as zero. So we assume that the optimal environmental effort level of construction organization should be a^*, not a'^*. Once the optimum effort level a^* has been determined by construction organization, project owner should select his supervision degree on environmental protection ${p_p}^*$ that can be calculated as from $\partial V\left(a^*\right)\big/\partial {p_p}^* = 0$:

$${p_p}^* = 1 - c[p_g\gamma(ka_0 + \delta\theta) + p_g\phi_g - \phi_p]/{p_g}^2\gamma^2k^2(1-p_a) \tag{45.3}$$

Similarly, the punishment standard set by project owner ${\phi_p}^*$ on the environmentally unfriendly behavior of construction organization can be derived from $\partial V\left(a^*, {p_p}^*\right)\big/\partial {\phi_p}^* = 0$, that is:

$${\phi_p}^* = p_g\gamma(ka_0 + \delta\theta) + p_g\phi_g - (1-p_a){p_g}^2\gamma^2k^2/c \tag{45.4}$$

Let $B = c[p_g\gamma(ka_0 + \delta\theta) + p_g\phi_g - \phi_p]/{p_g}^2\gamma^2k^2(1-p_a)$

$${Y_1}^* = {p_g}^2\gamma^2k^2(1-p_a)^2B^2/2c + (1-p_a)[(1-B)\phi_p + Bp_g\phi_g] + W_0 \tag{45.5}$$

45.3.2 Equilibrium of Government and Project Owner

Assume $a = a^*$, $p_p = p_p{}^*$, $Y_1 = Y_1{}^*$, the game model of government and project owner can be expressed as following.

$$\begin{aligned} &\max\ U = R - M(p_g) \\ st\quad &\max\ V = R_1 - p_g(1 - p_p{}^*)(1 - p_a)\eta - Y\ (IC) \\ &R_1 - p_g(1 - p_p{}^*)(1 - p_a)\eta - Y \geq V_0\ (IR) \end{aligned} \tag{45.6}$$

The participation constraint (IR) can be transformed into an equation based upon Kuhn-Tucker condition, as followed:

$$a_0{}^* = \frac{1}{k\gamma c(1 - p_a)}\begin{bmatrix} p_g\gamma^2k^2(1 - p_a)^2 + \frac{dM(p_g)}{dp_g}c \\ -(1 - p_a)(\gamma\delta\theta c + \phi_g c + k^2\gamma) \end{bmatrix} \tag{45.7}$$

$$\phi_g{}^* = \frac{1}{c(1 - p_a)}\begin{bmatrix} p_g\gamma^2k^2(1 - p_a)^2 + \dfrac{dM(p_g)}{dp_g}c- \\ (1 - p_a)(k\gamma ca_0 + \gamma\delta\theta c + k^2\gamma) \end{bmatrix} \tag{45.8}$$

$$\frac{dM(p_g)}{dp_g} = \frac{1}{c}\begin{Bmatrix} (1 - p_a)(\gamma kca_0 + \gamma\delta\theta c + \phi_g c + k^2\gamma)- \\ p_g\gamma^2k^2(1 - p_a)^2 \end{Bmatrix} \tag{45.9}$$

45.4 Strategic Options

45.4.1 Optimum Effort Level of Construction Organization

According to the above expression (45.2), the optimum effort level a^* that construction side paid to environmental protection is related to the monitoring intensity of government p_g and project owner p_p, loss factor of project owner's reputation γ, and the environmental impact factor k caused by the specific effort level selected for environmental protection.

Moreover, a^* is proportional to p_g, γ and k since $\partial a^*/\partial p_g > 0$, $\partial a^*/\partial \gamma > 0$, $\partial a^*/\partial k > 0$. While governmental supervision intensity rising, construction organization's negative behavior on environmental management will be more inclined to find out, as well as the possibility of project owner's reputation loss should be also increased. In addition, the more reputation the project owner loses, the more environmental impact generates. Therefore, rational construction party will still choose higher effort level for environmental protection. Option 1 will describe how to achieve the higher level in detail.

Option 1. In order to raise construction organization's effort level on environmental protection, such measures can be taken as intensifying governmental administration and increasing project owner's reputation loss while lacking of effective environmental supervision, enhancing the awareness of environmental consequence for construction teams.

The higher comprehensive quality attaches by the construction teams, the lower optimum effort level should be chosen, as $\partial a^*/\partial p_a < 0$. Thus, the high-quality construction team will pay less effort than the low-quality one aimed at achieving the same environmental standard. This discussion is summarized in Option 2.

Option 2. The comprehensive quality can partly reflect the environmental awareness of construction organizations, which is related to their effort level. Thus, to select the higher comprehensive quality team can be more beneficial to further project development, especially beneficial to environmental protection during project construction. Moreover, the environmental awareness of construction team can be raised correspondingly by relevant education and training.

The less cost spent by construction organization, the higher the effort level, as $\partial a^*/\partial c < 0$. Furthermore, the stronger supervision of project owner raises, the lower the efforts level construction teams select, as $\partial a^*/\partial p_p < 0$. Due to this higher monitoring intensity which results in higher environmental standard for construction project, it is more difficult for construction teams to start work, thus to increase their environmental budget. With the purpose of maximizing self-interest, rational construction organization will choose to reduce the effort level for environmental protection. The above discussions can be derived from Option 3.

Option 3. To reinforce the supervision intensity of project owner may not bring the positive effect for project construction, as the effort level cannot be improved only depending upon the project owner's supervision. However, construction organization will raise the effort level for environmental protection since the consideration of cost reduction.

45.4.2 Supervision of Project Owner and Government

According to the Eq. (45.3), following expression can be derived through the derivation of the comprehensive quality of construction party p_a and the punishment set by project owner ϕ_p.

$$\partial p_p{}^*/\partial p_a = -c\left[p_g\gamma(ka_0 + \delta\theta) + p_g\phi_g - \phi_p\right]/(1-p_a)^2 p_g{}^2\gamma^2k^2 \tag{45.10}$$

$$\partial p_p{}^*/\partial \phi_p = c/(1-p_a)p_g{}^2\gamma^2k^2 \tag{45.11}$$

From Eq. (45.11), we can find $\partial p_p{}^*/\partial\phi_p > 0$ which suggests project owner's supervision intensity is proportion to the punishment. The higher punishment is set, the more environmental supervision will be act on construction organization. If $\phi_p > p_g[\gamma(ka_0+\delta\theta)+\phi_g]$ in terms of Eq. (45.10), $\partial p_p{}^*/\partial p_a > 0$, which means the supervision intensity of project owner is proportional to the comprehensive quality of construction organization. This discussion can be demonstrated by Option 2, from which a^* is inverse proportion to p_a. As higher p_a is, construction team will make less effort to meet the required environmental standard of project. However, the probability of loosing effective environmental management will also be increased at the same time. Thus project owner should enhance the intensity of supervision.

Similarly, if $\phi_p < p_g[\gamma(ka_0 + \delta\theta) + \phi_g]$, $\partial p_p{}^*/\partial p_a < 0$ as the supervision intensity p_p is inverse proportion to the comprehensive quality p_a. When lower supervision intensity has been set, the loss of construction team's negative acts on environmental protection is less even if being found by project owner. Moreover, the environmental cost will be raised for the lower-quality construction team, who will take passive attitude on environmental management. Based on the relation between project owner's supervision intensity and construction organization's comprehensive quality, Option 4 is drawn as following.

Option 4. When the intensity of punishment ϕ_p is high enough and at least $\phi_p > p_g(\eta + \phi_g)$ ($\eta = \gamma(ka_0 + \delta\theta)$) has been satisfied, project owner should take higher supervision on high-quality construction team. On the contrary, when $\phi_p < p_g(\eta + \phi_g)$ has been satisfied, project owner should take higher supervision on low-quality construction team.

Following expression can be derived through the derivation of the specific environmental standard a_0 in terms of Eq. (45.3), which means the higher the environmental standard is set by the project construction, the lower intensity of supervision is determined by project owner.

$$\partial p_p{}^*/\partial a_0 = -c/(1 - p_a)p_g\gamma k < 0 \tag{45.12}$$

Moreover, a_0 can be expressed as followed by means of the solution of Eq. (45.9),

$$a_0 = \frac{dM(p_g)}{(1 - p_a)\gamma k d p_g} + \frac{k}{c}\left(p_g\gamma - p_a p_g\gamma - 1\right) - \frac{1}{k}\left(\frac{\phi_g}{\gamma} + \delta\theta\right) \tag{45.13}$$

It is found that $\partial a_0/\partial p_g > 0$ by derivation of governmental administrative intensity p_g, which can be demonstrated by the environmental standard is proportional to governmental administration. From expression (45.12), and (45.13), Option 5 is derived as followed.

Option 5. Once the higher environmental standard has been set up by the project construction, government will enhance administrative intensity accordingly and project owner will gradually reduce surveillance on construction organization. This further confirms that supervision of government and project owner can be interacted each other. Especially for some important engineering project that demands high environmental standard, government should play a leading role in environmental supervision.

45.4.3 Punishment of Project Owner and Government

According to the Eq. (45.4), following expression can be derived by the derivation of governmental administration on environmental protection p_g.

$$\partial\phi_p{}^*/\partial p_g = [\gamma(ka_0 + \delta\theta) + \phi_g] - 2p_g\gamma^2k^2(1 - p_a)/c \tag{45.14}$$

If $\eta+\phi_g > 2p_g\gamma^2k^2(1-p_a)/c$ and $\eta = \gamma(ka_0+\delta\theta)$, we can find $\partial\phi_p{}^*/\partial p_g > 0$, which indicates that once $\eta + \phi_g$ is above a specific value, the project owner's optimum punishment will increase along with governmental administration. The main reason to explain the above situation is that construction organization will be heavily fined by the negative action on environmental protection, under higher intensity of governmental administration. Thus, environmental cost spent by project owner will also be increased.

On the contrary, if $\eta+\phi_g < 2p_g\gamma^2k^2(1-p_a)/c$, $\partial\phi_p{}^*/\partial p_g < 0$, which indicates that once $\eta+\phi_g$ is below a specific value, the project owner's optimum punishment value is in inverse proportion to governmental administration. This situation can be expressed when the reputation loss caused by construction organization's negative action on environmental protection. And if the reputation loss will not constitute a substantial threat to project owner, the importance attached to environment will still be reduced. According to the Eq. (45.11), punishment reduction will give rise to supervision reduction and thus project owner can save the environmental cost. In order to ensure project owner to supervise efficiently on environmental protection during construction, the authors suggest that the factor of reputation loss should be satisfied certain condition as reflected in Option 6.

Option 6. It is found that the higher governmental supervision is, the higher the optimum punishment is set by project owner to guarantee the effective environmental management throughout the project construction as $\eta + \phi_g > 2p_g\gamma^2k^2(1 - p_a)/c$. Thus, $\eta > 2p_g\gamma^2k^2(1 - p_a)/c + \phi_g$ should be satisfied so as to help project owner pay enough attention to environmental protection during project construction.

Through the derivation of construction organization's comprehensive quality p_a, Eq. (45.8) can be transformed into following expression.

$$\partial\phi_g{}^*/\partial p_a = (2p_a - 1)p_g\gamma^2k^2/c(1 - p_a) \tag{45.15}$$

From Eq. (45.15), if $p_a > 0.5$, $\partial\phi_g{}^*/\partial p_a > 0$ can be derived which demonstrates the optimum governmental punishment on negative behavior of environmental protection is proportion to the comprehensive quality of construction organization. Conversely, if $p_a < 0.5$, $\partial \phi_g{}^*/\partial p_a < 0$ determines an inverse proportional relationship between governmental punishment and construction organization's comprehensive quality. Thus, the impact of punishment is varied by the environmental awareness of construction organization, which can be drawn in Option 7.

Option 7. Government should set high standard of economic punishment to promote environmental protection during project construction. For a higher-quality construction team, this high punishment policy will play a role in leading more active behavior to environmental management. On the other hand, the policy will aid a lower-quality construction team enhance self-awareness of environmental protection, thus to implement environmental management actively.

45.5 Conclusion and Further Study

Game theory provides a useful tool to the environmental governance of construction project and help construction parties improve their awareness of environmental protection. However, there are still some limitations in this study. First, the game model still has not been validated at this stage. Secondly, weight allocation for environmental governance, construction quality and period etc., have not been considered in the non-cooperative game model. In particular, the potentially societal benefit that may affect the governmental decision on environmental performances has been omitted. Thus, it is expected that the game theoretical model can be improved by using case studies in the near future.

References

Begum RA, Siwar C, Pereira JJ, Jaafar AH (2007) Factors and values of willingness to pay for improved construction waste management-a perspective of Malaysian contractors. Waste Manag 27(12):1902–1909

Begum RA, Siwar C, Pereira JJ, Jaafar AH (2009) Attitude and behavioral factors in waste management in the construction industry of Malaysia. Res Conserv Recy 53(6):321–328

Chuang SS, Lo CWH (2003) Evaluating sustainability in waste management the case of construction and demolition, chemical and clinical waste in Hong Kong. Res Conserv Recy 37(2):119–145

Kartam N, Al-Mutairi N, Al-Ghusain I, Al-Humoud J (2004) Environmental management of construction and demolition waste in Kuwait. Waste Manag 24(10):1049–1059

Mcdonald B, Smithers M (1998) Implementing a waste management plan during the construction phase of a project: a case study. Constr Manag Econ 16(1):71–78

Ofori G, Gang G, Briffett C (2002) Implementing environmental management systems in construction: lessons from quality systems. Build Environ 37(12):1397–1407

Shen LY, Tam WYV (2002) Implementation of environmental management in the HongKong construction industry. Int J Proj Manag 20(7):535–543

Tam VW, Tam CM (2008) Waste reduction through incentives: a case study. Build Res Inf 36(1):37–43

Teo MMM, Loosemore M (2001) Theory of waste behavior in the construction industry. Constr Manag Econ 19(7):741–751

Ye LP, Li W (2007) Probing into existing problem of construction trash disposal at home and its new model. Constr Mach Technol Manag 20(5):67–68

Chapter 46
Empirical Study on Stock Valuation Model Based on Multiple Linear Regression Analysis

Qing Li and Ai-min Li

Abstract Stock valuation is an important aspect of corporate investment activities. But the current methods used by assets appraisal institutions in our country are not quite reasonable. In this paper, a stock valuation modeling method based on multiple linear regression analysis is proposed. It attempts to reveal the intrinsic relationship between stock price and financial indicators of a listed company. By making an empirical case study of listed mechanical industry companies, this paper validates the feasibility and adaptability of this new modeling method, and tries to set up a reasonable stock valuation model to be taken as a reference for assets evaluation institutions.

Keywords Financial indicators • Multiple linear regression • Principal component analysis • Stock valuation

46.1 Introduction

Stock valuation theory has developed into a set of relatively complete and mature valuation methods and models after a long-term of exploring. It is widely used by European and American market in affairs of company investment, merger and reorganization, etc. Current stock valuation methods available are listed as follows: discounted cash flow method, relative value method, real option method and evaluation method based on assets (Christensen and Feltham 2009). In practical affairs of valuation in a mature market, discounted cash flow method and relative value method are the most popular ones (Zheng Zhong 2001).

Q. Li • A. Li (✉)
Department of Management Engineering, Zhengzhou University, Zhengzhou, China
e-mail: liaimin@zzu.edu.cn

E. Qi et al. (eds.), *The 19th International Conference on Industrial Engineering and Engineering Management*, DOI 10.1007/978-3-642-37270-4_46,

Assets appraisal institutions in China generally prefer evaluation method based on assets instead of relative value method. According to a survey made in 2003, only 3.47 % of assets evaluation institutions in China took relative value method as the main body of valuation method. On the contrary, more than 82.6 % of them used little or even did not use any relative valuation method. Among relative value methods applied in stock market, only P/E method is approachable in most cases (Hui Tian 2004).

There are three reasons that may explain such a situation. Firstly, assets evaluation industry in our country lacks in enough experience on research and application, which leads to a weak theoretical basis and poor professional capability. Secondly, the implementation of stock-market regulations is constrained by the current market environment. Thirdly, in addition to objective reasons of market environment, the poor application of income present value method and relative valuation method in our country is mainly due to the imperfection of existing evaluation methods and models (Lin Li 1999).

Aiming at these problems occurring in stock valuation process in Chinese stock market, this paper attempts to build a multiple linear regression model on theoretical basis of stock intrinsic value theory and capital market efficiency theory (Xiao-xia Xie 2011). At last, data of listed machinery companies from Shanghai Stock Exchange and Shenzhen Stock Exchange are introduced as research sample for empirical research. Through such a case study, feasibility and efficiency of the new modeling method are proved.

46.2 Methodology

46.2.1 Selection of Explanatory Variables

This paper selects 11 financial indicators related to five aspects of a company's comprehensive situation (Ho et al. 2012). Since stock price reflects the overall value of a company, it can be explained by the selected financial indicators, according to the theoretical basis mentioned above.

First of all, in order to reflect the aspect of profitability, a selection of earning per share (EPS), return on equity (ROE) and net profit/total business income (NP/TBI) are decided to represent both common stock revenue, and overall profitability (Shu Gong and Yong-li Chen 2011).

Secondly, concerning the aspect of debt paying ability, we select debt-to-assets ratio (DR) for long-term ability (Asvanunt et al. 2011). At the same time, three indicators for short-term ability are selected as current ratio (CT), quick ratio (QT) and supper quick ratio (SQT) (Brermam and Tamarowski 2000).

Thirdly, this paper chooses earnings per share growth rate (EGR) and operating income growth rate (OGR) as a group of indicators, which is to reflect the company's growth ability (Ruo-wei Ma and Qun-da Chen 2011).

In addition, quantity of dividend distribution makes great contribution to stock price as well (Su-long and Ahmed 2011). So we pick out cash and cash equivalents balance per share (CPS) to reflect the status of this aspect.

Finally, we come to one of the most important factors. That is the capital structure of a company (Ya-qin Gao and Yue Hou 2011). Attention has to be paid to a particular indicator, debt-to-assets ratio, for it is an indicator to mirror capital structure as well as debt paying ability. And current assets/total assets (CA/TA) is also included here (Omran and Pointon 2009).

46.2.2 Modeling Method

According to the theoretical basis introduced above, we take stock price as explained variable, while taking all the financial indicators as explanatory variables. Then we can set up a preliminary stock valuation model based on multiple linear regression theory:

$$\begin{aligned} SP = {} & \alpha + \beta_1 EPS + \beta_2 ROE + \beta_3 CT + \beta_4 QT + \beta_5 SQT \\ & + \beta_6 EGR + \beta_7 OGR + \beta_8 CPS + \beta_9 DR + \beta_{10} \frac{CA}{TA} \\ & + \beta_{11} \frac{NP}{TBI} \end{aligned} \tag{46.1}$$

Here, SP stands for stock price, while other variables have already been explained above.

And then, principle component analysis is necessary so as to summarize the majority of required information in a much more reasonable way (Zi-min Yin and Peng-fei Yu 2011).

At last, the final multiple linear regression model will be built.

46.3 Empirical Study

46.3.1 Description of Research Samples

Mechanical industry companies, listed at Shanghai Stock Exchange and Shenzhen Stock Exchange before April 2012, are taken as research samples. And the classification of listed companies refers to the standard of China International Trust and Investment (CITIC). All the data included in the empirical study are derived from RESSET/DB (http://www.resset.cn).

Removing ST shares and shares which fail to publish their annual financial information without delay; this paper picks out 93 listed companies. Their average

Table 46.1 Principal component analysis – Eigenanalysis of the correlation matrix

	Principle components			
	PC1	PC2	PC3	PC4
Eigenvalue	4.8010	2.8063	1.0931	0.8392
Proportion	0.436	0.255	0.099	0.076
Cumulative	0.436	0.692	0.791	0.867

Table 46.2 Principal component analysis – Coefficients of original variables

Variable	PC1	PC2	PC3	PC4
EPS	−0.218	0.442	−0.062	−0.428
ROE	−0.095	0.515	0.281	−0.297
CT	−0.392	−0.245	0.243	0.028
QT	−0.396	−0.240	0.227	0.028
SQT	−0.396	−0.239	0.215	0.030
EGR	−0.138	0.438	0.158	0.370
OGR	−0.160	0.329	−0.133	0.722
CPS	−0.290	0.012	−0.599	−0.212
DR	0.339	0.176	0.119	−0.076
CA/TA	−0.299	0.060	−0.536	0.022
NP/TBI	−0.378	0.173	0.239	−0.130

closing price in March 2012 is considered to be explained variable. Data of explanatory variables are picked from annual financial information published by the companies by the end of 2011.

46.3.2 Principle Component Analysis

To be prepared for regression, we firstly apply principal component analysis to the data. Now we get several principal components which are complementary related. The new variables reflect the majority of original information. The result is presented in Table 46.1.

The result shows that first four principal components provide a cumulative contribution up to 86.7 %. Therefore, original variables can be replaced by the new ones in the improved model which will be proposed later.

Furthermore, Table 46.2 presents the coefficients of every original variable constituting the main components. Obviously, PC1 mainly reflects debt paying ability, for coefficients of CT, QT, SQT and DR are relatively notable. So as to identify this character, PC1 is defined as debt paying ability factor (*DF*). Its approximate expression is described as Eq. (46.2). Here, η_1 stands for the remaining variables.

$$DF = -0.392CT - 0.396QT - 0.396SQT + 0.339DR + \eta_1 \tag{46.2}$$

Table 46.3 Regression analysis – SP versus DF, PF, CBF, GF

Predictor	Coef	SE Coef	T	P
Constant	15.0955	0.5283	28.58	0.000
DF	−2.5493	0.2424	−10.52	0.000
PF	2.4263	0.3170	7.65	0.000
CDF	−2.3314	0.5080	−4.59	0.000
GF	−2.5548	0.5798	−4.41	0.000

S = 5.09434
R-Sq = 70.4 %
R-Sq(adj) = 69.1 %

Table 46.4 Regression analysis – analysis of variance

Source	DF	SS	MS	F	P
Regression	4	5441.0	1360.3	52.41	0.000
Residual Error	88	2283.8	26.0		
Total	92	7724.8			

In the same way, PC2 mainly reflects profitability. Thus, it is defined as profitability factor (*PF*). Its approximate expression is described as Eq. (46.3). Here, η_2 stands for the remaining variables.

$$PF = 0.442EPS + 0.515ROE + 0.173\frac{NP}{TBI} + \eta_2 \tag{46.3}$$

The constitution of PC3 is a little more complicated than those of other factors. It mainly reflects both capital structure and dividend distribution. Thus, we define it as capital structure and dividend distribution factor (CDF). Its approximate expression is described as Eq. (46.4). Here, η_3 stands for the remaining variables.

$$CDF = -0.599CPS + 0.119DR - 0.536\frac{CA}{CT} + \eta_3 \tag{46.4}$$

PC4 mainly reflects the information of two growth ability indicators. So it is defined as growth capacity factor (GF). Its approximate expression is described as Eq. (46.5). Here, η_4 stands for the remaining variables.

$$GF = 0.370EGR + 0.722OGR + \eta_4 \tag{46.5}$$

From now on, we will use the four main factors instead of original 11 variables in the following study.

46.3.3 Stock Valuation Mode

Now we can ultimately build an improved regression model utilizing the new factors. The result is presented in Tables 46.3 and 46.4.

The final regression equation is described as Eq. (46.6).

$$SP = 15.1 - 2.55DF + 2.43PF - 2.33CDF - 2.55GF \tag{46.6}$$

46.4 Discussion

46.4.1 Significance of Factors

As shown in Table 46.3, all principal components perfectly passes t test, which means these factors have noble effect on stock price.

46.4.2 Adaptability of Regression Model

As shown in Table 46.4, the whole model perfectly passes F-test, which indicates stock price can be well determined by the four factors. We are also informed from the model that the coefficient of determination reaches the level of 70.4 %, and the adjusted coefficient of determination reaches the level of 69.1 %. These indicate the goodness-of-fit of this regression model reaches a desirable level. Such a conclusion proves the selection of basic financial indicators can explain the stock price efficiently.

46.4.3 Debt Paying Ability Factor

According to Eq. (46.2), a larger DF value indicates poorer debt paying ability. And the coefficient of DF is negative. That means poorer debt paying ability results in lower stock price.

46.4.4 Profitability Factor

According to Eq. (46.3), a larger PF value indicates stronger profitability. And the coefficient of PF is positive. That means the stronger profitability a company has, the higher its stock price will be, which corresponds with our general knowledge.

46.4.5 Capital Structure and Dividend Distribution Factor

According to Eq. (46.4), a larger CDF value indicates a less reasonable capital structure and less dividend distribution. And the coefficient of CDF is negative. That means a less reasonable capital structure brings about lower stock price. On the other hand, less dividend distribution leads to lower stock price.

46.4.6 Growth Capacity Factor

According to Eq. (46.5), a larger GF value indicates stronger growth ability. And the coefficient of GF is negative. So the superficial conclusion is that stronger growth ability results in lower stock price. It seems to be ridiculous at first glance. But actually, it is usually difficult for big companies to maintain a relatively high growth rate, due to their large scale and complex system. Thus, companies that keep a high growth rate are generally with a small scale (Li-gen Yao and Wei-shan Zhao 2010). And the stock prices of these small companies are normally low, due to their lack of financial support or comprehensive strength.

To sum up, this trend revealed by the model is acceptable as well as reasonable.

46.5 Conclusion

A common problem which exists in assets appraisal institutions in our country at present is the lack of an efficient and practical stock valuation modeling method. The two stock valuation methods most commonly used in European and American markets are unadapted for Chinese stock market. The discounted cash flow method requires explicit estimation of future dividends, cash flow of equity and residual income. However, dividend distribution mechanism in our country is still far from perfect. So efficient estimations are not available, which easily leads to serious distortion of the result. On the other hand, although relative value method is simple and easily understandable, its drawback of metrological specification insufficiency is obvious and unavoidable. While selecting analogical objects and indicators, subjective assumption will be strong enough to control the result.

This paper, aiming at the problems mentioned above, proposes a multiple linear regression modeling method for stock valuation based on financial indicators. And an empirical case study, on the basis of data of mechanical industry companies listed in Shanghai Stock Exchange and Shenzhen Stock Exchange, is made to prove the effectiveness and feasibility of this new modeling method.

References

Asvanunt A, Broadie M, Sundaresan S (2011) Managing, corporate liquidity: strategies and pricing implications. Int J Theor Appl Financ 14(3):369–406

Brermam MJ, Tamarowski C (2000) Investor relations, liquidity, and stock prices. J Appl Financ 12(4):26–37

Christensen PO, Feltham GA (2009) Equity valuation. Found Trends Account 4(1):1–112

Ho GTS, Ip WH, Wu CH, Tse YK (2012) Using a fuzzy association rule mining approach to identify the financial data association. Expert Syst Appl 39(10):9054–9063

Hui Tian (2004) The application of multiplier valuation method in equity valuation in stock market. Apprais J China 5:19–21 (in Chinese)

Li-gen Yao, Wei-shan Zhao (2010) An empirical study on growth of the small and medium listed enterprise in China. In: Proceedings of the 2nd international conference on information science and engineering, ICISE'2010, Hangzhou, China, 2010, pp 4.
Lin Li (1999) Research on enterprise valuation theory and application. PhD dissertation, Peking University (in Chinese)
Omran MM, Pointon J (2009) Capital structure and firm characteristics: an empirical analysis from Egypt. Rev Account Financ 8(4):454–474
Ruo-wei Ma, Qun-da Chen (2011) Research on factors that influence the growth of enterprises listed in GEM board. In: Proceedings of the 2011 international conference on business management and electronic information, BMEI'2011, Guangzhou, pp 633–636
Shu Gong, Yong-li Chen (2011) Value creation capacity, enterprise performance and corporate governance: evidence from Chinese listed companies on GEM, In: Proceedings of the fourth international conference on business intelligence and financial engineering, BIFE'2011, Wuhan, China, pp 381–385
Su-long Z, Ahmed PK (2011) Ownership structure, board governance, dividends and firm value: an empirical examination of Malaysian listed firms. Int J Busi Gov Ethic 6(2):135–161
Xiao-xia Xie (2011) Research on the explanation power of accounting information of annual report of listed companies of stock price. In: Proceedings of the 2011 international conference on business management and electronic information, BMEI'2011, Guangzhou, China, pp 588–592
Ya-qin Gao, Yue Hou (2011) The empirical research of relationship between capital structure and corporation value. In: Proceedings of the 2nd international conference on artificial intelligence, management science and electronic commerce, AIMSEC'2011, Dengleng, pp 3452–3456
Zheng Zhong (2001) Empirical research on company valuation theory and fundamental analysis. PhD dissertation, Xiamen University (in Chinese)
Zi-min Yin, Peng-fei Yu (2011) Factor analysis and evaluation on a foundation of comprehensive performance of listed companies. In: Proceedings of the international conference on computer and management, CAMAN'2011, Wuhan, China, pp 4

Chapter 47
The Research and Application of a Dynamic Dispatching Rule Selection Approach Based on BPSO-SVM for Semiconductor Production Line

Kuo Tian, Yu-min Ma, and Fei Qiao

Abstract Reasonable choice of scheduling strategies to optimize the process of production scheduling is an effective way to improve the economic benefit and market competitiveness of manufacturing enterprises. This paper proposes a BPSO-SVM-based dynamic scheduling rule selection approach for semiconductor production line. This approach combines with feature selection algorithm based on semiconductor production attributes and dispatching rule classification algorithm. It finds appropriate feature subsets and SVM parameters by feature selection algorithm and finds real-time optimal scheduling rules effectively under one better performance according to the status of the production line in a SVM classification model by classification algorithm. Finally, the approach is verified on Mini-fab, a typical model of semiconductor production line.

Keywords BPSO • Dynamic scheduling • Feature selection • Parameters optimization • SVM

47.1 Introduction

The essence of the scheduling optimization problem of semiconductor production line is to optimize scheduling strategies under related constraint conditions of production process and resources, so that it can achieve a better system performance

K. Tian (✉) • Y. Ma • F. Qiao
The School of Electronics & Information Engineering, Tongji University, Shanghai, China
e-mail: tiankuo5689@126.com

E. Qi et al. (eds.), *The 19th International Conference on Industrial Engineering and Engineering Management*, DOI 10.1007/978-3-642-37270-4_47,

index (Min Liu 2009).We can make full use of production data related with scheduling that is online or not in order to choose a better scheduling strategy quickly under a given condition of production. At the same time, we should remove some redundant data, which is a process of feature selection of production attributes in order to improve the scheduling efficiency.

This paper proposes a scheduling strategy selection approach based on semiconductor production attributes. This approach is based on historical data and uses support vector machine (SVM) (Jiawei Han and Micheline Kamber 2006; Muller et al. 2001; Cristianini and Shawe-Taylor 2000; Nello and Jhon 2004; Vapnik 1999) as a data mining tool. A process of finding optimal production attributes (feature subsets) is done by binary particle swarm optimization algorithm (BPSO) (Kennedy and Eberhart 1995, 1997; Xiaofeng Xie and Wenjun Zhang 2003; Liu and Qin 2004).We can get a better scheduling strategy under any given production status through a SVM classifier after optimization.

47.2 Description of Scheduling Strategy Selection Problem on Semiconductor Production Line

The production status of semiconductor production line is varying continuously, a real-time scheduling according to the change of production system status can get better system performance. The set of scheduling problem can be described as: {P,F,D} (Park et al. 1997), where P is the set of the performance indexes of semiconductor production line, F is the set of complete production attributes of semiconductor production line, D is the set of scheduling strategies which is meeting the optimal P under any given production status F.

When we take a dispatching decision-making, we should firstly select a proper subset of production attributes SF from the set of complete production attributes F through feature selection (FS) in order to improve the scheduling efficiency, then realize a fast mapping from SF to D applying one data mining algorithm, thus we can find an scheduling strategy to meet a better system performance index under any given subset of production attributes. The above process is shown as formula (47.1).

$$\mathrm{F} \underset{①}{\xrightarrow{\mathrm{FS}}} \mathrm{SF} \underset{②}{\xrightarrow{\mathrm{opt(P)}}} \mathrm{D} \tag{47.1}$$

Where, $\mathrm{SF} \subseteq \mathrm{F}$, 1 is the process of attribute feature selection, this process improve the scheduling efficiency; 2 implies the data mining process from the subset of production attributes, we can get a classification model to meet the requirement of real-time scheduling.

47.3 The Framework of Dispatching Rule Selection for Semiconductor Production Line

This paper is based on the historical data of semiconductor production line and applies simulation platform of semiconductor production scheduling, making a scheduling strategy selection through the principle of BPSO-SVM-based feature selection and classification, the process is as follows:

1. Establish training sample set: Get scheduling strategies to meet the best system performance indexes under different historical data status;
2. Get production attribute subset and SVM training parameters through the principle of BPSO-SVM-based feature selection and classification, form dynamic scheduling rule classifier based on SVM;
3. Set up testing sample set, get production scheduling strategies applying SVM classification algorithm and make an evaluation. If the prediction accuracy of optimal scheduling strategy based on feature subset is better than complete set, choose scheduling strategy dynamically applying SVM classification algorithm with above feature subset as selective element; otherwise, adjust historical production data and improve the process of feature selection and training parameters to obtain a better result.

The framework of dispatching rule selection for semiconductor production line based on BPSO-SVM is shown as Fig. 47.1.

47.4 Feature Selection and Classification Algorithm of Semiconductor Production Line

47.4.1 Design of SVM Algorithm

Suppose that the input mode set is described as: $D = \{x_i, y_i\}_{i=1}^{N}$, where x_i is the i-th sample of D and can be seen as the corresponding production attribute subset after feature selection, y_i is class label and can be seen as the dispatching strategy according to x_i, the plane equation for classification is:

$$\mathrm{W} \bullet \mathrm{X} + \mathrm{b} = 0 \tag{47.2}$$

Where W is the normal vector of hyperplane, b is the constant term of hyperplane. In production history database, each sample is corresponding to the best scheduling strategy under one production status. Firstly, we should establish a classification

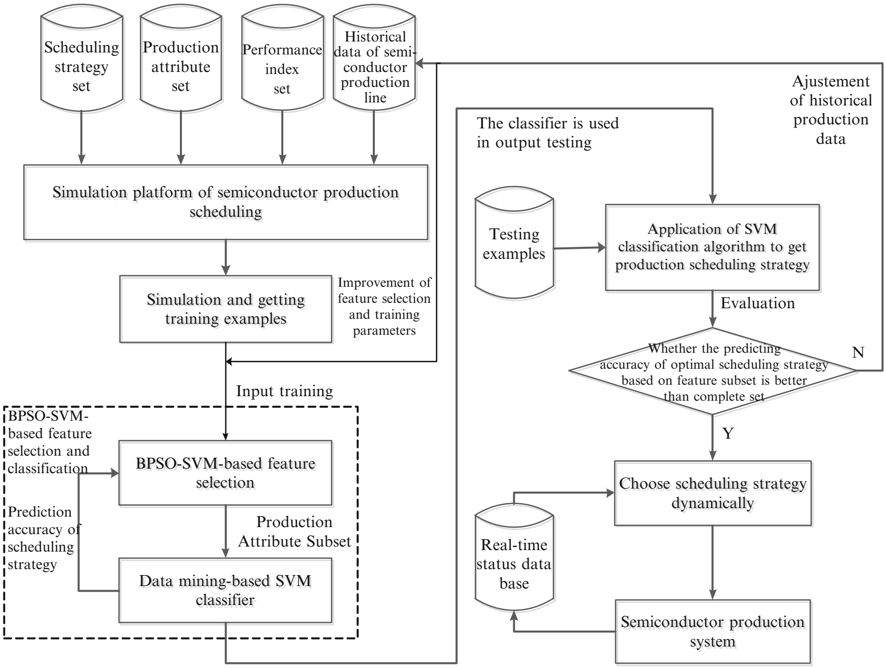


Fig. 47.1 Process of dispatching rule selection for semiconductor production line

hyperplane through D, and then we can get an optimal classification hyperplane f(x) by transforming a quadratic programming problem into its dual problem using Lagrange multiplier method. The above f(x) is shown in formula (47.3).

$$f(x) = \mathrm{Sgn}\left(\sum\nolimits_{i=1}^{N} \alpha_i y_i K(x_i, x) + b\right) \tag{47.3}$$

Where $K(x_i,x)$ is kernel function; α_i is the best Lagrange multiplier which is determined by quadratic programming; $\{x_i,y_i\}$ is training data, b is the threshold according to training samples. We can get the optimal scheduling strategy by plugging any x_i into f(x).

47.4.2 Design of BPSO Algorithm

1. Design of particle code based on production attribute and SVM parameters

A particle is represented in 0–1 coding mode: each bit represents an production attribute, “1” represents that the corresponding production attribute is selected for training and “0” is opposite, as shown in formula (47.4).

$$\{p_1, \ldots, p_{n_f}\} \tag{47.4}$$

Where p_i is a binary bit of particle, n_f is the characteristic number of feature selection. Considering to optimize training parameters simultaneously and the classification algorithm uses kernel function RBF, there are two training parameter segment kernel parameter γ and penalty factor C needed to be insert into formula (47.4), as is shown in formula (47.5).

$$\{p_1, \ldots, p_{n_f} | p_{n_f+1}, \ldots, p_{n_f+n_C} | p_{n_f+n_C+1}, \ldots, p_{n_f+n_C+n_\gamma}\} \tag{47.5}$$

Where n_c and n_γ are the numbers of binary digit of parameter C and γ respectively. The formula that transforms C or γ into their decimal form is shown in formula (47.6).

$$C\ (\text{or}\ \gamma) = \min + \frac{\max - \min}{2^l - 1} * d \tag{47.6}$$

Where max and min are their corresponding maximum and minimum decimal value respectively; l is the length of parameter bit string; d is the corresponding decimal value calculated from the binary value of the parameter.

2. Fitness function

The fitness function defined in this paper is shown in formula (47.7).

$$\text{Fitness} = \text{SVM}_{\text{Accuracy}} = \frac{N_C}{N_U} \tag{47.7}$$

Where N_C is the number of samples classified correctly in the test set, N_U is the number of samples in the test set, $\text{SVM}_{\text{Accuracy}}$ is the classification accuracy of SVM for each particle after training. It is got through K-fold cross validation (Anguita et al. 2009).

3. Conditions of algorithm termination

When the algorithm meets one condition below, the iteration stops:

1. The algorithm reaches the maximum iteration times.
2. After several consecutive iterations, the global best fitness does not change.

47.5 The Experiments Based on Mini-Fab

47.5.1 The Object of Experimental Research and the Selection of Production Attribute Subset

Mini-fab is a simplified semiconductor production line, which consists of three equipment groups including five equipments and six steps (Qidi Wu et al. 2006), as shown in Fig. 47.2.

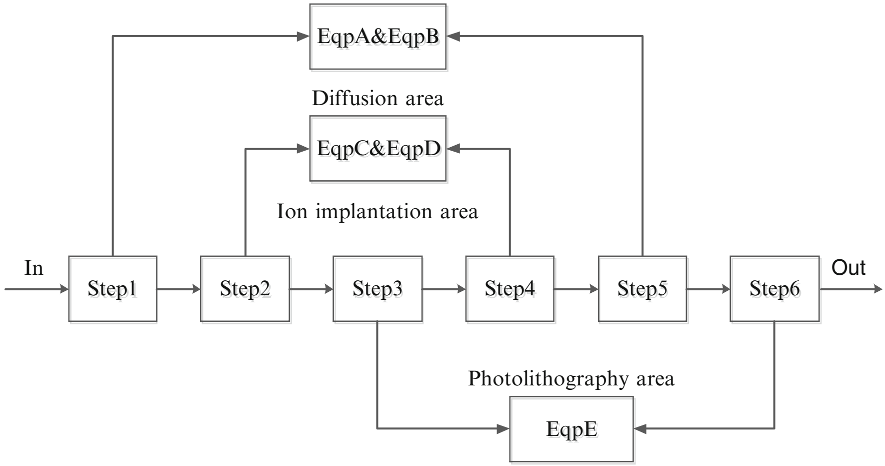


Fig. 47.2 Mini-fab model

The data sets of the experiment consist of production attribute set, scheduling strategy set and performance index set. They are expounded as follows:

1. Production Attribute set

This paper presents 26 production attributes (No. 1–26). They are summarized as:

{{Product Releasing attribute}, {Equipment failure and maintenance attribute}, {WIP attribute}, {Processing zone attribute}, {Production line attribute}}

2. Scheduling strategy set

There are five heuristic rules in the scheduling strategy set in this experiment: First Input First Output (FIFO), Earliest Due Date (EDD), Smallest Remaining Processing Time Policy (SRPT), Critical Ratio (CR), Fluctuation Smoothing Policy for Variance of Cycle Time (FSVCT).

3. Performance index set

There are six performance indexes included in this experiment: Mean Processing Cycle Time (MCT), Mean Day Productivity (MDayProd), Mean Day Move (MDay-Mov), Mean Day WIP (MDayWIP), On-time Delivery Rate (ODR), Utilization of Me (bottleneck equipment) (Util_Me).

This experiment is realized by SIEMENS Plant Simulation 9.0, randomly generates 300 sets of data related to the status of production line and evaluates the simulation results from six optimal performance indexes. The hardware and software environment of experiment are: Intel Core2 CPU with memory of 3.25G, C# programming in the platform of VS, LIBSVM (Hsu et al. 2003; Chang and Lin 2011).

Table 47.1 Prediction accuracy of optimal scheduling rule

Performance index	Attribute and algorithm category	Mean prediction accuracy	Comparison of mean prediction accuracy
MCT	TS_SVM	60 %	+17.22 %
	SS_BPSO-SVM	70.33 %	
MDayMov	TS_SVM	60.33 %	+14.92 %
	SS_BPSO-SVM	69.33 %	
MDayProd	TS_SVM	81.67 %	+3.26 %
	SS_BPSO-SVM	84.33 %	
MDayWIP	TS_SVM	81 %	+3.30 %
	SS_BPSO-SVM	83.67 %	
ODR	TS_SVM	60.33 %	+11.06 %
	SS_BPSO-SVM	67 %	
Util_Me	TS_SVM	34.67 %	+27.86 %
	SS_BPSO-SVM	44.33 %	

47.5.2 Experimental Parameter Settings

The experimental parameters are set as follows: Number of cross validation $k = 10$; Size of particle swarm $PN = 100$; Maximum iteration $N = 200$; Range of particle's velocity initialized $[-10,10]$; Maximum and minimum of SVM training parameters $C_{max} = 500$, $C_{min} = 0$, $\gamma_{max} = 1$, $\gamma_{min} = 0$; $n_C = 10$, $n_\gamma = 10$, $n_f = 26$; Inertia weight $w = 1$.

47.5.3 Experimental Results and Analysis

Table 47.1 shows the comparison of the test results between the complete production attribute set directly applying SVM algorithm and the production attribute subset after feature selection applying BPSO-SVM algorithm from six performance indexes.

As is shown in Table 47.1, compared with a complete production attribute set applying SVM algorithm, the method of feature selection and parameter optimization applying BPSO-SVM algorithm can get a good result of mean prediction accuracy from above six performance indexes. Especially for utilization of Me, the prediction accuracy increases by 27.86 %. Table 47.2 shows the production attribute subset after feature selection and its relevant number.

As is shown in Table 47.2, it removes a large amount of redundant attributes after feature selection, especially for mean processing cycle time and utilization of Me, both size of their production attribute subset decrease from 26 to 10. The mean size of production attribute subset is 11.5; the results show that BPSO-SVM algorithm has a good effect of feature selection. Table 47.3 shows the mean value of performance index of 30 samples under seven scheduling strategies.

Table 47.2 Production attribute subset after feature selection

Performance index	Production attribute subset after FS (No.)	Relevant number
MCT	2, 5, 6, 8, 11, 12, 20, 22, 23, 24	10
MDayMov	2, 4, 5, 6, 11, 13, 17, 22, 23, 24, 26	11
MDayProd	1, 4, 7, 11, 14, 15, 16, 19, 23, 25, 26	11
MDayWIP	1, 2, 4, 5, 6, 8, 9, 10, 11, 15, 17, 18, 19, 23, 26	15
ODR	4, 5, 6, 10, 12, 15, 16, 17, 21, 22, 23, 26	12
Util_Me	2, 6, 9, 10, 11, 12, 16, 21, 22, 24	10

Table 47.3 Comparison of performance indexes under different scheduling strategies

Scheduling Strategy	MCT (h)	MDayMOV (Step)	MDayProd (lot)	MDayWIP (lot)	ODR (%)	Util_Me (%)
SVM	114.91	103.14	17.48	11.29	91.88	93.97
BPSO-SVM	108.01	103.14	17.48	11.22	91.94	94.12
FIFO	275.19	101.45	13.31	11.29	89.97	93.59
EDD	120.55	102.03	15.60	9.50	89.19	93.99
SRPT	112.06	103.13	17.48	7.15	91.39	93.23
CR	137.75	102.68	17.19	7.44	90.12	93.27
FSVCT	197.13	102.01	16.99	7.59	91.09	93.83
Optimal Value	95.82	103.34	17.49	11.33	92.35	95.02
Prediction time (ms)	110	96	119	134	88	145

As is shown in Table 47.3, both SVM and BPSO-SVM dynamic scheduling strategies can get a better result from six performance indexes compared with other scheduling strategies. The BPSO-SVM dynamic scheduling strategy can get optimal results from the performance indexes of MCT, MDayMOV, ODR and Util_Me. Although the results of the performance indexes of MDayProd and MDayWIP are not the best compared with SVM strategy, we increase the prediction accuracy based on removing large amounts of redundant attributes. In addition, the mean prediction time under BPSO scheduling strategy is 115.33 ms; it shows a good requirement of real-time scheduling.

47.6 Conclusion

This paper puts forward a BPSO-SVM dynamic scheduling strategy selection algorithm based on production attributes to solve the scheduling strategy selection problem of semiconductor production line and the approach is validated through Mini-fab model. The experiment results show that the approach meets the requirement of real-time scheduling and improves the scheduling efficiency. Above all, the process of feature selection can not only reduce learning and training time of the classifier, but also improve the prediction accuracy of classification. Finally, we can find appropriate feature subset and SVM parameters and find optimal scheduling strategy effectively under a better performance index according to any given status of production line.

Acknowledgment This research was supported by Chinese National Natural Science Foundation (61034004), Science and Technology Commission of Shanghai (10DZ1120100, 11ZR1440400), Program for New Century Excellent Talents in University (NCET-07-0622) and Shanghai Leading Academic Discipline Project (B004).

References

Anguita D, Ghio A, Ridella S (2009) K–Fold cross validation for error rate estimate in support vector machines. In: Proceedings of the international conference on data mining (DMIN09), Miami, FL

Chang CC, Lin CJ (2011) LIBSVM: a library for support vector machines, ACM Trans Intell Syst Technol 2(3), Article 27

Cristianini N, Shawe-Taylor J (2000) An introduction to support vector machines. Cambridge University Press, Cambridge

Hsu CW, Chang CC, Lin CJ (2003) A practical guide to support vector classification. Available from: http://www.csie.ntu.edu.tw/~cjlin/papers/guide/guide.pdf

Jiawei Han, Micheline Kamber (2006) Data mining concepts and techniques. China Machine Press, Beijing

Kennedy J, Eberhart RC (1995) Particle swarm optimization. In: Proceedings of the IEEE international conference on neural networks, IEEE Service Center, Piscataway, 1995, 1942–1948

Kennedy J, Eberhart RC (1997) A discrete binary version of the particle swarm algorithm, In: Proceedings of the world multiconference on systemic, cybernetics and informatics 1997, IEEE Service Center, Piscataway, 1997, pp 4104–4109

Liu Y, Qin Z (2004) Feature selection with particle swarms, computational and information science. In: 1st international symposium, CIS 2004, Shanghai, China, December proceedings, LNCS3314, 2004, pp 425–430

Min Liu (2009) A survey of data-based production scheduling methods. Acta Automatica Sinica 35(6):785–801 (in Chinese)

Muller KR, Mika S, Ratsch G, Tsuda K, Scholkopf B (2001) An introduction to kernel-based learning algorithms. IEEE Trans Neural Netw 12(2):181–201

Nello C, Jhon ST (2004) An introduction to support vector machine. Publishing House of Electronics Industry, Beijing

Park SC, Raman N, Shaw MJ (1997) Adaptive scheduling in dynamic flexible manufacturing systems: a dynamic rule selection approach. IEEE Trans Robotic Autom 13(4):486–502

Qidi Wu, Qiao Fei, Li Li (2006) Semiconductor manufacturing system scheduling. Electronic Industry Press, Beijing

Vapnik VN (1999) An overview of statistical learning theory. IEEE Trans Neural Netw 10(5): 988–999

Xiaofeng Xie, Wenjun Zhang (2003) Overview of particle swarm optimization. Control Decis 18(2):129–133 (in Chinese)

Chapter 48
Reallocation of Resources to Preserve Relative Efficiencies: Inverse CCR Model

Saowanee Lertworasirikul and S.C. Fang

Abstract This paper is concerned with the inverse Data Envelopment Analysis (inverse DEA) when the production function exhibits constant returns to scale (inverse CCR). The inverse CCR problem is the reallocation of input resources to a particular decision-making unit (DMU) given that output values of the DMU are changed and relative efficiencies of all DMUs should remain the same. We focus on the problem where increases of some outputs and decreases of other outputs of the considered DMU can be considered at the same time. The proposed inverse CCR model is a multi-objective nonlinear programming problem (MONLP). We show that a Pareto solution to the MONLP can be obtained from solving a proposed linear programming model. We also investigate the relationship between the changes in input and output values of the particular DMU.

Keywords Efficiency analysis • Data envelopment analysis • Invert optimization

48.1 Introduction

Data envelopment analysis (DEA), as proposed by Charnes et al. (1978), is an effective approach for measuring relative efficiency levels of a set of comparable decision making units that have multiple inputs and multiple outputs. In DEA,

S. Lertworasirikul (✉)
Department of Product Development, Kasetsart University, Bangkok, Thailand
e-mail: saowanee.l@ku.ac.th

S.C. Fang
Department of Industrial & Systems Engineering, Faculty of Engineering,
North Carolina State University, Raleigh, NC, USA

E. Qi et al. (eds.), *The 19th International Conference on Industrial Engineering and Engineering Management*, DOI 10.1007/978-3-642-37270-4_48,

the so-called 'efficient frontier' or 'production frontier' is built as the envelope of all DMUs. The set of feasible activities or DMUs is called a production possibility set. An interesting problem based on DEA is how to allocate input resources to a particular DMU when the outputs of the DMU are increased with an assumption that the DMU maintains its current efficiency level with respect to other DMUs (Wei et al. 2000; Yan et al. 2002; Beasley 2003; Jahanshahloo et al. 2004; Hadi-Vencheh and Foroughi 2006; Li and Cui 2008; Alinezhad et al. 2007).

Reference (Wei et al. 2000) mentioned that the parameters in a DEA optimization model are actual quantities of inputs and outputs given by each concerned DMU; therefore, its inverse optimization problem can be considered for input or output analysis under the DEA. They called the problem based on DEA for the first time as "inverse DEA". In the model by Wei et al. (2000), they assumed that the changes in input and output values were nonnegative values. However, reference Hadi-Vencheh and Foroughi (2006) showed that the solution proposed by Wei et al. (2000) might fail in a special case, whereas their model could overcome this flaw.

The objective of this paper is to extend the inverse DEA by Wei et al. (2000) to the case where increases of some outputs and decreases of the other outputs of a particular DMU can be taken into account at the same time. The relationship between the changes in input and output values of a particular DMU in the inverse CCR problem is investigated. Then an approach to solve the inverse DEA model is developed to preserve efficiency levels of all DMUs at the same levels by reallocating inputs of the particular DMU.

The rest of this paper is organized as follows. In Sect. 48.2, DEA and inverse DEA models are shown. In Sect. 48.3, the relationship between input and output changes of a particular DMU in the inverse CCR is investigated. Section 48.4 presents a solution approach to the inverse CCR model with a numerical example. Finally, conclusions are provided in Sect. 48.5.

48.2 DEA and inverse DEA Model

48.2.1 Data Envelopment Analysis Model

DEA is a well-known technique for performance analysis of comparable DMUs. It has rapidly applied in several areas in both public and private sectors (Arnade 1994; Barros and Peypoch 2009; Chandra et al. 1998; Chilingerian 1994; Chen 1997).

The CCR model is the most frequently used (Charnes et al. 1978) and other models are extensions of the CCR model (Banker et al. 1984; Cooper et al. 2007). The envelopment form of the CCR model is in the following form.

$$
\begin{aligned}
&(\mathrm{DCCR}_0) \quad \text{minimize} \quad && \theta_0 \\
&\text{s.t.} \quad \sum_{i=1}^{n} \lambda_i x_{ji} \le \theta_0 x_{j0} && \mathit{for} \quad j = 1, \ldots, m \\
&\sum_{i=1}^{n} \lambda_i y_{ki} \ge y_{k0} && \mathit{for} \quad k = 1, \ldots, r \\
&\lambda_i \ge 0, \quad i = 1, 2, \ldots, n
\end{aligned}
$$

where $i = 1, \ldots, n$, $j = 1, \ldots, m$, $k = 1, \ldots, r$, x_{j0} is the input j of the considered DMU (DMU_0), x_{ji} is the input j of DMU_i, y_{k0} is the output k of DMU_0, y_{ki} is the output k of DMU_i, u_k is the weight of output k, λ_i is the coefficient of DMU_i, θ_0 is the relative efficiency level of DMU_0. The notation will be used throughout this paper.

48.2.2 Inverse CCR Model

Denote the particular DMU with current input and output values by DMU_0 and the DMU with the changes in input and output values by $\mathrm{DMU}_{0'}$ the inverse CCR problem is stated as follows:

For a group of DMUs with their relative efficiency levels of $\theta_1^*, \theta_2^*, \ldots, \theta_n^*$, which are obtained from solving their CCR models, if the values of certain outputs of a particular DMU (DMU_0) are changed from $\mathbf{y}_0$ to $\mathbf{y}_0 + \Delta\mathbf{y}_0 \ge \mathbf{0}, \Delta\mathbf{y}_0 \ne \mathbf{0}$, we want to find the minimum $\mathbf{x_0} + \Delta\mathbf{x_0}$, such that $\mathrm{DMU}_{0'}$ with new input and output values $(\mathbf{x_0} + \Delta\mathbf{x_0}, \mathbf{y_0} + \Delta\mathbf{y_0})$ and all other DMUs still have the same relative efficiency values.

From the above statement, the inverse CCR can be formulated as follows.

$$
\begin{aligned}
&(\mathrm{DRA}_0) \quad \text{minimize} \quad && \Delta\mathbf{x}_0 \\
&\text{s.t.} \quad \sum_{i=1}^{n} \lambda_i x_{ji} + \lambda_{0'}(x_{j0} + \Delta x_{j0}) \le \theta_0^*(x_{j0} + \Delta x_{j0}) && \mathit{for} \quad j = 1, \ldots, m \\
&\sum_{i=1}^{n} \lambda_i y_{ki} + \lambda_{0'}(y_{k0} + \Delta y_{k0}) \ge y_{k0} + \Delta y_{k0} && \mathit{for} \quad k = 1, \ldots, r \\
&x_{j0} + \Delta x_{j0} \ge 0 && \mathit{for} \quad j = 1, \ldots, m \\
&\lambda_{0'},\ \lambda_i \ge 0, \quad i = 1, 2, \ldots, n.
\end{aligned}
$$

Using $\Delta\mathbf{x}_0 = (\Delta x_{10}, \Delta x_{20}, \ldots, \Delta x_{m0})$ from the DRA_0 model, the relative efficiency level of DMU_l for $l = 1, 2, \ldots, n$ from solving the DRA_l model must be equal to θ_l^*.

$$
\begin{aligned}
(\text{DRA}_l) \quad \text{minimize} \quad & \theta_l \\
\text{s.t.} \quad & \sum_{i=1}^{n} \lambda_i x_{ji} + \lambda_{0'}(x_{j0} + \Delta x_{j0}) \le \theta_l x_{jl} \quad \textit{for} \quad j = 1, \ldots, m \\
& \sum_{i=1}^{n} \lambda_i y_{ki} + \lambda_{0'}(y_{k0} + \Delta y_{k0}) \ge y_{kl} \quad \textit{for} \quad k = 1, \ldots, r \\
& \lambda_{0'}, \lambda_i \ge 0, \quad i = 1, 2, \ldots, n.
\end{aligned}
$$

48.3 Relationship of Inputs and Outputs in the Inverse CCR Model

In this section, the relationship between the changes in input and output values of a particular DMU in the inverse CCR model is investigated such that the efficiency level of the DMU relative to other DMUs remains the same. We show that when some output values of DMU_0, $\Delta \mathbf{y}_0 \ge_p \mathbf{0}$, are increased, some components of $\Delta \mathbf{x}_0$ from the inverse CCR model can be negative, or equivalently, some components of $\mathbf{x}_0 + \Delta \mathbf{x}_0$, can be decreased, while other components of $\Delta \mathbf{x}_0$ can be positive or zero, or equivalently, other components of $\mathbf{x}_0 + \Delta \mathbf{x}_0$, can be increased or remain the same.

Note that $\Delta \mathbf{y}_0 \ge_p \mathbf{0}$, represents a set of inequalities $y_i \ge 0, i = 1, 2, \ldots, n$ with at least one strict inequality, $y_i > 0$.

Theorem 48.1 *Assume that the current efficiency level of* DMU_0 *with respect to other DMUs is* θ_0^*. *If the output values of* DMU_0 *are changed from* $\mathbf{y}_0$ *to* $\mathbf{y}_0 + \Delta \mathbf{y}_0 \ge \mathbf{0}$, $\Delta \mathbf{y}_0 \ne \mathbf{0}$, *to maintain the relative efficiency level of* DMU_0, *the input values of* DMU_0 *must be changed from* $\mathbf{x}_0$ to $\mathbf{x}_0 + \Delta \mathbf{x}_0$, *where* $\mathbf{x}_0 + \Delta \mathbf{x}_0$, *is obtained by solving the* DRA_0 *model. The values of* Δx_{j0} *for the input* j, $j = 1, \ldots m$ *depends on the value of* $\sum_{i=1}^{n} \lambda_i x_{ji} + \lambda_{0'} x_{j0}$ *and* $\theta_0^* x_{j0}$ *in the* DRA_0 *model as follows:*

1. *If* $\sum_{i=1}^{n} \lambda_i x_{ji} + \lambda_{0'} x_{j0} \le \theta_0^* x_{j0}$ for $j = 1, \ldots, m$, *then* $\Delta x_{j0} \le 0$,
2. *If* $\sum_{i=1}^{n} \lambda_i x_{ji} + \lambda_{0'} x_{j0} \ge \theta_0^* x_{j0}$ for $j = 1, \ldots, m$, *then* $\Delta x_{j0} \ge 0$.

Proof Assume that the optimal solution of DMU_0 from solving the DCCR_0 model is $(\theta_0^*, \lambda_i^*, i = 1, 2, \ldots, n)$ and assume that the output values of DMU_0 are changed from $\mathbf{y}_0$ to $\mathbf{y}_0 + \Delta \mathbf{y}_0$, $\Delta \mathbf{y}_0 \ge_p \mathbf{0}$. Let consider the DRA_0 model:

$$
\begin{aligned}
(\mathrm{DRA}_0)\quad & \text{minimize} \quad \Delta \mathbf{x}_0 \\
\text{s.t.}\quad & \sum_{i=1}^{n} \lambda_i x_{ji} + \lambda_{0'}(x_{j0} + \Delta x_{j0}) \le \theta_0^*(x_{j0} + \Delta x_{j0}) \quad \textit{for} \quad j = 1, \ldots, m \\
& \sum_{i=1}^{n} \lambda_i y_{ki} + \lambda_{0'}(y_{k0} + \Delta y_{k0}) \ge y_{k0} + \Delta y_{k0} \quad \textit{for} \quad k = 1, \ldots, r \\
& x_{j0} + \Delta x_{j0} \ge 0 \quad \textit{for} \quad j = 1, \ldots, m \\
& \lambda_{0'},\ \lambda_i \ge 0, \quad i = 1, 2, \ldots, n.
\end{aligned}
$$

By rearranging the first set of constraints in the DRA_0 model, the model is in the following form.

$$
\begin{aligned}
& \text{minimize} \quad \Delta \mathbf{x}_0 \\
\text{s.t.}\quad & \sum_{i=1}^{n} \lambda_i x_{ji} - (\theta_0^* - \lambda_{0'})\Delta x_{j0} \le (\theta_0^* - \lambda_{0'}) x_{j0} \quad \textit{for} \quad j = 1, \ldots, m \\
& \sum_{i=1}^{n} \lambda_i y_{ki} - (1 - \lambda_{0'})\Delta y_{k0} \ge (1 - \lambda_{0'}) y_{k0} \quad \textit{for} \quad k = 1, \ldots, r \\
& x_{j0} + \Delta x_{j0} \ge 0 \quad \textit{for} \quad j = 1, \ldots, m \\
& \lambda_{0'},\ \lambda_i \ge 0, \quad i = 1, \ldots, n.
\end{aligned}
$$

From the first set of constraints relating to inputs in the above model, if at the optimal solution $\sum_{i=1}^{n} \lambda_i x_{ji} \le (\theta_0^* - \lambda_{0'}) x_{j0}$ for $j = 1, \ldots, m$, then $(\theta_0^* - \lambda_{0'})\, \Delta x_{j0} \le 0$ and $\Delta x_{j0} \le 0$. But if $\sum_{i=1}^{n} \lambda_i x_{ji} \ge (\theta_0^* - \lambda_{0'}) x_{j0}$ for $j = 1, \ldots, m$, then $(\theta_0^* - \lambda_{0'})\, \Delta x_{j0} \ge 0$ and $\Delta x_{j0} \ge 0$.

Therefore, we can conclude that if $\sum_{i=1}^{n} \lambda_i x_{ji} + \lambda_{0'} x_{j0} \le \theta_0^* x_{j0}$ for $j = 1, \ldots, m$, then $\Delta x_{j0} \le 0$, while if $\sum_{i=1}^{n} \lambda_i x_{ji} + \lambda_{0'} x_{j0} \ge \theta_0^* x_{j0}$ for $j = 1, \ldots, m$, then $\Delta x_{j0} \ge 0$. □

From Theorem 48.1, it is different from what Wei et al. (2000) proposed. In the model by Wei et al. (2000), given an increase in the output values of DMU_0, $\Delta \mathbf{y}_0 \geq_p \mathbf{0}$, they assumed that all components of $\Delta \mathbf{x}_0$ from their inverse CCR model are equal or greater than 0. This is the reason why the solution by Wei et al. (2000) did not guarantee the efficiency result for input estimating.

In this paper we study the inverse CCR model where some outputs can be increased and other outputs can be decreased at the same time.

48.4 Solving the Inverse CCR Model

48.4.1 Solution Approach to Solve the Inverse CCR Model

To solve the inverse CCR model, we want to find the value of $\Delta\mathbf{x}_0 = (\Delta x_{10}, \Delta x_{20}, \ldots, \Delta x_{m0})$, which keeps the relative efficiency levels of all DMUs unchanged. However, solving the DRA$_0$ and DRA$_l$ models directly are not easy since they are in the form of MONLP. The MLDRA$_0$ model is proposed in this paper for the solution of the inverse CCR model. The MLDRA$_0$ model is in the form of a multi-objective linear programming model.

Theorem 48.2 *Assume that the current relative efficiency level of DMU$_0$ is θ_0^*. Given changes in the output values of DMU$_0$, $\Delta\mathbf{y}_0 \neq \mathbf{0}$, the relative efficiency levels of all DMUs ($l = 1, \ldots, n, 0'$) with respect to the set of DMU$_l$ ($l = 1, \ldots, n, 0'$) will remain unchanged if input values of DMU$_0$ are changed from* $\mathbf{x}_0$ to $\mathbf{x}_0 + \Delta\mathbf{x}_0$, *where* $\mathbf{x}_0 + \Delta\mathbf{x}_0$ *is obtained by solving the MLDRA$_0$ model.*

$$
\begin{aligned}
&(\mathrm{MLDRA}_0) \quad \text{minimize} \quad \Delta\mathbf{x}_0 \\
&\text{s.t.} \quad \sum_{i=1}^{n} \lambda_i x_{ji} \le \theta_0^*(x_{j0} + \Delta x_{j0}) \qquad \textit{for} \quad j = 1, \ldots, m \\
&\qquad \sum_{i=1}^{n} \lambda_i y_{ki} \ge y_{k0} + \Delta y_{k0} \qquad \textit{for} \quad k = 1, \ldots, r \\
&\qquad \lambda_i \ge 0, \quad i = 1, 2, \ldots, n
\end{aligned}
$$

where $\mathbf{W}^{\mathbf{T}} \in R^m$.

Proof The CCR model for a modified DMU$_{0'}$ relative to other DMUs ($l = 1, \ldots, n$) is the following DCCR$'_0$ model.

$$
\begin{aligned}
&(\mathrm{DCCR}'_0) \quad \text{minimize} \quad \theta_{0'} \\
&\text{s.t.} \quad \sum_{i=1}^{n} \lambda_i x_{ji} + \lambda_{0'}(x_{j0} + \Delta x_{j0}) \le \theta_{0'}(x_{j0} + \Delta x_{j0}) \quad \textit{for} \quad j = 1, \ldots, m \\
&\qquad \sum_{i=1}^{n} \lambda_i y_{ki} + \lambda_{0'}(y_{k0} + \Delta y_{k0}) \ge y_{k0} + \Delta y_{k0} \qquad \textit{for} \quad k = 1, \ldots, r \\
&\qquad \lambda_{0'}, \lambda_i \ge 0, \quad i = 1, 2, \ldots, n.
\end{aligned}
$$

The set of constraints in the DCCR'_0 model can be rearranged in the following form:

$$\sum_{i=1}^{n} \lambda_i x_{ji} \leq (\theta_{0'} - \lambda_{0'})(x_{j0} + \Delta x_{j0}) \qquad \textit{for} \quad j = 1, \ldots, m$$

$$\sum_{i=1}^{n} \lambda_i y_{ki} \geq (1 - \lambda_{0'})(y_{k0} + \Delta y_{k0}) \qquad \textit{for} \quad k = 1, \ldots, r$$

$$\lambda_{0'}, \lambda_i \geq 0, \; i = 1, 2, \ldots, n.$$

The objective function in the DCCR'_0 model is to minimize $\theta_{0'}$. Using $\Delta\mathbf{x}_0$ obtained from solving the MLDRA_0 model, an optimal solution can be obtained at $\lambda_{0'} = 0$ and the optimal objective value ($\theta_{0'}$) is equal to θ_0^*. Otherwise, $\Delta\mathbf{x}_0$ is not optimal for the MLDRA_0 model and can be reduced further. Therefore, the relative efficiency level of $\mathrm{DMU}_{0'}$ with respect to the set of DMU_l ($l = 1, \ldots, n, 0'$) will remain equal to θ_0^*.

For other DMUs, the dual form of DRA_l model can be written in the vector-metric PRA_l form as follows.

$$\begin{aligned}
\text{maximize} \quad & \mathbf{u}^{\mathbf{T}}\mathbf{y}_l \\
\text{s.t.} \quad & \mathbf{v}^{\mathbf{T}}\mathbf{x}_l = 1 \\
& -\mathbf{v}^{\mathbf{T}}\mathbf{x}_i + \mathbf{u}^{\mathbf{T}}\mathbf{y}_i \leq 0 \qquad \textit{for} \quad i = 1, \ldots, n \\
& -\mathbf{v}^{\mathbf{T}}(\mathbf{x_0} + \Delta\mathbf{x_0}) + \mathbf{u}^{\mathbf{T}}(\mathbf{y_0} + \Delta\mathbf{y_0}) \leq 0 \\
& \mathbf{u}, \mathbf{v} \geq \mathbf{0}
\end{aligned}$$

where $\mathbf{u}^{\mathbf{T}} = [u_1, u_2, \ldots, u_r]$, $\mathbf{v}^{\mathbf{T}} = [v_1, v_2, \ldots, v_m]$, $\mathbf{x}_i = \begin{bmatrix} x_{1i} \\ x_{2i} \\ \vdots \\ x_{ri} \end{bmatrix}$,

$\mathbf{y}_i = \begin{bmatrix} y_{1i} \\ x_{2i} \\ \vdots \\ y_{mi} \end{bmatrix}$, $\mathbf{x}_0 = \begin{bmatrix} x_{10} \\ x_{20} \\ \vdots \\ x_{r0} \end{bmatrix}$, $\Delta\mathbf{x}_0 = \begin{bmatrix} \Delta x_{10} \\ \Delta x_{20} \\ \vdots \\ \Delta x_{r0} \end{bmatrix}$ From the MLDRA_0 model, $(\mathbf{x_0} + \Delta\mathbf{x}_0, \mathbf{y_0} + \Delta\mathbf{y}_0) \in P$ where P is a production possibility set of all DMU_i, $i = 1, \ldots, n$. $P = \{(\mathbf{x}, \mathbf{y}) | \mathbf{x} \geq \mathbf{X}\boldsymbol{\lambda}, \mathbf{y} \leq \mathbf{Y}\boldsymbol{\lambda}, \boldsymbol{\lambda} \geq \mathbf{0}\}$, $\mathbf{X} = [x_{ji}]_{mxn}$, $\mathbf{Y} = [y_{ki}]_{rxn}$, and $\boldsymbol{\lambda} = (\lambda_i)_{n\times 1}$, $\boldsymbol{\lambda} \in R^n$.

When $(\mathbf{x_0} + \Delta\mathbf{x}_0, \mathbf{y_0} + \Delta\mathbf{y}_0) \in P$, then we have

$$
\begin{aligned}
-\mathbf{v}^{\mathbf{T}}(\mathbf{x_0} + \Delta\mathbf{x_0}) + \mathbf{u}^{\mathbf{T}}(\mathbf{y_0} + \Delta\mathbf{y_0}) &\le -\mathbf{v}^{\mathbf{T}}(\mathbf{X}\boldsymbol{\lambda}) + \mathbf{u}^{\mathbf{T}}(\mathbf{Y}\lambda) \\
&\le -\sum_{i=1}^{n} \mathbf{v}^{\mathbf{T}}\mathbf{x}_i\lambda_i + \sum_{i=1}^{n} \mathbf{u}^{\mathbf{T}}\mathbf{y}_i\lambda_i \\
&\le \sum_{i=1}^{n} \left(-\mathbf{v}^{\mathbf{T}}\mathbf{x}_i + \mathbf{u}^{\mathbf{T}}\mathbf{y}_i\right)\lambda_i
\end{aligned}
$$

From the PRA$_l$ model, $-\mathbf{v}^{\mathbf{T}}\mathbf{x}_i + \mathbf{u}^{\mathbf{T}}\mathbf{y}_i \le \mathbf{0}$ for $i = 1,\ldots,n$. Therefore, $-\mathbf{v}^{\mathbf{T}}(\mathbf{x_0} + \boldsymbol{\Delta}\mathbf{x_0}) + \mathbf{u}^{\mathbf{T}}(\mathbf{y_0} + \boldsymbol{\Delta}\mathbf{y_0}) \le \mathbf{0}$. This shows that $-\mathbf{v}^{\mathbf{T}}(\mathbf{x_0} + \boldsymbol{\Delta}\mathbf{x_0}) + \mathbf{u}^{\mathbf{T}}(\mathbf{y_0} + \boldsymbol{\Delta}\mathbf{y_0}) \le \mathbf{0}$ in the PRA$_l$ model is redundant and can be dropped out from the model without changing the solution set and the optimal objective value. In other words, the PRA$_l$ model is equivalent to the CCR model for DMU$_l$ before DMU$_0$ changes its output values. This implies that the relative efficiency levels of all DMU$_l$, $(l = 1, \ldots, n)$ remains unchanged. □

Theorem 48.3 *$\Delta\mathbf{x}_0 \in R^m$ obtained from solving the LDRA$_0$ model is a Pareto solution for the MLDRA$_0$ model.*

$$
\begin{aligned}
&(\text{LDRA}_0) \ \text{minimize} && \mathbf{W}^{\mathbf{T}}\Delta\mathbf{x}_0 \\
&\text{s.t.} && \sum_{i=1}^{n} \lambda_i x_{ji} \le \theta_0^*(x_{j0} + \Delta x_{j0}) \quad \textit{for} \quad j = 1, \ldots, m \\
& && \sum_{i=1}^{n} \lambda_i y_{ki} \ge y_{k0} + \Delta y_{k0} \qquad \textit{for} \quad k = 1, \ldots, r \\
& && \lambda_i \ge 0, \quad i = 1, 2, \ldots, n
\end{aligned}
$$

where $\mathbf{W}^{\mathbf{T}} \in R^m$.

Proof Assume that $\Delta\mathbf{x}_0^* \in R^m$ and $\boldsymbol{\lambda}^* = (\lambda_1^*, \ldots, \lambda_n^*)$ are the optimal solution from solving the LDRA$_0$ model but they were not Pareto solution to the MLDRA$_0$ model. There should be a possible $\overline{\Delta\mathbf{x}_0} \in R^m$ and $\overline{\boldsymbol{\lambda}} = (\overline{\lambda_1}, \ldots, \overline{\lambda_n})$ from the MLDRA$_0$ model where $\overline{\Delta\mathbf{x}_0} \le_p \Delta\mathbf{x}_0^*$, and thus $\mathbf{W}^{\mathbf{T}}\overline{\Delta\mathbf{x}_0} < \mathbf{W}^{\mathbf{T}}\Delta\mathbf{x}_0^*$, $\mathbf{W}^{\mathbf{T}} > \mathbf{0}$. Note that $\overline{\Delta\mathbf{x}_0} \le_p \Delta\mathbf{x}_0^*$ represents a set of inequalities $\overline{\Delta x_{j0}} \le \Delta x_{j0}^*, j = 1, 2, \ldots, m$ with at least one strict inequality, $\overline{\Delta x_{j0}} < \Delta x_{j0}^*$. Since the MLDRA$_0$ model and the LDRA$_0$ model have the same constraint sets, $\overline{\Delta\mathbf{x}_0} \in R^m$ and $\overline{\boldsymbol{\lambda}}$ are also the solution to the LDRA$_0$ model. This leads to a contradiction; therefore, $\Delta\mathbf{x}_0^* \in R^m$ and $\boldsymbol{\lambda}^*$ from the LDRA$_0$ model would also be a Pareto solution to the MLDRA$_0$ model. □

Table 48.1 Input and output data for the CCR model

	DMU				
Input/output data	1	2	3	4	5
Input 1	1	2	3	5	4
Input 2	1	4	1	4	3
Output 1	2	1	2	4	3

Table 48.2 The relative efficiency values of DMU_I and $\lambda_i, i = 1, \ldots, 5$

DMU_i	θ_i^*	λ_1	λ_2	λ_3	λ_4	λ_5
1	1	1	0	0	0	0
2	0.25	0.5	0	0	0	0
3	1	0	0	1	0	0
4	0.5	1.75	0	0.25	0	0
5	0.5	1.25	0	0.25	0	0

From Theorem 48.3, if we find any positive vector, $\mathbf{W}^{\mathbf{T}} \in R^m$, we would be able to find a Pareto solution for the MLDRA_0 model from solving the LDRA_0 model, which is a linear programming (LP) model. Therefore, there might be multiple Pareto solutions to the MLDRA_0 model.

48.4.2 Numerical Examples

The dataset is exhibited in Table 48.1 to show how to solve the inverse CCR model. After solving CCR models for all DMUs, their relative efficiency levels are given in Table 48.2. Only DMU_1 is fully efficient. DMU_3 is weakly or technical efficient because its input 1 can be reduced further to make this DMU fully efficient, even though $\theta_3^* = 1$ and it is on the production frontier. DMU_2, DMU_4 and DMU_5 are inefficient because their technical efficiencies are less than 1. Let consider DMU_2 with $\theta_2^* = 0.25$. Suppose that the output of DMU_2 is changed from 1 to 1.5 and let $\mathbf{W}^{\mathbf{T}} = (1,\ 1)$ for input weights. From Theorem 48.2 and 48.3, we can estimate the new input values by solving the following LDRA_0 model for DMU_2:

$$
\begin{aligned}
\text{minimize} \quad & \Delta x_{12} + \Delta x_{22} \\
\text{s.t.} \quad & \lambda_1 + 2\lambda_2 + 3\lambda_3 + 5\lambda_4 + 4\lambda_5 \leq 0.25(2 + \Delta x_{12}) \\
& \lambda_1 + 4\lambda_2 + \lambda_3 + 4\lambda_4 + 3\lambda_5 \leq 0.25(2 + \Delta x_{22}) \\
& 2\lambda_1 + \lambda_2 + 2\lambda_3 + 4\lambda_4 + 3\lambda_5 \geq 1.5 \\
& \Delta x_{12} \geq -2 \\
& \Delta x_{22} \geq -4 \\
& \lambda_1, \lambda_3, \lambda_4, \lambda_5 \geq 0.
\end{aligned}
$$

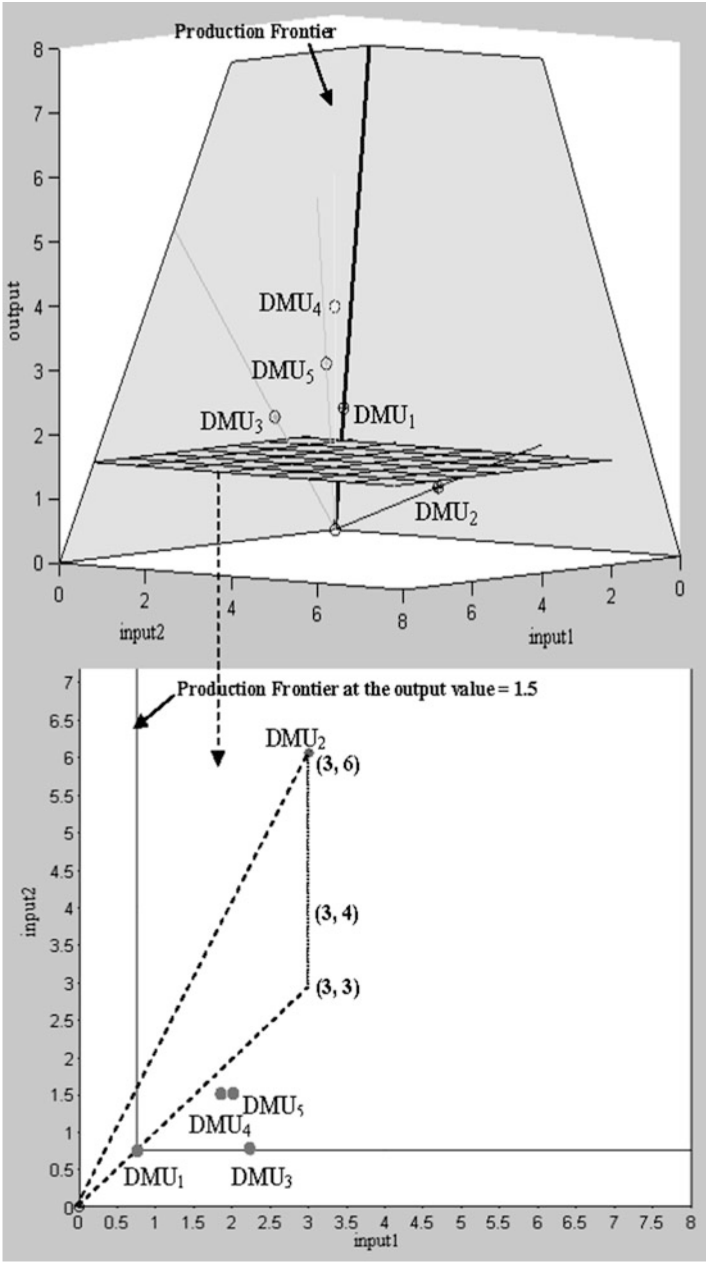


Fig. 48.1 The projections of all DMUs at the output value of 1.5

The optimal solution for this model is $\Delta x_{12} = 1, \Delta x_{22} = -1, \lambda_1 = 0.75, \lambda_2 = 0, \lambda_3 = 0, \lambda_4 = 0$ and $\lambda_5 = 0$.

When the output of DMU_2 is changed to 1.5, the input vector of DMU_2 is changed from $(2, 4)^{\mathrm{T}}$ to a new input vector $(3, 3)^{\mathrm{T}}$. Using the new input vector, the relative efficiency levels of all DMUs remain the same. The changes in the first and second input values of DMU_2 are positive and negative, respectively. This follows from Theorem 48.1, i.e., if $\sum_{i=1}^{5} \lambda_i x_{1i} + \lambda_{2'} x_{12} = (0.75)(1) + 0 = 0.75$, which is greater than $\theta_2^* x_{12} = (0.25)(2) = 0.5$, then $\Delta x_{12} \geq 0$. And if $\sum_{i=1}^{5} \lambda_i x_{2i} + \lambda_{2'} x_{22} = (0.75)(1) + 0 = 0.75$, which is less than $\theta_2^* x_{22} = (0.25)(4) = 1$, then $\Delta x_{22} \leq 0$.

If we use the approach by Wei et al. (2000), the results are $\Delta x_{12} = 1, \Delta x_{22} = 0, \lambda_1 = 0.75, \lambda_2 = 0, \lambda_3 = 0, \lambda_4 = 0$ and $\lambda_5 = 0$. The input vectors vector of DMU_2 is changed from $(2, 4)^{\mathrm{T}}$ to a new input vector $(3, 4)^{\mathrm{T}}$. The solution by Wei et al. (2000) did not give the best possible values of inputs. Figure 48.1 shows the production possibility set and the projections of all DMUs on the plane at the output value of 1.5. The efficiency value of the projection of DMU_2 on this plane is 0.25; however, it is projected on a weakly efficient part of the production frontier. The model proposed in this paper gives the new inputs (3, 3), whereas the model by Wei et al. (2000) gives the new inputs (3, 4). All points along the line connecting (3, 3) to (3, 6) give the same efficiency value; however, the Pareto-efficient point is (3, 3) because all inputs will be least consumed.

48.5 Conclusion

In this paper, we propose the inverse CCR problem to determine the lowest possible values of inputs for given values of outputs of a particular decision making unit (DMU) such that efficiency levels of all DMUs with respect to other DMUs remain unchanged. We show that given the increase in output values of DMU_0, $\Delta \mathbf{y}_0 \geq_p \mathbf{0}$, some components of $\Delta \mathbf{x}_0$ from the inverse CCR model can be negative, while other components of $\Delta \mathbf{x}_0$ can be positive or zero. However, in this paper we consider the problem when increases of some output values and decreases of the other output values can be taken into account simultaneously. To solve the inverse CCR model, which is in the form of MONLP, we need to find the value of $\Delta \mathbf{x}_0$ which keeps the relative efficiency levels of all DMUs unchanged. This can be done by solving the LDRA_0 model, which is in a linear programming form and can be solved by a standard LP solver.

Acknowledgments We gratefully acknowledge financial support from Thailand Research Fund (Grant no. MRG4980118) and Faculty of Agro-Industry, Kasetsart University, Thailand.

References

Alinezhad A, Makui A, Kiani Mavi R (2007) An inverse DEA model for inputs/outputs estimation with respect to decision maker's preferences: The case of Refah bank of IRAN. J Math Sci 1(1–2):61–70

Arnade CA (1994) Using data envelopment analysis to measure international agricultural efficiency and productivity. United Stated Department of Agriculture, Economic Research Service, Tech Bull1(36), 11831:1–30

Banker RD, Charnes A, Cooper WW (1984) Some models for estimating technical and scale efficiencies in data envelopment analysis. Manag Sci 30(9):1078–1092

Barros CP, Peypoch N (2009) An evaluation of European airlines' operational performance. Int J Prod Econ 122(2):525–533

Beasley JE (2003) Allocating fixed costs and resources via data envelopment analysis. Eur J Oper Res 147(1):198–216

Chandra P, Cooper WW, Li S, Rahman A (1998) Using DEA to evaluate 29 Canadian textile companies – considering returns to scale. Int J Prod Econ 54(2):129–141

Charnes S, Cooper WW, Rhodes E (1978) Measuring the efficiency of decision making units. Eur J Oper Res 2(6):429–444

Chen T-Y (1997) A measurement of the resource utilization efficiency of university libraries. Int J Prod Econ 53(1):71–80

Chilingerian JA (1994) Exploring why some physicians' hospital practices are more efficient: taking DEA inside the hospital. In: Charnes A, Cooper WW, Lewin AY, Seiford LM (eds) Data envelopment analysis: theory, methodology, and application. Kluwer Academic Publishers, Norwell, pp 167–194, ch. 9

Cooper WW, Seiford LM, Tone K (2007) Data envelopment analysis: a comprehensive text with models, applications, reference and DEA-solver software. Springer Science + Business Media LLC, Spring Street, pp 87–322

Hadi-Vencheh A, Foroughi AA (2006) A generalized DEA model for inputs/outputs estimation. Math Comput Model 43(5–6):447–457

Jahanshahloo GR, Lotfi FH, Shoja N, Tohidi G, Razavyan S (2004) Input estimation and identification of extra inputs in inverse DEA models. Appl Math Comput 156(2):427–437

Li X, Cui J (2008) A comprehensive Dea approach for the resource allocation problem based on scale economies classification. J Syst Sci Complex 21(4):540–557

Wei Q, Zhang J, Zhang X (2000) An inverse DEA model for inputs/outputs estimate. Eur J Oper Res 121(1):151–163

Yan H, Wei Q, Hao G (2002) DEA models for resource reallocation and production input/output estimation. Eur J Oper Res 136(1):19–31

Chapter 49
Auxiliary Process for User-Interface Design Based on Genetic Algorithm: Using Multimedia Software as an Example

Wei-Chang Yeh, Shang-Chia Wei, and Chin-Hao Yang

Abstract With the extensive application of human machine interaction, user-interface design is currently a significant issue in intelligent portable devices. This paper focuses on the user-interface design and proposes an efficient procedure for the key configuration in the user-interface of intelligent portable devices. In this paper, we assume a possible operational user-interface scenario and adopt revised genetic algorithms to generate an adequate key configuration and friendly user-interface in audio-visual software. The proposed methodology may accelerate the procedure of user-interface design, and easier to obtain user-interface patterns. The entire protocol has the potential to be applied to other related intelligent devices, and offer an alternative user-interface design in intelligent portable devices.

Keywords Genetic algorithms • Human machine interaction • Intelligent portable devices • User-interface design

49.1 Introduction

With the rapid development of the computer and its related technology, the wide use of the touch panel has not only increased its industrial use but has also resulted in its application to commercial requirements. The technology for human-machine interface in interactive software becomes increasingly important, as the need for touch panels grows and the field of application fields increases (Hall et al. 2004; Ahmed and Ashraf 2007; Karat et al. 1986; Ficarra et al. 2011). Recreational

W.-C. Yeh (✉) • S.-C. Wei • C.-H. Yang
Department Industrial Engineering and Engineering Management, National Tsing Hau University, Hsinchu, Taiwan
e-mail: yeh@ieee.org; d9534807@oz.nthu.edu.tw; s9934557@m99.nthu.edu.tw

E. Qi et al. (eds.), *The 19th International Conference on Industrial Engineering and Engineering Management*, DOI 10.1007/978-3-642-37270-4_49,

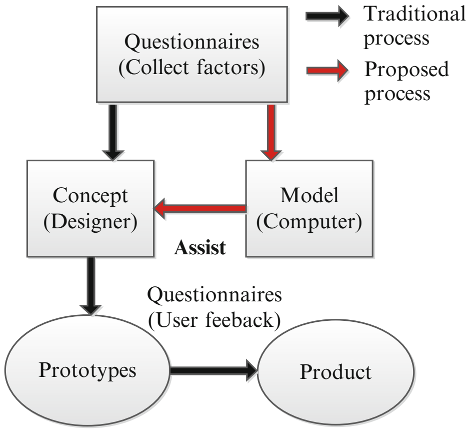


Fig. 49.1 The new concept for designing product process

software is one of the functions emphasized by intelligent portable device users, and for this reason, this project focuses on the interface design of entertainment software. However, different interactive designs of the human-machine interface will cause different perceptions about its use. Therefore, the design of the software interface will affect the usability of the software; poor configuration of the interface tends to cause inconvenience and shortcomings in the use of software (Kim 2010; Peer and Sharma 2008). Therefore, we survey users' operational habits and preferences by means of a questionnaire, and combine the basic norm of a human-machine interface with an optimized algorithm to design a configuration program of an entertainment interface that improves the friendliness and convenience of a touch software human-machine interface (Hall et al. 2004; Peer and Sharma 2008). This study presents a new concept for the design product process, as shown in Fig. 49.1. The traditional process of design is indicated by black arrows, which show that the designer is unable to fully consider all the complex elements. The red arrows indicate that a computer can consider complex effects among the various factors better than the human designer.

The goal of this research is to rationalize the operation to reduce cognitive loading, simplify the design to improve operational efficiency and better optimize the play space. To accomplish this goal, we aim for users to have a more comfortable and more intuitive operational experience; therefore, this work defines the possible results (design factors) from user questionnaires as an input (UI data) in the UI design model. The UI design model includes a specific algorithm like genetic algorithm (GA) that may contrive several UI patterns to satisfy important design factors. Then, designers choose and modify the UI patterns as UI prototypes. Finally, the adequate UI pattern for a specific product is obtained through user feedback using questionnaire responses.

49.2 Literature Review

49.2.1 User-Interface Design

We first discuss related studies about user-interface (UI) design or key configuration settings in human machine interaction, and then compare them with our work to present the main contribution.

A human-machine interface (HMI) can facilitate communication between users and machines. HMI is applicable to many scientific regions, such as PDA (Personal Digital Assist) or traditional circuit mechanical PLC (Programmable Logic Controller) and so on. A friendly HMI can reduce a user's learning time or detect operational faults for a mobile application in touch-screen devices, and it completes the work with minimum effort. In developing a user-friendly product, it is necessary to consider EC (Ergonomic Criteria), EUA (Elementary User Actions) and DF (Design Factors) (Kim 2010). Ahmed and Ashraf (Ahmed and Ashraf 2007) developed a UI (User-interface) prototype for a hotel management application based on a specific model. Lin and Wu (2010) presented an improved link analysis that can be applied to interfaces which require physical movements between the user and the interface and within touch screen and control panels.

49.2.2 Floorplanning

Optimizing key configuration of UI pattern is an HMI design problem that is similar to floorplanning (Chen et al. 2010; Nakaya et al. 2000). However, the key configuration setting in HMI is more intractable than in related problems such as the very large-scale integrated circuit (VLSI) floorplanning problem, because users' habits, software preferences and human machine interaction must also be considered. In the field of (VLSI), floorplanning is a general building block layout problem that determines the coarse placement of a given set of modules (Sait and Youssef 2004). The key configuration of floorplanning determines the performance, size, yield and reliability of VLSI chips. A module is a sub circuit whose shape is a rectangle with a given height and width in real number. A fIoorplanning of a set of modules is a non-overlapping placement of given modules. The minimum bounding rectangle of a floorplanning is called the chip. Each module has several terminals on its boundary which are interconnected with wires. Interconnections among terminals are specified by a net list. The floorplanning problem is to design M circuit's blocks on a chip with the minimum area and wire length. This problem is known to be NP-hard, and hence good heuristics are generally yielded (Chen et al. 2010; Nakaya et al. 2000; Sait and Youssef 2004; Bourbakis 2008). Gwee and Lim (1999) used a genetic algorithm with heuristic-based decoder for floorplanning in IC (Integrated Circuit) design for solving the AMI33 benchmark problem, and achieved a great result in terms of speed and reliability. Similarly, Tang and Sebastian

(2005) presented a GA for a non-slicing floorplanning problem using the O-tree representation. Experimental results confirm that the GA consistently produces better results than the deterministic algorithm.

This paper proposes a solution for the HMI design issue to improve operational efficiency, decrease the mental workload, and simplify interface design. We adopt a floorplanning model to transform into the HMI design model. We believe that the proposed concept concerning auxiliary UI design process could provide UI designer to obtain friendly UI rapidly and to reduce unnecessary labor cost.

49.3 Mathematical Formulation

49.3.1 Layout Principle

In HMI design problem, the UI design pattern is a code **X** that includes two dimensions, that is, the button sequence and button shape. The code **X** is converted into a UI design pattern mapping a UI layout according to the layout principle (Chen et al. 2010; Ho et al. 2004) described as follows:

1. Place the corresponding button of the first item bottom left.
2. Place the subsequent buttons from left to right
3. If a subsequent button to be placed over the assumed maximal width, the button is able to be placed to a new line above the current one.

A special constraint is attached to the second rule. If the button to be placed can be shifted a little to the right to reduce the height of floorplanning when placing buttons from left to right, the UI design model places the button in the lower position. For example, in Fig. 49.2, if a button is to be placed in the order left to right, it should be placed at point *a*. However, this arrangement is inappropriate because putting the button at point *b* would lower its height.

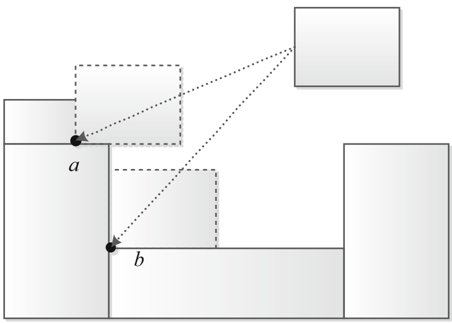


Fig. 49.2 Rules for placing buttons

49.3.2 Model Assumptions

To model this problem, the study makes the following assumptions:

1. All virtual buttons are in touch panel.
2. All virtual buttons are rectangular in shape, and there is no irregular shape in the touch panel.
3. The connection (*viz.* affinity) between two buttons is determined from questionnaire responses.
4. If there is empty space between one button and another in the final layout, the empty space will be converted to the button which is the most important among neighbor buttons, according to the connection (*viz.* affinity) of those buttons.

49.3.3 Evaluation of User-Interface Pattern

In general, the optimal objective is to minimize the UI area and to maximize affinity among all functional buttons. Considering the minimization of the UI area, the fitness is equal to the value of the area, and the fitness function of UI area $A(\mathbf{X})$ can be represented as

$$A(\mathbf{X}) = W^{*} \times H^{*}, \tag{49.1}$$

where W^{*} is the maximal width of a UI pattern $\mathbf{X}$, and H^{*} is the maximal height of UI pattern $\mathbf{X}$. The study aims to minimize $A(\mathbf{X})$ in the panel for users to watch the screen; therefore, the fitness value is as small as possible.

Affinity denotes a contact frequency between any two buttons so that we aim at the maximization of affinity in the UI pattern. For example, we may press the volume button after pressing the play button. If the play button neighbors the volume button, the total affinity will be a large value by formula (49.2). $E_{ij}(x)$ is the affinity between button i and button j, and $f_{ij}(x)$ is either 0 or 1 depending on whether affinity exists between button i and button j (formula (49.3)).

$$Af(\mathbf{X}) = \sum_{i \neq j} e_{ij} \times f_{ij}(x), \tag{49.2}$$

where

$$f_{ij}(x) = \begin{cases} 1, \text{ button } i \text{ is in contact with button } j. \\ 0, \text{ button } i \text{ is not in contact with button } j, \end{cases} \tag{49.3}$$

Table 49.1 Pseudo code for the GA

Alg.	Genetic Algorithm
Input	UI Data;
Output	UI pattern;
1:	*Initialization*;
2:	**Repeat**
3:	*Reproduction*;
4:	*Crossover*;
5:	*Mutation*;
6:	*Elitism*;
7:	**Until** termination condition is attained

The study considers the UI area and the affinity with reference to formula (49.1) and formula (49.2); we propose the affinity per unit area as final fitness function $F(\mathbf{X})$ that is maximum problem described as

$$F(\mathbf{X}) = \frac{Af(\mathbf{X})}{A(\mathbf{X})}. \tag{49.4}$$

49.3.4 GA for Optimal User-Interface Pattern

The HMI design problem is more intractable than typical NP-complete combinatorial optimization problems (COPs) such as the floorplanning problem (Gwee and Lim 1999); therefore, we apply genetic algorithm (GA) to optimize the new nonlinear COP. The GA is a stochastic search algorithm based on the process of natural selection and genetic recombination in biological evolution (de Castro and Von Zuben 2005), and the pseudo code for the GA is given in Table 49.1. The power behind a GA lies in its ability to explore many regions of the search space simultaneously in searching for an optimal floorplanning. The problem which the study seeks to solve is more difficult than traditional floorplanning, hence an improved genetic algorithm (GA) is applied to this question.

The details of the GA are described as follows:

1. *Initialization*: Produce all chromosomes.
2. *Reproduction*: Use roulette wheel to select the chromosomes for reproducing.
3. *Crossover*: Choose two chromosomes to change genes information by multi-point crossover.
4. *Mutation*: Choose two genes to change information on the chromosome by two-point mutation.
5. *Elitism*: Preserve half of the best-fit solutions selected from the current and former populations gives the population of the next generation.

49.4 Case Study and Experimental Result

49.4.1 Multimedia Software

Because the playback software features are too many to list, the study only lists ten functional buttons to substitute for the playback software, as shown in Table 49.2. All buttons are defined by collecting questionnaires concerning user survey for general video software.

This case defines thumb size as a basic block, because it is a convenient finger size on a touch panel. It is called button type 1, and the area of the button is approximately 1 cm^2 according to average thumb size. Then the size of the other buttons is the multiple of button type 1 and is generated as Table 49.3. Low bound (LB) and up bound (UB) indicates the range of the optional button type which is matched with the button sequence. Twelve button shapes can be chosen, and every button's shape is randomly generated of the range of the functional button's LB and UB. For example, the first button in Table 49.2 is a play and pause button; the button's LB is an eighth button type and the button's UB is a twelfth button type.

49.4.2 Experimental Result

Based on two kinds of parameter settings (*viz.* crossover rate and mutation rate), the study randomly selects five different UI patterns generated from genetic algorithm.

Table 49.2 General audio and video software functional buttons

No.	Group	Button	Diagram	LB	UB
1	Play	Play/Pause		8	12
2	Play	Stop		2	9
3	Play	Pre/Last		2	9
4	Play	Play slider		4	12
5	Basic	Switch/Zoom		8	10
6	Basic	Set up		1	8
7	Audio	Add/Remove		2	12
8	Audio	Mute/Volume		4	7
9	Audio	EQ mode		1	8
10	Audio	Play mode		1	8

Table 49.3 The virtual size of button

Type	1	2	3	4	5	6
Size	1	2h[a]	2v[b]	3h	3v	4h
Shape						
Type	7	8	9	10	11	12
Size	4v	4	6	8	9	5 h
Shape						

[a] *v* vertical
[b] h horizontal

Table 49.4 Comparison of different parameters in GA[a]

No.	Crossover rate	Mutation rate	Corr. coeff.
1	0.6	0.4	0.554
2	0.8	0.2	0.533

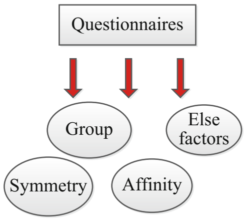
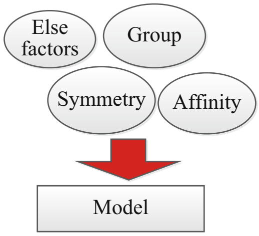
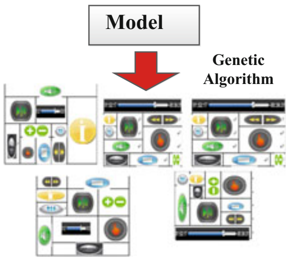
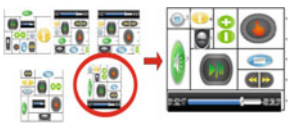
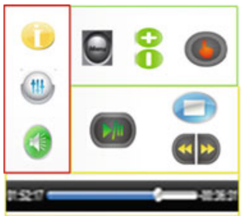
[a] Populations = 50, generations = 200

We asked participants to sort five UI patterns by their personal perceptions such as habits and preferences for general video software, and then placed the most favorable from left to right. Two hundred valid questionnaires were analyzed using Minitab, and the correlation coefficients between favor and fitness values are shown in Table 49.4. The genetic algorithm (GA) with two parameter settings has a medium relative relationship between favor and fitness values. We can claim that there is no significant difference between the parameter settings, and thus the parameter setting has little effect on the UI pattern. However, if we were to deliberate the ergonomic principles with users and designers, the relationship between favor and fitness value would be strong in the UI design model.

49.5 Conclusion

This study has presented genetic algorithm (GA) for the human-machine-interface (HMI) problem, which is significantly harder than the classical floorplanning problem. A key feature of this work is that it could assist a product designer to consider as many design principles as possible and generate different UI patterns quickly. The experiments show that our method could generate a number of friendly UI patterns which correspond substantially to the needs of product users. A designer could select the most adaptive one of those UI patterns and add the originality to the best UI pattern in order to create a unique product, as illustrated in Table 49.5. For a designer to design a UI pattern using traditional processes, there are likely to be

Table 49.5 The UI design process of as-is and to-be checklist

Procedure		Amendable checklist	
Stage 1 Collect factors (by designer)		AS-IS	Collect design factors about friendly UI concept from questionnaire responses, and pick out the significant factors
		TO-BE	(1) Consider more ergonomic factors
Stage 2 Mathematical model (by computer)		AS-IS	Use mathematical formulation to stand for significant factors as real as possible, and construct the model using GA.
		TO-BE	(2) The width of UI pattern is fixed.
Stage 3 Optimization (by computer)		AS-IS	Optimization of UI pattern considering interaction and exclusion effects of each button by GA, and collect several nice UI patterns.
		TO-BE	(3) Apply other heuristic algorithms to the model
Stage 4 Select adequate UI (by designer)		AS-IS	Choose an adequate UI pattern with minimal fitness value from many UI patterns
		TO-BE	(4) Add the real button to assist the virtual button
Stage 5 Modify and adjust UI (by designer)		AS-IS	The designer adopts aesthetic perception to make small placement adjustments based on the UI pattern
		TO-BE	(5) Can be applied to many types of equipment

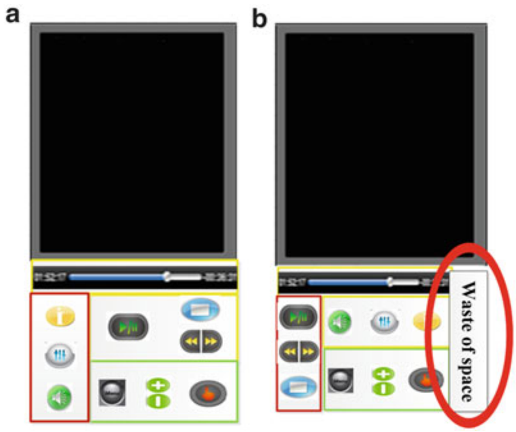


Fig. 49.3 Comparison of fixed screen size for different UI pattern. (**a**) Appropriate UI, (**b**) Inappropriate UI

Fig. 49.4 The real button assists with virtual buttons

unanticipated interaction effects as a result of the limitations of human thinking. This contribution of this work is that it applies a new method to the auxiliary UI design model, and the experimental results demonstrate that GA is useful for solving the HMI design problem proposed in this paper.

According to the AS-IS analysis and the TO-BE design in Table 49.5, we hope to improve the UI pattern model by considering the following amendable points in the future.

1. *Consider more ergonomic factors*: The study only considers factors that include superficial measure and buttons' affinity in our model. If more design factors

could be included into the model, the UI pattern generated by model would conform to the habits of the user. For example, the questionnaires show that the user prefers the Play-Slider-Bar to be placed at the bottom or top of the screen, and to occupy an entire row.

2. *The width of UI pattern is fixed*: The paper assumes that the width and height in the UI pattern will change. When the UI pattern is applied to the fixed screen on a cell phone, the method may generate a UI pattern with space that doesn't correspond to the size of the screen on the touch panel, as shown in Fig. 49.3. The area of UI pattern B would be smaller than that of UI pattern A, but it would also produce unsightly design patterns. Therefore, the width could be fixed, and the height changed by generation.
3. *Apply other heuristic algorithms to the model*: The GA is applied into this HMI design problem and has obtained a good experimental result. However, there are many heuristic algorithms which are probably better than the GA.
4. *Add the real button to assist the virtual button*: A friendly UI pattern for users would be achieved if real buttons existed alongside virtual buttons in the same UI pattern, as shown in Fig. 49.4.
5. *It could be applied to many types of equipment*: The study could be applied extensively in many cases such as Liquid Crystal Display (LCD), the UI pattern of a Notebook screen, the dashboard layout of car or motorcycle and the multi-operational interface of household appliances such as microwave, refrigerator and so on.

References

Ahmed S, Ashraf G (2007) Model-based user interface engineering with design patterns. J Syst Softw 80(8):1408–1422

Bourbakis NG (2008) A generic, formal language-based methodology for hierarchical floorplanning–placement. Comput Lang Syst Struct 34(1):25–42

Chen G, Guo W, Chen Y (2010) A PSO-based intelligent decision algorithm for VLSI floorplanning. Soft Comput 14(12):1329–1337

de Castro LN, Von Zuben FJ (2005) Recent developments in biologically inspired computing. Idea Group Publishing, Hershey, PA, USA

Ficarra FVC, Nicol E, Ficarra MC, Richardson L (2011) Advances in human-computer interaction: graphics and animation components for interface design. Lect Notes Comput Sci 6529:73–86

Gwee BH, Lim MH (1999) A GA with heuristic-based decoder for IC floor planning. Integr VLSI J 28(2):157–172

Hall L, Gordon A, Newall L, James R (2004) A development environment for intelligent applications on mobile devices. Exp Syst Appl 27(3):481–492

Ho SY, Ho SJ, Lin YK, Chu WCC (2004) An orthogonal simulated annealing algorithm for large floor planning problems. IEEE Trans Very Large Scale Integr (VLSI) Syst 12(8):874–877

Karat J, McDonald JE, Anderson M (1986) A comparison of menu selection techniques: touch panel, mouse and keyboard. Int J Man Mach Stud 25(1):73–88

Kim H (2010) Effective organization of design guidelines reflecting designer's design strategies. Int J Ind Ergon 40(6):669–688

Lin CJ, Wu C (2010) Improved link analysis method for user interface design-modified link table and optimisation-based algorithm. Behav Inf Technol 29(2):199–216

Nakaya S, Koide T, Wakabayashi S (2000) An adaptive genetic algorithm for VLSI floorplanning based on sequence-pair. In: Proceedings of the IEEE international symposium on circuits and systems (ISCAS), vol 3, Geneva, pp 65–68, 28–31 May 2000

Peer SK, Sharma DK (2008) Human computer interaction design with multi-goal facilities layout model. Comput Math Appl 56(9):2164–2174

Sait SM, Youssef H (2004) VLSI physical design automation: theory and practice. World Scientific, Singapore

Tang M, Sebastian A (2005) A genetic algorithm for VLSI floor planning using O-Tree representation. Lect Notes Comput Sci 3449:215–224

Chapter 50
A Mixed-Integer Linear Programming Model for an Integrated Planting, Harvesting, and Procuring in Pickled Ginger Industry

A. Ritvirool and A. Pantong

Abstract This paper addresses a production and procuring problem for fresh ginger rhizomes in a real-world pickled ginger industry. Planting and harvesting schedules affect fresh ginger rhizomes requirements for producing a pickled ginger. The management faces with the unbalance of fresh ginger rhizomes supply due to the lack of a decision support tool. Consequently, the requirements of the plant are not satisfied. A mixed-integer linear programming model has been developed for the problem to minimize the total cost, including cultivating cost, harvesting cost, inbound transportation cost, and purchasing cost. The numerical examples using the data from an industrial case study of pickled ginger plant are presented to illustrate how the proposed model could be applied to actual systems.

Keywords Ginger rhizomes • Planting • Harvesting • Procuring • Mixed-integer linear programming

50.1 Introduction

Ginger is the medicinal crop which is widely used as a spice in cooking. Currently, Japan and the United States are the major importing countries, whereas the major exporting countries are China and Thailand. Pickled ginger is used traditionally in Japanese and Chinese cuisine. The raw material of pickled ginger industry is fresh

---

A. Ritvirool (✉)
Department of Industrial Engineering, Naresuan University, Phitsanulok, Thailand
e-mail: apichair@nu.ac.th

A. Pantong
Department of Industrial Engineering, University of Phayao, Phayao, Thailand

E. Qi et al. (eds.), *The 19th International Conference on Industrial Engineering and Engineering Management*, DOI 10.1007/978-3-642-37270-4_50,

ginger rhizomes that are tender and low in crude fiber contents. The incoming fresh ginger rhizome lot is frequently not conforming to the requirements of the plant. Production and procurement planning of the fresh ginger rhizomes for processing pickled ginger is difficult. The production and purchasing managers need a decision support tool to help generate the optimal production and procurement plans which will lead to reduce costs.

The agri-food supply chain plays an important role in countries which export agricultural products. However, the lack of the scientific managements is the main problem of the enterprises which are now highly competitive. The mathematical models have been developed to help managing the agri-food supply chain in various agricultural industries. The mixed-integer linear programming (MILP) model has been used to solve a blending and shipping problem for bulk grain in Turkey. The blending, loading and transportation decisions could be integrated in one model to obtain an optimal solution (Bligen and Ozkarahan 2007). The planning problem of sugar cane supply for sugar industry in Germany has also been solved by the MILP model (Grunow et al. 2007). It is found that this model could be used for cultivation planning and harvest scheduling. Reference Ahumada and Villalobos (2009) reviews the application of planning models in agricultural supply chains. They focus on the field of production and distribution planning for agricultural perishable and non-perishable products. In addition, The MILP has been formulated for crop rotation schedule in Brazil (Santos et al. 2010) and developed for the production and logistics operations planning in semi-continuous food industries (Kopanos et al. 2012).

This paper aims to propose the MILP model for the fresh ginger rhizome production and procuring problem. The next section describes the problem in detail. Then, the mathematical model is formulated for the problem in Sect. 50.3. Section 50.4 presents the problem illustration using the data from a real industrial case study in Thailand. Then, the numerical results are presented in Sect. 50.5. Finally, in Sect. 50.6, some concluding remarks are given.

50.2 Problem Description

In Thailand, the best planting period for ginger rhizomes is summer before the rainy season begins. Planting in cool season, the crop needs irrigation. Furthermore, lack of precipitation and a water shortage increase the fiber content whereas the yield is decreasing. Harvesting should be done after 4–6 months of planting period. Crude fiber contents of ginger rhizomes increase continuously with rhizome maturity. Fiber development is exceedingly fast between 6 and 7 months after planting time (Kandiannan et al. 1996).

Fresh ginger rhizomes requirements for producing pickled ginger consist mainly of size and fiber contents. The rhizome is classified by crude fiber contents which have three classes consisting of G (low), H (medium), and K (high). Each lot of incoming fresh ginger rhizome to the factory is supplied from self-owned fields, contracted farmers' fields and selected suppliers' fields. Incoming lots of fresh ginger rhizomes from various sources have a combination of G, H, and K classes of fiber contents at different maturity stages. For example, the proportion of G, H, and K for the rhizomes age of 4 months after planting is 0.75, 0.18, and 0.07, respectively whereas those of 6 months after planting is 0.1, 0.61, and 0.29, respectively. The proportion of G decreases as the increment of rhizomes age. The supplier can also supply fresh premium ginger rhizomes, which have a high proportion of G, at all time of harvesting periods. However, the selling price is much higher than that of fresh regular ginger rhizomes. Therefore, the purchasing manager should specify the proportion of G, H, and K classes of fiber contents for incoming fresh ginger rhizome lots to match the pickled ginger factory requirements, evaluate and select the suppliers who can potentially supply ginger rhizomes conformance to the specification. Furthermore, the production manager should generate the production plan including rhizome production time, i.e., the time between planting and harvesting, for self-owned fields and contracted farmers' fields. Production time is a significant factor affecting the crude fiber contents and yield of ginger rhizomes. Purchasing and production managers should cooperate to allocate the number of rhizomes for cultivating self-owned fields, contracting with farmers and selecting suppliers. Both managers are faced with the fresh ginger rhizome production and procuring problem.

50.3 Mathematical Model

Notation

Sets

C Planting periods
D Growing periods
H Harvesting periods
I Cultivated fields
I_A Self-owned fields
I_B Contracted farmers' fields
J Rhizome suppliers
S Ginger rhizomes classes
T Truck size

Parameters

$GRat_s^d$ A proportion of ginger rhizomes class s and age d harvested in week h
$SRat_s^h$ A proportion of ginger rhizomes class s purchased in week h
$EXRat_s$ A proportion of premium ginger rhizomes class s purchased from supplier
$NFar_i$ Area of cultivated fields i (acre)
Yie_i Yield of cultivated field i (kg/acre)
$FCap$ The number of annual growing seasons (time/year)
$SCap_j$ Supplying capacity of the number of fresh ginger rhizomes purchased from supplier j (ton)
$TCap_t$ Loading capacity of truck size t (time/trip)
$EXCap^h$ Supplying capacity of the number of premium ginger rhizomes purchased in week h (ton)
$FCost_i$ Unit cultivating cost in fields i (baht)
$SCost_j$ Unit purchasing cost of ginger rhizomes ordered from supplier j (baht/ton)
$EXCost^h$ Unit purchasing cost of premium ginger rhizomes in week h (baht/ton)
$HCost$ Unit harvesting cost (baht/ton)
$TCost_t$ Unit inbound transportation cost of each truck size t (baht/ton/trip)
De_s A pickled ginger plant's annual demand for the fresh ginger rhizomes class s (ton)

Decision variables

$Far_i^{c,h}$ $= 1$, if field i is cultivated in week c and harvested in week h, 0 otherwise
$Gin_i^{c,h}$ The number of fresh ginger rhizomes planted in week c and harvested in week h at field i (ton)
$TGin^{c,h}$ The total number of fresh ginger rhizomes planted in week c and harvested in week h (ton)
$THar^h$ The total number of fresh ginger rhizomes harvested in week h (ton)
FG_s^h The total number of fresh ginger rhizomes class s supplied to plant in week h (ton)
Sup_j^h The number of fresh ginger rhizomes purchased from supplier j in week h (ton)
$EXSup^h$ The number of fresh premium ginger rhizomes purchased in week h (ton)
$NCar_t^h$ The number of trips of each truck size t in week h (trip)

The MILP model is formulated as follows:

$$\text{Minimize} \sum_{i \in I} \left(FCost_i \sum_{c \in C} \sum_{h \in H} Far_i^{c,h} + HCost \sum_{h \in H} THar^h \right) + \sum_{t \in T} \left(TCost_t \sum_{h \in H} NCar_t^h \right) + \sum_{j \in J} \left(SCost_j \sum_{c \in C} Sup_j^h + \sum_{h \in H} EXCost^h EXSup^h \right) \tag{50.1}$$

subject to

$$\sum_{c\in C}\sum_{h\in H} Far_i^{c,h} \leq FCap \quad \forall_{i\in I_A \cup I_B} \tag{50.2}$$

$$NFar_{iY_i}e_i Far_i^{c,h} = Gin_i^{c,h} \quad \forall_{c\in C, h\in H, i\in I_A\cup I_B} \tag{50.3}$$

$$\sum_{i\in I} Gin_i^{c,h} = TGin^{c,h} \quad \forall_{c\in C, h\in H} \tag{50.4}$$

$$THar^h = \sum_{c\geq 1}^{c\leq h} TGin^{c,h} \quad \forall_{h\in H} \tag{50.5}$$

$$THar^h = \sum_{t\in T} TCap^{c,h} NCar_t^h \quad \forall_{h\in H} \tag{50.6}$$

$$\begin{aligned} FG_s^h = {} & \sum_{c\geq 1}^{c\leq h} GRat_s^{d=h-c+16}\, TGin_h^c + \sum_{j\in J} SRat_s^{hSup_j^h} \\ & + EXRat_s EXSup^h \quad \forall_{h\in H, s\in S} \end{aligned} \tag{50.7}$$

$$FG_s^h \geq De_s^h \quad \forall_{h\in H, s\in S} \tag{50.8}$$

$$\sum_{h\in H} Sup_j^h \leq SCap_j \quad \forall_{j\in J} \tag{50.9}$$

$$EXSup^h \leq EXCap^h \quad \forall_{h\in H} \tag{50.10}$$

$$Gin_i^{c,h}, TGin^{c,h}, THar^h, FG_s^h, Sup_j^h, EXSup^h \geq 0 \quad \forall_{c\in C, h\in H, i\in I_A\cup I_B, s\in S, j\in J} \tag{50.11}$$

$$NCar_t^h \geq 0 \text{ and Integer} \quad \forall_{h\in H, t\in T} \tag{50.12}$$

$$Far_i^{c,h} \in \{0,1\} \quad \forall_{c\in C, h\in H, i\in I_A\cup I_B}. \tag{50.13}$$

The objective function (50.1) seeks to minimize the total cost of fresh ginger rhizome procurement. The first term in the objective function is cultivating cost, and the second term captures harvesting cost from self-owned and contracted farmers' fields. The third term is inbound transportation cost of loading ginger rhizome to the plant. The fourth and the fifth terms are costs of purchasing regular and premium fresh ginger rhizomes from suppliers, respectively.

Constraints (50.2) ensure that the number of annual growing seasons is limited to maintain the rhizome quality and yield for each growing season. The number of fresh ginger rhizomes should be allocated to self-owned fields and contract farmers' fields; it can then be formulated as Constraints (50.3) and (50.4). The number of fresh ginger rhizomes harvested each week can be expressed as Constraints (50.5). Inbound transportation of fresh ginger rhizomes from cultivated areas to the plant during harvesting periods is considered by determining the number of trips of each truck size as in constraints (50.6). The total number of all classes of fresh ginger rhizomes supplied to plant during harvesting periods can be determined as in Constraints (50.7). Constraints (50.8) ensure that the total number of all classes of fresh ginger rhizomes obtained satisfies plant's annual demand for the fresh ginger rhizomes. Constraints (50.9) and (50.10) ensure that the number of regular and premium fresh ginger rhizomes purchased from suppliers is not exceeded their supplying capacity. Constraints (50.11) and (50.12) are non-negativity conditions on the decision variables. Finally, constraints (50.13) ensure that the value of variables is binary integer.

50.4 Problem Illustration

Fresh ginger rhizome specification for producing pickled ginger consists of the crude fiber contents and size. According to the consumers' preferences, fresh ginger rhizomes used to produce pickled ginger are classified by the crude fiber content which has three classes, G, H, and K classes of fiber contents. Customer's demand is collected in advance from Japan and the US importers. Pickled ginger plant's annual expected demand for the fresh ginger rhizomes during harvesting periods is shown in Table 50.1. A planning horizon of 6 months is started from April to the end of September. The total periods are approximately 24 weeks. The planting period (c) is 2 month duration (8 weeks) started from April to May as well as harvesting periods (h) started from August to September.

The size of the test problem using data from a real industrial case study is shown in Table 50.2.

Table 50.1 Demand for fresh ginger rhizome (tons)

h	G	H	K
1	50	0	0
2	200	50	10
3	250	100	20
4	400	200	50
5	100	100	50
6	50	100	0
7	50	50	0
8	50	0	50
Total	1,150	600	180

Table 50.2 Size of the test problem

Number of planting periods	8
Number of growing periods	8
Number of harvesting periods	8
Number of cultivated fields	7
Number of self-owned fields	3
Number of contracted farmers' fields	4
Number of rhizome suppliers	5
Number of ginger rhizome classes	3
Number of truck size	3

Table 50.3 Parameter data

	$GRat_s^d$			$SRat_s^h$		
h,d	G	H	K	G	H	K
1	0.8	0.16	0.04	0.75	0.18	0.07
2	0.75	0.18	0.07	0.7	0.21	0.09
3	0.7	0.21	0.09	0.65	0.23	0.12
4	0.65	0.23	0.12	0.55	0.28	0.17
5	0.55	0.28	0.17	0.35	0.44	0.21
6	0.35	0.44	0.21	0.2	0.55	0.25
7	0.2	0.55	0.25	0.1	0.61	0.29
8	0.1	0.61	0.29	0.05	0.63	0.32

Table 50.4 Planting and harvesting plans ($Far_i^{c,h}$)

h	1	2	3	4	5	6	7	8
i	–	7	4,6	–	2	5	1	3
c	–	1	1	–	1	1	4	3
d	–	17	18	–	20	21	19	21

A proportion of ginger rhizomes class s and age d harvested in week h and a proportion of ginger rhizomes class s purchased in week h are shown in Table 50.3.

The number of total decision variables is 1,071 including binary, integer and non-integer variables of 448, 24, and 599, respectively and 571 linear constraints. The modeling language AMPL was used to model the MILP problem and solved with CPLEX 11.1.0. The computational test was performed on an Intel Core 2 Duo PC with a 2.53 GHz processor and 4 Gb of RAM. Solution time in CPU is 0.44 s.

50.5 Results

As seen in Table 50.4, planting is operated in all self-owned fields and Contracted farmers' fields (i) during the first month of planting periods (c). Harvesting is done almost every harvesting period (h) except the first and the fourth weeks of harvesting periods. The fresh ginger rhizome age (d) in each field is between 17 and 21 weeks. The results of the proposed model for production planning support previous studies

Table 50.5 Procuring plans (tons)

h	$THar^h$	Sup_j^h 1	2	3	4	$EXSup^h$
1	–	–	–	–	54.7	10
2	230	41	–	–	–	–
3	440	–	33	–	–	–
4	–	–	132.3	400	183.5	7
5	128	–	–	–	145.8	–
6	200	–	–	–	21.8	–
7	70	20.9	34.7	–	–	–
8	180	38.1	–	–	–	–
Total	1,248	100	200	400	405.8	17

Table 50.6 Fresh ginger rhizomes supplied to the plant (FG_s^h, tons)

h	G	H	K
1	50	10.7	3.9
2	201.2	50	19.8
3	329.5	100	43.6
4	400	201.1	121.8
5	121.4	100	52.4
6	74.4	100	47.5
7	51.1	50	24.5
8	64.9	103.2	50
Total	1,292.5	715	363.5

which have shown that ginger rhizomes should be harvested 4–5 months after planting (Kandiannan et al. 1996; Ravindran and Nirmal Babu 2005).

The total expected number of ginger rhizomes harvested from all fields ($THar^h$) is approximately 1,248 tons and ordering from the suppliers in the number of 1,122.8 tons as shown in Table 50.5. To satisfy the demand, regular ginger rhizomes are ordered from four suppliers (Sup_j^h) in the total number of 1,105.8 tons using purchasing plan shown in Table 50.5. In harvesting weeks 1 and 4, fresh premium ginger rhizomes ($EXSup^h$) are ordered from the supplier instead of harvesting from the cultivated field, in the total number of 17 tons. The total number of each class of fresh ginger rhizomes supplied to the plant during harvesting periods, as shown in Table 50.6, is an excessive supply over demand for the plant by 11.03, 16.08, and 50.48 % for classes G, H, and K, respectively. This assures that all demands are satisfy.

50.6 Conclusion

A mathematical model has been developed for a production and procuring problem as a decision support tool to deal with complex decisions. The results showed that the proposed mathematical model could be used to help the management generate the optimal planting, harvesting, and procurement plans for fresh ginger rhizomes which will lead to reduce costs.

Acknowledgment We are grateful to the staffs at the pickled-ginger processing plant in Phitsanulok, Thailand, for providing us with all data and information in the test case. This work was funded by Naresuan University, Thailand.

References

Ahumada O, Villalobos JR (2009) Application of planning models in the agri-food supply chain: a review. Eur J Oper Res 195:1–20
Bligen B, Ozkarahan I (2007) A mixed-integer linear programming model for bulk grain blending and shipping. Int J Prod Econ 107:555–571
Grunow M, Gunther HO, Westinner R (2007) Supply optimization for the production of raw sugar. Int J Prod Econ 110:224–239
Kandiannan K, Sivaraman K, Thankamani CK, Peter KV (1996) Agronomy of ginger – a review. J Spices Aromat Crop 5(1):1–27
Kopanos GM, Puigjaner L, Georgiadis MC (2012) Simultaneous production and logistics operations planning in semicontinuous food industries. Omega 40(5):634–650
Ravindran PN, Nirmal Babu K (2005) Ginger: the genus Zingiber. CRC Press, Boca Raton, pp 211–240
Santos LMR, Costa AM, Arenales MN, Santos RHS (2010) Sustainable vegetable crop supply problem. Eur J Oper Res 204(3):639–647

Chapter 51
Reliability Analysis for Mechanical Components Subject to Degradation Process and Random Shock with Wiener Process

Hui-bing Hao, Chun Su, and Zhong-zhou Qu

Abstract For mechanical component, there are usually several different processes which can cause the component to failure. In this paper, the reliability modeling is studied for the component which has two kinds of failure processes, i.e., degradation process and random shocks. The Wiener process is used to describe the degradation process, the cumulative effect due to random shock on the degradation process is considered, and the effect of shock is discussed. The parameters of the model are estimated by the maximum likelihood estimation method. A case study of fatigue crack growth is provided to illustrate the proposed model and method. The reliability assessment results are also compared with the method of normal distribution. The results show that considering the impact of shocks can obviously lower the reliability of the system. Thus, the effect that the shocks act on the degradation may not be neglected.

Keywords Degradation • Reliability analysis • Random shock • Wiener process

51.1 Introduction

Reliability is a probability that an item performs its required function under given conditions for a stated time (Birolini 2007). Traditionally, the assessment of system reliability is usually based on failure data. However, when the system is highly reliable, it is very difficult to obtain enough failure data. At this time, degradation

This work is supported by the Graduate Training Innovative Projects Foundation of Jiangsu Province, China under Grant No. CXLX12_0081.

H. Hao (✉) • C. Su • Z. Qu
Department of Industrial Engineering, School of Mechanical Engineering,
Southeast University, Nanjing, China
e-mail: haohuibing@163.com

E. Qi et al. (eds.), *The 19th International Conference on Industrial Engineering and Engineering Management*, DOI 10.1007/978-3-642-37270-4_51,

data can be an efficient way to estimate the reliability (Meeker et al. 2001). In the last decades, degradation data has played an important role in evaluating system reliability.

As we know that degradation, such as wear, erosion and fatigue, is very common for most mechanical system or components. It can be described by a continuous performance process in terms of time (Zuo et al. 1999; Robinson and Croder 2000; Tseng et al. 2009). In Ref. (Zuo et al. 1999), three kinds of methods for degradation data analysis are presented, including linear regression method, degradation path method, and stochastic process method. The degradation data need to be dealt with so as to fit a known distribution or stochastic process, and the maximum likelihood estimation method, Bayesian method and the least squares method are usually used to estimate the parameters (Robinson and Croder 2000; Tseng et al. 2009). However, in the above researches, the effect of shocking was not considered.

Shock is one of the major reasons that cause failure to mechanical components (Nakagawa 2007). On the one hand, the random shocks will speed up the degradation process. And on the other hand, the degradation process will make the system more vulnerable to the random shocks. Thus, the effect on the degradation due to shocks may not be neglected.

In recent years, the interactions between degradation and random shock process have drawn attentions from academia. Klutke and Yang (2002) derived an availability model for an inspected system subject to continuous degradation and shocks. Li and Pham (2005a, b) considered the reliability and maintenance model for two degradation processes and random shocks. Kharoufeh et al. (2006) derived the system lifetime distribution and the limiting average availability for a failure process involving both degradation and shocks. Deloux et al. (2009) proposed a predictive maintenance policy for a continuously deteriorated system with deterioration and random shocks. Lehman (2009) surveyed two classes of degradation-threshold-shock (DTS) models, including general DTS and DTS with covariates, in which system failure were caused by the competing failure between degradation and trauma. Wang et al. (2011) considered a reliability model by using fuzzy degradation data when degradation and shocks were involved.

Most studies among Ref. (Klutke and Yang 2002; Li and Pham 2005a, b; Kharoufeh et al. 2006; Deloux et al. 2009; Lehmann 2009; Wang et al. 2011) assumed that the degradation process and shock process are independent with each other. However, there exist correlations between the degradation process and the random shocks. Shocks can not only decrease the performance directly, but also speed up the degradation.

Additionally, the above studies used the continuous distribution function or the linear regression method to describe the degradation of the system. In fact, for some system the degradation indices are non-monotonic. In fact, existing methods cannot describe this phenomenon properly.

The Wiener process is a kind of stochastic process with independent increment, and it can be used to describe non-monotonic properties. Due to its good properties of modeling the degradation processes on analysis and computation, the Wiener process has been used in reliability fields by many researchers, e.g. Whitmore (1995), Wang (2011), and Chun Su and Ye Zhang (2010).

In this paper, we consider a combination model for systems subject to random shock and degradation process. The random shock can also cause an abrupt damage to the degradation process. The Wiener process is used to describe the degradation process, and the parameters are estimated by the maximum likelihood estimation method. A case study of fatigue crack growth is provided to illustrate the proposed model and method.

51.2 Reliability Analysis Considering Degradation and Shock

51.2.1 Analysis of Degradation Process

It is assumed that system degradation performance at time t is $W(t)$ and H is the failure threshold. When the degradation performance $W(t)$ exceeds the threshold H for the first time, the component is considered to be failed.

Degradation can be influenced by the random factors from the component itself and the environment. As a consequence, a good statistical model should take into account the sources of variation, and stochastic process is appropriate to describe the degradation of the system.

Suppose that the degradation process, $\{W(t), t \geq 0\}$, obeys a Wiener process:

$$W(t) = \mu t + \sigma B(t)$$

where μ is the drift parameter; σ is the diffusion parameter; $B(t)$ is the standard Brownian motion.

The Wiener process has the following properties:

1. $W(0) = 0$;
2. $W(t)$ has continuous sample paths;
3. $W(t)$ has independent increments;
4. $W(t) - W(s) \sim N(\mu(t-s), \sigma^2(t-s))$ for $t > s \geq 0$.

From the properties of the Wiener process, we can obtain that the degradation performance $W(t)$ is normally distributed as

$$W(t) \sim N(\mu t, \sigma^2 t)$$

The probability density function can be defined as

$$f(w) = \frac{1}{\sqrt{2\pi t}\sigma} \exp\left\{-\frac{(w-\mu t)^2}{2\sigma^2 t}\right\}$$

For the Wiener process has many good properties, it has been used by many researchers in reliability field. In this paper, the degradation path is assumed to follow a Wiener process.

51.2.2 Shock Process Analysis

It is known that many factors from the environment can bring shocks to the components. Random shock modeling has been extensively studied for the component under the external shock environments, such as sudden and unexpected usage loads and accidental dropping onto hard surfaces (Nakagawa 2007; Wang et al. 2011). Shocks caused by the random environmental factors usually follow Poison processes. In the literature, four categories of random shock models are considered: cumulative shock model, extreme shock model, run shock model and δ-shock model (Bai et al. 2011). In this paper, the cumulative shock model is considered.

Suppose that random shocks arrive according to a homogeneous Poisson process with rate λ. Let the random variable $N(t)$ denote the number of shocks until time t with the probability

$$P\{N(t) = n\} = \frac{(\lambda t)^n}{n!} e^{-\lambda t}, \;\; n = 0, 1, 2, \cdots$$

In addition, the shock damage sizes are used to measure the instant increase in the degradation and are assumed to be independent identically distributed random variables, denoted as Y_j for $j = 1, 2, \ldots, \infty$. The cumulative damage size of the degradation process due to random shocks until time t is given as

$$S(t) = \begin{cases} \sum_{j=1}^{N(t)} Y_j & \text{if } N(t) > 0 \\ 0 & \text{if } N(t) = 0 \end{cases}$$

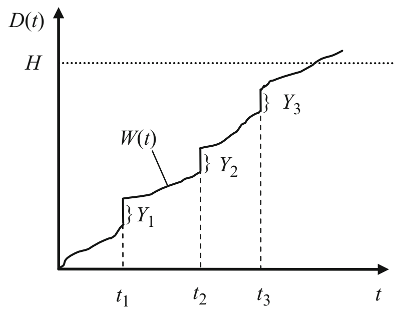
where $N(t)$ is the total number of shocks to the system that have arrived by time t.

51.2.3 Reliability Analysis for Combination Model

Shown as in Fig. 51.1, the total degradation damage is the cumulative effect of continuous degradation and sudden shocks. Obviously, the degradation is an aging process during field operation, and shock loads can cause additional abrupt damage, which contributes to the degradation process.

Now we establish reliability assessment model by considering a degradation process and the effect that shocks on the degradation process. The failure is defined as that the total degradation amount D(t) drops below certain failure threshold H.

Fig. 51.1 The degradation process under random shock



From Fig. 51.1, the total degradation performance including continuous degradation and shock damages can be expressed as

$$\begin{aligned} D(t) &= W(t) + S(t) \\ &= \mu t + \sigma B(t) + Y_1 + \cdots + Y_{N(t)} \end{aligned}$$

Based on the failure definition, the reliability at time t can be derived as

$$\begin{aligned} R(t) &= P(D(t) < H) \\ &= P(W(t) + S(t) < H) \end{aligned}$$

If the shocks arrive according to a Poisson process with rate λ, it can be obtained

$$\begin{aligned} R(t) &= P(W(t) + S(t) < H) \\ &= P\left(W(t) + \sum_{j=1}^{N(t)} Y_j < H\right) \\ &= \sum_{n=0}^{\infty} P\left(W(t) + \sum_{j=1}^{N(t)} Y_j < H \middle| N(t) = n\right) P\left(N(t) = n\right) \end{aligned}$$

Furthermore, it is supposed that the damage sizes due to shocks are independent identically distributed normal random variables with the distribution

$$Y_j \sim N(\mu_Y, \sigma_Y^2), \;\; j = 1, 2, \cdots$$

The corresponding probability density function is

$$g(y) = \frac{1}{\sqrt{2\pi}\sigma_Y} \exp\left\{-\frac{(y-\mu_Y)^2}{2\sigma_Y^2}\right\}$$

Then, we can get

$$S(t) = \sum_{j=1}^{n} Y_j \sim N(n\mu_Y, n\sigma_Y^2)$$

According to the properties of Wiener process, we can get

$$W(t) \sim N(\mu t, \sigma^2 t)$$

Thus,

$$W(t) + S(t) \sim N(\mu t + n\mu_Y, \sigma^2 t + n\sigma_Y^2)$$

And

$$\begin{aligned} P\left(W(t) + \sum_{j=1}^{N(t)} Y_j < H | N(t) = n\right) \\ = P\left(W(t) + \sum_{j=1}^{n} Y_j < H\right) \\ = \Phi(\frac{H - \mu t - n\mu_Y}{\sqrt{\sigma^2 t + n\sigma_Y^2}}) \end{aligned}$$

where $\Phi(\cdot)$ is the cumulative distribution function of a standard normally distributed variable.

Then the reliability function $R(t)$ can be derived as

$$R(t) = \sum_{n=0}^{\infty} \Phi(\frac{H - \mu t - n\mu_Y}{\sqrt{\sigma^2 t + n\sigma_Y^2}}) \cdot \frac{\exp(-\lambda t)(\lambda t)^n}{n!}$$

Thus the probability density function of the failure time, $f(t)$, is derived as

$$\begin{aligned} f(t) &= -\frac{dR(t)}{dt} \\ &= \sum_{n=0}^{\infty} \left(\phi(\frac{H - \mu t - n\mu_Y}{\sqrt{\sigma^2 t + n\sigma_Y^2}}) \cdot \frac{\exp(-\lambda t)(\lambda t)^n}{n!} \cdot \frac{\mu(\sigma^2 t + n\sigma_Y^2) + \frac{1}{2}\sigma^2 (H - \mu t - n\mu_Y)}{(\sigma^2 t + n\sigma_Y^2)^{3/2}} \right) \\ &\quad - \sum_{n=0}^{\infty} \Phi(\frac{H - \mu t - n\mu_Y}{\sqrt{\sigma^2 t + n\sigma_Y^2}}) \cdot \frac{\lambda \exp(-\lambda t)(\lambda t)^{n-1}(n - \lambda t)}{n!} \end{aligned}$$

where $\phi(\cdot)$ is the probability density function of a standard normally distributed variable.

51.3 Parameters Estimation of Reliability Model

It is supposed that there are N samples in the test, and each sample has M times of observations. Let W_i (t_j) be observation of the ith sample at the corresponding time t_j, i =1, 2, ..., N; j = 1, 2, ...,M. The degradation data for this model can be presented in the form

$$W_{N\times M}(t) = \begin{pmatrix} W_1(t_1) & \cdots & W_1(t_M) \\ \vdots & \ddots & \vdots \\ W_N(t_1) & \cdots & W_N(t_M) \end{pmatrix}$$

Let

$$\Delta W_i\left(t_j\right) = W_i\left(t_j\right) - W_i\left(t_{j-1}\right),\ t_0 = 0$$

Based on the independent increment property of the Wiener process, the random variable ΔWi(tj) has the following distribution

$$\Delta W_i\left(t_j\right) \sim N\left(\mu\Delta t_j, \sigma^2\Delta t_j\right),\ \Delta t_j = t_j - t_{j-1}$$

So, the probability density function of ΔW_i (t_j) is

$$f(\Delta W_i(t_j)) = \frac{1}{\sqrt{2\pi\Delta t_j}\sigma}\exp\left\{-\frac{(\Delta W_i(t_j) - \mu\Delta t_j)^2}{2\sigma^2\Delta t_j}\right\}$$

where i = 1, 2, ..., N, j = 1, 2, ..., M.

Thus the likelihood function of the samples is

$$\begin{aligned} L(\mu,\sigma) &= \prod_{i=1}^{N}\prod_{j=1}^{M} f(\Delta W_i(t_j)) \\ &= \prod_{i=1}^{N}\prod_{j=1}^{M} \frac{1}{\sqrt{2\pi\Delta t_j}\sigma}\exp\left\{-\frac{(\Delta W_i(t_j) - \mu\Delta t_j)^2}{2\sigma^2\Delta t_j}\right\} \end{aligned}$$

Taking the logarithm of the likelihood function and the log-likelihood function as

$$\ln L(\mu,\sigma) = \sum_{i=1}^{N}\sum_{j=1}^{M} \ln f(\Delta W_i(t_j))$$

Differentiating ln$L$ with respect to μ and σ, respectively, we can get

$$\hat{\mu} = \frac{\sum_{i=1}^{N}\sum_{j=1}^{M}\Delta W_i(t_j)}{\sum_{i=1}^{N}\sum_{j=1}^{M}\Delta t_j}$$

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^{N}\sum_{j=1}^{M}(\Delta W_i(t_j) - \hat{\mu}\Delta t_j)^2}{\sum_{i=1}^{N}\sum_{j=1}^{M}\Delta t_j}$$

51.4 Numerical Examples

Fatigue crack is a common degradation phenomenon for mechanical components. In this section, an example is given to illustrate the proposed models in Sects. 51.2 and 51.3. This example is based on the fatigue crack data of an alloy in (Lu and Meeker 1993). All samples had an initial crack length of 0.90 in.. Figure 51.2 shows the cumulative degradation of the fatigue crack data.

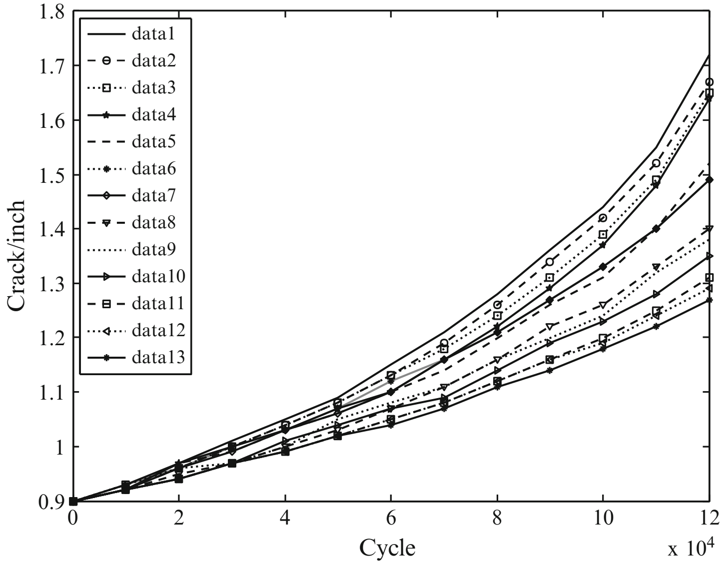


Fig. 51.2 The cumulative degradation of fatigue crack sizes

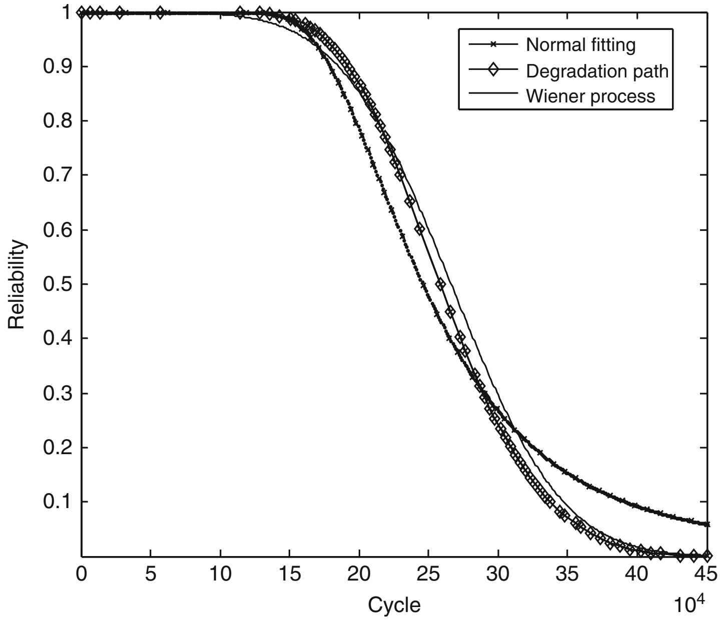


Fig. 51.3 Plot of degradation reliability using different methods

Wang et al. (2011) and some other researchers have used the data to analyze the reliability of the component. Here we will use the same data to illustrate the proposed model, and the results will also be compared with Wang's results.

Based on the degradation data, we can obtain the degradation increments data. According to the proposed estimated method in Sect. 51.3, the maximum likelihood estimations of μ and σ are obtained as:

$$\hat{\mu} = 4.78 \times 10^{-6}$$

and

$$\hat{\sigma}^2 = 0.77 \times 10^{-7}$$

The reliability curves using different estimation methods with the test data as shown in Fig. 51.3, and the details can be seen in Lu and Meeker (1993), Lu et al. (1997). But the effect of shocking is not considered in those paper. From Fig. 51.3, we can find that the reliability curves are nearly the same before12×10^4 -cycle by using different estimation methods, while after 12×10^4 cycle, the estimated reliability curves have a little difference.

It is supposed that the random shock process follows a Poisson process, and the shock damage size Y_j (j =1, 2, . . .) is assumed to follow a normal distribution, $Y_j \sim N(0.02 \text{ in.}, 0.1 \text{ in.})$; the failure threshold value $H = 2.0$ in., and the occurrence rate is $\lambda = 1.0$.

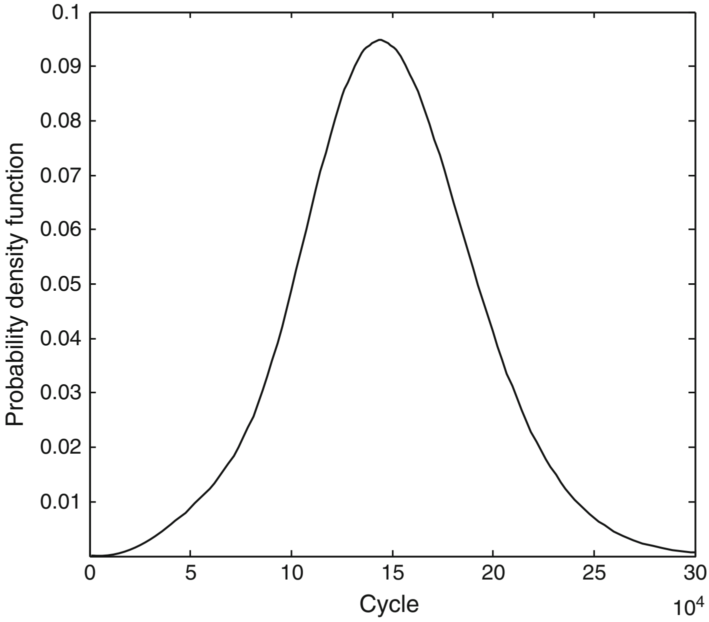


Fig. 51.4 Plot of failure time distribution

The probability density function $f(t)$ and the corresponding reliability function $R(t)$ are shown as in Figs. 51.4 and 51.5, respectively.

Additionally, in this paper we will discuss the impacts of shocks to the system reliability. If the damages caused by the random shocks are not considered, it is assumed that $Y_j = 0$, for $j = 1, 2, \ldots$. The reliability curve is plotted in Fig. 51.5, and the upper curve is reliability that the random shocks are not considered. Obviously, if we do not account for the effects of shocks, the reliability estimate will be higher. The lower curve is the reliability when the random shocks are accounted for.

As shown in the Fig. 51.5, the effects of shocks are not so significant before t reaches the 5×10^4 cycle, the two reliability curves nearly overlap before that. For the region that t is larger than 5×10^4 -cycle, the two curves begin to separate and the effects of shocks become larger.

Moreover, the degradation path is assumed to follow the normal distribution in (Wang et al. 2011), $W(t) \sim N(\mu(t), \sigma^2(t))$, where

$$\mu(t) = 0.8767 \exp(0.0398t)$$

and

$$\sigma^2(t) = 7.7602 \times 10^{-5} \exp(0.4646t).$$

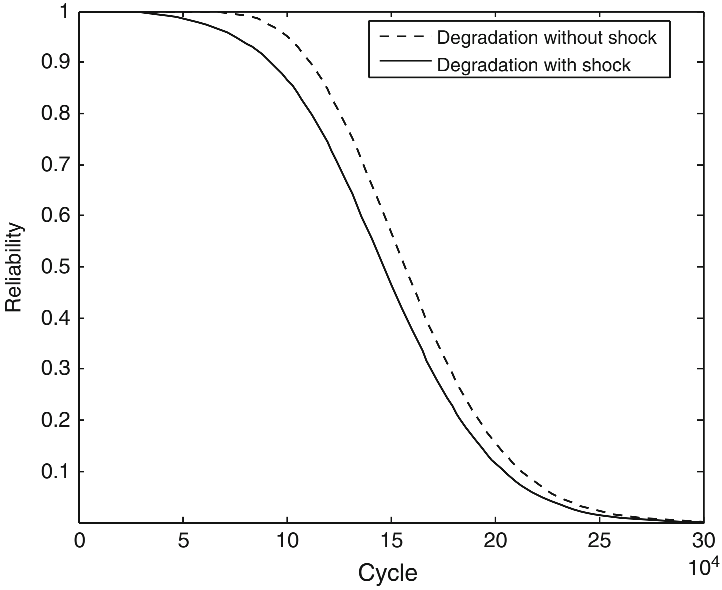


Fig. 51.5 Plots of reliability function under different case

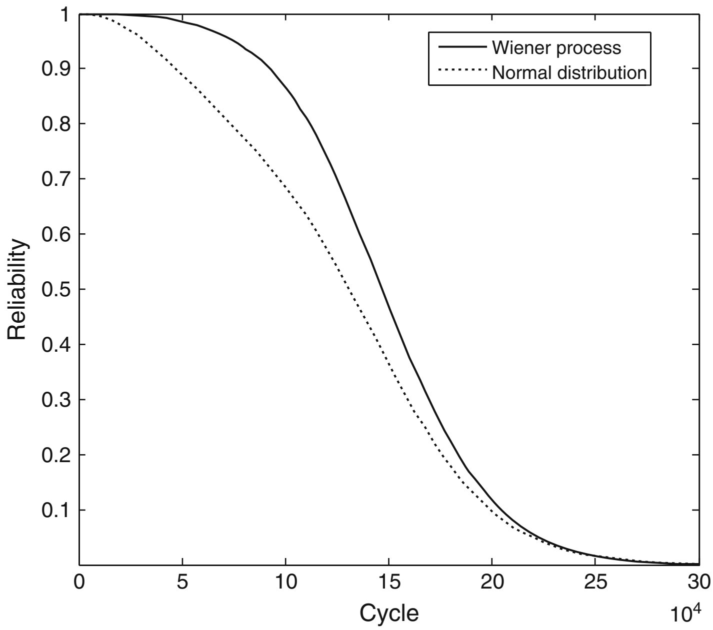


Fig. 51.6 Plots of reliability function under different methods

In order to compare the difference with different methods, the two reliability curves are plot in Fig. 51.6. The upper curve is the reliability under the Wiener process and the lower curve is the reliability under the normal distribution.

From the Fig. 51.6, it can be found that the reliability has a little difference under the Wiener process and normal distribution. Thus, Wiener process is effective for the reliability modeling considering degradation and random shock.

51.5 Conclusion

In this paper, we establish a reliability model for components subject to random shock and degradation process. The random shock can caused an abrupt damage to the degradation process. The Wiener process is used to describe the degradation process, and the maximum likelihood estimation method is used to estimate the parameters of the model.

The case study shows that shocks have obviously impact on the reliability of the component, and the reliability is higher when the impact of shocks is not considered.

References

Bai JM, Li ZH, Kong XB (2011) Generalized shock models based on a cluster point process. IEEE Trans Reliab 55(3):542–550

Birolini A (2007) Reliability engineering theory and practice, 5th edn. Springer, New York/Berlin/Heidelberg

Chun Su, Ye Zhang (2010) System reliability assessment based on Wiener process and competing failure analysis. J Southeast Univ 26(4):405–412

Deloux E, Castanier B, Berenguer C (2009) Predictive maintenance policy for a gradually deteriorating system subject to stress. Reliab Eng Syst Saf 94(2):418–431

Kharoufeh JP, Finkelstein DE, Mixon DG (2006) Availability of periodically inspected systems with Markovian wear and shocks. J Appl Probab 43(2):303–317

Klutke GA, Yang Y (2002) The availability of inspected systems subject to shocks and graceful degradation. IEEE Trans Reliab 51(3):371–374

Lehmann A (2009) Joint modeling of degradation and failure time data. J Stat Plan Inference 139(5):1693–1706

Li WJ, Pham H (2005a) Reliability modeling of multi-state degraded systems with multi-competing failures and random shocks. IEEE Trans Reliab 54(2):297–303

Li WJ, Pham H (2005b) An inspection-maintenance model for systems with multiple competitive processes. IEEE Trans Reliab 54(2):318–327

Lu CJ, Meeker WQ (1993) Using degradation measures to estimate a time-to-failure distribution. Technometrics 35(2):161–174

Lu CJ, Park J, Yang Q (1997) Statistical inference of a time-to-failure distribution derived from linear degradation data. Technometrics 39(4):391–400

Meeker WQ, Doganaksoy N, Hahn GJ (2001) Using degradation data for product reliability analysis. Qual Prog 34(6):60–65

Nakagawa T (2007) Shock and damage models in reliability theory. Springer, London

Robinson ME, Croder MJ (2000) Bayesian method for a growth-curve degradation model with repeated measures. Lifetime Data Anal 6(4):357–374

Tseng ST, Balakrishnan N, Tsai C-C (2009) Optimal step stress accelerated degradation test plan for Gamma degradation process. IEEE Trans Reliab 58(4):611–618

Wang X (2011) Wiener processes with random effects for degradation data. J Multivar Anal 101(2):340–351

Wang ZL, Huang HZ, Du L (2011) Reliability analysis on competitive failure processes under fuzzy degradation data. Appl Soft Comput 11(3):2964–2973

Whitmore GA (1995) Estimating degradation by a Wiener diffusion process subject to measurement error. Lifetime Data Anal 1(3):307–319

Zuo MJ, Jiang RY, Yam RCM (1999) Approaches for reliability modeling of continuous-state devices. IEEE Trans Reliab 48(1):9–18

Chapter 52
A New Constructive Cost Model for Software Testing Project Management

Yun-feng Lu and Yong-feng Yin

Abstract Software testing is becoming a relatively independent activity of software development process. To estimate the cost of software testing is an important part of project management. The researchers have focused on the software cost and treated software testing cost as a part of the total software development cost for several years. Few researchers consider software testing as a separate process to estimate its' cost. This paper proposed a new constructive cost model for software testing project management (CCMST) which contains the main cost divers of software testing. The cost drivers included in this model is more comprehensive than former models. Application case study proved the model is usable and valid.

Keywords Constructive • Cost • Project management • Software testing

52.1 Introduction

There are a variety of software development processes. Traditional software development process is composed of four essential activities: analysis, design, implementation and testing (Pressman 2001). Software crisis has still handicapped the development of software from 1960s. The software projects are always over budget, behind schedule and unreliable (Ming-Shu et al. 2007). In 1995, The Standish Group published the CHAOS REPORT, the research shows a staggering 31.1 % of projects will be cancelled before they ever get completed. Further results indicate 52.7 %

This research is sponsored by Aeronautical Science Foundation of China (No. 20095551025) and the Fundamental Research Funds for the Central Universities (No.YWF-11-03-Q-114).

Y. Lu (✉) • Y. Yin
School of reliability and system engineering, Bei Hang University, Beijing, China
e-mail: lyf@buaa.edu.cn

E. Qi et al. (eds.), *The 19th International Conference on Industrial Engineering and Engineering Management*, DOI 10.1007/978-3-642-37270-4_52,

of projects will cost 189 % of their original estimates (The Standish Group 1995). In the year 2004, only 29 % of software projects in large enterprises succeeded, 53 % were "challenged" (significantly over budget and schedule), and 18 % failed to deliver any usable result (The Standish Group 2004). The cost of software development grows rapidly as the size and complexity of software increasing. Software testing is an important activity of software development. It has been estimated that approximately 40–50 % of the total amount of software development resource is testing-resource (Yamada et al. 1995). With the development of the software industry, software testing becomes a relatively independent process. The software development organization generally is equipped with independent department of software testing. Sometimes, independent third-party software testing is required as a precondition for software acceptance. To estimate the cost of software testing is an important part of project management, and becoming more and more important.

52.2 Present Research of Software Testing Cost Estimation

Since the 1960s, the research of software cost estimation has been developed about 50 years. Researchers have proposed a variety of software cost models (Ming-Shu et al. 2007), among of them the COCOMO model (Boehm 1981, 1984, 2000; Boehm et al. 1995) is widely accepted and used. In these models, software testing is treated as a part of the software development process. The model which regards software testing as a separate process is few. The cost drivers of software testing are clearly different with those of the other activities of software development. Therefore it is not appropriate to directly use software cost model as the cost model for software testing. So it is necessary to construct the cost model for software testing according to the characteristics of software testing. The researchers have proposed some methods to estimate the cost of software testing. But the methods have not pay sufficient attention to the cost drivers of software testing (Kadry 2011; Engel and Last 2007; Silva et al. 2010; Wan and Shu 2008; Yang et al. 2007). There are researchers who learned from the COCOMO II model (Boehm 2000) proposed a cost estimation model of software function test (Song et al. 2010). But this model does not recognize the exponential factors and only defines four usual effort multipliers. Referring to the COCOMO II model and according to the experience of a large number of software testing projects, this paper proposed a more comprehensive constructive cost model for software testing project management named CCMST.

52.3 COCOMO II Model

The COCOMO II model uses sets of multiplicative and exponential cost drivers to adjust for project, target platform, personnel and product characteristics. The set of multiplicative cost drivers are called Effort Multipliers (EM). The nominal weight assigned to each EM is 1.0. If a rating level has a detrimental effect on effort, then its corresponding multiplier is above 1.0. Conversely, if the rating level reduces the effort then the corresponding multiplier is less than1.0. The exponential cost drivers, called Scale Factors and represented by the E exponent, account for the relative economies or diseconomies of scale encountered as a software project increases its size. A constant is used to capture the linear effects on effort with projects of increasing size. The estimated effort for a given size project is expressed in person months (PM), see (52.1) (Boehm 2000).

$$\mathrm{PM} = \mathrm{A} \times \mathrm{Size}^{\mathrm{E}} \times \prod_{i=1}^{n} \mathrm{EM_i} \tag{52.1}$$

E exponent includes five scale factors which are precedentedness, development flexibility, architecture/risk resolution, team cohesion and process maturity. The factors' numerical ratings SF are summed and used to determine a scale exponent. E via (52.2), B = 0.91.

$$\mathrm{E} = \mathrm{B} + 0.01 \times \sum_{j=1}^{5} \mathrm{SF_i} \tag{52.2}$$

There are 17 multiplicative cost drivers used in the COCOMO II Post-Architecture model. The cost drivers are RELY, DATA, CPLX, RUSE, DOCU, TIME, STOR, PVOL, ACAP, PCAP, PCON, APEX, PLEX, LTEX, TOOL, SITE, and SCED. They are grouped into four categories: product, platform, personnel, and project (Boehm 2000).

52.4 Constructive Cost Model for Software Testing

As the COCOMO II, CCMST uses a set of multiplicative and exponential testing cost drivers to adjust for product, project, environment, and personnel characteristics. The set of multiplicative cost drivers are called Testing Effort Multipliers (TEM). The nominal weight assigned to each TEM is 1.0. If a rating level has a detrimental effect on effort, then its corresponding multiplier is above 1.0. Conversely, if the rating level reduces the effort then the corresponding multiplier is less than 1.0. The exponential testing cost drivers, called Scale Factors and represented by the C exponent, account for the relative economies or diseconomies of scale encountered when a software testing project increases its size. D constant is used to capture the linear effects on effort with projects of increasing size. Testing size is expressed in Size. The estimated cost for a given size project is expressed in testing person months (TPM), see (52.3).

$$\mathrm{TPM} = \mathrm{D} \times \mathrm{Size}^{\mathrm{C}} \times \prod_{i=1}^{n} \mathrm{TEM_i} \tag{52.3}$$

Table 52.1 Rating scheme for the CCMST scale factors

Scale factors (TF)	Very low	Low	Nominal	High	Very high
Precedentedness (PREC)	Thoroughly unprecedented	Largely unprecedented	Somewhat unprecedented	Generally familiar	Largely familiar
Architecture/risk resolution (RESL)	Little (20 %)	Some (40 %)	Often (60 %)	Generally (75 %)	Mostly (90 %)
Team cohesion (TEAM)	Very difficult interactions	Some difficult interactions	Basically cooperative interactions	Largely cooperative	Highly cooperative
Process maturity (PMAT)	No specification	Some specification	Basic specification	Enough specification	Complete specification

C exponent includes four scale factors which are precedentedness, architecture /risk resolution, team cohesion, and process maturity. The factors' numerical ratings TF are summed and used to determine a scale exponent. C via (52.4), $H = 0.92$.

$$C = H + 0.01 \times \sum_{j=1}^{4} TF_i \tag{52.4}$$

There are 15 multiplicative cost drivers used in the CCMST model. They are grouped into four categories: product, project, environment, and personnel.

52.4.1 Size Factor

In fact, testing size is the most important testing cost drivers. The lines of code of the software under test are easy to be obtained. For the CCMST model, testing size equals to the lines of code in thousands (KLOC) of the software under test. The lines of code exclude comment lines.

52.4.2 Scale Factors

Each scale factors TF has five rating levels which are respectively: very low, low, nominal, high, and very high. The weights assigned to each rating levels are obtained from the software testing experts' advice. Table 52.1 provides the rating levels for the CCMST scale factors. The definition of precedentedness, architecture /risk resolution and team cohesion is basically consistent with the definition in the COCOMO II model (Boehm 2000). The difference is that all of the scale factors are for the software testing project.

Table 52.2 Categories of testing effort multipliers

Product	TCON-Testing Content
	DOCU-Documentation Match to Testing Needs
Project	SDOC-Software Documents' Quality
	SWCP-Software Complexity
	SICP-Software Interface Complexity
	SSTB-Software Stability
	SCED-Required Testing Schedule
	SITE-Multisite Testing
Environment	TOOL-Efficiency of Testing Tools
	COOP-the Cooperation of Software Developers
Personnel	TCAP-Software Testing Engineer Capability
	TEXP-Testing Experience
	LEXP-Language Experience
	PEXP-Platform Experience
	PCON-Personnel Continuity

For precedentedness, if the software under test and testing requirements are similar to the precedent, the rating will be higher. For architecture/risk resolution, if there are a series of evaluation activities in the process of testing, and appropriate resolution are adopted to eliminate risks, RESL will be higher. For team cohesion, if the testing engineers, project managers, coordinators and other projects stakeholders can cooperate with each other as a team to carry out software testing work, TEAM will be higher. For process maturity, if the software testing organization sets up a complete test system and test specifications, testing can be carried out under the test system and specification, PMAT will be higher.

52.4.3 Testing Effort Multipliers

As indicated in Table 52.2, the 15 testing Effort Multipliers are grouped into four categories. Each TEM has five rating levels which are very low, low, nominal, high, and very high. Each rating level has a weight. Table 52.3 provides the TEM rating scales. The following subsections discuss the treatment of these testing effort multiplier cost drivers.

1. *TCON-Testing content*: Testing Content is a very important testing effort Multipliers. A complete software testing routine contains unit testing, integrated testing, validation testing, system testing, including software document and code inspection. If the testing content is reduced, TCON will reduce, on the contrary increase.
2. *DOCU-Documentation Match to Testing Needs:* The rating scale for the DOCU cost driver is evaluated in terms of the suitability of the project's test documentation. The rating scale goes from Very low (many needs uncovered) to Very high (very excessive for needs).

Table 52.3 Testing effort multiplier cost driver ratings for the CCMST

TEM	Very low	Low	Nominal	High	Very high
TCON	System testing	Integrated testing Validation testing System testing Including software document inspection and code inspection,	Unit testing Integrated testing Validation testing System testing Including software document inspection and code inspection,	Excessive testing content	Very excessive testing content
DOCU	Many needs uncovered	Some needs uncovered	Right-sized to needs	Excessive for needs	Very excessive for needs
SDOC	No essential document or very not detailed	Not enough or not detailed	Enough and detailed	Excessive for needs and very detailed	Very excessive for needs and very detailed
SWCP	Very simple control, computational, device, data and user interface management operations	Simple control, computational, device, data and user interface management operations	Nominal control, computational, device, data and user interface management operations	Complex control, computational, device, data and user interface management operations	Very complex control, computational, device, data and user interface management operations
SICP	Very little interface and data	A little interface and data	Nominal interface and data	Lots of interface and data	Mass interface and data
SSTB	Lots of SRS modification, software is not stable	Some SRS modification, software is not very stable	Little SRS modification, software is stable	Very little SRS modification, software is stable	No SRS modification, software is very stable

SCED	60 %	80 %	100 %	130 %	150 %
SITE collocation	International	Multi-city or multi-company	Same city or metro, area, company	Same building or complex	Fully collocated
SITE communications	Some phone, mail	Individual phone, fax	Email or other electronic communication	Wideband electronic communication, occasional video conference.	Interaction multimedia
TOOL	Very simple and very inefficient	Simple and inefficient, little integration	Essential, moderately integrated	Mature tools, efficient, moderately integrated	Mature tools, very efficient, well integrated
COOP	Very difficult interactions	Some difficult interactions	Basically cooperative interactions	Largely cooperative	Highly cooperative
TCAP	15th percentile	35th percentile	55th percentile	75th percentile	90th percentile
TEXP	<1 project	1 project	2 projects	4 projects	6 projects
PEXP	≤2 months	6 months	1 year	3 years	6 years
LEXP	≤2 months	6 months	1 year	2 years	3 years
PCON	50 %/year	30 %/year	20 %	10 %	0 %

3. *SDOC-Software documents' quality:* The software documentation is the basis of software testing. Full and detailed documentation is an important factor which influences the test efficiency. The cost driver stresses the importance of software documentation. If the rating scale is Very low or Low, it will be almost impossible to perform test or the sufficiency of software testing cannot be guaranteed. In this situation, software developers should be required to improve the documentation, in order to improve the sufficiency and efficiency of testing.
4. *SWCP-Software complexity:* Testing effort has a positive relationship with the software complexity. The software complexity is related with control operations, computational operations, device-dependent operations, data management operations and user interface management operations. All these factors should be considered when selecting the rating level of *SWCP*.
5. *SICP-Software interface complexity:* In fact, software interface complexity essentially reflects the software complexity. It is separated from the SWCP mainly on account of that the interface complexity has a more tremendous impact than other software complexity factors. When selecting the rating level, the number of software interface and the amount of data should be considered.
6. *SSTB-Software stability*: During the test, if the software requirements specifications (SRS) changes, it may cause the software code to be changed and increase the testing effort. If the software is running unstably, it indicated that there are more bugs. Testing and regression testing effort will increase. When selecting the rating level of SSTB, it should be noticed that the changes of SRS is more important.
7. *SCED-Required testing schedule:* The SCED is the schedule constraint imposed on the test project. The rating level is determined according to the percentage relative to the nominal schedule.
8. *SITE-Multisite testing:* The frequency of multisite developments and third-party testing are increasing, so multisite testing become more and more frequent. Determining its cost driver rating involves the assessment and averaging of two factors: site collocation (from fully collocated to international distribution) and communication support (from surface mail and some phone access to full interactive multimedia).
9. *TOOL-Efficiency of Testing Tools:* The efficiency of testing tools directly affects the efficiency of the software testing. This cost driver considers maturity, integration and efficiency of test tools. The testing tools mentioned here include test software and test equipment.
10. *COOP-the cooperation of software developers:* With the good cooperation of software developers, the efficiency of software testing can be improved. Because testers need developers to confirm bugs, modify bugs and jointly solve the problems emerged in the process of testing in time.
11. *TCAP-Software testing engineer capability:* Personnel capability has strong influence on software testing productivity. When determining the rating level, the attention should be paid to the tester's capability as a team member rather

than as an individual. Ability, efficiency, communication and collaboration capabilities should be considered. TCAP does not consider the experience of the tester which classified by TEXP, LEXP and PEXP.

12. *TEXP-Testing experience:* The cost driver depends on the test experience of a test team. The more similar projects the team has completed, the higher level. Very low level means the team has done less than one similar project. Very high level means the team has done more than six similar projects.
13. *PEXP-Platform experience:* The rating scale for the PEXP cost driver is evaluated in terms of the team's experience of using required test tools and equipment.
14. *LEXP-Language experience:* The LEXP cost driver depends on the team's experience of the program language of the software under test. Very low level means the team has less than 2 months' experience. Very high level means the team has more than 6 years' experience. If the testing content does not include unit testing, code inspection and the other White-Box testing content, then the LEXP will not affect the testing effort and the LEXP level should be nominal.
15. *PCON-Personnel continuity:* The rating scale for PCON is in terms of the project's annual personnel turnover: from 0 to 50 %.

52.5 Application Case Study

The CCMST could be seen as a framework. For specified organizations, (52.3), (52.4), Table 52.1 and Table 52.3 should be calibrated according to their own historical project data. For Table 52.1 and Table 52.3, each rating level corresponds to a value. If a rating level isn't a standard level of the tables, the corresponding value can be obtained by interpolation method. We choose one software evaluation center to verify the usability and validity of the model.

52.5.1 Determine the Parameters of CCMST

Firstly we pick 66 completed software testing projects and the relevant data. Secondly we determine the parameters of CCMST using the least square method and the linear regression. The constant $H = 0.92$, $D = 0.38$. Tables 52.4 and 52.5 provide the parameters of the scale factors and the testing effort multipliers.

52.5.2 Estimate the Effort of the Projects

In order to prove the usability and validity of the model, we choose three projects and the scale of the projects' size is different. Firstly we confirm the rating level of the scale factors and the testing effort multipliers, and then get the corresponding

Table 52.4 The parameters of the scale factors

Scale factors (TF)	Very low	Low	Nominal	High	Very high
Precedentedness (PREC)	4.52	3.02	2.11	1.80	1.28
Architecture/risk resolution (RESL)	4.21	2.89	2.17	1.43	1.12
Team cohesion (TEAM)	4.28	3.34	2.73	1.22	1.03
Process maturity (PMAT)	4.30	3.27	2.86	2.33	1.92

Table 52.5 The parameters of the testing effort multipliers

TEM	Very low	Low	Nominal	High	Very high
TCON	0.37	0.61	1.00	1.13	1.18
DOCU	0.83	0.90	1.00	1.08	1.12
SDOC	1.72	1.37	1.00	0.91	0.88
SWCP	0.76	0.84	1.00	1.19	1.43
SICP	0.78	0.86	1.00	1.21	1.52
SSTB	2.37	1.62	1.00	0.94	0.86
SCED	1.48	1.18	1.00	1.00	1.00
SITE	1.25	1.11	1.00	0.91	0.82
TOOL	1.38	1.17	1.00	0.86	0.73
COOP	1.22	1.08	1.00	0.93	0.91
TCAP	1.37	1.21	1.00	0.85	0.74
TEXP	1.18	1.09	1.00	0.87	0.71
PEXP	1.15	1.08	1.00	0.96	0.91
LEXP	1.18	1.10	1.00	0.94	0.90
PCON	1.28	1.12	1.00	0.93	0.87

value. Secondly we estimate the effort (TPM) according to (52.3) and (52.4). After the projects are completed, we calculate the actual effort (TPM_ACT) and the deviation. DEVIATION = 100 * (TPM-TPM_ACT)/TPM_ACT%. The relevant data are listed in Table 52.6.

All the absolute deviations are less than 20 % which are 14.03, 6.40 and 9.49 %. The deviations are within the bounds of reason. The deviation of project 1 is 14.03 % which is relatively larger. After analyzing the test process, we find the main reason. When the project started, the documents of the software was of poor quality and the software ran unstably, so the level of SDOC and SSTB was low. But the software developers improved the software and the documents as quickly as possible, so the two factors are bigger than the reality during the main stages of the project.

52.6 Conclusion

With the development of the software industry, software testing becomes a relatively independent activity of software development process. The researchers have not pay sufficient attention to the cost drivers of software testing. This paper proposed a new constructive cost model for software testing project management, which

Table 52.6 The results of the application case

Cost drivers and results	Project 1	Project 2	Project 3
SIZE	8.2	43.0	78.8
PREC	3.02	2.11	1.80
RESL	2.89	2.17	2.17
TEAM	2.73	1.22	1.22
PMAT	2.86	3.27	2.33
TCON	1.00	0.61	0.61
DOCU	0.95	1.00	1.00
SDOC	1.50	1.37	0.91
SWCP	1.00	1.43	1.19
SICP	1.00	1.21	1.21
SSTB	1.70	1.00	1.00
SCED	1.00	1.18	1.3
SITE	0.82	0.82	0.82
TOOL	1.00	1.00	1.00
COOP	1.08	1.00	1.00
TCAP	0.85	0.85	0.94
TEXP	1.15	1.09	0.71
PEXP	0.94	1.08	0.91
LEXP	0.90	0.90	0.90
PCON	0.87	0.87	0.93
TPM	4.93	18.44	12.70
TPM_ACT	4.54	19.70	11.60
DEVIATION	14.03 %	−6.40 %	9.49 %

contains the main cost divers of software testing. The model also provided the rating levels. The CCMST is a more comprehensive cost model for software testing than former models. The application case study proves that the model is usable and valid. Currently available historical project data for calibrating CCMST is still not sufficient, in order to further refine and calibrate the model, an extensive data collection and analysis effort is necessary.

References

Boehm B (1981) Software engineering economics. Prentice Hall, Englewood Cliffs

Boehm B (1984) Software engineering economics. IEEE Trans Softw Eng SE-10(1):200–217

Boehm B (2000) Software cost estimation with COCOMO II. Prentice Hall, Englewood Cliffs

Boehm B, Clark B, Horowitz E, Westland C (1995) Cost models for future software life cycle processes: COCOMO 2.0. Ann Softw Eng 1(1):57–94

Engel A, Last M (2007) Modeling software testing costs and risks using fuzzy logic paradigm. J Syst Softw 80:817–835

Kadry S (2011) A new proposed technique to improve software regression testing cost. Int J Secur Appl 5:45–58

Ming-Shu L, Mei H, Da Y, Feng-Di S, Qing W (2007) Software cost estimation method and application. J Softw 18(4):775–795 (in Chinese)

Pressman RS (2001) Software engineering: a practitioner's approach, 5th edn. The MCGRAW-Hill Companies, Inc, New York, pp 18–20

Silva DGE, Jino M, De Abreu BT (2010) Machine learning methods and asymmetric cost function to estimate execution effort of software testing. In: 3rd international conference on software testing, verification and validation, ICST 2010, Paris, 7 Apr 2010–9 Apr 2010, pp 275–284

Song W, Yang GX, Cai LZ (2010) A cost estimation model of software function test and its application. Comput Appl Softw 27(6):32–34 (in Chinese)

The Standish Group (1995) CHAOS report, Boston

The Standish Group (2004) 2004 the 3rd quarter research report, Boston

Wan Y, Shu F (2008) Estimation model for software testing effort based on value. Comput Eng Des 29:544–546 (in Chinese)

Yamada S, Ichimori T, Nishiwaki M (1995) Optimal allocation policies for testing-resource based on a software reliability growth model. Math Comput Modell 22:295–301

Yang S, Kan H, Yu B (2007) A model of cost estimation and control for software test automation. Acta Electronica Sinica 35:1588–1591 (in Chinese)

Chapter 53
Correction Coefficient of Catalytic Combustion Gas Detector

Lu Pan, Yan-yan Wang, Zhi-hong Sun, Da-hu Zhang, and Lin Jiang

Abstract Catalytic combustion gas detectors are widely applied in air detection of dangerous places, such as petrochemistry industries and coal mines. Correction coefficient is a significant fact to using correctly. In this article, the author discussed the reasons why the correction coefficients are similar but different, explained the principle of selecting correction air, analyzed the influence of correction coefficients on measurement, summarized the problems in using and finally put forward some suggestions.

Keywords Air detection • Combustible and explosive • Correction coefficient

53.1 Introduction

The catalytic combustion gas detector is precision, stable, sensitive and inexpensive so is widely used in safety testing. As it is showed as the volume ratio of lower explosive limit (% LED), so it is called LEL detector. However, this detector has its limitations. It is not analysis instrument, unable to distinguish gas species, so needs correction coefficient to revise when used to measure different kinds of gases (Feng Wei 2002).

L. Pan (✉) • Y. Wang • Z. Sun
Department of foundation, Air Service School, Xuzhou, China
e-mail: zhuanglangxuzhou@yahoo.cn

D. Zhang
Department of Aviation Oil supplies, Air Service School, Xuzhou, China

L. Jiang
Department of Training, Air Service School, Xuzhou, China

E. Qi et al. (eds.), *The 19th International Conference on Industrial Engineering and Engineering Management*, DOI 10.1007/978-3-642-37270-4_53,

53.2 Methodology

In this article, theoretical analysis and practical application are used to analyze the difference, rationality and error of correction coefficient synthetically. Graphs and tables are used to contrast the measurement results of different gases.

53.3 Results

53.3.1 *Definition of Correction Coefficient*

The correction coefficient is the quotient of actual volume ratio (%LEL) of the test gas to readings in the detector (which is based on correction gases). The larger the correction coefficient, the more obviously the reading deviates from the actual concentration.

Usually the manufacturers will take into account the correction coefficient. Some manufacturers use shift design on the dial and display the actual concentration of different gases in detector directly. Some only give the correction coefficient, and users need to calculate by oneself after measurement. The readings multiplied by the correction factor are the actual concentration (%LEL).

Different gases react in the Wheatstone bridge with the different catalytic combustion heat, so the correction coefficients vary. Before use, you should ask manufacturers for the correction coefficient, avoid "giving a false alarm when safe" or "lack of alarm when danger" (Liu chang xue 2011).

53.3.2 *The Similarity of the Correction Coefficient*

The same catalytic combustion gas detector can measure a variety of combustible gas roughly, and the results are similar, the correction coefficients are similarity. It is closely related to the detection principle of catalytic combustion heat and the similarity of ultimate combustion heat of different combustible gases.

1. Catalytic combustion detectors determine the concentration by detecting the combustion heat of combustible gas.

Catalytic combustion gas detectors detect neither the volume concentration, nor volume ratio of the lower explosion limit (% LEL), but the combustion heat the gas release in the Wheatstone bridge.

Although the lower explosive limit (LEL) of different gases vary, such as LEL of methane is 5.3 % and LEL of gasoline 1.4 %, but the combustion heat that the gases released at 100 %LEL concentration (ultimate combustion heat, QLEL) is basically the same. Except for a few gases, the ultimate combustion heat of combustible gas is about 41.9–54.4 kJ (Liu xue 2011).

Table 53.1 Combustion heat of common combustible gas

Combustible gas	Molecular formula	Actual concentration in LEL (%)	Heat of combustion (KJ/mol)	Q_{1mol} (KJ)
Methane	CH4	5.0	889.5	44.3
Ethane	C2H6	3.0	1,558.3	46.8
Propane	C3H8	2.1	2,217.9	46.4
Isobutene	i-C4H10	1.8	2,866.2	51.4
Pentane	C5H12	1.4	3,524.6	49.3
Hexane	C6H14	1.2	4,159.1	49.7
Heptanes	C7H16	1.1	4,812.4	53.1

Q_{1mol} – combustion heat which 1 mol combustible gas (22.4 L) release at lower explosive limit

Table 53.1 show that the ultimate combustion heat of ethane is similar to propane. Based on ultimate combustion heat, if combustion heat up to about 46.4 kJ, we can consider approximately that concentration of the test gas reach the lower explosive limit (100 %LEL),and we would not need to know the type of gas or specific concentration of this test gas at the lower explosion limit.

2. Linearity

Study found that for most gas in the scope of the lower explosion limit, the combustion heat grows with the concentration in a linear relationship. The higher the concentration (%LEL) is, the more the heat is.

For example, the combustion heat at 50 %LEL is half of the ultimate combustion heat at lower explosive limit. If we know the ratio of combustion heat to the ultimate combustion heat of the test gas, we can get the concentration of the gas. So catalytic combustion gas detector set ultimate combustion heat as the standard, and determine the concentration of the test gas by comparing combustion heat (Zhang Kai Lan and Xie huai biao 2011).

3. Broad spectrum

Set ultimate combustion heat of the gas which is just test as the standard, this method is accurate, but as a result one detector can only measure one gas, detectors cannot used widely.

Since the ultimate combustion heat of different gases is basically identical, we can choose a representative gas (such as methane) as correction gas, compare combustion heat of test gas with ultimate combustion heat of correction gas, we can get the concentration of the test gasapproximately. Therefore, in theory, with one detector we can measure different gases roughly. It shows a strong broad-spectrum character.

4. Reading unity

For different gas, although the volume concentration is not alike, but as long as the ultimate combustion heat is alike, lower explosive limit percentage concentration measured by catalytic combustion principle is basically the same.

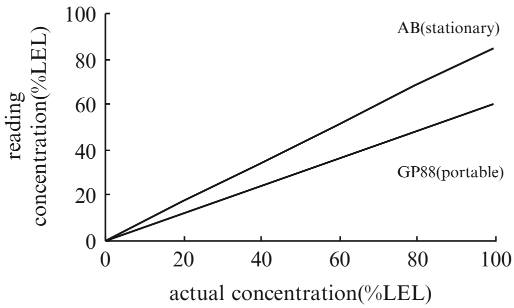


Fig. 53.1 Calibration curve of gasoline (Calibrated by isobutene)

Therefore, with lower explosive limit percentage concentration (%LEL) as a measurement unit, gases of disparate volume concentration get a bridge of reading unity.

53.3.3 The Difference of the Correction Coefficient

In practical applications, for certain gas the combustion heat of flameless combustion in the Wheatstone bridge after catalyst at a lower temperature differ from that of complete combustion, it even show significant difference between various gases. So, when the calibration gas differs from the test gas, we must use correction coefficient to revise. The reasons are as follows:

1. Although the ultimate combustion heat (QLEL) between different gases is similar, but it still exist differences (Table 53.1), change from about 41.9–54.4 kJ.
2. The same gas at the same concentration show different readings on different detectors (Hong junxiong and Zhu zhaoyou 2011).

Because the catalyst materials and production process differ between different manufacturers, catalytic temperature will be different, therefore, for the same gas at the same concentration, the combustion heat on different types of sensor bridge is not only less than that of complete combustion, but also differ from one another. The phenomenon that combustion heat deviates from the complete combustion heat is caused by the incomplete catalysis and incomplete combustion.

Figure 53.1 shows the calibration curve of gasoline on different types of gas detector which is calibrated by isobutene and produced by Riken Corporation in Japan. Although with the same calibration gas and the same test gas, the calibration curves are different. It shows that the same concentration gasoline release different catalytic combustion heat in the two sensors, correction coefficients of gasoline in the two sensors are different, fixed sensor is bigger than portable sensor.

3. The same concentration of multiple gases showed different readings in the same detector.

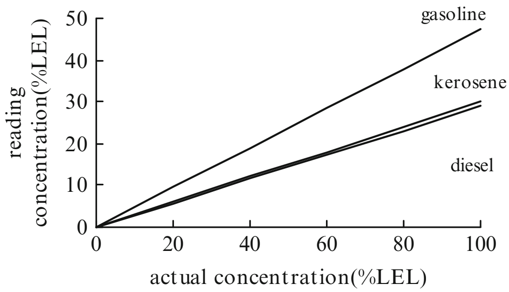


Fig. 53.2 Correction curve of different fuel oils (Calibrated by methane)

The flash point and volatile of different gases are not alike, so the ratio of reaching the Wheatstone bridge and catalytic combustion are different. Methane is the most active, and most likely to reach the surface of the bridge. But it is more difficult for gas with high flash point, such as flash point of gasoline is −40 °C, kerosene is 38 °C and diesel is 45 °C, to spread through the fire protection metal mesh of LEL sensor, so it is less for reaching the bridge, and come into being catalytic combustion, so the heat of catalytic combustion is also less (Zhou Juli and Zheng Daoming 2011).

Figure 53.2 shows the calibration curves of different fuel oils. The same concentration of gasoline, diesel and kerosene are measured with the same detector of RAR Company (calibrated by methane). Correction coefficient of gasoline is bigger than kerosene, and kerosene is bigger than diesel.

So catalytic combustion heats in the electric bridge vary between different gases and it leads to very distinct correction factor. It is related to gas species, bridge and producing process of sensor.

53.3.4 *Rationality of Correction Coefficient*

The reasonable correction coefficient rests with the application of detectors and the choice of calibration gases (Zhen Jie 2011). The application scope is determined before leaving the factory, therefore, choose of calibration gas is significant

$$\begin{aligned}&\text{correction coefficient}\\&=\frac{\text{actual concentration of test gas}}{\text{reading concentration of test gas}}\\&=\frac{\text{ultimate catalytic combustion of correction gas}}{\text{ultimate catalytic combustion of test gas}}\end{aligned} \tag{53.1}$$

For same detector, it is certain for ultimate catalytic combustion heat of one test gas. The correction coefficient will differ along with the correction gas.

Table 53.2 Relative correction coefficient of flammable and explosive sensor

Type of combustible gas	Calibrated by methane	Calibrated by propane	Calibrated by pentane
Hydrogen	0.91	0.59	0.45
Methane	1.00	0.67	0.50
Propane	1.54	1.00	0.77
n-butane	1.67	1.11	0.83
n-pentane	2.00	1.33	1.00
n-hexane	2.20	1.33	1.00
Octane	2.25	1.67	1.25
Unleaded gasoline	1.67	1.11	0.83

Equation (53.1) shows that the more ultimate combustion heat of correction gas, the more correction coefficient, and the more meter reading concentration deviate from actual concentration of test gas. So selecting the gas whose ultimate catalytic combustion heat is close to most gases as the correction gas is good for reducing correction coefficient (close to one) and make readings close to actual concentration (Wang Bing Bing and Liu Xiao Bo 2010).

At first the catalytic combustion detector is designed for detecting methane. Now methane detection still exists in coal mine and petrochemical enterprises abundantly. Therefore, a lot of detectors use methane as the calibration gas. However, methane is toxic, flammable and explosive, it cause inconvenience to the calibration work. In addition, ultimate combustion heat of methane is bigger than most combustible gases, so the readings are small. Therefore the alarms do no alarm, and hide safety problems.

By analysis (Table 53.2), for the same test gas, the readings in the detector calibrated by methane will be lower than actual value, and will be higher calibrated by pentane and will be relatively close calibrated by propane. Therefore, for many cases without methane, propane is a more suitable calibration gas (Wang Haiying 2012).

Of course, isobutene is also a common calibration gas. Before using, it is recommended to ask manufacturers the type of calibration gas and the correction coefficient concerned.

53.4 Discussion

53.4.1 Error Analysis of Measurement

In actual use, without using or misusing the correction coefficient result in the deviation of the readings and a lot of danger, examples are followed.

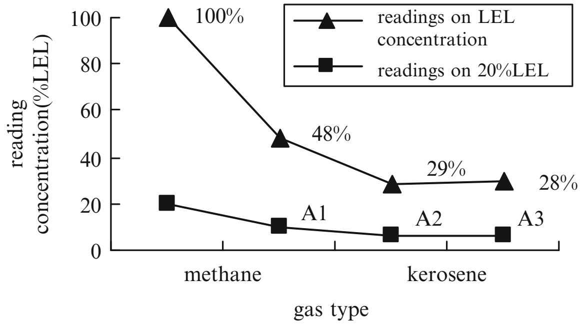


Fig. 53.3 Correction scale of fuel oil (Calibrated by methane)

Table 53.3 Correction coefficients of LEL and TC/LEL sensor of RAE company (Calibrated by methane)

Gas name	Concentration in LEL (%)	Correction coefficient	Reading (%LEL)
Methane	5.0	1	100
Gasoline	1.3	2.1	48
Jet fuel (JP-4,JP-5,JP-8)	0.7	3.4	29
Diesel	0.8	3.3	30

Example 53.1 measure gasoline, kerosene, diesel with RAE LEL detector (calibrated by methane), and compare the differences before and after using correction coefficient (Fig. 53.3).

The detector is calibrated by methane, and the correction coefficients are all greater than 1, so the readings are smaller than actual concentration (Xiao Lin Ni 2012).

Before using the correction coefficient, the reading of 20 % alarm lines in detector is far more smaller than the actual concentration of the test gas, and respectively it correspond to 42 %LEL of the gasoline (A1 in Fig. 53.3), 68 %LEL of kerosene (A2 in Fig. 53.3), and 66 %LEL of the diesel (A3 in Fig. 53.3), and so that will lead to serious consequences of “not giving alarm in danger”.

After using the correction coefficient, the reading of 9.6 % of full scale corresponds to 20 %LEL of the gasoline, ought to alarm. Similarly, 6 and 5.8 % correspond to 20 %LEL of the gasoline and the Kerosene respectively. Details are list in Table 53.2.

The detector is calibrated by methane, and the readings are smaller than actual concentration (Lu Yue Jun 2012). So the readings without correction coefficient will be paralysis of the user. Particularly, pay attention to set alarm limit less than 20 %LEL (Table 53.3), and avoid accidents.

Example 53.2 measurement analysis of the mixed gas of unknown proportion.

Table 53.4 Meter reading and real value of gas, kerosene and diesel in fuel and air mixture

Gas name	Reading (%LEL)	Actual concentration (%LEL)	Pass alarm line
Gasoline	12	25.2	Yes
Kerosene	2	6.8	No
Diesel	2	6.6	No
Mixed gas	16	/	No

There are pipelines to various oil tanks in pumping stations, so there may be unknown proportion gas mixture of gasoline, kerosene and diesel (Shi Yin Lin and Fu Yan Feng 2011). Measuring the gas mixture in the pumping station by LEL detector of RAE Company calibrated by methane and without using correction coefficient. Gas compositions and proportions are shown in Table 53.4.

The gas mixture contained gasoline gas which concentration is greater than 20 % LEL, and the detector ought to alarm. But the detector was calibrated by methane and without using correction coefficient, so the meter reading was only about 16 % and the detector would not alarm even in danger (Zhong Hong Li and Han Wei 2012).

Of each mixture, there will be a kind of the most hazardous compound to control the alarm point of the whole mixture. If we determined the gas which lower explosive limit by detector reading is lowest among the gas mixture (which correction coefficient is the biggest or of which sensitivity is the smallest), we can set a relative accurate alarm point of all the compounds. In this case, of detector reading the alarm line of kerosene is the lowest, so we should Take kerosene as standard, and set 5.8 % of full-scale as alarm line of the gas mixture (Zhao Ai Jun 2012).

53.4.2 Error Analysis of Use and Maintenance

The correction coefficient is not static. The catalytic combustion gas detection is calibrated relatively, and the correction coefficient is likely changed because of improper use.

1. Detector exposed to the environment of sulfur or lead, acidic gases or high concentrations (close to or exceed the lower explosive limit) for a long time will lead to catalyst poisoning, and sensitivity decrease and the combustion heat going down.
2. Pump tube blockage, leakage etc. will make the test gas reduce and reading small (Ma Ju Mei and Yang Wen 2011).

Above two cases will lead to the correction coefficient increase, and bringing the risk of “not alarm when danger”. Simultaneously, the user cannot find the problem by calibration using calibration gas. Even if found, we cannot repair, and need to carry the detector to the manufacturer for inspection and repair regularly. Usually manufacturers provide that LEL detector need an overhaul at 6 months, reach term of service at 3 years. When it is about to the time limit, we should pay attention to examine and repair particularly.

53.5 Conclusion

LEL detector is the clairvoyance and electronic nose in kinds of dangerous places. We should only use of the correction coefficient properly; the detector can really play the role of the protection.

References

Feng Wei (2002) A brief exposition of correction combustible gas detector by isobutene. Safe 2:6–7

Hong junxiong, Zhu zhaoyou (2011) Design for combustible gas detection and alarm device. Hunan Agric Mach 7:49–50

Liu chang xue (2011) Discussion on the application of poisonous and harmful gas detector. Sci Technol Enterp 12:183

Liu xue (2011) First exploration of reconnaissance installation and maintenance of inflammable gas alarm. Jilin Lab Prot 10:38–39

Lu Yue Jun (2012) Uncertainty evaluation of indication error of measurement by combustible gas detector alarm. Technol Rich Wizard 5:313–314

Ma Ju Mei, Yang Wen (2011) Study on the LEL sensor poisoning of combustible gas alarm. Technol Wind 19:178

Shi Yin Lin, Fu Yan Feng (2011) Principle and application of combustible gas alarm. Brand Stand 24:48

Wang Bing Bing, Liu Xiao Bo (2010) CO gas concentration detection technology in industrial process. Ind Saf Environ Prot 36:61–63

Wang Haiying (2012) Research on flammable gas alarm circuit. Trop Agric Eng 36:31–32

Xiao Lin Ni (2012) Instructional design of combustible gas sensors. Teach Appar Exp 2:13–14

Zhang Kai Lan, Xie huai biao (2011) Inquisition of calculation decimation for combustible gas alarm. China New Technol Prod 19:149

Zhao Ai Jun (2012) Test of combustible gas detection and alarm device. China New Technol Prod 2:137

Zhen Jie (2011) Algorithm research of mixing hazardous gas detection. Comput Simul 28:413–416

Zhong Hong Li, Han Wei (2012) Uncertainty assessment on indication by combustible gas detection and alarm. China Metrol 1:87–88

Zhou Juli, Zheng Daoming (2011) Detection and prevention of harmful gases in underground engineering. Sichuan Hydropower Gener 30:17–20

Chapter 54
Identification of Risk Factors for Casualty in Building Fires

Chun-hua Yu, Li-ping Tian, and Zeng-chang Zhao

Abstract In this paper, risk factors of casualty in building fires which can be used in risk assessment have been identified by two methods. One is event tree analysis (ETA) method; the other is fault tree analysis (FTA) method. Combined with Fire Statistical Yearbook of China and literatures, the reasons of casualty in building fires were analyzed with 109 cases happened in China from 2001 to 2010. Six types of risk factors have been seek out from a list of basic matters which had been produced through the cases used the two methods above, and these factors will be adapt to the modern complex building constructions. The results will have important implications for fire prevention planning and fire management in the future.

Keywords Building fires • Casualty • Event tree analysis • Fault tree analysis • Risk identification

54.1 Introduction

In recent decades, urban population, the city area and urbanization level increase very quickly in China. Many large commercial buildings, high-rise buildings and underground buildings come forth. These constructions are characterized by large space, dense crowd and heavy fire load which can give off amount of poisonous smoke when burning. Crowded people could be injured or killed simultaneously

C. Yu (✉) • L. Tian
School of Management, Hebei University of Technology, Tianjin, China
e-mail: yuchunhua2007@163.com

Z. Zhao
Langfang Fire Detachment, Langfang, China

E. Qi et al. (eds.), *The 19th International Conference on Industrial Engineering and Engineering Management*, DOI 10.1007/978-3-642-37270-4_54,

once fire happens. Therefore, reasonable fire safety design is very important to ensure life safety. In China, during the past years, fire safety design and management must comply with prescriptive codes, such as National Standards for Fire-Safety Design of Buildings (GBJ16-87) (Fire Safety Bureau, Ministry of Public Security 1993; Ministry of Construction, Ministry of Public Security 1999). However, GBJ16-87 is a prescriptive code which is not suitable to the fire safety design of those new-style, complex and special buildings. Thus, performance-based fire protection design is employed for fire safety design in buildings. China has carried out some work on performance-based fire protection design and developed it for fire safety design practice. In general, fire risk assessment is the foundation of performance-based fire protection design, and appropriate performance-based fire protection design cannot be well implemented without reasonable fire risk assessment. In this paper, six types of factors have been identified to apply to the fire risk assessment, they are factors related to the building status, factors related to fire protection system, factors related to evacuation system, factors related to personal traits, factors related to fire safety management and factors related to external rescue groups.

54.2 Methodology

Some scholars have made preliminary research for fire risk factors and factor categories (Brennan and Thomas 2001; Mavis Duncanson et al. 2002; Holborn et al. 2003). The main purpose of these scholars is to demonstrate the power of statistical analysis. There are a large range of models that simulate the evacuation of a building, from simple models that simulate only the movement of people within the building, to very complex models that attempt to incorporate human behavior (Kuligowski 2003; Olenick and Carpenter 2003). The main content of this paper is to research risk factors of casualty in building fires using the safety analysis method of event tree and fault tree.

54.2.1 Basic Principles of Event Tree Analysis

Event tree analysis (ETA) for risk analysis was first used in the atomic energy field and gradually extended to other fields (Beim and Hobbs 1997). ETA represents a logic combination of various events that may follow from an initiating event. The initiating event of the tree uses dichotomous situations, i.e., success/failure (true/false or yes/no) to convey the event consequence in different branches of the tree (Lees 2005). Each individual path that is followed by the different branches finally identifies the possible outcome events by developing an event-consequence model. In risk analysis, the event-consequence model and the outcome events are successively used in pre-incident application, to test the incident precursors and post-incident application, and to identify the possible outcome events for an accidental event (AIChE 2000). The ET approach was used in order to compute the risk to which occupants of a building may be subjected if a fire breaks out in that building (Frantzich 1998; Magnusson et al. 1996).

54.2.2 Basic Principles of Fault Tree Analysis

Fault tree analysis (FTA) is developed using standard logic formulations. A brief analysis is also conducted to determine the occurrence probability of the TOP event. A fault tree demonstrates the causal relationship among basic failures (basic events) that contribute to a predetermined system failure (top event). To complete the construction of a fault tree manually for a complicated system, it is necessary to first understand how the process works. Subsequently, the causal relationship among the basic events and top event is expressed using a logic symbol such as AND-gate or OR-gate. The first step in fault tree construction is the selection of a top event. The user must then determine the logical dependency between a basic event and the top event. Fault trees can be constructed for individual units or the whole unit of a process plant. However, in computer-aided fault tree synthesis the whole tree is decomposed into small individual units, which are connected to lower level elements by a logical fault tree diagram. The accuracy of fault tree analysis depends on the available information on the logical dependency of the subsystems, and the failure data of the basic events. FTA has been applied successfully in the field of fire safety (Daniela Hanea and Ben Ale 2009).

54.2.3 Causal Process Analysis of Casualty in Building Fires

Casualties (injuries and fatalities) occur relatively infrequently in fires that are notified to fire brigades. Also, most fires result in very few injuries and casualties and moreover the great majority of casualties occur in these fires. Injuries and fatalities that occur as a consequence of fires in buildings may be viewed from several perspectives. Most injuries and fatalities are the result of relatively frequent occurrences-particular types of fires and situations that usually result in no casualties but occasionally in one or two injuries or fatalities. Thus, the study adopts qualitative methods to research risk factors. The paper chooses 109 fire cases happened in China from 2001 to 2010 according to Fire Statistical Yearbook of China (Fire Safety Bureau, Ministry of Public Security 2011) and relevant literature (Geng Huimin 2011) at first. Secondly, an event tree and three son-fault trees are built up according to buildings fire disaster's occurrence and development rule. The rule of the model is using failure events of ETA as top events of FTA respectively. Risk factors are from four models —the event tree of casualty in building fires, the fault tree of failure in discovering fire early, the fault tree of failure in self-rescue, the fault tree of failure in emergency-rescue by firemen.

54.3 Results

54.3.1 *Event Tree Analysis of Casualty in Building Fires*

Usually, when the fire breaks out, the trapped persons should make sure the fire occur at early stage and then escape. If they fail to discover fire at early stage, they should try to rescue by themselves. If they fail to rescue themselves, they have to rely on external rescue groups.

If external rescue fails, they might die in fire. Discovering fire at early stage, self-rescue and external emergency-rescue are three lines of defense preventing casualty from building fires. Previous process is shown in Fig. 54.1.

54.3.2 *Fault Tree Analysis of Failure in Discovering Fire at Early Stage*

People need to get information outside to confirm fire has broken out. Personal reasons and external reasons result in failing to discover fire at early stage. Personal reasons are physiological defect, drunken sleeping and lacking of alert, etc. External reasons are faults of security guards, alarm facilities problems and unusual fire point, etc. Previous process is shown in Fig. 54.2.

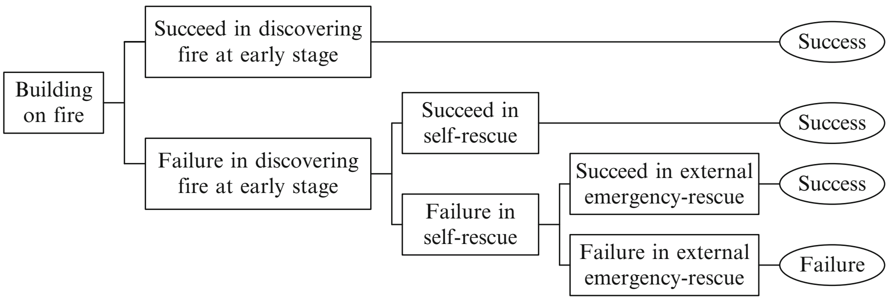


Fig. 54.1 Event tree of casualty in building fires

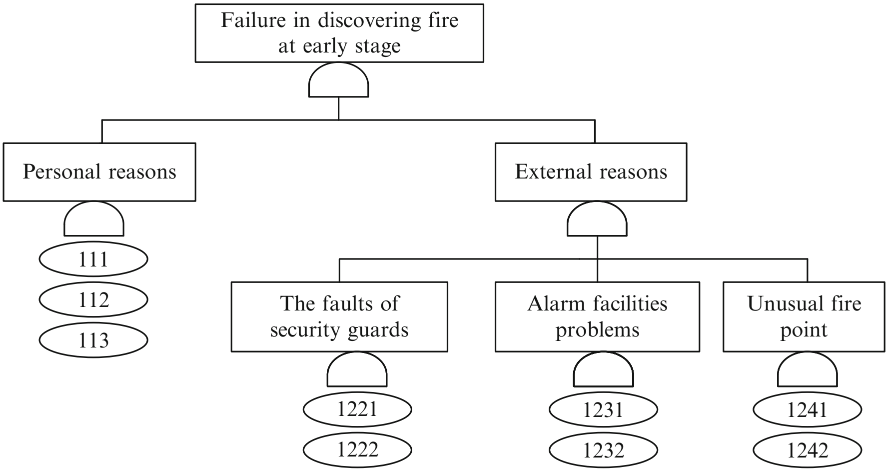


Fig. 54.2 Fault tree of failure in discovering fire at early stage

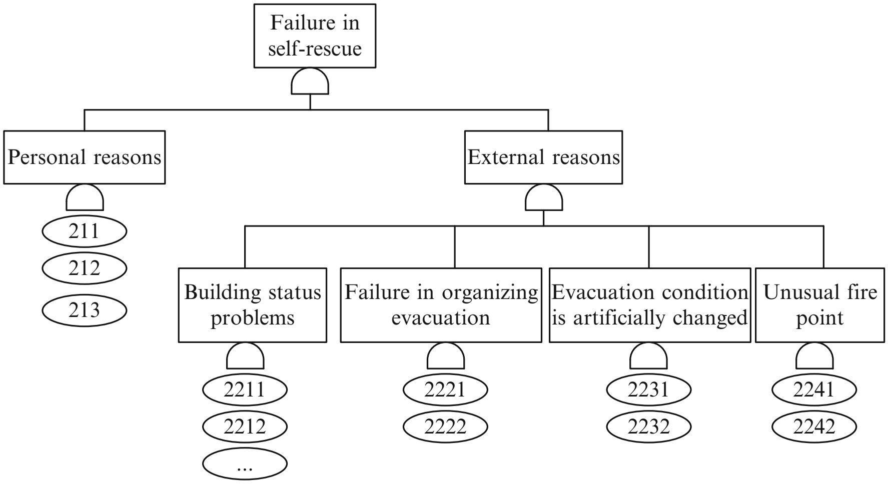


Fig. 54.3 Fault tree of failure in self-rescue

54.3.3 Fault Tree Analysis of Failure in Self-rescue

Trapped people fail to rescue themselves attributes to personal and external reasons. Personal reasons are failure in looking for escape route, failure in taking security measures and returning to fire for property, etc. External reasons are building status problems, failure in organizing evacuation, evacuation condition being artificially changed and unusual fire origin, etc. Process of failure in self-rescue is shown in Fig. 54.3.

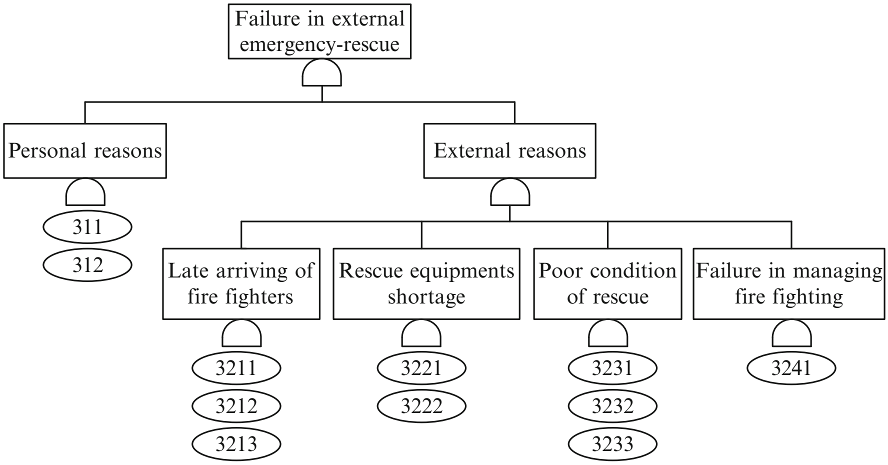


Fig. 54.4 Fault tree of failure in external emergency-rescue

54.3.4 Fault Tree Analysis of Failure in External Emergency-Rescue

Trapped people fail to rescue themselves by external emergency attributes to personal and external reasons, too. Personal reasons of trapped people are failure in sending SOS signals for help and frangibility of mind, etc. External reasons are late arriving of firefighters, rescue equipments shortage, poor condition of rescue and failure in managing fire fighting. Process of failure in emergency-rescue by firemen is shown in Fig. 54.4.

54.3.5 Basic Matters of Casualty in Building Fires

Three fault tree analysis of 109 fire cases demonstrates the causal relationship among basic failures (basic events) that contribute to a list of basic matter of casualty in building fires, as shown in Table 54.1.

54.4 Conclusion

The main evaluation index of casualty in building fires are building status, fire protection system, evacuation system, personal traits, fire safety management and external rescue groups (Ying Huang 2009). Basic matters of Table 54.1 are divided into six aspects. The results are shown in Fig. 54.5.

Table 54.1 List of basic matters of casualty in building fires

Code	Basic matter
111	Drunken sleeping
112	Physiological defect
113	Lacking of alert
1221	Security guards fail to discover fire in time
1222	Security guards fail to send warning signal in time
1231	No alarm facilities
1232	Malfunction of alarm facilities
1241	Fire point is not easy to discover
1242	Noisy environment
211	Failure in looking for escape route
212	Failure in taking security measures
213	Returning to fire for property
2211	Fire resistance rating don’t meet standards
2212	Fire load is very large
2213	Evacuation passageways don’t meet national standards
2214	Safety exits don’t meet standards
2215	Smoke facilities are below standards
2216	Fire-fighting facilities are below standards
2217	Unreasonable building structure
2221	Delay of evacuation time
2222	Wrong way of evacuation
2231	Block of evacuation passage
2232	Closed doors and windows
2241	Escape path is on fire
2242	Fire exit is on fire
311	Failure in sending SOS signals for help
312	Frangibility of mind
3211	Alarm delay
3212	The building is far from the fire brigade
3213	Malfunction of vehicles
3221	Lack of rescue equipments
3222	Rescue equipments are in poor quality
3231	Obstacles block the rescue space
3232	Narrow rescue space
3233	Lack of extinguishers
3241	Failure in managing fire fighting

The results of this paper may have important implications for fire prevention planning and fire management. Analysis on real case shows that the method of event tree analysis and fault tree analysis could explain each factor causing casualty in the building fire logically, and provide feasible referable criteria for preventing the building fire and reduction of casualty to the minimum. Within these risk factors, people and fire departments should focus on events that are statistically likely to have higher risk of casualty in building fires.

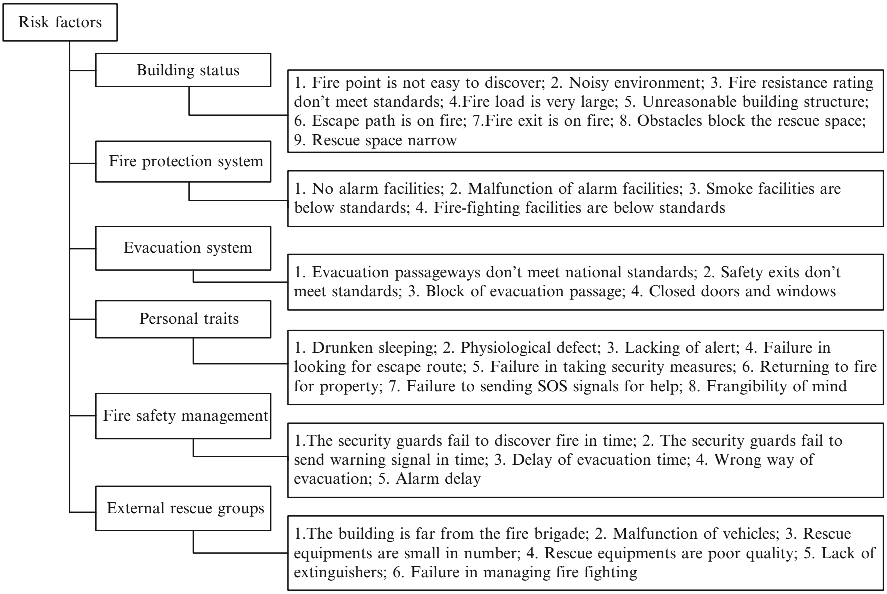


Fig. 54.5 Risk factors index system of casualty in building fires

References

AIChE (2000) Guidelines for chemical process quantitative risk analysis, 2nd edn. AIChE, New York

Beim GK, Hobbs BF (1997) Event tree analysis of lock closure risks. J Water Resour Plan Manag ASCE 123:137–198

Brennan P, Thomas IR (2001) Injuries and fatalities in fires. In: A continuum fifth Asia-Oceania symposium on fire science and technology, Newcastle, Australia, pp 351–361

Daniela Hanea, Ben Ale (2009) Risk of human fatality in building fires: a decision tool using Bayesian networks. Fire Saf J 44(5):704–710

Fire Safety Bureau, Ministry of Public Security (1993) Compilation of standards and technology of fire safety design. Planning Press, Beijing (in Chinese)

Fire Safety Bureau, Ministry of Public Security (2011) Fire statistical yearbook of China (2001–2010). China personnel press, Beijing

Frantzich H (1998) Risk analysis and fire safety engineering. Fire Saf J 31(4):313–329

Geng Huimin (2011) Typical case analysis for fire casualty accident. Tianjin Science and Technology Translation Publishing Company, Tianjin (in Chinese)

Holborn PJ, Nolan PF, Golt J (2003) An analysis of fatal unintentional dwelling fires investigated by London fire brigade between 1996 and 2000. Fire Saf J 38(1):1–42

Kuligowski E (2003) The evaluation of performance-based design process for hotel building: the comparison of two egress models. PhD dissertation, Department of Fire Protection Engineering, University of Maryland, College Park, Washington DC, America, pp 37–111

Lees FP (2005) Loss prevention in the process industries, 3rd edn. Butterworths Press, London, pp 9/05–9/122

Magnusson SE, Frantzich H, Harada K (1996) Fire safety design based on calculations: uncertainty analysis and safety verification. Fire Saf J 27(4):305–334

Mavis Duncanson, Alistair Woodward, Papaarangi Reid (2002) Socioeconomic deprivation and fatal unintentional domestic fire incidents in New Zealand 1993–1998. Fire Saf J 37(2):165–179

Ministry of Construction, Ministry of Public Security (1999) Building construction fire safety engineering assessment clause no. 30. Amendment to GBJ16-87. Planning Press, Beijing (in Chinese)

Olenick SM, Carpenter DJ (2003) An updated international survey of computer models for fire and smoke. J Fire Prot Eng 13:87–110

Ying Huang (2009) Research on fire risk assessment and safety management methods of public building. PhD dissertation, Xi'an University of Architecture and Technology, Xian, China, pp 36–39 (in Chinese)

Chapter 55
Data-Driven Based Key Performance Index Residual Generation and Its Application on Complex Electrical Equipment

Zhi-gang Yao, Li Cheng, and Yu-lei Wang

Abstract Motivated by the increasing needs for key performance index related fault detection in complex electrical equipments, this paper proposes the subspace aided data-driven robust fault detection technique. The main idea is to use the original test data to identify the residual generators firstly, and then make use of performance indices to design of robust residuals which are robustness to non-quality variables and sensitivity to quality variables. Robust and robust reduced order residual generations are proposed, and finally the proposed methods are certified by application on complex electrical equipment.

Keywords Data-driven • Fault detection • Key performance index • Residual generator

55.1 Introduction and Problem Formulation

Complex electrical equipments deteriorate with time consuming and usage, naturally require frequent service and maintenance. However, efficient fault detection and isolation (FDI) in complex electrical equipments which contain large amount

Z. Yao (✉)
Department of Optical and Electronic, Shijiazhuang Mechanical Engineering College, Shijiazhuang, China
e-mail: xyhk_yzg@sina.com

L. Cheng
Department of Mathematic, Lishui University, Lishui, China
e-mail: chenglilily@sina.com

Y. Wang
Institute for Automatic Control and Complex Systems, University of Duisburg-Essen, Duisburg, Germany
e-mail: yulei.wang@uni-due.de

E. Qi et al. (eds.), *The 19th International Conference on Industrial Engineering and Engineering Management*, DOI 10.1007/978-3-642-37270-4_55,

sensors is of utmost important. All these sensors play a significant role in collecting information and monitoring system operation, and are considered the least reliable components of the complex electrical equipments. If any fault is not detected immediately, there would have catastrophic consequences.

During the recent few decades theoretical and experimental research has shown a large number of creative methods in literature to detect and diagnose several sorts of faults. The precondition of FDI is residual generation, which is of special interest for a large number of theory researchers and engineering application (Roth et al. 2011; Frank and Ding 1997; Naik et al. 2010). Model based methods which based on a well-established mathematical model are available for the design of FDI systems (Chen and Patton 1999; Ding 2008; Gertler 1998; Ye et al. 2002; Ding et al. 2011; Tan and Edwards 2002; Isermann 2011; Qin 2006). Alternatively, subspace identification methods (SIM) are an attractive tool, which obtain a state space model directly from test data (Qin 2006; Ding et al. 2009; Wang et al. 2011). As a result, there are only three steps from original data to the implementation of FDI. Recently, reference (Ding et al. 2009) proposed the subspace method aided data-driven design of fault detection and isolation system, reference (Wang et al. 2011) proposed a data-driven based approach to design a robust FDI system. The above-mentioned two approaches make the design of FDI system become more convenient and reduce the on-line computation.

The relationship between performance consistency and equipment maintaining has become a widespread concern due to the worldwide increasing market competition and system safety and reliability, and to design a reliable fault detection system is also a challenge task in practice. The key performance indicator (KPI) related fault detection cover a wide range of requirements form complex electrical equipment application. And it's as early as possible to detect any sensor faults to ensure system operation normally. The engineers may pay much attention to the relationship between the measurable variables and final outputs, some of which are the so-called KPI in this contribution. So from the application of view, we shall develop an efficient data-driven method to design a robust KPI fault detection system and improve the FDI performance by the construction of residuals. Followed by the preliminaries, Sect. 55.3 presents the data-driven construction of two types of robust residual generators, the effectiveness of algorithms on application is verified in Sect. 55.4.

55.2 Preliminaries Related Robust Residual Generation

In this paper, we suppose that test data sets including system input and output records are available, and we consider KPI sensor fault detection problem for complex equipments described by

$$x(k+1) = Ax(k) + B_1(y_1(k) + f_1(k)) + B_2(y_2(k) + f_2(k)) \tag{55.1}$$

$$\theta(k) = Cx(k) + D_1(y_1(k) + f_1(k)) + D_2(y_2(k) + f_2(k)) \tag{55.2}$$

where $x(k) \in R^n$ denotes the state vector, $y_1(k) \in \mathrm{R}^{l_1}$ and $y_2(k) \in \mathrm{R}^{l_2}$ denotes the KPI-correlated and non-KPI-correlated output vectors, respectively. $\theta(k) \in R^m$ is the KPI vector, $f_1(k) \in \mathrm{R}^{l_1}$ the relevant fault vectors to be detected and $f_2(k) \in \mathrm{R}^{l_2}$ is described as the disturbances against KPI-FDI results. Suppose that A, B_1, B_2, C, D_1, D_2 are unknown and constant matrices. Let us use the following data structures:

$$X(i) = [x(i), x(i+1), \cdots, x(i+N-1)] \in R^{n \times N}$$

Similarly, $\Theta(i) \in R^{m \times N}$, $Y_1(i) \in R^{l_1 \times N}$, $Y_2(i) \in R^{l_2 \times N}$, $F_1(i) \in R^{l_1 \times N}$, $F_2(i) \in R^{l2 \times N}$.

Further denote, $Y_{sp}^1, Y_{sp}^2, Y_{sf}^1, Y_{sf}^2, \Theta_{sp}, \Theta_{sf}, F_{sf}^1, F_{sf}^2$, and $Z_{sp} = [\Theta_{sp}^T, \mathrm{Y}_{sp}{}^{1,\mathrm{T}}, \mathrm{Y}_{sp}^{2,\mathrm{T}}]^{\mathrm{T}} \in R^{sp(l+m) \times N}$, and $Z_{sf} = [\Theta_{sf}^T, \mathrm{Y}_{sf}^{1,\mathrm{T}}, \mathrm{Y}_{sf}^{2,\mathrm{T}}]^{\mathrm{T}}$.

Then (55.1), and (55.2) can be brought into

$$\Theta_s = \Gamma_s X(i) + H_{s,1} Y_s^1 + H_{s,2} Y_s^2 + H_{s,1} F_s^1 + H_{s,2} F_s^2 \tag{55.3}$$

$$\Gamma_s = \begin{bmatrix} C \\ \vdots \\ CA^{s-1} \end{bmatrix} \quad H_{s,1} = \begin{vmatrix} D_1 & 0 & \cdots & 0 \\ CB_1 & D_1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ CA^{s-2}B_1 & \cdots & CB_1 & D_1 \end{vmatrix}$$

$$H_{s,2} = H_{s,1|B_1 = B_2,\ D_1 = D_2}$$

where the subscript N denotes the number of test data, the subscripts p, f denote, respectively, the "past" and the "future" data, and s_p, s_f stand for some integers and are assumed for simplicity with the same length s and thus uniform the symbol subscript s.

The design of a parity space based residual generator consists of in solving the equation

$$\alpha_s \Gamma_s = 0 \tag{55.4}$$

where $\alpha_s = [\alpha_{s,0}, \cdots, \alpha_{s,s}]$ is the so-called parity vector delivered from its parity space $\Gamma_s^{\perp}$. It follows then the construction of the residual generators as

$$\begin{aligned} r(k) &= \alpha_s(\theta_s(k) - H_{s,1} y_s^1(k) - H_{s,2} y_s^2(k)) \\ &= \alpha_s(H_{s,1} f_s^1(k) + H_{s,2} f_s^2(k)) \end{aligned} \tag{55.5}$$

where $y_s^1(k) = [\mathrm{y}_1^{\mathrm{T}}(\mathrm{k-s}), \cdots, \mathrm{y}_1^{\mathrm{T}}(\mathrm{k})]^{\mathrm{T}} \in R^{sl_1}$, $y_s^2(k) \in R^{sl_2}$, $\theta_s(k) \in R^{sm}$, $f_{a,sf} \in R^{sfl}$ and $f_{s,sf} \in R^{sfm}$.

In order to achieve a robust FDI, the residual signals should be further designed not only to be sensitive to fault signal f_1, but also keep robust to f_2 simultaneously. Referring Ding et al. (1999), we creatively denote the following performance index

$$J = \min_{\alpha_s \in \Gamma_s^{\perp}} \frac{\alpha_s H_{s,2} H_{s,2}^T \alpha_s^T}{\alpha_s H_{s,1} H_{s,1}^T \alpha_s^T} \tag{55.6}$$

In our contribution, the formula (55.6) is the basis of the robust residual generation technique. As a result, (55.6) will be widely used in the design of our robust fault detection system.

The solution of optimization problem (55.6) is equivalent to the solution of the generalized eigenvalue-eigenvector problem

$$l_s(\Gamma_{\mathrm{s}}^{\perp}\mathrm{H}_{\mathrm{s},2}\mathrm{H}_{\mathrm{s},2}^{\mathrm{T}}\Gamma_{\mathrm{s}}^{\perp,\mathrm{T}} - \lambda_{\mathrm{s}}\Gamma_s^{\perp} H_{s,1} H_{s,1}^T \Gamma_s^{\perp,T}) = 0 \tag{55.7}$$

The minimum eigenvalue of λ_s is the optimal value of J and the corresponding eigenvector of l_s is for building the optimal parity vector. Furthermore, different decisions on λ_s and l_s depend on the application of the designed robust fault detection and are our key study in this paper.

55.3 Data-Driven Construction of Essential Types of Robust Residual Generators

For the construction of parity space based residual generators, α_s, $\alpha_s H_{s,1}$ and $\alpha_s H_{s,2}$ are necessary. We also note that the parity vectors belong to the left null space of Γ_s, denoted by $\Gamma_s^{\perp}$. So, the identification of $\Gamma_s^{\perp}$, $\Gamma_s^{\perp} H_{s,1}$ and $\Gamma_s^{\perp} H_{s,2}$ would allow a direct construction of a parity space based residual generator. Based on this reason, we have designed one type of residual generator for sensors, which consists of two KPI sub-generators for robust fault detection (KPI-RFD) and robust reduced order fault detection (KPI-RrFD) in the next two subsections.

55.3.1 Design of Robust Residual Generators

In this subsection, we will propose the KPI-RFD problem in form of an algorithm.

Algorithm 55.1 (KPI-RFD)

- Generate data sets Z_{sf}, Z_{sp} and construct $\frac{1}{N} Z_{sf} Z_{sp}^T$
- Do a SVD on $\frac{1}{N} Z_{sf} Z_{sp}^T$, and then we have $\frac{1}{N} Z_{sf} Z_{sp}^T = U_z diag([\sigma_{\mathrm{z},1}, \sigma_{\mathrm{z},2}])\mathrm{V}_{\mathrm{z}}^{\mathrm{T}}$ where $U_z = [\mathrm{U}_{\mathrm{z},1}, \mathrm{U}_{\mathrm{z},2}] \in \mathrm{R}^{\eta_z s \times \eta_z s}$, $U_{z,2} = [\mathrm{U}_{\mathrm{z},21}^{\mathrm{T}}, \mathrm{U}_{\mathrm{z},22}^{\mathrm{T}}, \mathrm{U}_{\mathrm{z},23}^{\mathrm{T}}]^T$, $U_{z,21}^T \in R^{\eta \times m s_f}$, $U_{z,22}^T \in R^{\eta \times l_1 s_f}$, $U_{z,23}^T \in R^{\eta \times l_2 s_f}$, $\sigma_{z,2} = 0 \in R^{\eta \times \eta}$, $\eta_z = m + l_1 + l_2$, $\eta = ms - n$.

- Set $\Gamma_s^{\perp} = U_{z,21}^T$, $\Gamma_s^{\perp} H_{s,1} = -U_{z,22}^T$, and $\Gamma_s^{\perp} H_{s,2} = -U_{z,23}^T$
- Compute minimum eigenvalues and their eigenvectors $\lambda_{s,\min}$ and $l_{s,\min}$ according to (55.7)
- Set a threshold $\beta > 0$ such that $\lambda_{s,\min} \leq \beta$, otherwise enhances the parity order and go to the algorithm again.

Note that robust parity vector for KPI faults is of the form

$$\alpha_s = l_{s,\min}\Gamma_s^{\perp}, \alpha_s H_{s,1} = l_{s,\min}\Gamma_s^{\perp} H_{s,1}, \alpha_s H_{s,2} = l_{s,\min}\Gamma_s^{\perp} H_{s,2} \tag{55.8}$$

and its residual generator is equivalent to (55.5).

Set the freedom of robust residual $\eta = ms - n > 0$, which means one of existence conditions is the integer $s > n/m$. As demonstrated in (Wang and Qin 2006), to ensure that $\Gamma_s^{\perp}$ and $\Gamma_s^{\perp} H_{s,u}$ are available, Z_{sf} and Z_{sp} should be selected such that $rank([X^T(i), Y_{sf}^T]^T Z_{sp}^T = n + s(l_1 + l_2)$. Furthermore, Z_{sp} is to eliminate the impact of the system noises, it is well-known that SVD has been widely accepted as the foundation of subspace theory and rank condition describes an input excitation condition.

55.3.2 Design of Robust Reduced Order Residual Generators

From Algorithm 55.1 we know s is generally larger than n. However, from the viewpoint of practice, aiming at reducing the online computation, it is better to find a suboptimal but acceptable robust parity vector which has a lower order residual generator, whose idea can be explained by supposing that $\alpha_s \in \Gamma_s^{\perp}$ is a parity vector and satisfies $\alpha_s = [\alpha_{s_\mathrm{d}}, 0, \cdots, 0]$ with $\alpha_{s_d} \in R^{(s_d+1)m}$ and $s_d + 1 < s$. Then we have

$$\alpha_s H_{s,1} = [\alpha_{s_d} H_{s_d,1}, 0, \cdots, 0],\ \alpha_s H_{s,2} = [\alpha_{s_d} H_{s_d,2}, 0, \cdots, 0] \tag{55.9}$$

As a result, we are able to construct a residual generator as

$$r_d(k) = \alpha_{s_d}\left(\theta_{s_d}(k) - H_{s_d,1} y_{1,s_d}^1(k) - H_{s_d,2} y_{s_d}^2(k)\right) \tag{55.10}$$

whose order s_d is smaller than s.

Based on this purpose, using Algorithm 55.1 we identify $\Gamma_s^{\perp}$, $\Gamma_s^{\perp} H_{s,1}$ and $\Gamma_s^{\perp} H_{s,2}$ first. Then do a QR decomposition of $\Gamma_s^{\perp} V$ with V the left-right reversal of $I_{ms\times ms}$, which leads to

$$\Gamma_s^{\perp} V = [Q_1 R_1, Q_2]$$

where $R_1 \in R^{\eta\times\eta}$ an upper triangular and $Q_1 \in R^{\eta\times\eta}$ an orthonormal columns. Left multiplying matrix Q_1^{-1} to $\Gamma_s^{\perp}$, we have

$$Q_1^{-1}\Gamma_s^{\perp} = [R_1, Q_1^{-1} Q_2] V^{-1} = [Q_3, \Psi]$$

with $Q_3 \in R^{\eta\times n}$ and an upper triangular $\Psi \in R^{\eta\times\eta}$ as

$$\Psi = \begin{bmatrix} \Psi_{1,1} & \cdots & \Psi_{1,\eta-1} & \Psi_{1,\mu} \\ \vdots & \iddots & \Psi_{2,\eta\text{-}1} & 0 \\ \Psi_{\eta\text{-}1,1} & \iddots & \iddots & \vdots \\ \Psi_{\eta,1} & 0 & \cdots & 0 \end{bmatrix}$$

Combined with Q_3, the order of parity vectors $Q_1^{-1}\Gamma_{s_f}^{\perp}$ has been rearranged from bottom to up.

At last, we propose Algorithm 55.2 to design a robust parity vector into the form of $\alpha_s = [\alpha_{s_d}, 0, \cdots, 0]$ for the purpose of the robust order reduction.

Algorithm 55.2 (KPI-RrFD)

- Do a QR decomposition of $\Gamma_s^{\perp} V = [\mathrm{Q_1R_1, Q_2}]$
- Compute $Q_1^{-1}\Gamma_s^{\perp} = [\mathrm{R_1, Q_1^{-1}Q_2]V^{-1}} = [Q_3, \Psi]$
- Select backward i rows with $i \in [1, \eta]$ in Q_3 and Ψ together to construct a new parity space denoted by $\Gamma_{s_d}^{\perp} \in R^{i\times ms_f}$, $s_d = n + i$ and the corresponding $\Gamma_{s_d}^{\perp} H_{s,1}$, $\Gamma_{s_d}^{\perp} H_{s,2}$
- Substitute them into (55.7) and solve the minimum eigenvalues $\lambda_{s_d,\min}$ and their eigenvectors $l_{s_d,\min}$
- Check that $\lambda_{s_d,\min} \le \beta$ otherwise update the parity order i.

Note that robust isolation parity vectors for KPI faults are of form

$$\alpha_{iso,j} = l_{iso,j,\min} \bar{P}_{s,j} \Gamma_{s_f}^{\perp} \tag{55.11}$$

The dynamics of the parity space residual generators given in (55.5) is governed by

$$\begin{aligned} r_{iso}(k) &= [r_{iso,1}, \cdots, r_{iso,l_1}] \\ &= [\alpha_{\mathrm{iso},1}, \cdots, \alpha_{\mathrm{iso},l_1}](\theta_{\mathrm{s}}(\mathrm{k})\text{-}\mathrm{H_{s,1}y_s^1(k)}\text{-}\mathrm{H_{s,2}y_s^2(k)}) \end{aligned}$$

55.4 Application on Complex Electrical Equipment

55.4.1 Description on VMMC

Vehicle-Mounted Manipulator Control System (VMMC) is well equipped with a series of measurement sensors, which plays an important role in insuring the complex system operating in a normal condition. The VMMC is an extremely complex equipment to achieve certain particular tasks under normal conditions in special circumstance. The schematic layout of VMMC is illustrated by Fig. 55.1, which consists of several computers and sensors. What we are most concerned

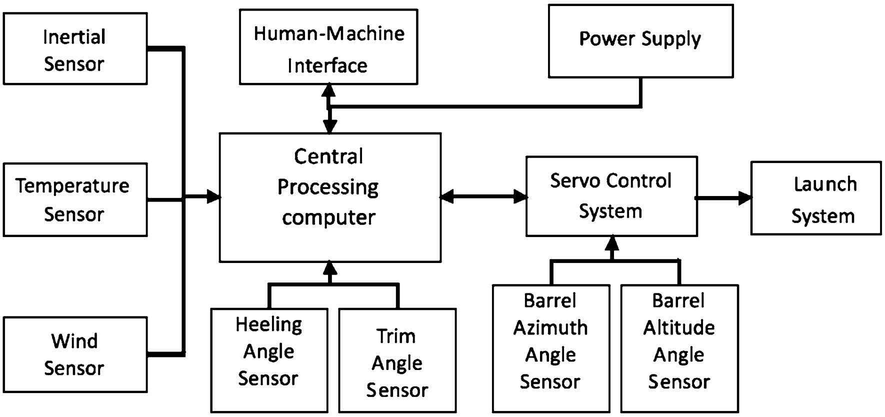


Fig. 55.1 Schematic layout of the vehicle-mounted manipulator control system

about are the desired output location and angle of the mechanical arm, which are the KPIs in this paper. The output performance affected by sensors is too complex to modeling but can be ordered by the impact to KPIs.

The greatest impact on the output is the angle of system delivery platform which generate the data from two sensors: heeling angle device and trim angle equipment.

55.4.2 Robust Fault Detection

Setting the value of s as 9 an N as 1,000, the output and KPI output block hankel matrices are built as $Y^1_{s_f}, Y^1_{s_p} \in R^{20\times 1000}$, $Y^2_{s_f}, Y^2_{s_p} \in R^{60\times 1000}$ and $\Theta^2_{s_f}, \Theta^2_{s_p} \in R^{20\times 1000}$. Parity subspace $\Gamma_s^{\perp} \in R^{10\times 20}$ and its corresponding matrices $\Gamma_s^{\perp} H_{s,1} \in R^{10\times 20}$ and $\Gamma_s^{\perp} H_{s,2} \in R^{10\times 60}$ are identified as shown in (55.8) and each robust parity vector $\alpha_s \in R^{1\times 20}$ is solved by the Algorithm 55.1 and Algorithm 55.2 so that the final form of the parity relation residual generators for KPI index can be realized.

It is well-known that the residual signal must be evaluated to reduce the influence of noises and improve the sensitivity to sensor faults. The GLR based evaluation method is widely used for threshold computation and decision making. The threshold is selected as the false alarm rate of 1 %. The benchmark here is simulated with two type tests including two different fault cases: the ramp change and the sine wave are all analyzed and testified the proposed two algorithms, respectively.

The first type of fault is ramp type change with a slow rate 0.002, in heeling angle sensor, which occurs after the 100 sample of normal operation. The fault-free case in shown in Fig. 55.2 and the threshold for residual evaluation is shown by a dash line. As seen in Fig. 55.3, the RFD residual successfully detect the ramp fault. Next we make use of the RrFD residual to detect the sine wave fault occurred in trim angle sensor. Assume it occurs after the 100 sample as well. As we show in the Figs. 55.4 and 55.5, both figures testify the availability of KPI-RrFD residual in this case.

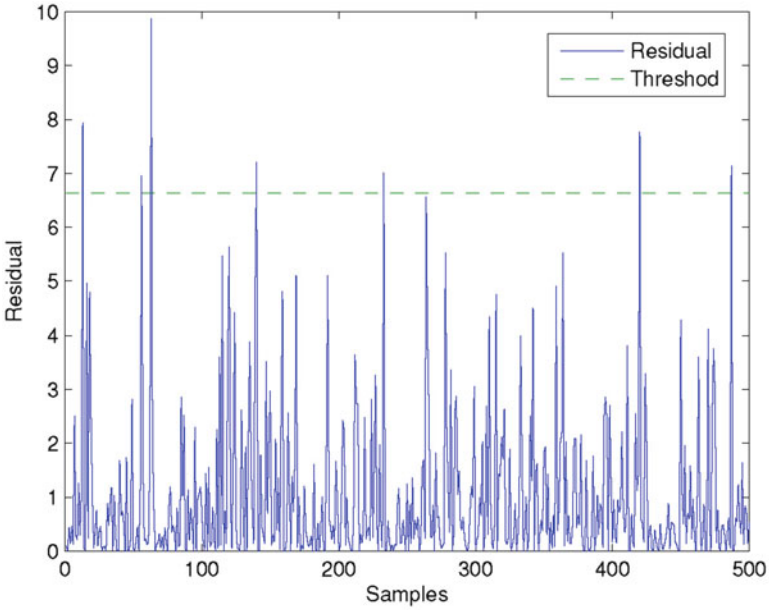


Fig. 55.2 KPI-RFD in normal case

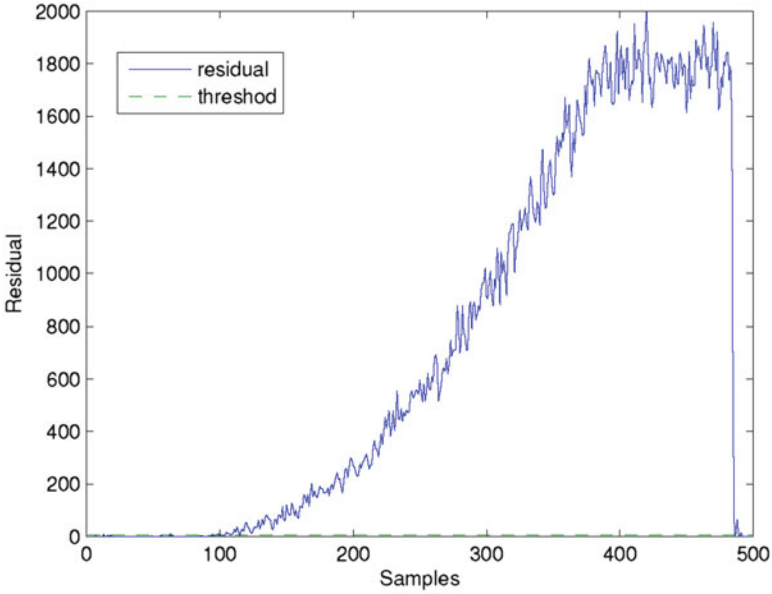


Fig. 55.3 KPI-RFD in fault case

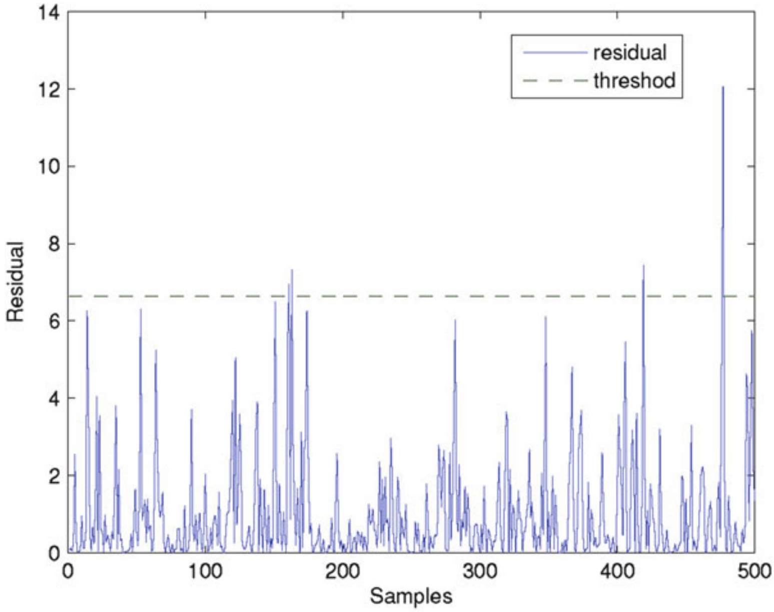


Fig. 55.4 KPI-RrFD in normal case

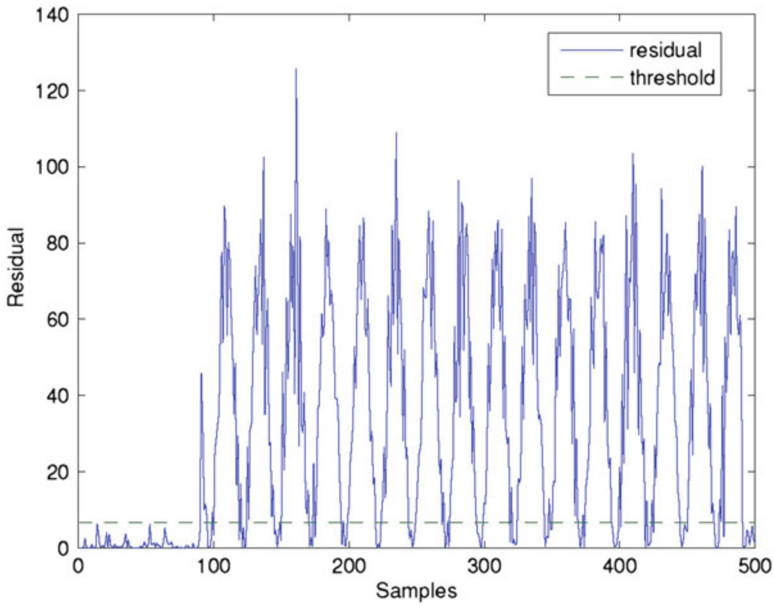


Fig. 55.5 KPI-RrFD in fault case

55.5 Conclusion

In this contribution, we present a new approach to design a robust fault detection system, which based on data-driven quality related parity space and residual generation methods, also without using KPI models and parameters. The core idea is to utilize performance indices to a direct and uniform design of robust FDI systems. Aiming at detection to sensor faults two algorithms for residual generators are presented. Our developed approach has been successfully applied to the VMMC system and justified the effectiveness.

Acknowledgment This work was supported by national natural science foundation of China (NSFC), No. 60905051, and the department of education, Zhejiang Province, No. Y201120498.

References

Chen J, Patton RJ (1999) Robust model-based fault diagnosis for dynamic systems. Kluwer Academic Publishers, Boston

Ding SX (2008) Model-based fault diagnosis techniques design schemes algorithms and tools. Springer, Berlin

Ding SX, Guo L, Jeinsch T (1999) A characterization of parity space and its application to robust fault detection. IEEE Trans Autom Contr 44:337–343

Ding SX, Zhang P, Naik A, Ding E, Huang B (2009) Subspace method aided data-driven design of fault detection and isolation systems. J Process Contr 19:1495–1510

Ding SX, Zhang P, Jeinsch T, Ding EL, Engel P, Gui W (2011) A survey of the application of basic data-driven and model-based methods in process monitoring and fault diagnosis. In: 18th IFAC World Congress, Milano, pp 12380–12388

Frank P, Ding SX (1997) Survey of robust residual generation and evaluation methods in observer-based fault detection systems. J Process Contr 7:403–424

Gertler J (1998) Fault detection and diagnosis in engineering systems. Marcel Dekker, Berlin

Isermann R (2011) Fault-diagnosis application (Model-based condition monitoring: actuators, drives, machinery, plants, sensors, and fault-tolerant systems). Springer, Berlin/Heidelberg

Naik AS, Yin S, Ding SX, Zhang P (2010) Recursive identification algorithm to design fault detection systems. J Process Contr 20:957–965

Qin S (2006) An overview of subspace identification. Comput Chem Eng 30:1502–1513

Roth M, Lesage JJ, Litz L (2011) The concept of residuals for fault localization in discrete systems. Contr Eng Pract 18(5):978–988

Tan CP, Edwards C (2002) A robust sensor fault reconstruction scheme using sliding mode observers applied to a nonlinear aero-engine model. In: Proceedings of the American control conference, Anchorage, pp 5086–5091

Wang J, Qin S (2006) Closed-loop subspace identification using the parity space. Automatica 42:315–320

Wang YL, Ma GF, Ding SX, Li C (2011) Subspace aided data-driven design of robust detection and isolation systems. Automatica 47:2474–2480

Ye H, Wang GZ, Ding SX, Su HY (2002) An IIR filter based parity space approach for fault detection. In: 15th Triennial World Congress, Barcelona, Spain

Chapter 56
Evolutionary Stability Analysis of Behavior Among Partners Under a Default Punishment Mechanism

Rong Ding and Yan-ming Sun

Abstract Behavior among the alliance partners can be divided into two kinds of reciprocity and opportunism. Based on evolutionary game theory and methods, it constructs Prisoner's Dilemma evolutionary game model for behavior selection among partners under a default punishment mechanism, and analyzes the asymptotic stability of the behavioral strategies evolution in different punishment situations. The results show that: When liquidated damages are large enough to compensate for both the net loss of deception in reciprocity and the difference between the temptation earnings and the cooperation gains, it will be unprofitable to take opportunistic behavior, and both the two sides will adopt the win-win strategy of reciprocity behavior as evolutionary stable strategy. Finally, the numerical simulation analysis is taken to verify the correctness of the conclusion. The conclusion can be used as the reference of the cooperation agreement established among the partners.

Keywords Default punishment mechanism • Evolutionary stable strategy • Liquidated damages • Partnership selection

---

Supported by "the National Natural Science Foundation of China" (*No. 50675069*).
Supported by "the Fundamental Research Funds for the Central Universities in China" (*No. 2011ZM0036*)

R. Ding (✉) • Y. Sun
School of Business and Administration, South China University of Technology, Guangzhou, China
e-mail: dingrong_hb@yahoo.cn; meymsun@scut.edu.cn

E. Qi et al. (eds.), *The 19th International Conference on Industrial Engineering and Engineering Management*, DOI 10.1007/978-3-642-37270-4_56,

56.1 Introduction

In order to respond to certain external market opportunities, a temporary alliance formed by task-driven, geographically dispersed partners can be seen as collaboration community by the enterprise with common interest. As the alliance has a temporary, instability and complexity, the free-rider phenomenon and moral hazard caused by the opportunistic behavior of partners become the main reasons affecting the alliance performance.

During the process of cooperation, there must be comprehensive and effective mechanisms of cooperation. Nowak (2006a) summed up the five most typical mechanisms to explain cooperative behavior: kin selection, direct reciprocity, indirect reciprocity, group selection, spatial structure. In the study of direct or indirect reciprocity mechanism, the current academic focus is using empirical methods to study the control mechanisms, trust mechanisms, information sharing mechanism, the benefits and risks of coordination mechanisms, policy incentives and oversight mechanisms among partners (Das and Teng 2001; Inkpen and Currall 2004; Hacklin et al. 2006; Wen 2007; Granot and Yin 2008; Lin et al. 2009; Xiang Baohua and Ren Xinjian 2007). The above researches provide an important basis for the management of the alliance, among them, the majority of the oversight mechanisms focus on the external third-party policy incentives and government regulation, lack of concerning on the coalition formation process of the endogenous factors. Dyer (1997) think that a formal contract is an important means to control the behavior of partners, the contract specifies the responsibilities and obligations of the parties, when other side(s) taking the opportunistic behavior, it can give the legal rights of parties to protect their own interests. To facilitate both sides can fulfill contractual obligations normally, prior agreement will usually punish the defaulting party to compensate for compliance with the contract, therefore, on the basis of the recognition and inclusion of enterprises to pursue the best interests of the selfish characteristics, how to determine amounts of liquidated damages to effectively regulate and guide opportunistic behaviors of enterprises is the core issue.

The existing literatures have shown that behavior among partners has the characteristics of a typical Prisoner's Dilemma game (Luo Jianfeng 2012; Axelrod and Hamilton 1981; Nowak and Sigmund 1993; Nowak 2006b; Miehael and Hauert 2005). Therefore, based on evolutionary game theory and methods, and reflections on the above problems, to address how to determine amounts of liquidated damages of punishment, it constructs Prisoner' Dilemma evolutionary game model for behavior selection among partners under a default punishment mechanism. The results can be used as the reference of the cooperation agreement established among the partners.

56.2 Description and Building of the Model

In order to simplify the description of the model, it builds a coalition of the two enterprises (enterprises 1 and 2) game model, which is also applicable to describe the coalition of enterprises (enterprises between n). Model assumes are as follows: Both sides have bounded rationality, they generally can not immediately find the optimal strategy in repeated game process, according to the information of strategy, payoff and their own situation, on the basis of imitation and learning to adjust their strategy continuously. The strategy selection set of both sides is {reciprocity, opportunism}. Reciprocal behavior represents adhere to cooperate with each other, mutually beneficial in the alliance process; Opportunistic behavior means inactive cooperation with each other and free-riding, or betrayal at any time when there is more temptation in the alliance process. Assuming the alliance's initial cost of cross-enterprise communication and bargaining cost for the differences of opinion is C_1, the sum of resource investment, occupation and consumption as well as recessive loss caused by the market failure is C_2, cost-sharing coefficient is $\beta,\ 0 \le \beta \le 1$; If both sides take reciprocal behavior, they not only can gain profit $R_1,\ \ R_2$, but also excess returns $\alpha\Delta R$ and $(1-\alpha)\Delta R$ by using the resources of both complementary strengths. Once this tasks can be completed, at least in a certain period, there will be a degree of technology or service monopoly, both sides can gain an excess return. The net income will respectively be $R_1+\alpha\Delta R-\beta(C_1+C_2)$, $R_2+(1-\alpha)\Delta R-(1-\beta)(C_1+C_2)$ the excess income distribution coefficient α clearly defined in the contract, $0 \le \alpha \le 1$. If both sides take opportunistic behavior, both sides will have nothing, their net income will respectively be $-\beta C_1$ and $-(1-\beta)C_1$; Considering the situation of one party taking reciprocal behavior, and the other taking opportunistic behavior, the party taking reciprocal behavior will be losses, net loss of deception (usually negative) is $-\beta(C_1+C_2)$ or $-(1-\beta)(C_1+C_2)$, the other party taking opportunistic behavior will get betrayal temptation income T_1 or T_2, the corresponding temptation net income is $T_1-\beta C_1$ or $T_2-(1-\beta)C_1$. Assume that both two sides focused on the long-term relationship, prior agree on amounts of liquidated damages P in the form of contractual agreements to prevent the emergence of opportunistic behavior, the party taking opportunistic behavior should be punished, to compensate the losses of the party taking reciprocal behavior. The payoff matrix of both sides under a default punishment mechanism is shown in Table 56.1. In the real economic environment, the behavior between the partners has

Table 56.1 The matrix for behavior game among partners under a default punishment mechanism

Enterprise1,2	Reciprocity	Opportunism
Reciprocity	$R_1+\alpha\Delta R-\beta(C_1+C_2)$, $R_2+(1-\alpha)\Delta R-(1-\beta)(C_1+C_2)$	$-\beta(C_1+C_2)+P$, $T_2-(1-\beta)C_1-P$
Opportunism	$T_1-\beta C_1-P$, $-(1-\beta)(C_1+C_2)+P$	$-\beta C_1, -(1-\beta)C_1$

the typical characteristics of the Prisoner' Dilemma, therefore, Table 56.1 satisfy the following relationship, Simplified to:

$$T_1 > R_1 + \alpha\Delta R - \beta C_2 > 0 \tag{56.1}$$

$$T_2 > R_2 + (1-\alpha)\Delta R - (1-\beta)C_2 > 0 \tag{56.2}$$

Assume that in the long-term cooperation, the density for enterprise 1 to taking the reciprocal behavior is x, and opportunistic behavior density is $1-x,\ 0 \le x \le 1$; the density for enterprise 2 to take the reciprocal behavior is y, and opportunistic behavior density is $1-y,\ 0 \le y \le 1$; Derived from Table 56.1, when the enterprise 1 taking reciprocal behavior and opportunistic behavior, the fitness U_{ic}, U_{id} and average fitness $\overline{U_i}$ are respectively shown as follows:

$$\begin{aligned} U_{ic} &= y\left[R_1 + \alpha\Delta R - \beta(C_1 + C_2)\right] + (1-y)\left[-\beta(C_1 + C_2) + P\right] \\ U_{id} &= y(T_1 - \beta C_1 - P) + (1-y)(-\beta C_1) \\ \overline{U_i} &= xU_{ic} + (1-x)U_{id} \end{aligned}$$

Similarly, when the enterprise 2 taking reciprocal behavior and opportunistic behavior, the fitness U_{jc}, U_{jd} and average fitness $\overline{U_j}$ are respectively shown as follows:

$$\begin{aligned} U_{jc} &= x[R_2 + (1-\alpha)\Delta R - (1-\beta)(C_1 + C_2)] + (1-x)[-(1-\beta)(C_1 + C_2) + P] \\ U_{jd} &= x[T_2 - (1-\beta)C_1 - P] + (1-x)[-(1-\beta)C_1] \\ \overline{U_j} &= yU_{jc} + (1-y)U_{jd} \end{aligned}$$

The replicated dynamic equations are respectively shown as follows:

$$\begin{aligned} F(x) = \frac{dx}{dt} &= x(U_{ic} - \overline{U_i}) \\ &= x(1-x)[y(R_1 + \alpha\Delta R - T_1) - \beta C_2 + P] \end{aligned} \tag{56.3}$$

$$\begin{aligned} F(y) = \frac{dy}{dt} &= y(U_{jc} - \overline{U_j}) \\ &= y(1-y)\{x[R_2 + (1-\alpha)\Delta R - T_2] - (1-\beta)C_2 + P\} \end{aligned} \tag{56.4}$$

When $F(x) = 0, F(y) = 0$,there are five evolutionary equilibrium points in Table 56.1: O (0, 0), A (1, 0), B (0, 1), C (1, 1),D (d1, d2), $d_1 = \frac{-(1-\beta)C_2 + P}{T_2 - [R_2 + (1-\alpha)\Delta R]}$, $d_2 = \frac{-\beta C_2 + P}{T1 - (R_1 + \alpha\Delta R)}$.

56.3 Evolution Stability Analysis of the Model

Evolutionary stable strategy (ESS) (Wang Xianyu and Xiao Yuming 2008) is used to describe the stability status of game evolutionary process. Firstly, calculate Jacobian matrix of Table 56.1 by the method proposed by literatures (Nagarajan and Sosic 2007, 2009) shown in formula (56.5), then calculate the value of Matrix determinant det J and trace $tr\ J$ of the formula (56.5) shown in Table 56.2. The local stability of the evolution equilibrium point in Table 56.1 can be judged by the value of det J and $tr\ J$ and their symbol. In the situation of different parameter values, it will analyze the asymptotic stability of each evolution equilibrium point.

$$\begin{bmatrix} (1-2x)[y(R_1+\alpha\Delta R-T_1)-\beta C_2+P] & x(1-x)(R_1+\alpha\Delta R-T_1) \\ y(1-y)[R_2+(1-\alpha)\Delta R-T_2] & (1-2y)\{x[R_2+(1-\alpha)\Delta R-T_2]-(1-\beta)C_2+P \end{bmatrix} \tag{56.5}$$

Situation 56.1

$$0 < P < (1-\beta)C_2 < T_2 - [R_2 + (1-\alpha)\Delta R - (1-\beta)C_2],$$
$$0 < P < \beta C_2 < T_1 - (R_1 + \alpha\Delta R - \beta C_2), d_1 < 0, d_2 < 0.$$

In this situation, $D(d_1, d_2) \notin (0, 1)$, $D(d_1, d_2)$ will not be the evolution equilibrium point. There are only four evolution equilibrium points: O (0, 0), A (1, 0), B (0, 1), C (1, 1), C (1, 1) is an unstable point, A (1, 0), B (0, 1) are saddle points, O (0, 0) is the only ESS. The result of evolutionary stability is shown in Table 56.3.

The evolution phases diagram between partners behavior game is shown as Fig. 56.1a shows. Only when both sides have adopted the strategy of opportunistic behavior, the result of evolution is ESS. This shows that although the two sides signed a binding agreement with liquidated damages, but the amounts of liquidated

Table 56.2 Matrix determinant and trace value of evolutionary equilibrium point

Equilibrium point	det J	trJ
O (0, 0)	$(-\beta C_2 + P) \cdot [-(1-\beta)C_2 + P]$	$-C_2 + 2P$
A (1, 0)	$(\beta C_2 - P) \cdot$ $[R_2 + (1-\alpha)\Delta R - T_2 - (1-\beta)C_2 + P]$	$R_2 + (1-\alpha)\Delta R$ $- T_2 + (2\beta - 1)C_2$
B (0, 1)	$(R_1 + \alpha\Delta R - T_1 - \beta C_2 + P) \cdot$ $[(1-\beta)C_2 - P]$	$(R_1 + \alpha\Delta R - T_1)$ $+ (1 - 2\beta)C_2$
C (1, 1)	$(R_1 + \alpha\Delta R - T_1 - \beta C_2 + P) \cdot$ $[R_2 + (1-\alpha)\Delta R - T_2 - (1-\beta)C_2 + P]$	$-(R_1 + R_2 + \Delta R$ $- T_1 - T_2 - C_2 + 2P)$
D (d1, d2)	$-\mathrm{d}_1(1-\mathrm{d}_1)(R_1 + \alpha\Delta R - T_1) \cdot$ $\mathrm{d}_2(1-\mathrm{d}_2)[R_2 + (1-\alpha)\Delta R - T_2]$	0

Table 56.3 Evolutionary stability result under Situation 56.1

Equilibrium point	det J symbol	trJ symbol	Equilibrium result
O (0, 0)	+	−	ESS
A (1, 0)	−	Uncertain	Saddle point
B (0, 1)	−	Uncertain	Saddle point
C (1, 1)	+	+	Unstable point

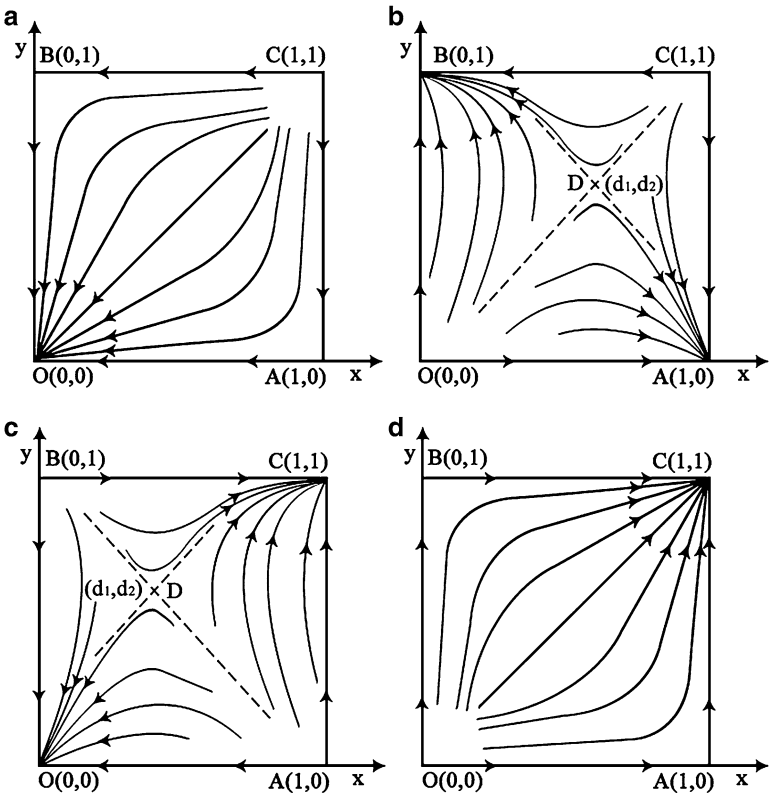


Fig. 56.1 Game evolutionary phase diagram under Situation 56.1–56.4

damages P are too small, not enough to compensate for the net loss of deception when taking reciprocal behavior, also the difference between the temptation earnings and the cooperation gains. Therefore, it eventually could not get rid of a typical prisoner's Dilemma. This default punishment mechanism has not played role of constraining opportunistic behavior and promoting mutually beneficial cooperation.

Situation 56.2

$$(1-\beta)C_2 < P < T_2 - [R_2 + (1-\alpha)\Delta R - (1-\beta)C_2],$$
$$\beta C_2 < P < T_1 - (R_1 + \alpha\Delta R - \beta C_2), D(d_1, d_2) \in (0, 1).$$

Table 56.4 Evolutionary stability result under Situation 56.2

Equilibrium point	det J symbol	trJ symbol	Equilibrium result
O (0, 0)	+	+	Unstable point
A (1, 0)	+	−	ESS
B (0, 1)	+	−	ESS
C (1, 1)	+	+	Unstable point
D (d1, d2)	−	Empty	Saddle point

There are five evolution equilibrium points: O (0, 0), A (1, 0), B (0, 1), C (1, 1), $D(d_1, d_2)$.O (0, 0), C (1, 1) are unstable points, A (1, 0), B (0, 1) are ESS, and $D(d_1, d_2)$ is a saddle point. The result of evolutionary stability is shown in Table 56.4.

The evolution phases diagram between partners behavior game is shown as Fig. 56.1b shows. As the amounts of liquidated damages P gradually increased, both sides have strong motivation to change strategy and aspirations. The stable point of the system turns from O (0, 0) in Situation 56.1 into A (1, 0) and B (0, 1) in Situation 56.2.

On the one hand, the amounts of liquidated damages are sufficient to compensate for the net loss of deception when taking reciprocal behavior, it is $P > \max\{\beta C_2, (1-\beta)C_2\}$. Enterprises will get more benefits when having chosen a reciprocal behavior strategy, resulting in the increase of the probability of taking reciprocal behavior strategy. But on the other hand, the amounts of liquidated damages are not enough to compensate for the difference between the temptation earnings and the cooperation gains, and is $P < \min\{T_2 - [R_2 + (1-\alpha)\Delta R - (1-\beta)C_2], T_1 - (R_1 + \alpha\Delta R - \beta C_2)\}$. It will still get more benefits when choosing opportunistic behavior strategy, resulting in the increase of the probability of taking opportunistic behavior strategy.

In this situation, the probability of saddle point D (d1, d2) deviating from O (0, 0), the worst solution for Pareto, is gradually increased. The area of ADBC is reducing, the power of enterprise cooperation is increasing, and the system eventually evolves into a typical hawk dove game. This means an increasing in the amount of liquidated damages have positive significance for punishing opportunistic behavior, and promoting mutually beneficial cooperation between enterprises.

Situation 56.3

$$T_2 - [R_2 + (1-\alpha)\Delta R - (1-\beta)C_2] < P < (1-\beta)C_2,$$
$$T_1 - (R_1 + \alpha\Delta R - \beta C_2) < P < \beta C_2, D(d_1, d_2) \in (0, 1).$$

There are five evolution equilibrium points: O (0, 0), A (1, 0), B (0, 1), C (1, 1), $D(d_1, d_2)$. O (0, 0), C (1, 1) are ESS, A (1, 0), B (0, 1) are unstable points, and $D(d_1, d_2)$ is a saddle point. The result of evolutionary stability is shown in Table 56.5.

Table 56.5 Evolutionary stability result under Situation 56.3

Equilibrium point	det J symbol	trJ symbol	Equilibrium result
O (0, 0)	+	−	ESS
A (1, 0)	+	+	Unstable point
B (0, 1)	+	+	Unstable point
C (1, 1)	+	−	ESS
D (d1, d2)	−	Empty	Saddle point

Table 56.6 Evolutionary stability result under Situation 56.4

Equilibrium point	det J symbol	trJ symbol	Equilibrium result
O (0, 0)	+	+	Unstable point
A (1, 0)	−	Uncertain	Saddle point
B (0, 1)	−	Uncertain	Saddle point
C (1, 1)	+	−	ESS

The evolution phases diagram between partners behavior game is shown as Fig. 56.1c shows. As the amounts of liquidated damages P continue to increase, the stable point of the system turn from A (1, 0) and B (0, 1) into O (0, 0) and C (1, 1).

On the one hand, the amounts of liquidated damages are sufficient to compensate for the difference between the temptation earnings and the cooperation gains, that is $P > \max\{T_2 - [R_2 + (1-\alpha)\Delta R - (1-\beta)C_2], T_1 - (R_1 + \alpha\Delta R - \beta C_2)\}$. Enterprises will get more benefits when choosing reciprocal behavior strategy, resulting in the increase of the probability of taking reciprocal behavior strategy; On the other hand, the amounts of liquidated damages are not enough to compensate for the net loss of deception, that is $P < \min\{\beta C_2, (1-\beta)C_2\}$. Enterprises will still get more benefits when choosing opportunistic behavior strategy, resulting in the increase of the probability of taking opportunistic behavior strategy.

In this situation, what state the system will evolve into depends on P. $P > \max\{T_2 - [R_2 + (1-\alpha)\Delta R + 2(1-\beta)C_2/2, T_1 - [R_1 + \alpha\Delta R] + 2\beta C_2/2\}$, for both sides, the probability of taking reciprocal behavior strategy is more than 50 %. The probability of saddle point D (d1, d2) closing to C (1, 1), the best solution for Pareto, is gradually increased. The area of ADBC will continue to reduce, the power of enterprise cooperation continues to increase, eventually evolved into a typical stag game, and this means that a default punishment mechanism is very active in promoting the mutually beneficial cooperation between enterprises.

Situation 56.4

$$P \geq T_2 - [R_2 + (1-\alpha)\Delta R - (1-\beta)C_2], P \geq (1-\beta)C_2,$$
$$P \geq T_1 - (R_1 + \alpha\Delta R - \beta C_2), P \geq \beta C_2, D(d_1, d_2) \notin (0, 1).$$

There are four evolution equilibrium points: O (0, 0), A (1, 0), B (0, 1), C (1, 1). Only point C (1, 1) is ESS, A (1, 0) and B (0, 1) are saddle points, O (0, 0) is unstable point. The result of evolutionary stability is shown in Table 56.6.

The evolution phases diagram between partners behavior game is shown as Fig. 56.1d shows. Only when both sides have adopted reciprocal behavior strategy, the result of evolution is ESS. Compared with the above three Situations, the amounts of liquidated damages significantly increase, which are more than the difference between the temptation earnings and the cooperation gains, and are also more than the net loss of deception when taking reciprocal behavior strategy, resulting the high returns for taking opportunistic behavior strategy are severely punished away. Therefore, the probability of system evolves from the initial instability point O (0, 0), A (1, 0), B (0, 1) to C (1, 1) increases, this means that a default punishment mechanism play a good role in ensuring the smooth development of mutually beneficial cooperation between enterprises.

56.4 Numerical Experiments and Simulation

The above have made theoretical modeling and evolutionary stability analysis of behavior between partners under a default punishment mechanism, to be more intuitive, it will make further study through numerical experiments and simulation. The horizontal axis t denotes the simulation time [0, 10], consider each game as a simulation cycles, the simulation is finished when the game has run the ten-cycle; the vertical axis P_c denotes the collaborators density [0, 1] under the conditions of the evolutionary stable state with random role. According to the realities and assumptions of the model, in order to facilitate study of the problem, assume that $\alpha = \beta = 0.5$, C_1 is negligible. Meanwhile, in order to minimize the impact of random numbers on the numerical results, in the same initial conditions, it makes computational experiment 100 times respectively and then averages the results. As can be seen, the initial collaborators density under conditions of random role does not affect the final evolution stable result of alliance, and only affect the speed to reach the evolutionary stable state.

On the basis of a number of simulation experiments and results comparing under Situation 56.1, it concludes that: For any given initial cooperation density (x, y), the both game sides evolve to take opportunistic behavior strategy, and verifying Situation 56.1. When $P = 1$, $C_2 = 4$, $T_1 = 5$, $R_1 + \alpha\Delta R = 3$, and the initial value is (0.9, 0.4), evolution trends are shown in Fig. 56.2a. Similarly, when $P = 3$, $C_2 = 4$, $T_1 = 5$, $R_1 + \alpha\Delta R = 3$, and the initial value is (0.7, 0.3), evolution trends of Situation 56.2 are shown in Fig. 56.2b; When $P = 3$, $C_2 = 8$, $T_1 = 3$, $R_1 + \alpha\Delta R = 5$, and the initial value is (0.8, 0.6), evolution trends of Situation 56.3 are shown in Fig. 56.2c. When $P = 5$, $C_2 = 4$, $T_1 = 4$, $R_1 + \alpha\Delta R = 3$, and the initial value is (0.2, 0.4), evolution trends of Situation 56.4 are shown in Fig. 56.2d.

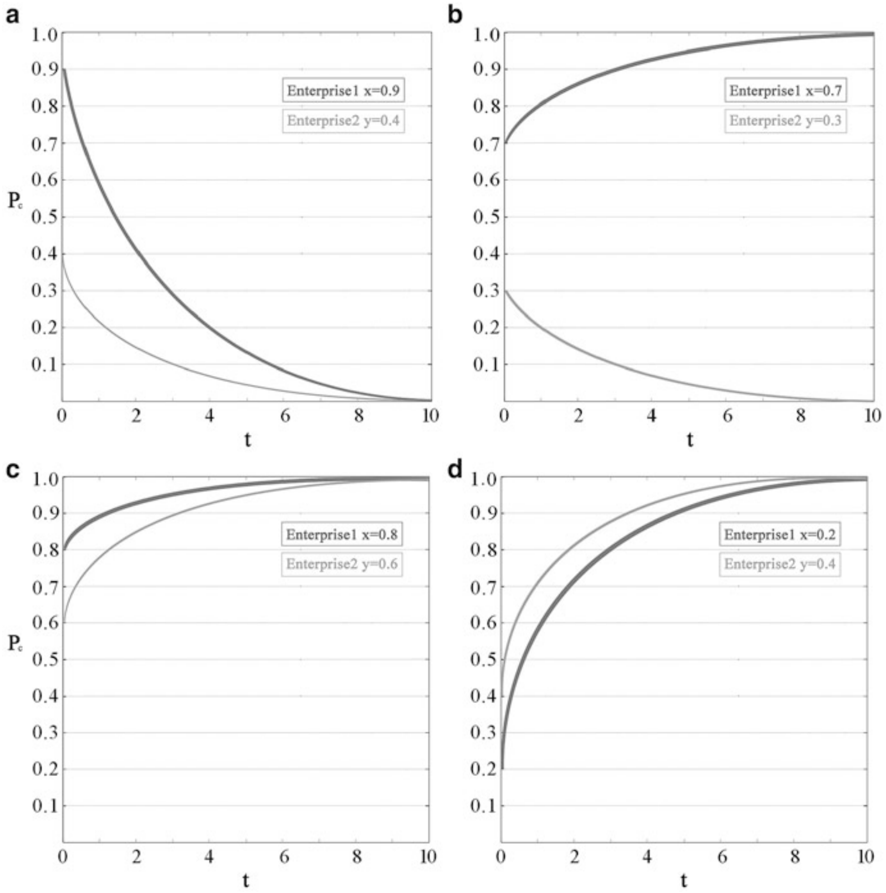


Fig. 56.2 A simulation result under Situation 56.1–56.4

56.5 Conclusion

Based on evolutionary game theory and methods, it constructs Prisoner's Dilemma evolutionary game model for behavior selection among partners under a default punishment mechanism, makes in-depth discussion of the asymptotic stability of the behavioral strategies evolution in different punishment situations between partners, and gives a deeper and more intuitive explanation by the numerical simulation. The results show that: When liquidated damages are large enough to compensate for both the net loss of deception in reciprocity and the difference between the temptation earnings and the cooperation gains, it will be unprofitable to take opportunistic behavior, and both the two sides will adopt the win-win strategy of reciprocity behavior as evolutionary stable strategy.

This can be interpreted as that: before beginning cooperation, the enterprises should specify the reasonable amounts of liquidated damages according to the potential earning of both sides in advance, to prevent opportunistic tendencies in the process of cooperation, facilitate both sides to take the win-win strategy of reciprocal behavior, protect cooperation agreement is carried out smoothly, and lay the foundation for long-term viability of the partnership. In addition, the sensitivities of some other parameters, such as cost-sharing coefficient and excess earning distribution coefficient, are not analyzed deeply, which can be done further research and expansion.

References

Axelrod R, Hamilton WD (1981) The evolution of cooperation [J]. Science 211(4489):1390–1396

Das Tk, Teng BS (2001) Trust, control, and risk in strategic alliances: an integrated framework [J]. Organ Stud 22(2):251–283

Dyer J (1997) Effective interfirm collaboration: how firms minimize transaction costs and maximize transaction value [J]. Strateg Manag J 18(7):535–556

Granot D, Yin S (2008) Competition and cooperation in decentralized push and pull assembly system [J]. Manag Sci 54(4):733–747

Hacklin F, Marx TC, Fahrni F (2006) Strategic venture partner selection for collaborative innovation in production systems: a decision support system-based approach [J]. Int J Prod Econ 104(1):100–112

Inkpen AC, Currall SC (2004) The co-evolution of trust, control and learning in joint ventures [J]. Organ Sci 15(5):586–599

Lin Z, Yang HB, Arya B (2009) Alliance partners and firm performance: resource complementary and status association [J]. Strateg Manag J 30(9):921–940

Luo Jianfeng (2012) The default punishment mechanism of cooperation among enterprises based on evolutionary game theory [J]. Syst Eng 30(1):27–31 (in Chinese)

Miehael D, Hauert C (2005) Models of cooperation based on the Prisoner's dilemma and the snowdrift game [J]. Ecol Lett 8(7):748–776

Nagarajan M, Sosic G (2007) Stable farsighted coalitions in competitive markets [J]. Manag Sci 53(1):29–45

Nagarajan M, Sosic G (2009) Coalition stability in assembly models [J]. Oper Res 57(1):131–145

Nowak MA (2006a) Five rules for the evolution of cooperation [J]. Science 314(5805):1560–1563

Nowak MA (2006b) Five rules for the evolution of cooperation [J]. Science 314(5805):1560–1563

Nowak MA, Sigmund KA (1993) Strategy of win stay, lose shift that outperforms tit-for-tat in the Prisoners-Dilemma game [J]. Nature 364(6432):56–58

Wang Xianyu, Xiao Yuming (2008) Game theory and its application [M]. Science Press, Beijing (in Chinese)

Wen BL (2007) Factors affecting the correlation between interactive mechanism of strategic alliance and technological knowledge transfer performance [J]. J High Technol Manag Res 17(2):139–155

Xiang Baohua, Ren Xinjian (2007) A game simulation research on the relationship of the selection of enterprises' coopetition behavior, time and the Pay-off [J]. Syst Eng Theor Pract 27(7):34–40 (in Chinese)

Chapter 57
Stability and Hopf Bifurcation of a Four-Dimensional Dynamic Economic System

Hong-liang Tu and Jun-hai Ma

Abstract In this paper, first of all, a four-dimensional dynamic economic system is established. First, choosing the saving rate as the bifurcation parameter, the stability and Hopf bifurcation of the system are studied. Then, a example of simulation of the model is used to prove the derived results in Sect. 57.2. Furthermore, a time-delayed feedback is added to the dynamic finance system. Choosing the delay as the bifurcation parameter, the local stability and the existence of Hopf bifurcation of the model with delay are researched. Stability changes and Hopf bifurcation happens while the delay passes through a critical value. We can see that increasing the delay can lead the dynamic economic system to fluctuate. Then, a numerical example is taken to confirm the theoretical results obtained in Sect. 57.3. Finally, some conclusions are made. This paper has an important theoretical and practical significance.

Keywords Dynamic economic system • Hopf bifurcation • Stability • Time delay • Time-delayed feedback

57.1 Introduction

Recently, dynamic economics have become more and more prevalent in the mainstream economics. This effect is quite profound, and influences both micro- and macroeconomics. Especially, economic dynamics is interest in delay differential equations, because some phenomena in the economy may not be portrayed with linear or non-linear differential equations. In engineering, biology economy and social sciences, quite a lot of problems can be described with the help of delay

H. Tu (✉) • J. Ma
Department of Management and Economics, Tianjin University, Tianjin 300072, People's Republic of China
e-mail: tuhongliang00@126.com; mjhtju@yahoo.com.cn

E. Qi et al. (eds.), *The 19th International Conference on Industrial Engineering and Engineering Management*, DOI 10.1007/978-3-642-37270-4_57,

differential equations. In this paper, we build up a four dimensional dynamic economic system and study how the saving rate and the time delay affect the stability of the dynamic economic system.

Hopf bifurcations of dynamical system with delay have already been studied by many researches (Fanti 2004; Gao and Ma 2009; Li et al. 2002; Yan and Li 2006; Chen 2007; Zhang and Wei 2004; Szydlowski and Krawiec 2005; Ma and Gao 2007). But, these papers are considered population dynamics model, and two or three dimensional finance system model with delays.

In Cai (2005), Cesare and Sportelli (2005) and Neamtu (2007), the researchers have studied the IS-LM model with delay taxation, and shown that the delay tax collection can lead to all kinds of dynamic characteristics.

Gao and Ma (2009), Ma et al. (2007), and Chen (2008a, b) have introduced a dynamic finance model which is made up of three first-order differential equations. The model describes three variables on the variation of time: the interest rate x, the investment demand y, and the prize index z. The factors effect on the change of the interest rate x mainly stem from the contradictions of the investment market and the structural adjustment of the price. The change rate of the investment demand y is in proportion to the rate of investment and an inversion with the investment cost and interest rates. The change of the prize index z, one facet, is affected by the inflation rates, the other facet, is determined by the contradiction between supply and demand in the commercial markets. Gao and Ma (2009), Ma et al. (2007), and Chen (2008a, b) have reported a simplified finance model as follow:

$$\begin{cases} \dot{x} = z + (y - a)x, \\ \dot{y} = 1 - by - x^2, \\ \dot{z} = -x - cz. \end{cases} \tag{57.1}$$

Where a is the amount of saving, b is the per investment cost, and c is the demand elasticity in the commercial market.

So far as I know, there is few literature on four dimensional dynamic economic system including the change of price index. As price index closely contact with inflation, studying economic system including the change of price index has an important theoretical and practical significance. Base on the IS-LM model which has been studied in Cai (2005), Cesare and Sportelli (2005), Neamtu (2007) and the model of (57.1) which has been researched in Gao and Ma (2009), Ma et al. (2007), and Chen (2008a, b), we build up a four-dimensional finance system as follow:

$$\begin{cases} \dot{Y}(t) = \alpha_1[I(t) - sY(t)] - \alpha_2[-R(t) - cP(t)]Y(t), \\ \alpha_1 > 0, \alpha_2 > 0, s \in (0,1), \\ \dot{R}(t) = \beta_1 P(t) + (I(t) - sY(t))R(t), \\ \dot{P}(t) = -R(t) - cP(t), \\ \dot{I}(t) = \gamma_1 \{\alpha_1[I(t) - sY(t)] - \alpha_2[-R(t) - cP(t)]Y(t)\} \\ -\nu I(t) - R^2(t)I(t), \gamma_1 > 0, \nu > 0. \end{cases} \tag{57.2}$$

with $Y(t)$ as production, $I(t)$ as investment demand, s as saving rate, $R(t)$ as interest rate, $P(t)$ as prize index, and $c > 0$ is the demand elasticity of the commercial market, $\nu > 0$ is the per investment cost. The change of $Y(t)$ mainly comes from two facets: one, is in proportion to the surplus between investment and savings, the other, is in proportion to the reverse change of $P(t)$. The factors that affect the change of $R(t)$ mainly come from two aspects: one is the contradictions of the surplus between investment and saving in the investment market, the other is the structural adjustment of price. The change of $P(t)$: one facet, is influenced by the inflation rate, the other facet, is affected by the contradiction between supply and demand in the commercial market. The change rate of $I(t)$ mainly stems from two aspects: first, is in proportion to the change of $Y(t)$, second, is an inversion with the investment cost and interest rate. The four-dimensional dynamic finance system is much more complicated than two-dimensional or three-dimensional dynamic finance system, so it is very interesting to study the four-dimensional dynamic finance system in detail.

The structure of this paper is as follows. Choosing the saving rate as the bifurcation parameter, we study the stability and Hopf bifurcation of the equilibrium point of system (57.2) in Sect. 57.2. Furthermore, a numerical example is used to confirm the derived results. In Sect. 57.3, the local stability and Hopf bifurcation of system (57.2) with time-delayed feedback are investigated. Then, simulation of a specific version of the model is given, in which show the existence and nature of the period solutions. In the end, the conclusions of this paper are made.

57.2 The Stability and Hopf Bifurcation of System (57.2)

We will investigate the effects of the saving rate s on system (57.2) in this section. The stability of the equilibrium point and the existence of Hopf bifurcation will be considered when the saving rate s as parameter.

We can see that (0, 0, 0, 0) is the only one non-negative equilibrium point of system (57.2).

The Jacobi matrix of system (57.2) at equilibrium (0, 0, 0, 0) as follow:

$$J = \begin{pmatrix} -\alpha_1 s & 0 & 0 & \alpha_1 \\ 0 & 0 & \beta_1 & 0 \\ 0 & -1 & -c & 0 \\ -\gamma_1\alpha_1 s & 0 & 0 & \gamma_1\alpha_1 - \nu \end{pmatrix} \tag{57.3}$$

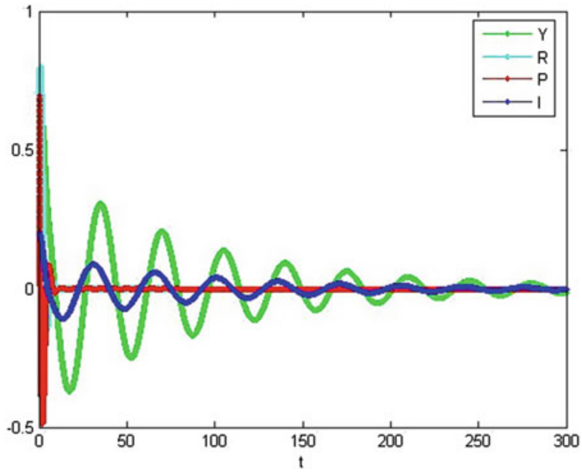
The character equation of system (57.2) and by simple calculation, we can derive

$$|\lambda E - J| = 0,$$

$$(\lambda^2 + c\lambda + \beta_1)[\lambda^2 + (\nu - \gamma_1\alpha_1 + \alpha_1 s)\lambda + \nu\alpha_1 s] = 0. \tag{57.4}$$

According to Ruoth-Hurwitz criteria, the follow theorem can be got.

Fig. 57.1 $s = 0.22$



Theorem 57.1 **(I)** *If* $s > \frac{1}{\alpha_1}(\gamma_1\alpha_1 - \nu)$, *system* (57.2) *is stable near the equilibrium point. It means that the economy can develop quickly and healthily when the saving rate s is kept at a appropriate level.*

(II) *If* $s < \frac{1}{\alpha_1}(\gamma_1\alpha_1 - \nu)$, *system* (57.2) *is unstable near the equilibrium point.*

(III) *If* $s = \frac{1}{\alpha_1}(\gamma_1\alpha_1 - \nu)$ *and* $\gamma_1\alpha_1 - \nu > 0$, *there is a Hopf bifurcation happens on system* (57.2) *at the equilibrium point.*

Finally, numerical simulations of system (57.2) will be shown to justify the Theorem 57.1.

Choosing $\alpha_1 = 0.98, \alpha_2 = 0.85, \beta_1 = 0.95, c = 0.92, \gamma_1 = 0.35, \nu = 0.15$, system (57.2) becomes into:

$$\begin{cases} \dot{Y}(t) = 0.98[I(t) - sY(t)] - 0.85[-R(t) - 0.92P(t)]Y(t), \\ \qquad\qquad , s \in (0, 1), \\ \dot{R}(t) = 0.95P(t) + (I(t) - sY(t))R(t), \\ \dot{P}(t) = -R(t) - 0.92P(t), \\ \dot{I}(t) = 0.35\{0.98[I(t) - sY(t)] - 0.85[-R(t) - 0.92P(t)]Y(t)\} \\ \qquad\qquad - 0.15I(t) - R^2(t)I(t). \end{cases} \tag{57.5}$$

Then, we consider the system (57.5) at the equilibrium point (0, 0, 0, 0).

First, we choose $s = 0.22$, which satisfies the condition of **Theorem 57.1 (I)**, we can see that its solution is asymptotically stable such as the corresponding waveform plots shown in Fig. 57.1.

Fig. 57.2
s = 0.1969387755102

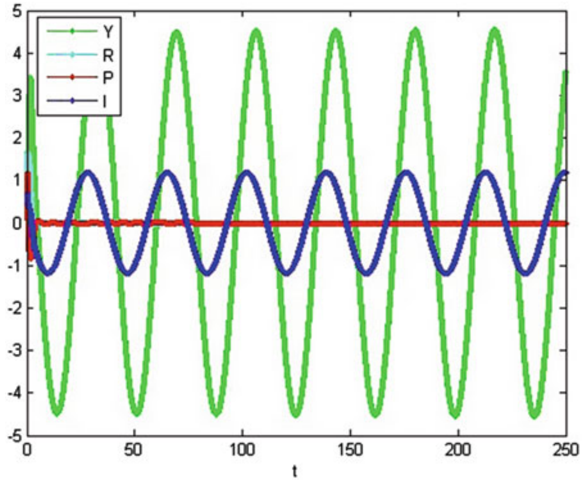
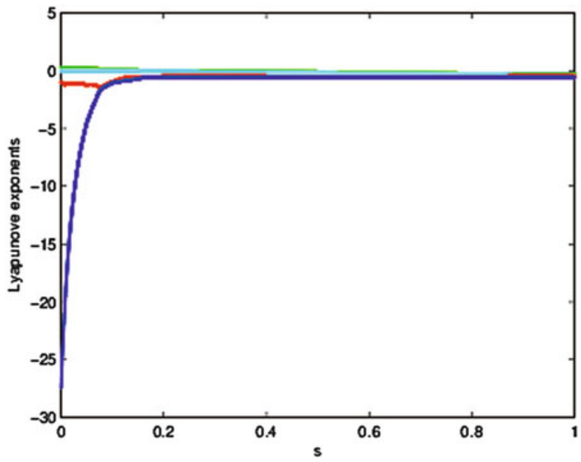


Fig. 57.3 Lyapunove exponents' spectrum of system (57.5)



Then, we choose s = 0.1969387755102, which satisfies the condition of Theorem 57.1 (III). There is a Hopf bifurcation happened on the system (57.2) at the equilibrium point as the corresponding waveform plots shown in Fig. 57.2.

Finally, through analysis of Lyapunove exponents of system (57.5) as show in Fig. 57.3. Two Lyapunove exponents less than zero and two Lyapunove exponents greater than zero, and the equilibrium becomes unbounded for $s \in (0, 0.15)$. Three Lyapunove exponents less than zero and one Lyapunove exponents greater than zero, and the equilibrium becomes fluctuate for $s \in (0.15,\ 0.1969)$. For example, the equilibrium becomes fluctuate when s = 0.17 as the corresponding waveform plots shown in Fig. 57.4. Four Lyapunove exponents less than zero and the equilibrium becomes stabile for $s \in (0.1969,1)$.

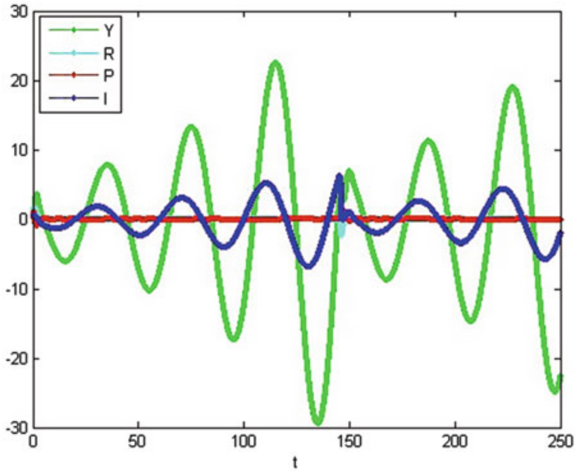


Fig. 57.4 s = 0.17

57.3 The Model with Time-Delayed Feedback

In this section, the effects of the time delay on the system (57.2) will be investigated. The delay usually exists in the realistic economic system. For predigesting the production $Y(t)$, time-delayed feedback is only added to the production $Y(t)$. Time-delayed feedback is added to the first equation of system (57.2), and the following new system can be obtained:

$$\begin{cases} \dot{Y}(t) = \alpha_1[I(t) - sY(t)] - \alpha_2[-R(t) - cP(t)]Y(t) \\ \qquad + k_1[Y(t) - Y(t-\tau)], \qquad \alpha_1 > 0, \alpha_2 > 0, s \in (0,1), \\ \dot{R}(t) = \beta_1 P(t) + (I(t) - sY(t))R(t), \\ \dot{P}(t) = -R(t) - cP(t), \\ \dot{I}(t) = \gamma_1\{\alpha_1[I(t) - sY(t)] - \alpha_2[-R(t) - cP(t)]Y(t)\} \\ \qquad - \nu I(t) - R^2(t)I(t), \gamma_1 > 0, \nu > 0. \end{cases} \tag{57.6}$$

We can see that (0, 0, 0, 0) is the only one non-negative equilibrium point of system (57.6).

The character equation of system (57.6), and by simple calculation, we can derive:

$$\begin{aligned} &|\lambda E - A| = 0, \\ &(\lambda^2 + c\lambda + \beta_1)[\lambda^2 + (\nu - \gamma_1\alpha_1 + \alpha_1 s - k_1 + k_1 e^{-\lambda\tau})\lambda \\ &\quad + \nu(\alpha_1 s - k_1 + k_1 e^{-\lambda\tau}) + \gamma_1\alpha_1 k_1(1 - e^{-\lambda\tau})] = 0 \end{aligned} \tag{57.7}$$

where

$$A = \begin{pmatrix} -\alpha_1 s + k_1 - k_1 e^{-\lambda\tau} & 0 & 0 & \alpha_1 \\ 0 & 0 & \beta_1 & 0 \\ 0 & -1 & -c & 0 \\ -\gamma_1\alpha_1 s & 0 & 0 & \gamma_1\alpha_1 - \nu \end{pmatrix}.$$

Furthermore, we only require considering the root distribution of the following transcendental equation:

$$\begin{aligned} &\lambda^2 + (\nu - \gamma_1\alpha_1 + \alpha_1 s - k_1 + k_1 e^{-\lambda\tau})\lambda \\ &\quad + \nu(\alpha_1 s - k_1 + k_1 e^{-\lambda\tau}) + \gamma_1\alpha_1 k_1(1 - e^{-\lambda\tau}) = 0 \end{aligned} \tag{57.8}$$

Let $\lambda = \omega i$, $\tau = \tau$, and substituting this into Eq. (57.8), then Eq. (57.8) becomes into

$$\begin{aligned} &-\omega^2 + [\nu - \gamma_1\alpha_1 + \alpha_1 s - k_1 + k_1(\cos\omega\tau - i\sin\omega\tau)]\omega i \\ &+ \nu[\alpha_1 s - k_1 + k_1(\cos\omega\tau - i\sin\omega\tau)] \\ &+ \gamma_1\alpha_1 k_1[1 - (\cos\omega\tau - i\sin\omega\tau)] = 0 \end{aligned} \tag{57.9}$$

Separating the real and imaginary parts of Eq. (57.9), the following equations can be obtained

$$\begin{aligned} -\omega_0{}^2 + \nu(\alpha_1 s - k_1) + k_1\alpha_1\gamma_1 + k_1\sin\omega_0\tau_0 - k_1(\alpha_1\gamma_1 - \nu)\cos\omega_0\tau_0 = 0 \\ \omega_0(\alpha_1 s - k_1 + \nu - \alpha_1\gamma_1) + k_1(\alpha_1\gamma_1 - \nu)\sin\omega_0\tau_0 + k_1\omega_0\cos\omega_0\tau_0 = 0 \end{aligned} \tag{57.10}$$

Through simple calculation, the following equations can be got

$$\begin{aligned} &\cos\omega\tau \\ &= \frac{k_1\omega^2(-\alpha_1 s + k_1 - \nu + \alpha_1\gamma_1) + k_1(\alpha_1\gamma_1 - \nu)[-\omega^2 + \nu(\alpha_1 s - k_1) + k_1\alpha_1\gamma_1]}{k_1{}^2(\alpha_1\gamma_1 - \nu)^2 + (k_1\omega)^2}, \\ &\sin\omega\tau \\ &= \frac{k_1\omega(\alpha_1\gamma_1 - \nu)(-\alpha_1 s + k_1 - \nu + \alpha_1\gamma_1) + k_1\omega[\omega^2 - \nu(\alpha_1 s - k_1) - k_1\alpha_1\gamma_1]}{k_1{}^2(\alpha_1\gamma_1 - \nu)^2 + (k_1\omega)^2}. \end{aligned} \tag{57.11}$$

According to the identity $\sin^2\omega\tau + \cos^2\omega\tau = 1$, we have

$$k_1{}^2\omega^6 + b_2\omega^4 + b_1\omega^2 + b_0 = 0, \tag{57.12}$$

where

$$\begin{aligned}
b_0 &= (k_1\alpha_1\gamma_1 - \nu k_1)^2[k_1\alpha_1\gamma_1 + \nu(\alpha_1 s - k_1)]^2 - {k_1}^4(\alpha_1\gamma_1 - \nu)^4,\\
b_1 &= 2{k_1}^2(\alpha_1\gamma_1 - \nu)[\nu(\alpha_1 s - k_1) + k_1\alpha_1\gamma_1][(-\alpha_1 s + k_1 - \nu\\
&\quad + \alpha_1\gamma_1) - (\alpha_1\gamma_1 - \nu)] + {k_1}^2[(\alpha_1\gamma_1 - \nu)(-\alpha_1 s + k_1 - \nu + \alpha_1\gamma_1)\\
&\quad - \nu(\alpha_1 s - k_1) - k_1\alpha_1\gamma_1)]^2 - 2{k_1}^4(\alpha_1\gamma_1 - \nu)^2,\\
b_2 &= 2{k_1}^2[(\alpha_1\gamma_1 - \nu)(-\alpha_1 s + k_1 - \nu + \alpha_1\gamma_1) - \nu(\alpha_1 s + k_1)\\
&\quad - k_1\alpha_1\gamma_1)] + {k_1}^2[(\nu - \gamma_1\alpha_1) + (-\alpha_1 s + k_1 - \nu + \alpha_1\gamma_1)]^2 - {k_1}^4.
\end{aligned}$$

Denote $\zeta = \omega^2$, (57.12) becomes

$$ {k_1}^2\zeta^3 + b_2\zeta^2 + b_1\zeta + b_0 = 0. \tag{57.13}$$

Let

$$\varphi(\zeta) = {k_1}^2\zeta^3 + b_2\zeta^2 + b_1\zeta + b_0. \tag{57.14}$$

Since $\frac{Lim}{\zeta\to+\infty}\varphi(\zeta) = +\infty$,, we can conclude that Eq. (57.13) has at least one positive real root if $b_0 < 0$. The roots of (57.13) can be calculated when $\alpha_1, \alpha_2, \beta_1, c, \gamma_1, \nu, k_1$ of the system (57.6) are given. Without loss of generality, supposing, that it has three positive roots, denoted by $\zeta_1, \zeta_2, \zeta_3$, respectively. Then Eq. (57.12) have three positive roots $\omega_i = \sqrt{\zeta_i}, i = 1, 2, 3$.

In view of (57.11), we have

$$\begin{aligned}
\tau_i^k = \frac{1}{\omega_i}\Bigg\{ar\cos\Bigg[&\frac{1}{{k_1}^2(\alpha_1\gamma_1 - \nu)^2 + (k_1\omega_i)^2}(k_1{\omega_i}^2(-\alpha_1 s + k_1 - \nu + \alpha_1\gamma_1)\\
&+ k_1(\alpha_1\gamma_1 - \nu)(-{\omega_i}^2 + \nu(\alpha_1 s - k_1) + k_1\alpha_1\gamma_1))\Bigg] + 2k\pi\Bigg\}.
\end{aligned} \tag{57.15}$$

where $i = 1, 2, 3, \quad k = 0, 1, 2, \ldots$, thus $\pm\omega_i$ are a pair of purely imaginary roots of Eq. (57.8) with τ_i^k.

Define

$$\tau_0 = {\tau^0}_{i_0} = \frac{\min}{i = 1, 2, 3}\tau_i^0, \quad \omega_0 = \omega_{i_0}. \tag{57.16}$$

Taking the derivative of τ in (57.8), it is easy to obtain:

$$\begin{aligned}
2\lambda\frac{d\lambda}{d\tau} &+ (\nu - \gamma_1\alpha_1 + \alpha_1 s - k_1 + k_1 e^{-\lambda\tau})\frac{d\lambda}{d\tau}\\
&+ k_1(\lambda + \nu - \gamma_1\alpha_1)\left(-\lambda - \tau\frac{d\lambda}{d\tau}\right).
\end{aligned} \tag{57.17}$$

From (57.17), by simple computation, we can derive

$$re\left[\left(\frac{d\lambda}{d\tau}\right)|_{\lambda=\omega_0 i,\tau=\tau_0}\right] = \frac{d_{11}d_{13} + d_{12}d_{14}}{{d_{13}}^2 + {d_{14}}^2}, \tag{57.18}$$

where

$$\begin{aligned}
d_{11} &= k_1\omega_0[(\nu - \gamma_1\alpha_1)\sin\omega_0\tau_0 - \omega_0\cos\omega_0\tau_0],\\
d_{12} &= k_1\omega_0[(\nu - \gamma_1\alpha_1)\cos\omega_0\tau_0 + \omega_0\sin\omega_0\tau_0],\\
d_{13} &= \nu - \gamma_1\alpha_1 + \alpha_1 s + k_1(-1 + \cos\omega_0\tau_0)\\
&\quad - k_1\tau_0[(\nu - \gamma_1\alpha_1)\cos\omega_0\tau_0 + \omega_0\sin\omega_0\tau_0],\\
d_{14} &= 2\omega_0 - k_1\sin\omega_0\tau_0 - k_1\tau_0[\omega_0\cos\omega_0\tau_0\\
&\quad - (\nu - \gamma_1\alpha_1)\sin\omega_0\tau_0].
\end{aligned}$$

Clearly, if ${d_{13}}^2 + {d_{14}}^2 \neq 0$, we can conclude that

$$sign\left[re\left(\frac{d\lambda}{d\tau}\right)\Bigg|_{\lambda=\omega_0,\tau=\tau_0}\right] = sign\left[re\left(\frac{d\lambda}{d\tau}\right)^{-1}\Bigg|_{\lambda=\omega_0,\tau=\tau_0}\right].$$

Now, a result can be quoted from Ruan and Wei (2003) to analyze (57.8). Expressed as follow:

Lemma 57.1 *The exponential polynomial*

$$\begin{aligned}
P(\lambda, e^{-\lambda\tau_1}, .., e^{-\lambda\tau_m}) &= \lambda^n + p_1^{(0)}\lambda^{n-1} + \ldots + p_{n-1}^{(0)}\lambda + p_n^{(0)}\\
&\quad + (p_1^{(1)}\lambda^{n-1} + \ldots + p_{n-1}^{(1)}\lambda + p_n^{(1)})e^{-\lambda\tau_1} + \ldots +\\
&\quad (p_1^{(m)}\lambda^{n-1} + \ldots + p_{n-1}^{(m)}\lambda + p_n^{(m)})e^{-\lambda\tau_m}.
\end{aligned} \tag{57.19}$$

where $\tau_i \geq 0 (i = 1, 2, \ldots, m)$, *and* $p_j^{(i)} (i = 0, 1, 2, \ldots, m; j = 1, 2, \ldots, n)$ *are constants. For* $(\tau_1, \tau_2, \ldots, \tau_m)$ *vary, the sum order of the zero of* $P(\lambda, e^{-\lambda\tau_1}, .., e^{-\lambda\tau_m})$ *on the open right half plane can change only if a zero appears on or cross the imaginary axis.*

Theorem 57.2 *Assume the following conditions are satisfied:*

(P1) *if* $\alpha_1 s - \gamma_1\alpha_1 + \nu > 0$, Eq. (57.8) *has four roots with negative real parts, and system* (57.6) *is stable near the equilibrium point when* $\tau = 0$.

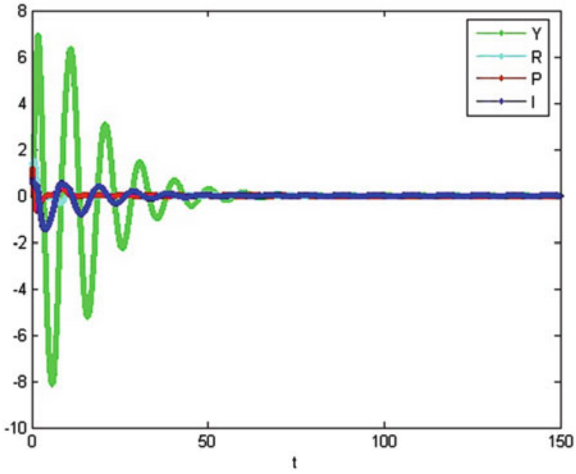
(P2) $re(\frac{d\lambda}{d\tau}) \neq 0$.

Thus the following results can be obtained:

(I) *For* Eq. (57.6), *its solution is asymptotical stable for*

$$\tau \in [0, \tau_0);$$

Fig. 57.5 $\tau = 0.105 < \tau_0$



(II) *For* Eq. (57.6), *its solution will become unstable if* $\tau > \tau_0$.
(III) Eq. (57.6) *occurs a Hopf bifurcation for the solution when* $\tau = \tau_0$.

That is to say, system (57.6) has a branch of periodic solutions near $\tau = \tau_0$.

Finally, numerical simulations of system (57.6) are used to prove the Theorem 57.2.

We choose $\alpha_1 = 0.98, \alpha_2 = 0.85, k_1 = 0.78, s = 0.25, \beta_1 = 0.95, c = 0.92, \gamma_1 = 0.35, \nu = 0.15$, system (57.6) becomes into

$$\begin{cases} \dot{Y}(t) = 0.98[I(t) - 0.25Y(t)] - 0.85[-R(t) \\ \qquad\qquad - 0.92P(t)]Y(t) + 0.78(Y(t) - Y(t-\tau)), \\ \dot{R}(t) = 0.95P(t) + (\dot{I}(t) - 0.25Y(t))R(t), \\ \dot{P}(t) = -R(t) - 0.92P(t), \\ \dot{I}(t) = 0.35\{0.98[I(t) - sY(t)] - 0.85[-R(t) \\ \qquad\qquad - 0.92P(t)]Y(t)\} - 0.15I(t) - R^2(t)I(t). \end{cases} \tag{57.20}$$

The next part, we consider system (57.20), the equation of (57.12) has two positive solutions $\zeta_{10}^{+} = 0.17117953544781, \zeta_{20}^{+} = 0.65091748067160$.

According to (57.15), we obtain $\tau_1^k = 1.15375402769439 + 3.07258609484024$ kπ, (k = 0,1,2,3, ...), $\tau_2^k = 36.36824488071419 + 11.68363960544670$ kπ, (k = 0,1,2,3, ...). From the definition of (57.16), we can see that $\tau_0 = 1.15375402769439$.

First of all, choosing $\tau = 0.105 < \tau_0$, Fig. 57.5 shows the corresponding waveform plots of system (57.20). By Theorem 57.2, we can see that the solution is asymptotically stable.

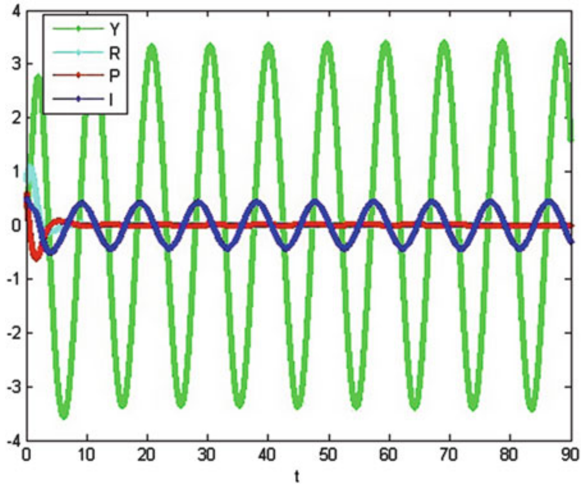


Fig. 57.6 $\tau = \tau_0 = 1.15375402769439$

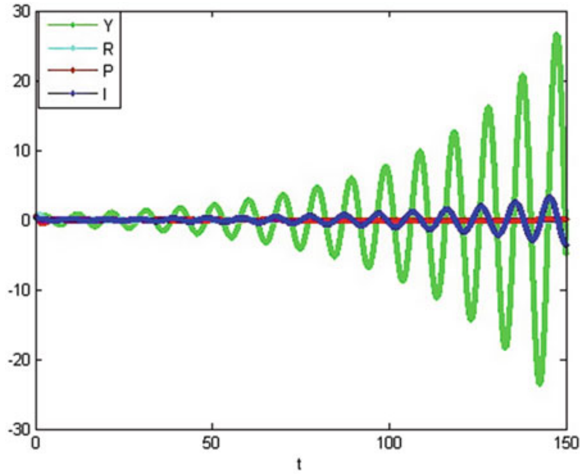


Fig. 57.7 $\tau = 1.2 > \tau_0$

Then, we choose $\tau = \tau_0 = 1.15375402769439$, Hopf bifurcation occurs. System (57.20) change from an equilibrium state to a cycle state, the variable $R(t)$ and $P(t)$ are still stable, but the variables Y(t) and I(t) are period, as waveform plots shown in Fig. 57.6.

Finally, we choose $\tau = 1.2 > \tau_0$, the corresponding waveform plots is shown in Fig. 57.7. We can see that system (57.20) undergo a Hopf bifurcation as shown in Fig. 57.7, R(t) and P(t) are still stable, but the variables Y(t) and I(t) are larger and larger.

57.4 Conclusion

In this paper, first of all, base on the IS-LM model which has been studied in Cai (2005), Cesare and Sportelli (2005), Neamtu (2007) and the model of (57.1), which has been researched in Gao and Ma (2009), Ma et al. (2007), and Chen (2008a, b), a four-dimensional economic system has been established. First, choosing the saving rate s as the bifurcation parameter, the stability and Hopf bifurcation have been studied. With the decreasing of the saving rate s, the fluctuation of the economic system is more and more fierce. Whereas if the saving rate s keeps at a proper high level, the energy of the system will be dissipated. This is coincident with the actual situation. It means that the saving rate s must keep at a proper high level to make sure the quick and healthy development of the economy. Then, a numerical example has been given to confirm derived results in Sect. 57.2. For predigest the production $Y(t)$, a time-delayed feedback is added to the first equation of system (57.2). Furthermore, the effect of time delay on the economic system has been researched. Taking the time delay τ as the bifurcation parameter, it is found that the equilibrium point will lose its stability and Hopf bifurcation arises if the time delay τ passes a certain critical value.

References

Cai J (2005) Hopf bifurcation in the IS-LM business cycle model with time delay. Electron J Differ Equ 15:1–6

Cesare LD, Sportelli M (2005) A dynamic IS-LM model with delayed taxation revenues. Chaos Soliton Fract 25(1):233–244

Chen Y (2007) Stability and Hopf bifurcation analysis in a three-level food chain system with delay. Chaos Soliton Fract 31(3):683–694

Chen WC (2008a) Dynamics and control of a financial system with time-delayed feedbacks. Chaos Soliton Fract 37(4):1198–1207

Chen WC (2008b) Nonlinear dynamics and chaos in a fractional-order financial system. Chaos Soliton Fract 36(5):1305–1314

Fanti L (2004) Fiscal policy and tax collection lags: stability, cycles and chaos. Riv Int Sci Econ Commer 51:341–365

Gao Q, Ma J (2009) Chaos and Hopf bifurcation of a finance system. Nonlinear Dyn 58(1–2):209–216

Li Z, Tang Y, Hussein S (2002) Stability and Hopf bifurcation for a delay competition diffusion system. Chaos Soliton Fract 14(8):1201–1225

Ma J, Gao Q (2007) Analysis and simulation of chaotic character of business cycle model. DCDIS Ser B 14:310–315

Ma J, Sun T, Wang Z (2007) Hopf bifurcation and complexity of a kind of economic system. Int J Nonlinear Sci Numeri Simul 8:347–352

Neamtu M (2007) Hopf bifurcation in a dynamic IS-LM model with time delay. Chaos Soliton Fract 34(2):519–530

Ruan S, Wei J (2003) On the zeros of transcendental functions with applications to stability of delay differential equations with two delays. Dyn Contin Discret Impuls Syst Ser A Math Anal 10:863–874

Szydlowski M, Krawiec A (2005) The stability problem in the Kaldor-Kalecki business cycle model. Chaos Soliton Fract 25(2):229–305

Yan X, Li W (2006) Hopf bifurcation and global periodic solutions in a delayed predator-prey system. Appl Math Comput 177(1):427–445

Zhang C, Wei J (2004) Stability and bifurcation analysis in a kind business cycle model with delay. Chaos Soliton Fract 22(4):883–896

Chapter 58
Spline Estimation for a Class of Time Series Variance Model

Xin-qian Wu, Wan-cai Yang, and Shu-hong Zhang

Abstract A class of nonparametric variance model with weakly stationary linear innovation process is considered in this paper. Based on polynomial spline method, optimal global rate of convergence of the estimator of nonparametric variance function is obtained. The methodology is illustrated by simulation and real data examples.

Keywords Global convergence • Linear innovation process • Nonparametric variance model • Spline estimation

58.1 Introduction

Heteroscedasticity is often encountered in many fields such as economics and finance. A time-varying nonparametric variance model is the form of

$$Y_i = \sigma(x_i)\,\varepsilon_i, \quad i = 1, \cdots, n, \tag{58.1}$$

where $x_i = i/n$ are fixed design points, Y_i are response variables, $\{\varepsilon_i\}$ is an unobserved random variable sequence with mean 0 and variance 1, $\sigma^2(x) > 0$ is an unknown smoothing function and is referred to as variance function.

Many authors studied the estimate of variance function under independent sequence $\{\varepsilon_i\}$. See, for example, Müller and Stadtmüller (1987), Hall and Carroll (1989), Brown and Levine (2007), Wang et al. (2008), Cai et al. (2009), and Boente et al. (2010). The above works focus on kernel smoothing and difference sequence methods. However, the independence condition is not always appropriate

X. Wu (✉) • W. Yang • S. Zhang
School of Mathematics and Statistics, Henan University of Science and Technology, Luoyang, People's Republic of China
e-mail: wuxinqian1001@163.com

E. Qi et al. (eds.), *The 19th International Conference on Industrial Engineering and Engineering Management*, DOI 10.1007/978-3-642-37270-4_58,

in applications. Under the strongly mixing dependent sequence $\{\varepsilon_i\}$, Dahl and Levine (2006) constructed a local linear estimator of $\sigma^2(x)$ based on the difference approach and obtained mean square convergence and asymptotic normality. Guo and Koul (2007) discussed the asymptotic theories of kernel estimator of $\sigma^2(x)$ under long memory errors. Wu et al. (2009a), and Wu and Yang (2009) discussed global and uniform rate of convergence of the spline estimator of variance function, and gave the bias and variance.

In this paper, we consider (58.1) with general weakly stationary linear innovation process

$$\varepsilon_i = \sum_{j=-\infty}^{\infty} \varphi_j e_{i-j}, \tag{58.2}$$

where $\sum_{j=-\infty}^{\infty} \left|\varphi_j\right| < \infty$, and $\{e_i\}$ is a sequence of independent and identically distributed (i.i.d.) random variables with mean 0 and variance σ_e^2. Here the α -mixing condition is not needed for the innovation sequence. Based on the polynomial spline estimate of variance function $\sigma^2(x)$ constructed in Wu et al. (2009a) and Wu and Yang (2009), we discuss the optimal global rate of convergence and give simulation results. Also, an application to the Shanghai composite index series is analyzed.

58.2 Polynomial Spline Estimation

It follows from (58.1) that

$$Y_i^2 = \sigma^2(x_i) + \sigma^2(x_i)\left(\varepsilon_i^2 - 1\right). \tag{58.3}$$

Note that $\sigma^2(x)$ can be approximated well by a spline function $\sum_{s=1}^{K} \beta_s B_s(x) \in S_{m,k}$, where $S_{m,k}$ is a corresponding space of polynomial splines on the interval $D = [0, 1]$ with degree m, the number of equally spaced interior knots $k-1$ and a set of basis functions $B_s(x)$, β_s are constants, and $K = m + k$. So, the least squares polynomial spline estimator of $\sigma^2(x)$ in Wu et al. (2009a) and Wu and Yang (2009) is defined as

$$\hat{\sigma}^2(x) = \sum_{s=1}^{K} \hat{\beta}_s B_s(x), \tag{58.4}$$

where

$$\hat{\beta} = \left(\hat{\beta}_1, \cdots, \hat{\beta}_K\right)^{\mathrm{T}} = (\mathrm{X}^{\mathrm{T}}\mathrm{X})^{-1}\mathrm{X}^{\mathrm{T}} Z. \tag{58.5}$$

Here $Z = (Z_1, \cdots, Z_n)^{\mathrm{T}} = \left(Y_1^2, \cdots, Y_n^2\right)^{\mathrm{T}}$, X is a $(n \times K)$ -matrix with the i-th row and the s-th column element being $B_s(x_i)$.

58.3 Main Results

For any function $w(x)$, let

$$\|w\|_2 = \left\{ \int_D w^2(x) dx \right\}^{1/2}.$$

The following assumptions are needed for our theoretical results (Wu et al. 2009a, 2011; Wu and Yang 2009; Burman 1991).

(i) The function $\sigma(x)$ is q time differentiable and the q th derivative of $\sigma(x)$ satisfies the following Hölder condition:

$$\left|\sigma^{(q)}(x) - \sigma^{(q)}\left(x'\right)\right| \le c\left|x - x'\right|^{v}, x, x' \in D,$$

where $c > 0$ is a constant, $0 < v \le 1$, $p = q + v > 0.5$.

(ii) $K = O(n^r)$, $0 < r < 1$.

(iii) $\sum_{j=-\infty}^{\infty} \left|\varphi_j\right| < \infty$. It implies that $\sum_{j=-\infty}^{\infty} \varphi_j^2 < \infty$ and $\sum_{j=-\infty}^{\infty} \varphi_j^4 < \infty$.

(iv) $E(e_i) = 0$, $E\left(e_i^2\right) = \sigma_e^2$, $0 < \sigma_e < \infty$, and $E\left(|e_i|^4\right) < +\infty$ $(i = 1, \cdots, n)$.

Theorem *If $m \ge p$, the assumptions (i)–(iv) hold and $\sigma_e^2 \sum_{j=-\infty}^{\infty} \varphi_j^2 = 1$, then*

$$\left\|\hat{\sigma}^2 - \sigma^2\right\|_2 = O_P\left((K/n)^{1/2} + K^{-p}\right). \tag{58.6}$$

Furthermore, $\left\|\hat{\sigma}^2 - \sigma^2\right\|_2 = o_P(1)$.

Proof Let

$$\Lambda = \mathrm{diag}\left(\sigma^2(x_1), \cdots, \sigma^2(x_n)\right), \eta = (\eta_1, \cdots, \eta_n)^{\mathrm{T}},$$

where $\eta_i = \varepsilon_i^2 - 1$. Write $c_1, c_2, c_3, \cdots$ are generic positive constants. According to the line of the proofs of Theorem in Wu et al. (2009a) and Theorem 3.1 in Wu et al. (2009b), it suffices to show that

$$\eta^{\mathrm{T}} \Lambda \mathrm{X} \mathrm{X}^{\mathrm{T}} \Lambda^{\mathrm{T}} \eta = O_P(n). \tag{58.7}$$

Obviously,

$$
\begin{aligned}
&E\left(\eta^{\mathrm{T}}\Lambda \mathrm{X}\mathrm{X}^{\mathrm{T}}\Lambda^{\mathrm{T}}\eta\right)\\
&= E\left\{\sum_{s=1}^{K}\left[\sum_{i=1}^{n} B_s\left(x_i\right)\sigma^2\left(x_i\right)\eta_i\right]^2\right\}\\
&= \sum_{s=1}^{K}\sum_{i=1}^{n} E\left\{B_s^2\left(x_i\right)\sigma^4\left(x_i\right)\eta_i^2\right\}\\
&\quad + 2\sum_{s=1}^{K}\sum_{1\le i<j\le n} E\left\{B_s\left(x_i\right)\sigma^2\left(x_i\right)\eta_i B_s\left(x_j\right)\sigma^2\left(x_j\right)\eta_j\right\}\\
&= I_1 + 2I_2.
\end{aligned}
$$

It follows from (58.9) and (58.10) in Wu et al. (2011) that

$$
\begin{aligned}
E\left(\eta_i^2\right) &\le E\left(e_i^4\right)\sum_{j=-\infty}^{\infty}\varphi_j^4 + \left(\sigma_e^2\sum_{j=-\infty}^{\infty}\varphi_j^2\right)^2\\
&\quad + 2\sigma_e^2\sum_{j=-\infty}^{\infty}\varphi_j^2 + 1 = E\left(e_i^4\right)\sum_{j=-\infty}^{\infty}\varphi_j^4 + 4 < \infty
\end{aligned}
$$

and

$$
\begin{aligned}
\left|E\left(\eta_i\eta_j\right)\right| &\le \left[\mathrm{Var}\left(e_0^2\right) + \sigma_e^4\right]\sum_{l=-\infty}^{\infty}\varphi_l^2\varphi_{j-i+l}^2\\
&\quad + \sigma_e^2\sum_{l=-\infty}^{\infty}\left|\varphi_l\varphi_{j-i+l}\right|.
\end{aligned}
$$

Note that

$$
\sum_{s=1}^{K} B_s^2(x) \le 1 \text{for all } x \in D.
$$

It follows from Assumptions (i) and (iv) that

$$
\begin{aligned}
I_1 &= \sum_{s=1}^{K}\sum_{i=1}^{n} B_s^2\left(x_i\right)\sigma^4\left(x_i\right)E\left(\eta_i^2\right)\\
&\le c_1 n\max_{x\in D}\ \sigma^4(x) = O(n).
\end{aligned}
$$

By Assumptions (i), (iii) and (iv), we have

$$
\begin{aligned}
|I_2| &\le \max_{x\in D} \sigma^4(x) \sum_{1\le i<j\le n} \left|E\left(\eta_i\eta_j\right)\right| \\
&\quad \cdot \left\{ \left[\sum_{s=1}^{K} B_s^2(x_i) \right]^{1/2} \left[\sum_{s=1}^{K} B_s^2\left(x_j\right) \right]^{1/2} \right\} \\
&\le c_2 \sum_{1\le i<j\le n} \left\{ \left[\operatorname{Var}\left(e_0^2\right) + \sigma_e^4 \right] \sum_{l=-\infty}^{\infty} \varphi_l^2 \varphi_{j-i+l}^2 \right. \\
&\quad \left. + \sigma_e^2 \sum_{l=-\infty}^{\infty} \left|\varphi_l \varphi_{j-i+l}\right| \right\} \\
&\le c_3 \sum_{i=1}^{n} \sum_{l=-\infty}^{\infty} \varphi_l^2 \sum_{j=-\infty}^{\infty} \varphi_{j-i+l}^2 \\
&\quad + c_4 \sum_{i=1}^{n} \sum_{l=-\infty}^{\infty} |\varphi_l| \sum_{j=-\infty}^{\infty} \left|\varphi_{j-i+l}\right| = O(n).
\end{aligned}
$$

Thus,

$$
E\left(\eta^{\mathrm{T}} \Lambda \mathrm{X} \mathrm{X}^{\mathrm{T}} \Lambda^{\mathrm{T}} \eta\right) \le I_1 + 2\,|I_2| = O(n).
$$

Therefore, (58.7) holds. This completes the proof of the theorem.

By (58.6), we can obviously get the optimal global rate of convergence of the estimator of variance function $\sigma^2(x)$ under the appropriate order of K. See the following corollary in details.

Corollary *Under the conditions in the above theorem and*

$$
K = O\left(n^{1/(2p+1)}\right),
$$

then

$$
\left\|\hat{\sigma}^2 - \sigma^2\right\|_2 = O_P\left(n^{-p/(2p+1)}\right). \tag{58.8}
$$

58.4 Simulation

The data are generated from (58.1) with

$$
\begin{aligned}
\sigma(x) &= 0.4 \sin(\pi x) + 0.01 \quad \text{for } x \in [0, 1], \\
\varepsilon_i &= 0.5\varepsilon_{i-1} + e_i
\end{aligned}
$$

and e_i is an i.i.d. sequence with common distribution, i.e.,

$$P\left(e_i = -\sqrt{3}/2\right) = P\left(e_i = \sqrt{3}/2\right) = 0.5.$$

So that $E\,(e_i) = 0$ and $\sigma_e^2 = 3/4$. ε_i can be rewritten as a general linear process as follows:

$$\varepsilon_i = \sum_{j=-\infty}^{\infty} \varphi_j e_{i-j}, \quad \varphi_j = \begin{cases} 0, & j \le -1, \\ 2^{-j}, & j \ge 0. \end{cases}$$

It is well known that ε_i is absolutely continuous and does not satisfy the strong mixing condition (Fan and Yao 2006).

Taking $\varepsilon_0 = 0$ and cubic spline with the B-spline basis which is used to construct the spline estimate of $\sigma^2(x)$. By the Corollary, K is selected from

$$\left[0.1n^{1/(2q+3)}, 3n^{1/(2q+1)}\right]$$

with $q = 2$ based on AIC and BIC, i.e.,

$$\begin{aligned} \text{AIC} &= \ln\left(\text{RSS}/n\right) + 2K/n, \\ \text{BIC} &= \ln\left(\text{RSS}/n\right) + \ln(n) \cdot K/n, \end{aligned} \tag{58.9}$$

where

$$\text{RSS} = \sum_{i=1}^{n} \left(Y_i^2 - \hat{\sigma}^2\left(x_i\right)\right)^2$$

is the residual sum of squares. The performance of the function estimator is assessed by

$$\text{RASE} = \left\{ n_{\text{grid}}^{-1} \sum_{j=1}^{n_{\text{grid}}} \left[\hat{\sigma}^2\left(x_{\text{grid},j}\right) - \sigma^2\left(x_{\text{grid},j}\right)\right]^2 \right\}^{1/2}, \tag{58.10}$$

where

$$\left\{x_{\text{grid},j}, j = 1, \cdots, n_{\text{grid}}\right\}$$

is a set of equally spaced grid points where the function is evaluated, and the number of grid points $n_{\text{grid}} = 101$.

In simulation runs, sample sizes are $n = 100$, 200 and 400, the number of replications is 100. The means and standard errors (in the parentheses) of RASEs

Table 58.1 Means and standard errors (in parentheses) of RASEs for the estimates of $\sigma^2(x)$

	Spline fit		
n	AIC	BIC	Local linear fit
100	0.0249 (0.0104)	0.0248 (0.0086)	0.0255 (0.0096)
200	0.0189 (0.0101)	0.0179 (0.0078)	0.0199 (0.0086)
400	0.0123 (0.0063)	0.0111 (0.0058)	0.0143 (0.0052)

of the spline estimate of $\sigma^2(x)$ are reported in Table 58.1. The mean RASEs of the estimate decrease as sample size increases, conforming to the main results. Also, the spline fits with BIC selected knot numbers give overall smaller mean RASEs. Figure 58.1 presents the plots of the spline estimated curves with BIC selected knot numbers, in which the solid curve represents the true function, the dotted, dash-dotted and dashed curves correspond to the average, minimum and maximum fitted functions respectively. Visually, the estimate errors decrease as sample size increases.

In order to comparing kernel method with the spline method, Table 58.1 also gives the similar results of the local linear estimate with Epanechnikov kernel and cross-validation bandwidth rule. It shows that the spline fits give overall smaller mean RASEs. Moreover, the polynomial spline method enjoys great computational efficiency. It takes about 1 min to run 100 simulations with $n = 400$ on a Pentium 4 PC. However, it takes the local linear method about 45 min to do the same thing.

58.5 Application

We consider the opening price series $\{Y_{0,i}\}$ of the Shanghai composite index from July 1, 2010 to June 30, 2011, a total of 243 observations (Fig. 58.2). The data can be obtained at the website of http://vip.stock.finance.sina.com.cn/corp/go.php/vMS_MarketHistory/stockid/000001/type/S.phtml. We are interested in the return series $\{Y_i\}$ with 242 data points (Fig. 58.3), where

$$Y_i = Y_{0,i} - Y_{0,i-1}.$$

We fit a nonparametric variance model (58.1). The variance function $\sigma^2(x)$ is estimated by cubic B-spline. The optimal K is automatically selected according to the method in Sect. 58.4. Figure 58.4 gives the estimated $\sigma^2(x)$. The residual sequence $\{\hat{\varepsilon}_i\}$ is displayed in Fig. 58.5, where $\hat{\varepsilon}_i = Y_i / \hat{\sigma}(x_i)$. The sample mean is 0.0297 and the sample variance is 1.0375. By the run test and the sample autocorrelation function and partial autocorrelation function, $\{\hat{\varepsilon}_i\}$ can be regarded as a stationary white noise process.

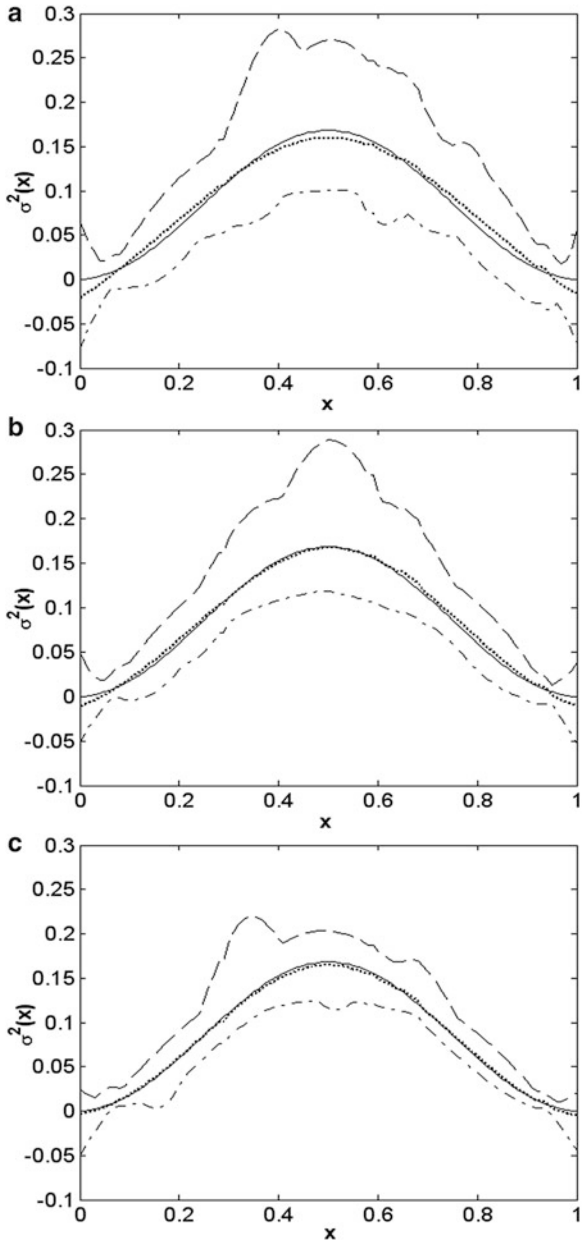


Fig. 58.1 Spline estimated curves with BIC (**a**) $n = 100$, (**b**) $n = 200$, (**c**) $n = 400$

Fig. 58.2 Shanghai composite index

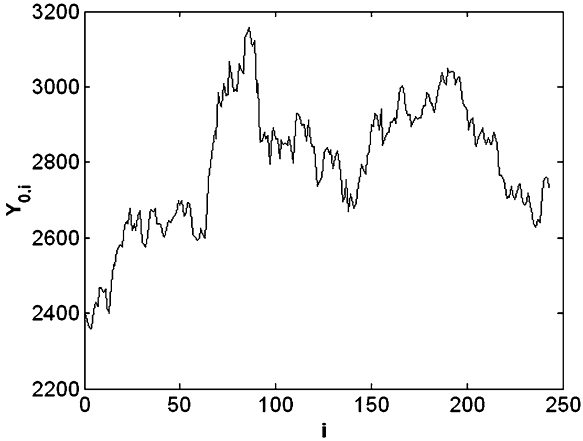


Fig. 58.3 Return data

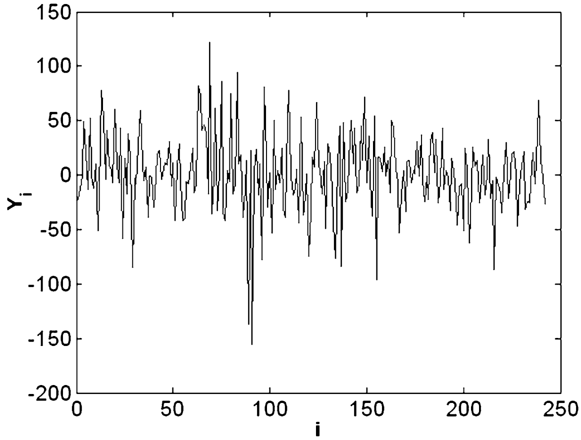
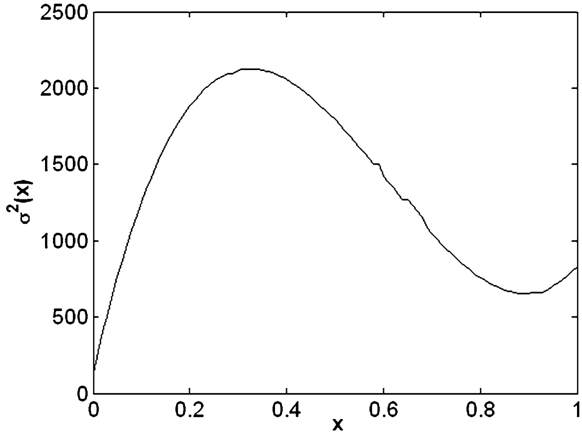


Fig. 58.4 Variance function



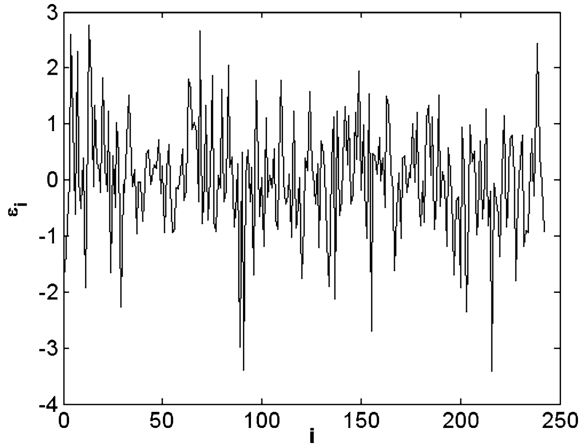


Fig. 58.5 Residual sequence

58.6 Conclusion

In this paper, we discussed the global rate of convergence of polynomial spline estimator of variance function in a class of nonparametric variance model with weakly stationary linear innovation process. The rate can attain to the optimal one under mild conditions, see Stone (1982). Simulation result shows that the spline method outperforms the local linear method.

Acknowledgment This research was partially supported by the Development Program of Science and Technology of Henan Province under Grant 092300410178, the Natural Science Foundations of Education Department of Henan Province under Grants 2010B110009 and 2008A110005, and the Scientific Research Foundation for Doctoral Program of Henan University of Science and Technology under Grant 09001322.

References

Boente G, Ruiz M, Zamar RH (2010) On a robust local estimator for the scale function in heteroscedastic nonparametric regression. Stat Probab Lett 80(15–16):1185–1195

Brown LD, Levine M (2007) Variance estimation in nonparametric regression via the difference sequence method. Ann Stat 35(5):2219–2232

Burman P (1991) Regression function estimation from dependent observations. J Multivar Anal 36(2):263–279

Cai TT, Levine M, Wang L (2009) Variance function estimation in multivariate nonparametric regression with fixed design. J Multivar Anal 100(1):126–136

Dahl CM, Levine M (2006) Nonparametric estimation of volatility models with serially dependent innovations. Stat Probab Lett 76(18):2007–2016

Fan J, Yao Q (2006) Nonlinear time series: nonparametric and parametric methods. Science Press, Beijing, pp 67–77

Guo H, Koul HL (2007) Nonparametric regression with heteroscedastic long memory errors. J Stat Plan Inference 137(2):379–404

Hall P, Carroll RJ (1989) Variance function estimation in regression: the effect estimation of the mean. J R Stat Soc Ser B 51(1):3–14

Müller HG, Stadtmüller U (1987) Estimation of heteroscedasticity in regression analysis. Ann Stat 15(2):610–625

Stone CJ (1982) Optimal global rates of convergence for nonparametric regression. Ann Stat 10(4):1040–1053

Wang L, Brown LD, Cai TT, Levine M (2008) Effect of mean on variance function estimation in nonparametric regression. Ann Stat 36(2):646–664

Wu X, Yang W (2009) Spline estimation for a nonparametric variance model under dependent innovations. In: Conference proceedings of the 2009 international Institute of Applied Statistics Studies, Sydney Australia, Aussino Academic Publishing House, vol II, pp 2737–2741

Wu X, Feng A, Tian P (2009a) Spline estimation for nonparametric variance models under dependent observations. J Henan Univ Sci Technol (Nat Sci) 30(5):93–96, 104 (in Chinese)

Wu X, Tian Z, Wang H (2009b) Polynomial spline estimation for nonparametric (auto-) regressive models. Studia Scientiarum Mathematicarum Hungarica 46(4):515–538

Wu X, Yang W, Li S (2011) Spline estimation for nonparametric variance models with linear process innovations. In: Proceedings of the 2011 international conference on consumer electronics, communications and networks, IEEE Press 6:5251–5254

Chapter 59
Application of GM (1, N)-Markov Model in Shanghai Composite Index Prediction

Wan-cai Yang, Xin-qian Wu, Er-xin Zhang, and Guang Zhu

Abstract In order to overcome the limitations of little used information and low accuracy for single stock market prediction model and the limitations of exponential trend for GM(1,1)-Markov combination forecast model, GM(1, N)-Markov model is suggested in this paper. Positive analysis is done for Shanghai composite index (monthly closing price). The results show that the established GM (1, 3)-Markov model outperforms the GM (1, 1) model and the GM (1, 1)-Markov model.

Keywords Combination forecast • GM (1, N) • Markov chain • Shanghai composite index

59.1 Introduction

With the development of social economy, stock has become an important tool for people to invest and manage finance. Stock market can be regarded as a grey system, in which some information is known and other information is unknown. But grey forecast is suitable for some objects with short time, a few data, little volatility and long-term tendency, and its forecast tendency is a smooth curve. For some data sequences with large stochastic volatility, the grey system usually gives bad fitting and low forecast accuracy. Note that transition probability matrix in Markov chain theory can reflect the influence extent of stochastic factors, and is suitable for dynamic process with large stochastic volatility. It can remedy the limitation of grey forecast. At the same time, grey system can remedy the weaknesses of un-follow-up effect and stationary which are used in Markov chain. If a grey system model is used

W. Yang (✉) • X. Wu • E. Zhang • G. Zhu
School of Mathematics and Statistics, Henan University of Science and Technology, Luoyang, People's Republic of China
e-mail: wancaiy-001@163.com

E. Qi et al. (eds.), *The 19th International Conference on Industrial Engineering and Engineering Management*, DOI 10.1007/978-3-642-37270-4_59,

to fit data and find its change tendency, the limitation of Markov forecast can be amended, see Tang et al. (2007). Thus, Combining grey system model with Markov chain has attracted much attention. At present, the popular combination model is GM (1, 1)-Markov model (Tang et al. 2007; Chen and Duan 2002; Tao et al. 2007; Li et al. 2007a, b; Liu et al. 2012; Liu 2011).

However, GM (1, 1)-Markov model often gives the result of exponential growth tendency, which fails to reflect the true dynamic characteristic of some real data. In this paper, GM (1, 1)-Markov model is extended to GM (1, N)-Markov model by adding some related influent factors. An application to Shanghai composite index (monthly closing price) is also analyzed.

59.2 GM (1, N)-Markov Model

59.2.1 GM (1, N) Model

Let

$$X_1^{(0)} = \left(x_1^{(0)}(1), x_1^{(0)}(2), \cdots x_1^{(0)}(n)\right)$$

be a sequence of system characteristic data. While

$$\begin{cases} X_2^{(0)} = \left(x_2^{(0)}(1), x_2^{(0)}(2), \cdots x_2^{(0)}(n)\right) \\ X_3^{(0)} = \left(x_3^{(0)}(1), x_3^{(0)}(2), \cdots x_3^{(0)}(n)\right) \\ \qquad\qquad \vdots \\ X_N^{(0)} = \left(x_N^{(0)}(1), x_N^{(0)}(2), \cdots x_N^{(0)}(n)\right) \end{cases}$$

are the related factor sequences. Write

$$X_i^{(1)} = \left(x_i^{(1)}(1), x_i^{(1)}(2), \cdots x_i^{(1)}(n)\right)$$

is the 1-AGO sequence of $X_i^{(0)}$ $(i = 1, 2, \cdots, N)$ and

$$Z_1^{(1)} = \left(z_1^{(1)}(2), z_1^{(1)}(3), \cdots, z_1^{(1)}(n)\right)$$

is the close neighbour average generation sequence of $X_1^{(1)}$. Then,

$$x_1^{(0)}(k) + a z_1^{(1)}(k) = \sum_{i=2}^{N} b_i x_i^{(1)}(k) \tag{59.1}$$

is called a GM(1,N) model, where $-a$ is a coefficient of system evolution, $b_i x_i^{(1)}(k)$'s are drive terms, b_i's are drive coefficients.

Let $\hat{a} = (a, b_2, \cdots, b_N)^T$,

$$B = \begin{bmatrix} -z_1^{(1)}(2) & x_2^{(1)}(2) & \cdots & x_N^{(1)}(2) \\ -z_1^{(1)}(3) & x_2^{(1)}(3) & \cdots & x_N^{(1)}(3) \\ \cdots & \cdots & \cdots & \cdots \\ -z_1^{(1)}(n) & x_2^{(1)}(n) & \cdots & x_N^{(1)}(n) \end{bmatrix}, Y = \begin{bmatrix} x_1^{(0)}(2) \\ x_1^{(0)}(3) \\ \vdots \\ x_1^{(0)}(n) \end{bmatrix}.$$

Then, the least squares estimate of parameter $\hat{a}$ is defined as

$$\hat{a} = (B^T B)^{-1} B^T Y. \tag{59.2}$$

The approximate time response expression of GM (1, N) model (59.1) is the form of

$$\hat{x}_1^{(1)}(k+1) = \left(x_1^{(1)}(0) - \frac{1}{a} \sum_{i=2}^{N} b_i x_i^{(1)}(k+1) \right) e^{-ak} + \frac{1}{a} \sum_{i=2}^{N} b_i x_i^{(1)}(k+1), \tag{59.3}$$

where $x_1^{(0)}(0) = x_1^{(0)}(1)$. The corresponding degressive reduction expression is the form of

$$\hat{x}_1^{(0)}(k+1) = \hat{x}_1^{(1)}(k+1) - \hat{x}_1^{(1)}(k), \tag{59.4}$$

where $\hat{x}_1^{(0)}(k)$'s are simulated values.

GM (1, N) model was used in some fields (Wang 2007; Zhu and Wang 2010), but selection of variables and n is important. Based on multivariate regression theory, the value of N can be determined by the number of significant variables which influence Shanghai composite index, see Li and Wu (2010). For other methods, see also Yin and Luo (1999), etc.

59.2.2 Dividing State, Constructing Transition Probability Matrix and Testing "Unfollow-up Effect"

For the stochastic process varying with time, we should consider a "A ship rises with the tide" dynamic division criterion, which construct the relation between predicted values and simulated values. But they should be coincide with the elementary change trend of time series, see Tao et al. (2007). Because relative values vary less than residual values, relative values are main and residual division is auxiliary for division of states. States dividing are

$$E_i \in [\otimes_{1i}, \otimes_{2i}]\,(i = 1, 2, \cdots, m),$$

where state boundary grey elements

$$\otimes_{1i} = \alpha_{1i}\hat{x}_1^{(0)}(t),$$
$$\otimes_{2i} = \alpha_{2i}\hat{x}_1^{(0)}(t),$$

α_{1i} and α_{2i} are relative values of state division. Relative value is obtained by actual value dividing forecast value, i.e.,

$$\frac{x_1^{(0)}(t)}{\hat{x}_1^{(0)}(t)}. \tag{59.5}$$

The number of state, m, should be decided by some requirements of concrete object, e.g., the purpose, sample number, precision, etc. In this paper, the division criterion includes that every state has sample data and state space is divided as much as possible after unfollow-up effect is tested. Thus, the accuracy can be improved.

Let f_{ij} $(i, j \in E)$ denote the frequency from state i to state j by a step transition. Then, one step transition probability is that

$$p_{ij} = \frac{f_{ij}}{\sum_{i=1}^{m}\sum_{j=1}^{m} f_{ij}}.$$

The value of the sum of the j row elements divided by the sum of all elements in the transition frequency matrix, denoted by $p_{\cdot j}$, is called the marginal probability, i.e.,

$$p_{\cdot j} = \frac{\sum_{i=1}^{m} f_{ij}}{\sum_{i=1}^{m}\sum_{j=1}^{m} f_{ij}}.$$

When n is large enough, statistics

$$\chi^2 = 2\sum_{i=1}^{m}\sum_{j=1}^{m} f_{ij}\left|\ln\frac{p_{ij}}{p_{\cdot j}}\right| \tag{59.6}$$

is distributed according to χ^2 distribution with degree of freedom $(m-1)^2$. The value of statistics χ^2 can be computed from (59.6). Given the significance level α, the value of quantile $\chi^2_\alpha\left((m-1)^2\right)$ can be obtained from the χ^2 distribution table. If

$$\chi^2 > \chi^2_\alpha\left((m-1)^2\right), \tag{59.7}$$

then the series $\{x_i\}$ has the property of "unfollow-up effect". Otherwise, it can not be regarded as Markov chain.

The following calculation is on the basis of "unfollow-up effect" (Meyn and Tweedie 1993; Xia 2005).

59.2.3 Computing Predicted Value

After the state transition matrix of forecast object is constructed, the initial distribution π_0 is decided by the state of the last one in historical data. Based on multiplying k step transition matrix $P(k)$ by π_0, we can obtain the state probability of the k step predicted value, where

$$P(k) = P^k.$$

In the end, according to the state probability of the k step predicted value, we can obtain the expectation of the states for the k step relative value, i.e.,

$$E_{\text{rv}} = 0.5\sum_{i=1}^{m}(\otimes_{1i} + \otimes_{2i}) \times \pi_{ki}, \tag{59.8}$$

where π_{ki} stands for the i-th state probability of relative values. Because there is approximately equal probability for the k-th relative forecast value, this algorithm is not suitable for computing the interval forecast of relative values. In other words, it is only suitable for computing relative forecast values. Therefore,

$$\begin{aligned}\text{the final predicted value of object} = {} & \text{relative forecast value} \\ & \times \text{GM}(1, \text{N}) \text{ forecast value.}\end{aligned} \tag{59.9}$$

For the details of theory of grey system and its applications (Liu et al. 2004).

59.3 Application

We choose the Shanghai composite index (monthly closing price) between September 2008 and September 2010 as historical data, and choose the Shanghai composite index between October 2010 and March 2011 as verification data.

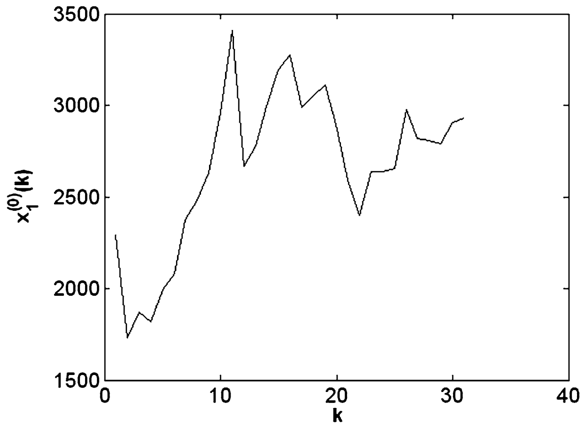


Fig. 59.1 Shanghai composite index

There is a total of 31 observations, see Fig. 59.1. The data can be obtained at the website of http://app.finance.ifeng.com/hq/stock_monthly.php?code=sh000001.

59.3.1 Establishing GM (1, N) Model

According to the result of multivariable regression analysis for correlated variables of influencing Shanghai composite index in Li and Wu (2010) and the similar method of avoiding multiple collinearity in regression analysis (see He and Liu 2007), we choose two main factors, exchange rate and ICBC closing price, from correlated variables of influencing Shanghai composite index. Accordingly, we choose the CNY/USD middle exchange rate and the closing price for ICBC stock price between September 2008 and September 2010 as correlated variable sequences, i.e.,

$$\begin{cases} X_2^{(0)} = \left(x_2^{(0)}(1), x_2^{(0)}(2), \cdots x_2^{(0)}(n)\right) \\ X_3^{(0)} = \left(x_3^{(0)}(1), x_3^{(0)}(2), \cdots x_3^{(0)}(n)\right) \end{cases},$$

where $X_2^{(0)}$ and $X_3^{(0)}$ stand for the sequences of CNY/USD middle exchange rate and closing price for ICBC stock price, respectively. There is a total of 25 observations in each sequence (Figs. 59.2 and 59.3). The data can be obtained at the website of

http://www.safe.gov.cn/model_safe/index.html
and
http://vol.stock.hexun.com/Close/share/HistoryDetail.aspx?rank=3&code=601398,

respectively. Thus, a GM (1, 3) model will be established.

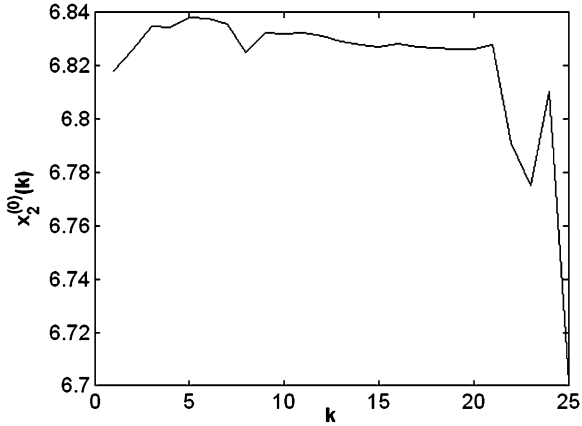


Fig. 59.2 CNY/USD middle exchange rate

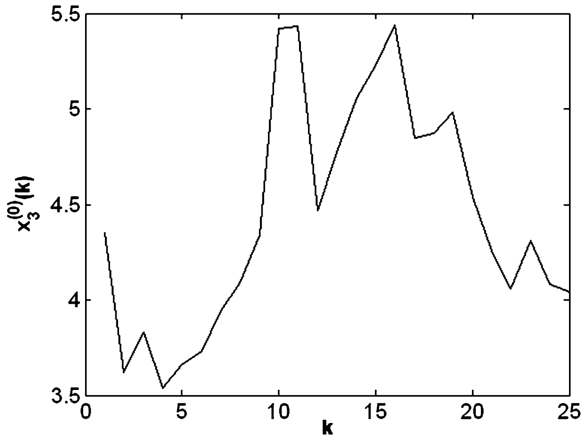


Fig. 59.3 Closing price for ICBC stock price

According to the method in Sect. 59.2, we have

$$\hat{x}_1^{(1)}(t)\,(k+1) = \left(x_1^{(1)}(0) - \frac{1}{a}\sum_{i=2}^{N} b_i x_i^{(1)}\,(k+1)\right)e^{-ak} + \frac{1}{a}\sum_{i=2}^{N} b_i x_i^{(1)}\,(k+1), \tag{59.10}$$

where

$$x_1^{(1)}(0) = x_1^{(0)}(0), a = 0.822696,$$
$$b_2 = -0.255108, b_3 = 536.697678.$$

Table 59.1 Dividing state criterion

Number		State
i	Relative Value Interval	E
1	[0.748, 0.8468]	$\left[0.748\hat{x}^{(0)}(t), 0.8468\hat{x}^{(0)}(t)\right]$
2	(0.8468, 0.9456]	$\left(0.8468\hat{x}^{(0)}(t), 0.9456\hat{x}^{(0)}(t)\right]$
3	(0.9456,1.0444]	$\left(0.9456\hat{x}^{(0)}(t), 1.0444\hat{x}^{(0)}(t)\right]$
4	(1.0444,1.1432]	$\left(1.0444\hat{x}^{(0)}(t), 1.1432\hat{x}^{(0)}(t)\right]$
5	(1.1432, 1.242]	$\left(1.1432\hat{x}^{(0)}(t), 1.242\hat{x}^{(0)}(t)\right]$

59.3.2 State, Transition Matrix and "Unfollow-up Effect"

Here we consider five states $E_1, E_2, E_{3,} E_4, E_5$. The interval length of every state is 1/5 of ratio range, see Table 59.1.

After the states is known, we can construct one-step state transition probability matrix P by one-step state transition probability formula. Here,

$$P = \begin{bmatrix} \frac{2}{3} & \frac{1}{3} & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & \frac{1}{17} & \frac{14}{17} & \frac{1}{17} & \frac{1}{17} \\ 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Since the data we choose comes from the period from the financial crisis to the economic recovery, the stock market fluctuated intensely, and there existed some "outliers", therefore, some lines in the transition probability matrix only have one probability value. Here we do not modify them in order to predict the stock mark with big volatility.

The value of statistics χ^2 is 28.2182. Because

$$28.2182 > \chi^2_{0.05}((m-1)^2) = \chi^2_{0.05}((5-1)^2) = \chi^2_{0.05}(16) = 26.2962,$$

the relative value sequence for the closing price of Shanghai composite index and its simulated value has Markov property.

59.3.3 Predicted Value

Let the relative value for the index's true value and its predicted value in the end of September 2010 be the initial state. The initial distribution is

$$\pi_1 = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 \end{pmatrix}.$$

Table 59.2 Predicted values for several models

Time	True Value	GM(1,1)	GM(1,1) Markov	GM(1,3) Markov
2010/10	2978.84	3078.64	2643.01	2884.17
2010/11	2820.18	3118.03	2726.98	2770.49
2010/12	2808.08	3157.92	2805.04	2800.85
2011/1	2790.69	3198.33	2879.39	2809.13
2011/2	2905.05	3239.25	2948.11	2816.73
2011/3	2928.11	3280.69	3011.79	2967.22

Table 59.3 Prediction accuracy

	GM(1,1)	GM(1,1) Markov	GM(1,3) Markov
Average relative error	10.75 %	3.7 %	1.71 %
Volatility of error (Variance of residual)	9632.078	12.719	1.213
Maximum relative error	14.61 %	11.27 %	3.18 %

Thus, the absolute distribution in the end of the next month is

$$\pi_2 = \pi_1 \cdot \mathrm{P} = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 \end{pmatrix},$$

by which the expectation of the relative value in the end of October 2010 is

$$(0.762 + 0.8778)/2 = 1.0938,$$

and the predicted value is $1.0938\hat{x}^{(0)}(26)$.

Similarly, by

$$\pi_n = \pi_1 \cdot \mathrm{P}^{(n-1)},$$

we can calculate the absolute distribution and obtain the relative value in the end of the corresponding month, that is to say,

$$\begin{aligned} \text{the final predicted value of object} = \text{relative forecast value} \\ \times \,\mathrm{GM}(1,3) \text{ forecast value.} \end{aligned} \tag{59.11}$$

The results are shown in Tables 59.2, 59.3 and Fig. 59.4. 1–6 in Fig. 59.4 stand for the corresponding months between October 2010 and March 2011.

It can be seen from Tables 59.2, 59.3 and Fig. 59.4 that the GM(1,3)-Markov model gives overall smaller relative errors and volatilities.

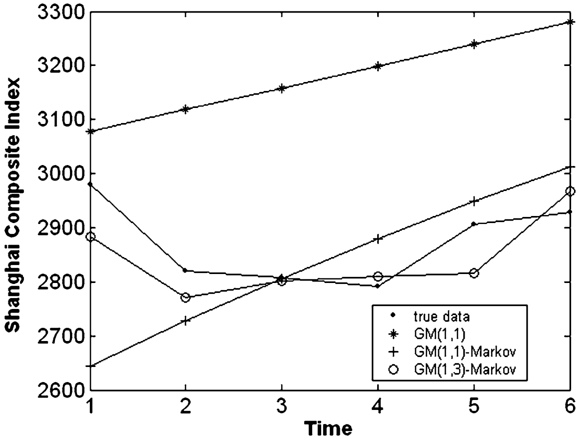


Fig. 59.4 True values and predicted values

59.4 Conclusion

In this paper, GM (1, N)-Markov model is suggested to forecast stoke price. We establish a GM (1, 3)-Markov model for Shanghai composite index. The results show that the presented model has higher accuracy in prediction.

Certainly, there are many methods to divide the state space. A special case is only considered in this paper. More thorough discussion is beyond of the content. Furthermore, the variables in GM (1, 3)-Markov model is selected by regression method, which results in a heavy burden of computation. How to find a good method for selecting variables will be discussed in the future.

Acknowledgment This research was partially supported by the Development Program of Science and Technology of Henan Province under Grant 092300410178, the Natural Science Foundations of Education Department of Henan Province under Grants 2010B110009 and 2008A110005, and the Scientific Research Fundation for Doctoral Program of Henan University of Science and Technology under Grant 09001322.

References

Chen H, Duan J (2002) Application of Gray-Markov model in stock price prediction. Econ Prob 24(8):37–39 (in Chinese)

He X, Liu W (2007) Applied regression analysis, 2nd edn. China Renmin University Press, Beijing, pp 170–179 (in Chinese)

Li Y, Wu G (2010) On influence factors of Shanghai composite index. China Mark 17(14):82–84 (in Chinese)

Li GD, Yamaguchi D, Nagai M (2007a) Application of GM (1, 1)-Markov chain combined model to China's automobile industry. Int J Ind Syst Eng 2(3):327–347

Li GD, Yamaguchi D, Nagai M (2007b) A GM(1,1) –Markov chain combined model with an application to predict the number of Chinese international airlines. Technol Forecast Soc Chang 74(8):1465–1481

Liu C (2011) Study of rainfall prediction model based on GM (1, 1)-Markov chain. In: 2011 international symposium on water resource and environmental protection (ISWREP), vol 1, pp 744–747

Liu S, Dang Y, Fang Z (2004) Grey system theory and its applications, 3rd edn. Science Press, Beijing, pp 125–158 (in Chinese)

Liu Y, Lin W, Yang MZ (2012) Application of improved GM (1, 1)-Markov chain model of business order prediction. Appl Mech Mater 143–144:468–472

Meyn SP, Tweedie RL (1993) Markov chains and stochastic stability. Springer, New York

Tang N, Gui Y, Li B (2007) Application of Gray-Markov model in analysis of stock index. In: Proceedings of the fifth annual conference on uncertainty, Wuhan, pp 195–198 (in Chinese)

Tao Y, Xu J, Li S (2007) Predict gas emissing quantity of mining coal face with improved grey Markov model. J China Coal Soc 32(4):391–395 (in Chinese)

Wang W (2007) Application of grey system GM (1, N) model to predicting spring flow. In: IEEE international conference on grey systems and intelligent services (GSIS 2007), pp 448–452

Xia L (2005) Research of Markov chain prediction method and its application on hydrology series. Doctoral dissertation, Hohai University, Nanjing, China, pp 31–32

Yin Z, Luo L (1999) Variable selection and error analyzing of grey model (1, n). Syst Eng Theor Pract 32(11):81–83, 120, November 1999 (in Chinese)

Zhu C, Wang B (2010) A study on GM (1, N) model for teaching evaluation. Adv Mater Res 121–122:341–345

Chapter 60
Research on Fixed-Point Maintenance Support Simulation of Mechanization Army Division Based on HLA

Chuang Gu, Xiao-ming Du, Ji-wei Cai, Bing Liu, and Gui-qi Wang

Abstract Modeling and simulation of the fixed-point maintenance support system of mechanization army division based on HLA was advanced to the requirement of evaluation of maintenance support efficiency. In order to increase correctness, reusability and interoperability of the system, the conceptual modeling technology, HLA technology and VV&A technology are used to develop the simulation system. Firstly, composition and running process of fixed-point maintenance support system of mechanization army division are analyzed in detail, and the conceptual model, mathematical model, index system and evaluation model of maintenance support efficiency are established. Secondly, the system frame and total running flow of the simulation system based on HLA are designed, and a VV&A process model for Federation Development and Execution Process is introduced. At last, the simulation system of fixed-point maintenance support efficiency evaluation of mechanization army division is developed to simulate the materiel maintenance support activities under the task background of the mechanization army division's attack warfare.

Keywords Conceptual model • Fixed-point support simulation • HLA • VV&A

60.1 Introduction

There are Petri, Agent, HLA and special simulation soft methods used to simulate materiel maintenance support system at present. As a universal technical frame, HLA is used widely for its reusability and interoperability.

C. Gu (✉) • X. Du • J. Cai • B. Liu • G. Wang
Department of Equipment Command and Administration, Ordnance Engineering College, Shijiazhuang, China
e-mail: guchuang2009chuang@163.com

E. Qi et al. (eds.), *The 19th International Conference on Industrial Engineering and Engineering Management*, DOI 10.1007/978-3-642-37270-4_60,

Literature Yan-zhong Zhang and Shi-ying Li (2006), Dong-nanWang (2006) and Jun-jie LIU and Ya-lin Ji (2006) discussed the problems about design, development and application of maintenance support simulation system based on HLA. Literature Jian-bing Tang (2007) and Yan-ping Fan and Yan-qi Bai (2007) discussed activities, process, content and methods of simulation based on HLA.

At present, there are many researchers about the concept modeling, materiel maintenance support simulation and evaluation of materiel maintenance support efficiency. However, the researchers about how to use concept modeling technology, HLA technology and VV&A technology together to develop maintenance support simulation system based on HLA are few, which is very important to increase correctness, reusability and interoperability of simulation system.

Therefore, this paper discuss the problem of concept modeling, mathematic modeling, VV&A process model for Federation Development and Execution Process, evaluation of fixed-point maintenance support efficiency of mechanization army division, design and development of the simulation system of fixed-point maintenance support efficiency evaluation of mechanization army division.

60.2 Mechanization Army Division's Equipment Supportment System

Mechanization army division's equipment support system is composed of division equipment organs, division warehouses, branch repair elements, and subordinate regimental equipment support systems.

60.2.1 The Maintenance Support Flow of Mechanization Army Division in the Wartime (Fig. 60.1)

60.2.2 The Fixed-Point Maintenance Support Flow of Mechanization Army Division in the Wartime

The process of mechanization army division's maintenance support in wartime generally includes support force receiving support tasks, choosing deployment region for support institution, disposing and unfolding support forces, rescuing and repairing damaged equipments, evacuating damaged equipments and diverting support forces (Wu-kui 2003).

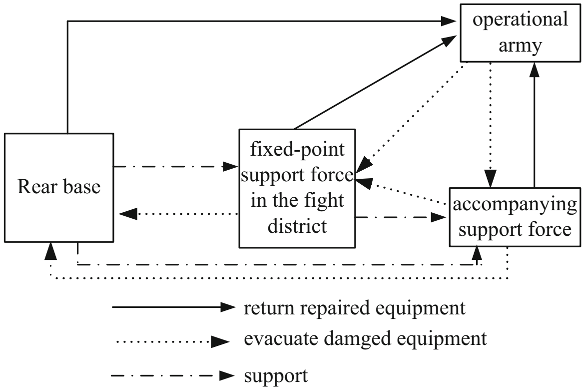


Fig. 60.1 The maintenance support flow of mechanization army division

60.2.3 Analysis of the Fixed-Point Maintenance Support Flow of Mechanization Army Division's in the Wartime

After fixed-point maintenance groups take over the damaged equipments delivered by furl groups, they evaluate the damaged equipments to determine the failure at first and decide maintenance strategies and repairable components. As soon as the damaged equipments are repaired, they should be delivered to operational troops. When the repair groups are busy, the damaged equipments wait in line according to the principle that first arrive, first repair, and repair lightly damaged equipments at first, and then repair damaged equipments.

60.3 Modeling of Mechanization Army Divsion's Fixed-Point Suport Process

60.3.1 Modeling of Mechanization Army Division's Fixed-Point Equipment Support Conceptual Models

The modeling of military conceptual models is the first time to abstract the realistic world. Military concept models are the structured normative description of military entities, actions and expected goals. Conceptual models provide a better understanding of the system in early stages of development. It also could increase quality of requirements, enhance communication between users and developers and be used in verification and validation of software products (Pace 2000).

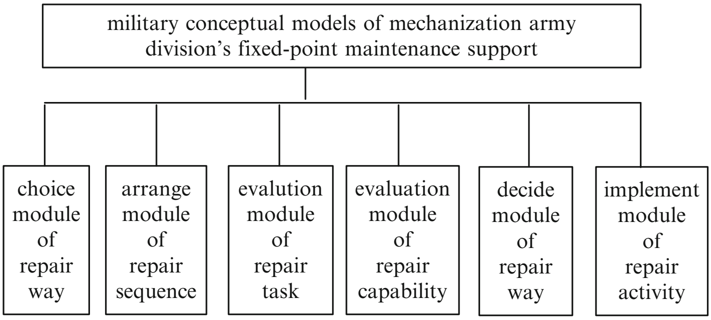


Fig. 60.2 Constitutive module of military concept model

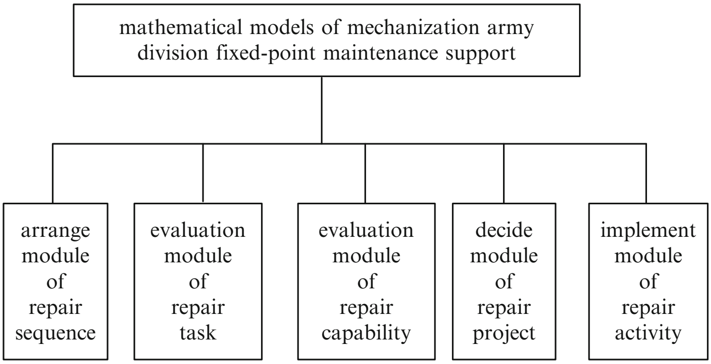


Fig. 60.3 Constitutive module of mathematics model

Military conceptual models are divided into six modules by analyzing and abstracting fixed-point maintenance support activities in the wartime as in Fig. 60.2.

60.3.2 Modeling of Mechanization Army Division's Fixed-Point Equipment Support Mathematical Models

Mathematical model is used to describe mutual relation and changing course of actual system with mathematics expression, logic rule and data. They are usually formed by some equations or algorithms. Mathematical models of mechanization army division's fixed-point maintenance support are divided into five modules by making quantitative analysis on military conceptual models of fixed-point equipment maintenance as in Fig. 60.3.

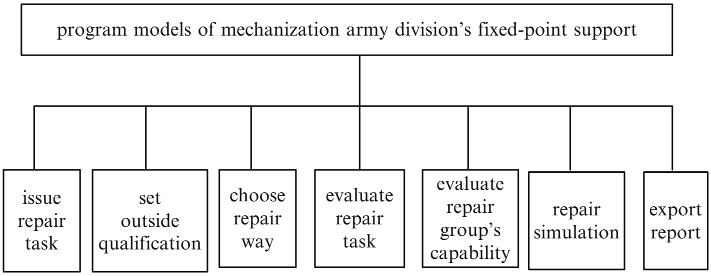


Fig. 60.4 Constitutive module of program model

60.3.3 Modeling of Mechanization Army Division's Fixed-Point Equipment Support Program Model

Mechanization army division's fixed-point support program models are described by computer language, which are executable final model products on the base of military conceptual models and mathematical models.

Program models of mechanization army division's fixed-point support are divided into seven modules as in Fig. 60.4.

60.4 Efficincy Evaluation of Fixed-Pont Maintenance Support

60.4.1 Influence Factor of Maintenance Support Efficiency

Fixed-point maintenance support efficiency is affected by many factors, such as reserved amounts of spare parts, organization and command of maintenance support activities, distribution scheme of support force, geography environment of disposition region and harassment caused by enemy.

60.4.2 Index Architecture of Evaluation

In order to evaluate the efficiency of fixed-point maintenance support organization, task finish time, proportion of finished task, repaired number of damaged equipment, evacuation proportion of damaged equipment and working load of repair group are chosen from the factor of finished task repair ability, working load of repair group (Liu Wei and Xi-sheng Jia 2010; Si-gao 2005). Index architecture of maintenance support efficiency evaluation is as follows in Fig. 60.5.

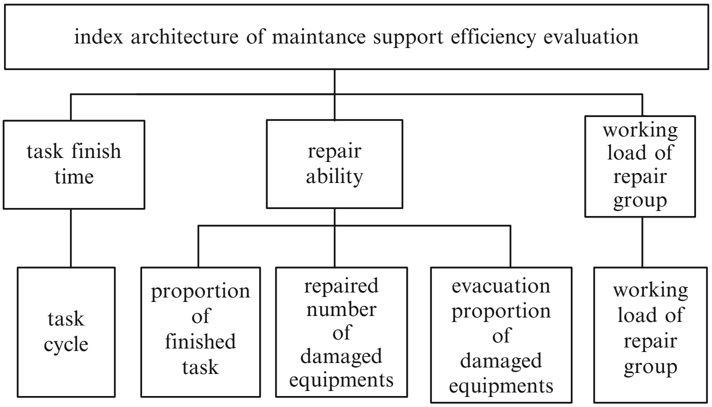


Fig. 60.5 Index architecture of maintenance support efficiency evaluation

60.4.3 Mathematics Model of Maintenance Support Efficiency Evaluation

Task finish time represents the time of the whole process from damaged equipments' packed off to damaged equipments' laid back, including damaged equipments' round-trip time from damaged spot to fixed-point (T1), repair preparation time (T2), repair waiting time (T3) and repair operation time (T4).

$$T(h) = T1 + T2 + T3 + T4$$

Proportion of finished task represents the ratio of number of the repaired equipments and total damaged equipments that are sent to fixed-pointed support organization.

$$R = R1/R2 \times 100\%$$

Repaired number of damaged equipment represents number of the repaired equipments that could carry out the task of fighting and supporting.

Evacuation proportion of damaged equipment represents the ratio of number of damaged equipment sent to rear repair organization and number of damaged equipment sent to the fixed-pointed group.

$$M = M1/M2 \times 100\%$$

Working load of repair group represents the ratio of repair group's effective working time and total working time.

Working load rate of repair group: it could be calculated with following formula.

$$Nr = (N1 + N2 + N3 + N4)/(D \times N5)$$

In the above formula, N1 represents repair preparation time of maintenance organization. N2 represents operation time of maintenance organization. N3 represents operation time of damaged equipment by maintenance organization. N4 represents repair defensive time of maintenance organization. N5 represents daily working time of maintenance organization. D represents setting up time of maintenance organization.

60.5 Design of Mechanization Army Ivision's Fixed-Point Equipment Support Simulation System

60.5.1 Choose of Federal Members

The major work of the equipment support simulation based On HLA is to construct federal members, build federal, design SOM/FOM, develop models or reform the interface of application systems by adopting the service of RTI and time synchronism advancing pattern, and so on.

In the equipment support simulation, simulation system needs to simulate actual battlefield environmental, battlefield information and damage effects of both confrontation sides. According to demand of simulation demands and simulation scale, the number of object class and interaction class should be small to reduce inner function structural and data interaction of simulation federal members. Business members, simulation monitoring members and battlefield environmental members are chosen according to federal members' option principles. Federal members' option principles have the principle of army structure, key activities, function independency, interaction frequency, top-down analysis.

60.5.2 Architecture of Mechanization Army Division's Equipment Support Simulation System Based on HLA

① artificial intervention member
② management member of assumption
③ peration monitoring member
④ data recording member
⑤ simulation displaying member

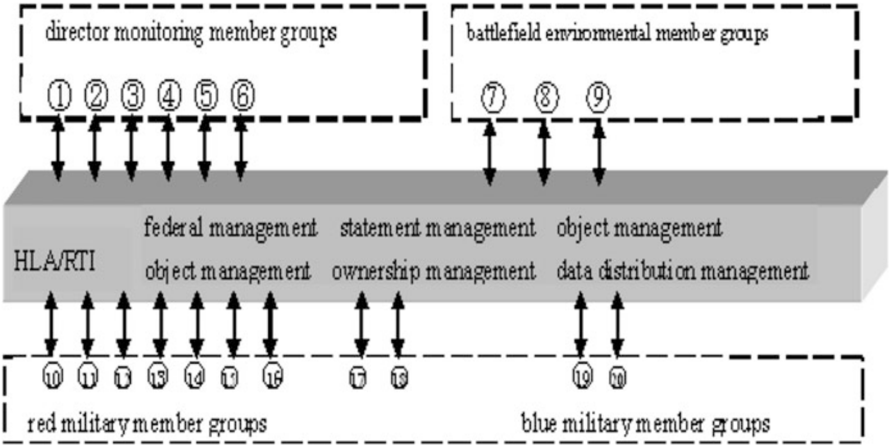


Fig. 60.6 Architecture of equipment support simulation system

(6) efficiency evaluation member
(7) geography environmental member
(8) meteorology environmental member
(9) electromagnetism environmental member
(10) euipment support command institution member
(11) maintenance force member
(12) support guarantee member
(13) munition support member
(14) conveyance force member
(15) service defensive member
(16) quipment use member
(17) superior equipments command institution member
(18) operational army member
(19) command institution member
(20) operational army member (Fig. 60.6)

60.5.3 Federation Development and Execution Process Based on VV&A

Federation Development and Execution Process provide universal development steps for the development and execution of federal members. FEDEP specify

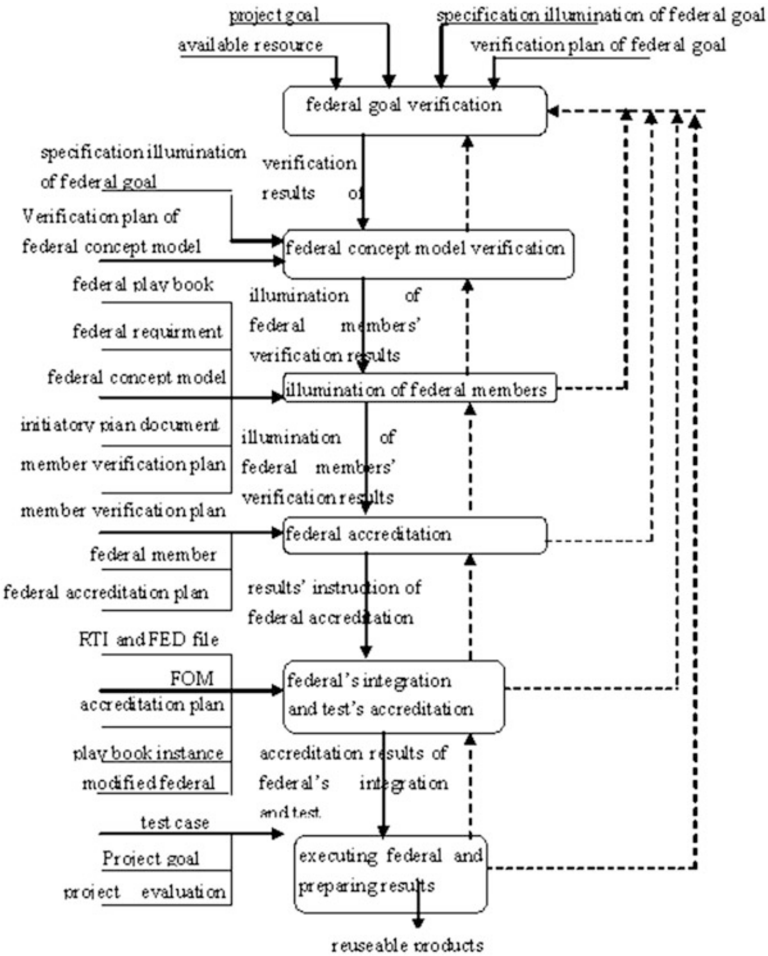


Fig. 60.7 Federation development and execution process based on VV&A

management and organization of the federation system's development from aspects of simulation system's demand, design, coding, test and acceptance. FEDEP version 1.5 abstracts FEDEP to six basic steps.

VV&A is a set of scientific methods used in the field of military modeling and simulation field. The object of VV&A is to ensure the accuracy and credibleness of modeling. It is accordant with object of FEDEP. They are used to meet military application demands, advance and control development and application of simulation system to satisfy user with facility of the realistic world's analog and simulation results' creditability. Therefore, federation development and execution process based on VV&A methods provide guarantee for creditability and usability of the simulation models, which improve the accuracy and creditability. FEDEP based on VV&A shows in Fig. 60.7 (Xin-ping 2004).

60.6 Development of Fixed-Point Maintenance Support Simulation System

The development tool of simulation system includes Microsoft Visual Studio, Microsoft SourceSafe, RTI, Sybase Power Designer 12.0 and Tau/Architect Add-on.

60.6.1 Composing of Simulation System

The simulation system could be divided into of five parts. They are RTI, director monitoring member group, battlefield environmental member group, Red military member group and blue military member group.

RTI is system's key communication parts. All federal members realize mutual communications, mutual connection and mutual operation on RTI.

Director monitoring member group include artificial intervention member, management member of assumption, operation monitoring member, data recording member, simulation displaying member and efficiency evaluation member.

Battlefield environmental member group include geography environmental member, meteorology environmental member and electromagnetism environmental member.

Red military member group include red military operational army federal members, superior equipments support command institution, local equipments support command institution, equipment repairing and rescuing federal member, materials supply member, service defensive federal member, damaged equipments' take-over federal member and support guarantee federal member.

Blue military member group include blue military operational army federal member and the blue military federal member of headquarters.

60.6.2 Running Flow of Simulation System (Fig. 60.8)

60.7 Simulation of Mechanization Army Division's Fixed-Point Equipment Support

60.7.1 Assumption of Simulation

The fixed-point maintenance support activities are simulated and analyzed under the task background of a mechanization army division's attack warfare. Simulation's input parameter are as follows: Fighting beginning time, fight duration, combat troops' circumstances, battlefield's circumstances, enemy's circumstances, disturb level of enemy's firepower, number of reserved repair equipments and simulating step width.

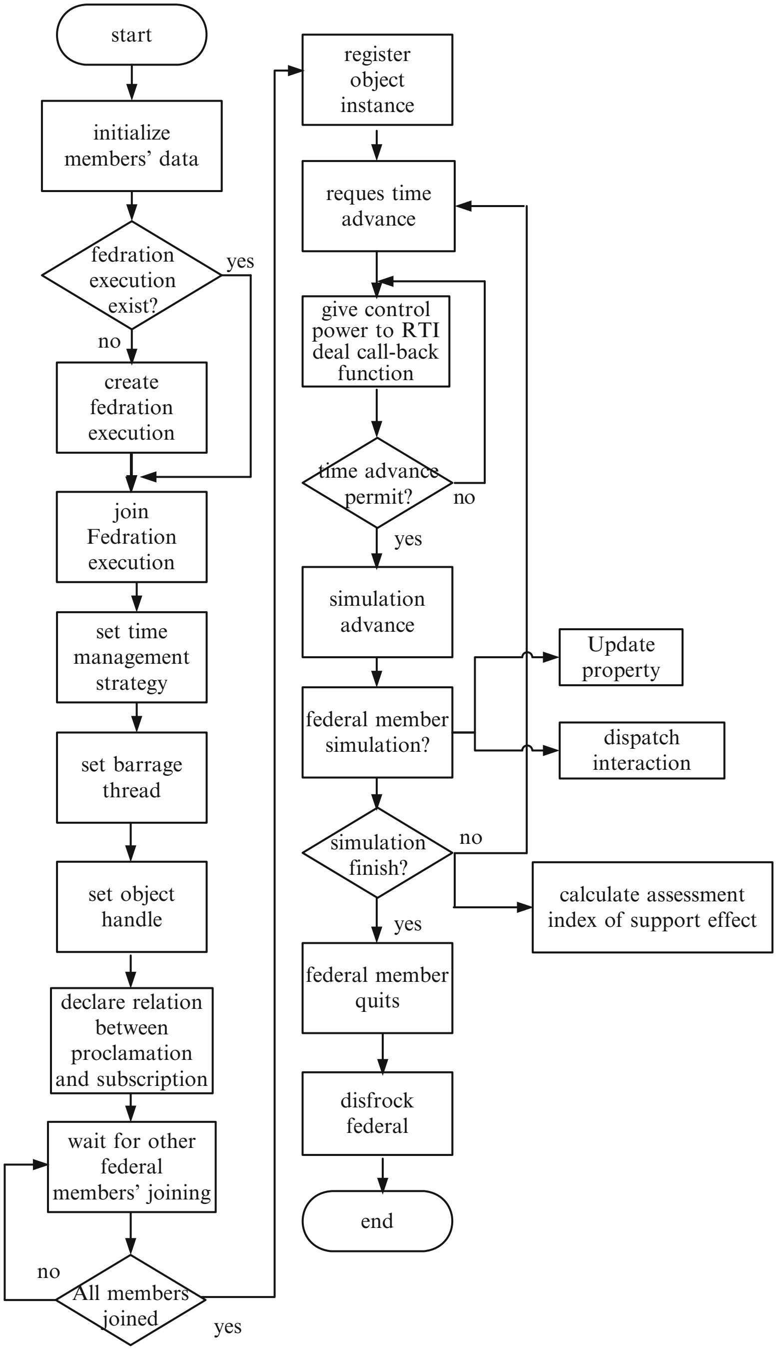


Fig. 60.8 Flow chart of simulation system's running process

Fig. 60.9 Program interface of simulation system

Fig. 60.10 Showing interface of simulation's results

60.7.2 Displaying Interface of Simulation System

60.8 Conclusion

Conceptual model and mathematical model of fixed-point maintenance support efficiency evaluation simulation of mechanization army division, index system and evaluation model of maintenance support efficiency evaluation are established after analyzing the composition and running process of maintenance support system of mechanization army division.

According to the characteristic of fixed-point maintenance support and object of simulation system, frame of simulation system and total running flow are designed. A VV&A process model is used for Federation Development and Execution Process. At last, the fixed-point maintenance support simulation efficiency evaluation systems of mechanization army division based on HLA are realized (Figs. 60.9 and 60.10).

References

Dong-nanWang (2006) Development of maintenance support simulation system based on HLA. Comput Digit Eng 34(5):123–126

Jian-bing Tang (2007) VV&A process of warfare simulation based on HLA. Comput Eng 33(14):254–256

Jun-jie LIU, Ya-lin Ji (2006) Simulation of tactical materiel support system based on HLA. Sci Technol Eng 6(24):3966–3969

Liu Wei, Xi-sheng Jia (2010) Research of wartime equipment maintenance support effectiveness and assessment index system. J Ordnance Eng College 22(6):1–4

Pace DK (2000) Conceptual model development for C^4ISR simulations. In: Proceedings of the 5th international command and control research and technology symposium, vol 10. Beijing, pp 60–62

Si-gao Zhu (2005) Study on the operational effectiveness criteria system design of a missile weapon system. Master dissertation, Huazhong University of Science and Technology

Wu-kui Zhao (2003) Equipment support science. People's Liberation Army Press, Beijing, pp 56–58

Xin-ping Cao (2004) Verification, validation and accreditation research on HLA simulation systems. PhD dissertation, National University of Defense Technology

Yan-ping Fan, Yan-qi Bai (2007) Research on improvementation of FEDEP based on VV&A. Comput Simul 21(7):254–256

Yan-zhong Zhang, Shi-ying Li (2006) Research on the development of maintenance support simulation system based on HLA. Syst Simul Technol 22(8):290–292

Chapter 61
Evolution of Industrial Ecosystem with Government's Intervention: Integration of Evolutionary Game Model into Multi-agent Simulation

Kai-fang Zheng, Su-ling Jia, and Hong-sen Wang

Abstract Government plays important role in the evolution of Industrial Ecosystem. To explore the influence of government's intervention on IE, an evolutionary game model with adjustment parameters, which represent the intervention actions of government, is established, and then is integrated into multi-agent simulation system. The evolution process of the IE is simulated and analyzed by the simulation with different value of adjustment parameters. The results show that the influence can be different according to the level of government's intervention, and certain level of intervention actions can guarantee the evolution process of IE towards "Eco-efficiency".

Keywords Evolutionary game • government intervention • industrial ecosystem • multi-agent simulation

61.1 Introduction

The industrial ecosystem (IE) theory is an important theoretical direction for the formation and evolution of eco-industry. IE, also called industrial ecology, studies the material and energy flows through industrial systems by analogy with the structure and mechanism of natural ecosystems. IE, as the study of interaction between industrial activities and their surrounding environment, aims to evaluate and reduce the harmful influence of industry on natural environment (Graedel et al. 1995).

K. Zheng (✉) • S. Jia • H. Wang
School of Management and Economics, Beihang University, Beijing, P. R. China
e-mail: mercy.zkf@gmail.com

E. Qi et al. (eds.), *The 19th International Conference on Industrial Engineering and Engineering Management*, DOI 10.1007/978-3-642-37270-4_61,

However, as Boons and Baas (1997) point out, differ from biological ecosystems, the evolution of IE towards greater eco-efficiency is not a spontaneous process; instead, it needs intentional action in IE. It demands the coordination of activities of economic actors as well as governmental agencies, to make sure its direction towards eco-efficiency (Boons and Baas 1997).

The influence of government intervention on the evolution process of IE can be found in many practical cases and theoretical researches. In the case of Kalundborg, Denmark, the governmental regulatory system plays the role of guider, establishing a flexible, cooperative relationship between government and the regulated industries and leading firms to focus their energies on finding creative ways to become more environmentally benign (Ehrenfeld and Gertler 1997). While in the case of Guitang Group in China, the regional government is acting as a facilitator in this IE, focusing on sustainable development, which is vital to establish the long-term symbiotic relationships between the Guitang Group and the sugarcane farmers (Zhu and Cote 2004). Boons and Baas (1997) discuss the implications from asymmetrical dependency relations between the organizations involved in different types of IE, point out that for an IE defined by a geographical boundary, regional governments should be available as authoritative coordination institutions (Boons and Baas 1997). Davis et al. (2009) prove by the simulation results that the most effective way to reduce CO_2 emissions of firms in an IE is to impose a high CO_2 tax rate.

This chapter considers the influence of government's intervention measures on the evolution process of IE. An evolutionary game model is established to imitate the learning and evolution process of firms in an IE. Then the multi-agent simulation of this model is conducted to explore the evolution process under government intervention. The results of the simulation can describe the different evolutionary routes of IE under different level of government intervention.

61.2 Evolutionary Game Model

Classical game theory essentially requires that all of the players make rational choices. On the contrary, evolutionary game theory shows that players can be rational, or even making explicit decisions (Easley and Kleinberg 2010). It is only required that players have a strategy, and the results of the game will test how good that strategy is. In this chapter, all firms in an IE can be seen as a single large population (Guo and Tian 2006), within which the individual firms are repeatedly drawn at random from a large population to play a symmetric two-person pure strategy game (Weibull 1996).

The theory described above is introduced to establish the evolutionary game model in this section, to explore the evolution of IE.

Fig. 61.1 Payoff matrix without government intervention

		Firm A	
		M	E
Firm B	M	i_1, i_1	i_2, i_3
	E	i_3, i_2	i_4, i_4

61.2.1 Assumptions of the Evolutionary Game Model

1. Assume that all the firms in the IE form a single population, in which they share the similar characteristics (for example, the same kinds of products, the same production scale, the same customer, etc.).
2. The firms in this model are bounded rational. They can only find their optimal strategy after constantly imitation and learning, eventually reaching its equilibrium.
3. There are two operation and production strategies for the firms in the model, the "Maximized Profits" strategy (*M* strategy) and "Eco-friendly" strategy (*E* strategy). "Maximized Profits" strategy means firms care nothing but their own profits, expanding their production without limit and ignoring the possible threat to the environment. Contrarily, firms that choose "Eco-friendly" strategy should control their production within a proper level and take advantage of new technology, in order to reduce their emissions and pollution to the environment. "Eco-friendly" strategy means more investment on new technology and less benefit from expanding scale of production. Generally speaking, without intervention from outside, the benefits of firms with "Eco-friendly" strategy are less than those of firms with "Maximized Profits" strategy.
4. Government does not participate in the game directly. It affects the system by adjusting the payoff matrix of the model. More specifically, the benefits of the *M* strategy will be discounted by government while that of *E* strategy will get extra subsidies from government.

61.2.2 Payoff Matrix Without Government Intervention (Bai et al. 2011)

Each firm in the IE has only two pure strategies. The strategy set of the players is $S = \{s_1, s_2\}$, with s_1 representing the *M* strategy and s_2 for the *E* strategy. The payoff matrix is shown in Fig. 61.1.

In the Fig. 61.1, *M* represents the "Maximized profits" strategy, and *E* represents the "Eco-friendly" strategy. i_1, i_2, i_3, i_4 are the payoffs, or the benefits, that the firms can gain when they select different strategies in the game.

Assuming $\rho_t(M)$ is the proportion of firms that use the *M* strategy at moment t, then the expected benefit of the firm that use M strategy at that moment is:

$$\mathrm{E_t}(M) = \rho_t(M)i_1 + [1 - \rho_t(M)]\,i_2 \tag{61.1}$$

Fig. 61.2 Adjusted payoff matrix with government intervention

		Firm A: M	Firm A: E
Firm B	M	i_1(1-a), i_1(1-a)	i_2(1-a), i_3(1+b)
	E	i_3(1+b), i_2(1-a)	i_4(1+b), i_4(1+b)

The expected benefit of firm that use E strategy at moment t is

$$\mathrm{E_t}(E) = \rho_t(M)i_3 + [1 - \rho_t(M)]\, i_4 \tag{61.2}$$

According to (61.1) and (61.2), at moment t, the average expected benefit of firm in the market is:

$$\overline{\mathrm{E}}_t = \rho_t(M)\, \mathrm{E_t}(M) + [1 - \rho_t(M)]\ \mathrm{E_t}(E) \tag{61.3}$$

61.2.3 Adjusted Payoff Matrix

As the authoritative organizations and policy maker in the IE, the government must supervise the behavior of the firms in the system. In the game model, two parameters, a and b, are introduced to represent the government's reaction to M and E strategy that the firms choose. The measures of the government intervention are described as followed:

For the firms that use strategy M, their benefits will be taxed at a certain proportion a $(0 \le a < 1)$.

For those that use strategy E, they will get a tax drawback, accounting for a certain percentage b of their benefits $(0 \le b < 1)$.

Consequently, the payoffs of the firm under different strategy are revised to $i_1(1-a)$, $i_2(1-a)$, $i_3(1+b)$ and $i_4(1+b)$ respectively.

So when the government intervention is introduced into the evolutionary game model, the payoff matrix is adjusted into (Fig. 61.2):

Based on the adjusted payoff matrix, the expected benefit of the firm that use M strategy and E strategy, and the average expected benefit at moment t are also changed:

$$\mathrm{E_t}(M) = \rho_t(M)i_1\,(1-a) + [1 - \rho_t(M)]\, i_2\,(1-a) \tag{61.4}$$

$$\mathrm{E_t}(E) = \rho_t(M)i_3\,(1+b) + [1 - \rho_t(M)]\, i_4\,(1+b) \tag{61.5}$$

$$\overline{\mathrm{E}}_t = \rho_t(M)\, \mathrm{E_t}(M) + [1 - \rho_t(M)]\ \mathrm{E_t}(E) \tag{61.6}$$

61.2.4 Learning Algorithm

The firms in the industrial ecosystem will find their optimum strategies (in the evolutionary game theory, the optimum strategy is the most suitable strategy for the environment (Weibull 1996)) after continuous imitation and learning, adapting themselves to the environment better. In the model, the learning process is implemented through learning algorithm, which can be described as (Bai et al. 2011):

$$S_{(m,\ n)t+1} = \left(S_{(m,\ n)t},\ S_{(x,\ y)t},\ R_{(m,\ n)t},\ E_t, \bar{E}_t\right) \tag{61.7}$$

The firm located in (m, n) and the firm in (x, y) meet and carry out a game at moment t. $S_{(m,\ n)t}$ is the strategy of the firm in (m, n) at that time, and $S_{(x,\ y)t}$ is that of the firm in (x, y). $R_{(m,\ n)t}$ is the benefit of firm in (m, n) from the game at moment t, while E_t represents the expected benefit of the same firm at the same time. $\bar{E}_t$ is the average expected benefit of all the firms in the system at moment t. $S_{(m,\ n)t+1}$ represents the strategy that the firm in (m, n) will use at the next moment (moment $t+1$), which will be decided by $S_{(m,\ n)t}$, $S_{(x,\ y)t}$, $R_{(m,\ n)t}$, E_t and $\bar{E}_t$ according to the following rules:

When $S_{(m,\ n)t} = M$, if $R_{(m,\ n)t} < \min\{E_t,\ \bar{E}_t\}$, then $S_{(m,\ n)t+1} = E$ (the strategy at the next moment changes into E). If not, $S_{(m,\ n)t+1} = S_{(m,\ n)t} = M$ (the strategy at moment $t+1$ keeps the same with that at moment t).

When $S_{(m,\ n)t} = E$, if $R_{(m,\ n)t} < \min\{E_t,\ \bar{E}_t\}$, then $S_{(m,\ n)t+1} = M$. If not, $S_{(m,\ n)t+1} = S_{(m,\ n)t} = E$.

Generally speaking, if $R_{(m,\ n)t} < \min\ \{E_t,\ \bar{E}_t\}$ at moment t, then the firm placed in (m, n) changes its strategy at moment $t+1$. If not, it keeps the same strategy at the next moment.

61.3 Multi-agent Simulation of The Evolutionary Game Model

61.3.1 Design of Agents

The agents in the model are the firms, which have the following attributes:

1. *Market space.* Market space is the external environment where the agents (firms) exist and evolve, described as a two-dimensional grid in the model. The agents (firms) are created in the market space, located by the (x, y) coordinate on the grid.

2. *Operation and production strategy* that the agents chose. "0" indicates the "Maximized Profits" strategy, whereas "1" indicates the "Eco-friendly" strategy.
3. *The current actual benefit*. The current actual benefit of a firm is the outcome of its current strategy based on the payoff matrix in the model. The possible current benefits of an firm could be $i_1(1-a)$, $i_2(1-a)$, $i_3(1+b)$ or $i_4(1+b)$, according to the current strategies of the firm and its opponent.
4. *The expected benefit*. The expected benefit of a firm at a certain moment can be calculated by Eqs. (61.4) or (61.5).
5. *The average expected benefit*. It is the average expected benefit for all the firms in the market, calculated by Eq. (61.6).
6. *The adaptability and competitive advantage of firms*. Firms' adaptability is their ability to adapt themselves to the environment (market), by changing or keeping their choice of strategies to find optimum (the most suitable) strategy for the environment until the whole system gets to the stable equilibrium. Firms decide to choose or keep their strategies according to their competitive advantage, which is measured by their benefits.

61.3.2 Design of Interaction Mechanism

1. *Creativity of new firms*

The firms are created with a certain probability in the two-dimensional market space when initializing the simulation model. This multi-agent model focuses on the influence of government policy on the competitive advantage and the strategy choice of firms in the market. At the very beginning of the establishment of the firms, their attributes of strategy are assigned a value "0" or "1" randomly to illustrate their strategy choice. Then, the agents are placed in the cells of the two-dimensional grid randomly. Each cell can contain one agent only, and it is also possible that some cells are empty.

2. *The learning and interaction activities of the agents*

The interaction mechanism of the agents from moment t to moment $t+1$ consists of three steps: choosing opponent, carrying out a game and evaluating the benefit of game.

Step1. *Finding the opponent of the game*

At the very beginning of the game, for an agent (firm) at (m, n), it must choose one agent from the surrounding 8 cells as the game opponent from moment t to moment $t+1$.

It is possible for some agents, that there are no other agents in the surrounding cells. This means that these agents are isolated. The isolated agents do not participate in the game at moment t, they choose a surrounding cell and move to the new position instead. At the beginning of the next moment, they search for their neighbor agents in the new position. If they still do not have any neighbors, repeat that again until they find their neighbors.

Step2. *Carrying out a game*

At moment t, a firm (agent) carries out a game with its opponent according to the learning algorithm designed in Sect. 61.2.4, and calculates its actual benefit based on the payoff matrix. The firms do not know which strategy is the most suitable for the given condition of the environment at the beginning of simulation until continuous learning and adaptation.

Step3. *Benefit evaluation and strategy decision*

After a game finishing, each firm will evaluate its benefit and decide whether change its strategy or not. The rule or standard of evaluation has been described in Sect. 61.2.3. The expected benefits and the average expected benefits can be indicated as the formula (61.4, 61.5, and 61.6).

3. *Evolution of the whole system*

The agents (firms) in the model participate in the game repeatedly, evolving continuously according to the given interaction mechanism, until the whole system reaches to its stable equilibrium point.

61.3.3 Initial Settings and Simulation Implement

There are some initial settings needed to be done at the very beginning of simulation.

1. The initial market space sets 100×100. The firms are created by 0.05 probabilities in the cells of market space, with their initial strategy which is set '0' or '1' at random.
2. Considering the assumption of classical game theory and model, in the program of this simulation, assign $i_1 = 5$, $i_2 = 6$, $i_3 = 3$, $i_4 = 4$.
3. The simulation is conducted repeatedly with different values of parameters (a, b), in order to explore how the equilibrium state will be. More specifically, with different value of (a, b), different conditions of government intervention are imitated in the simulation implement and their influence will be shown in the outcomes of the simulation.

According to the evolutionary game model and learning mechanism described in the preceding sections, the simulation model is conducted with the Swarm-2.2-java and Eclipse-SDK-3.7.2 integrated development platform (Liu and Chen 2009).

61.4 Simulation Results

61.4.1 The Results in the Absence of Intervention

First, consider the situation without the influence of government, with parameters $(a, b) = (0, 0)$.

Figure 61.3 gives a hint that the evolution route of IE in the absence of government intervention, showing how the percentage of two kinds of firms changes with the simulation time. The data of firms with *M* strategy are represented by blue curves while those of *E* strategy are represented in yellow.

It can be seen that the proportions of firms with *M* strategy and *E* strategy are nearly the same at beginning, then the percentage of *M* strategy firms gradually increases whereas that of *E* strategy firms decreases. At about t = 55, the system reaches its equilibrium state, and about 85 % of firms choose *M* strategy and 15 % of them choose the *E* strategy.

61.4.2 The Results Under Different Intervention Conditions

Set parameters (a, b) with value (0.1, 0.1), (0.3, 0.4), (0.5 0.5), (0.5, 0.9) and (0.9, 0.5) respectively, and the results are quite different.

1. *(a, b) = (0.1, 0.1)*

The outcomes under this situation do not make differences comparing with the result when $(a, b) = (0, 0)$. At about t = 50 the system reaches its equilibrium state, where the percentages of two kinds of firms are approximately 85 % (*M* strategy) and 15 % (*E* strategy) respectively.

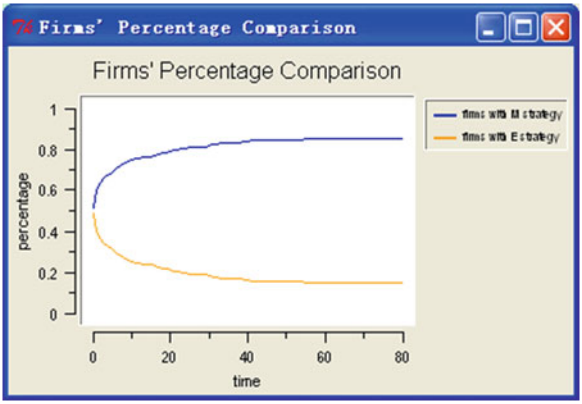


Fig. 61.3 Comparison of the proportion of firms with different strategy

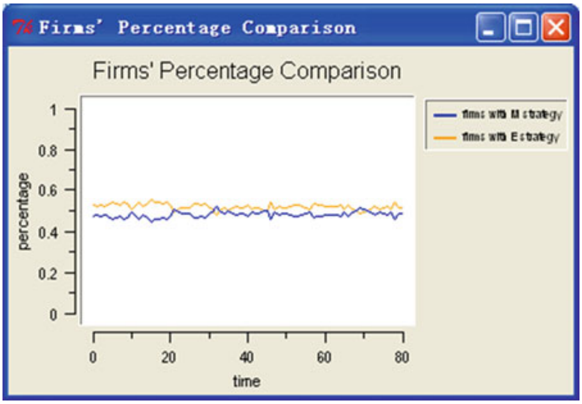


Fig. 61.4 Comparison of the proportion of firms when (a, b) = (0.3, 0.2)

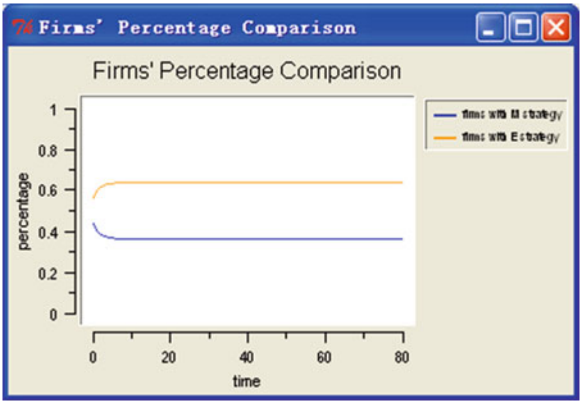


Fig. 61.5 Comparison of the proportion when $(a, b) = (0.5, 0.5)$

2. *(a, b) = (0.3, 0.2)*

Under this situation, the system maintains at an interesting dynamic equilibrium state. As can be seen from the Fig. 61.4, the proportions of both kinds of firms almost keep at the same level as they present at the beginning, fluctuating slightly during the whole simulation period. Also, the firms with *E* strategy have a little higher proportion than those with *M* strategy for most of the time.

3. *(a, b) = (0.5 0.5), (0.5 0.9), (0.9, 0.5)*

Figure 61.5 shows the change of proportions when $(a, b) = (0.5, 0.5)$. The percentage of firms with *E* strategy increases from about 55 to 63 % gradually, while that of *M* strategy reduces to 37 % from about 45 %. The stable equilibrium comes at $t = 7$.

The results when $(a, b) = (0.5, 0.9)$ (Fig. 61.6) and (0.9, 0.5) (Fig. 61.7) are almost the same, and also much similar to the result when $(a, b) = (0.5, 0.5)$. The system reaches its dynamic equilibrium when $t = 33$, and the proportions at the dynamic equilibrium are 67 % (*E* strategy) and 33 % (*M* strategy)

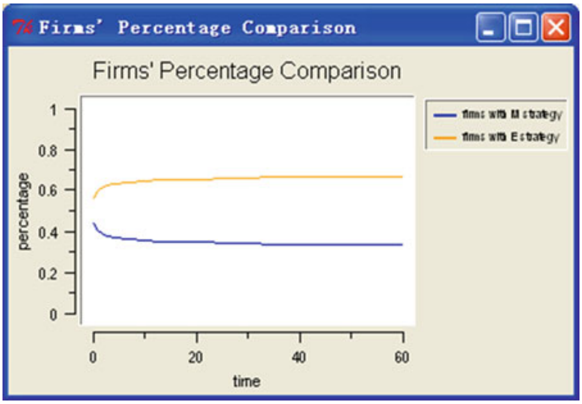


Fig. 61.6 Comparison of the proportion when $(a, b) = (0.5, 0.9)$

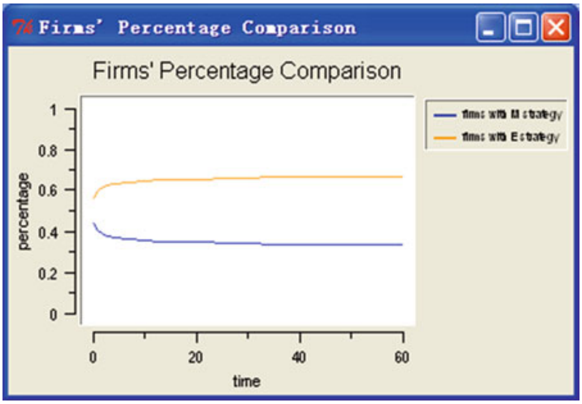


Fig. 61.7 Comparison of the proportion when $(a, b) = (0.9, 0.5)$

61.5 Conclusions

This chapter establishes an evolutionary game model to explore the impact of government's intervention on the evolution of IE. The different value of (a, b) represents different level of government's intervention. The results of simulation show that the governmental intervention has influence on the evolution of IE, and the impacts on the evolution can be quite different under the different value of (a, b).

Without the intervention of government ($a = 0$, $b = 0$), because of the better benefit M strategy can gain according to the pay-off matrix, increasing number of firms in the IE will choose M strategy. As in Fig. 61.3, in that case, percentage of firms with M strategy dominates in the IE, which means the evolution direction is totally opposite to the original goal of "Eco-efficiency".

Different level of government's intervention leads to the different results of the evolution process. The results can be generally divided into three categories:

1. The adjustment to the benefits of the game is not effective enough to change the evolution direction. For example, when $(a, b) = (0.1, 0.1)$, there are still 85 % firms that choose the *M* strategy in the equilibrium state.
2. When the adjustment reaches to a certain level, the intervention can prevent the situation from deteriorating even further. However, as shown in Fig. 61.4, it cannot guarantee the evolution toward "eco-efficiency", but only maintains the existing situation of IE.
3. When $(a, b) = (0.5, 0.5)$, $(0.5, 0.9)$ and $(0.9, 0.5)$, the adjustment makes the evolution direction toward the original aim of IE effectively. Based on the given conditions and learning mechanism, when system reaches its stable equilibrium state, the percentage of firms with *E* strategy dominates in the IE, which is nearly twice as much as that of *M* strategy (Figs. 61.5, 61.6, and 61.7).

References

Bai B, Deng XQ, Gao DH (2011) Multi-agent modeling and simulation for the evolution of organizational routines from the CAS and EGT perspective. J Inf Comput Sci 8(3):495–502

Boons FAA, Baas LW (1997) Types of industrial ecology: the problem of coordination. J Cleaner Prod 5(1–2):79–86

Davis C, Nikolic I, Dijkema GPJ (2009) Integration of life cycle assessment into agent-based modeling. J Ind Ecol 13(2):306–325

Easley D, Kleinberg J (2010) Networks, crowds, and markets: reasoning about a highly connected world. Cambridge University Press, New York, Ch. 7

Ehrenfeld J, Gertler N (1997) Industrial ecology in practice the evolution of interdependence at Kalundborg. J Ind Ecol 1(1):67–79

Graedel TE, Allenby BR, American Telephone and Telegraph Company (1995) Industrial ecology. Prentice Hall, Englewood Cliffs

Guo ZD, Tian JX (2006) Study on the co-evolution mechanism of industrial ecology engineering based on multi-agent asymmetric evolutionary game theory. In: Proceeding of 2006 international conference on management science and engineering. Institute of Electrical and Electronics Engineers Inc, Lille, pp 1907–1910

Liu Z, Chen YJ (2009) Swarm for java: simulation and implementation techniques. China Machine Press, Beijing (in Chinese)

Weibull JW (1996) Evolutionary game theory. The MIT Press, Cambridge, MA

Zhu QH, Cote RP (2004) Integrating green supply chain management into an embryonic eco-industrial development: a case study of the Guitang Group. J Cleaner Prod 12(8–10):1025–1035

Chapter 62
The Impact of Major Event on Regime Switching of Financial Market Volatility Spillover: The Case of the 2011 Japanese Earthquake

Lu Wang, Jiong-lou Xu, and Mao Li

Abstract Natural disasters may inflict significant regime switching of financial market volatility spillover. Using Japan and other four world's major stock market indexes, this chapter examines if any regime switching occurred across financial markets after the 2011 Japanese earthquake based on copula model. The results indicate that strengthened cross-markets regime switching with significant evidence of volatility spillover are noticeable for Japan-Hong Kong, Japan-US and Japan-China pairs. Every national/regional stock market is found to suffer on the effect by the 2011 Japanese earthquake.

Keywords Copula • major event • regime switching • volatility spillover

62.1 Introduction

The study of volatility spillover among international financial markets has attracted the interest of researchers and professionals both in theoretical and empirical field (King et al. 1994; Diebold and Yilmaz 2009; Malik and Ewing 2009). Aiming to control risks they face, portfolio managers and regulators need to take into account the volatility spillover of assets in the international financial markets, for example.

A volatility spillover occurs when changes in price volatility in one market produce a lagged impact on volatility in other markets, over and above local effects (Milunovich and Thorp 2006). Volatility spillover patterns appear to be widespread in financial markets. There is significant evidence for spillovers between equity markets, futures contracts, bond markets, equities and exchange rates, exchange rates, various industries, commodities, size-sorted portfolios, and swaps.

L. Wang (✉) • J. Xu • M. Li
Department of Statistics, Southwest Jiaotong University, Chengdu, China
e-mail: wanglu@home.swjtu.edu.cn

E. Qi et al. (eds.), *The 19th International Conference on Industrial Engineering and Engineering Management*, DOI 10.1007/978-3-642-37270-4_62,

Macroeconomic policy makers and investors are not only concerned about the existence of the volatility spillover but even more about sudden breaks in volatility spillover, for example the breaks caused by currency crises. Such breaks could affect the economy through a change in capital flows or in real linkages among markets, such as trade. They may lower diversification benefits from international investing and change investors' behavior after the break (Hitt and Hoskisson 1994; Forbes and Rigobon 2002; Ang and Bekaert 2002).

Just as the normal distribution is inadequate for modeling univariate time series, so the bivariate normal distribution is not suitable for modeling the relationship between two assets (Li 2000). As well as the asset returns not being normally distributed, their comovements may not be adequately captured by correlation coefficients. For example, marginal distributions tend to be characterized by fat tails and the probability of two markets both exhibiting a relatively high movement (in the same direction) may be much higher than can typically be captured with a bivariate normal distribution. A copula models the relationship between two independent variables of their marginal distributions. It does so by means of a joint distribution function with standard uniform margins. Hence it gives the probability that the observations in two series are below certain quantiles.

Natural disasters have inflicted serious damage on human life, property and economy. The recent Japanese earthquake has resulted in many damages that make the disaster particularly significant. On March 11, 2011, a 9.0 magnitude undersea mega thrust earthquake hit Tohoku in Japan. This powerful shock triggered a Tsunami that struck coastlines across the east of the country, leaving thousands death and damaging considerable property. Owing to the leading role of Japan in the world economy, this huge earthquake disaster caused the turbulent to Japan's financial market. At the same time, the global financial markets were also affected by the earthquake, which led to regime switching of volatility spillover in the international financial markets. Investors should take a more cautious and proactive strategy facing changes in the markets.

Therefore the goal of this chapter is to examine whether any regime switching of volatility spillover has occurred after the Japanese earthquake based on copula model. In other words, we seek to provide evidence as to whether such a disaster has increased the interdependence among financial assets in different countries/regions.

The remainder of this chapter is organized as follows. Section 62.2 looks at some methodological aspects concerning copulas, data is presented in Sect. 62.3, we provide empirical analysis in Sect. 62.4 before concluding with Sect. 62.5.

62.2 Copulas: Basic Concepts

Each multivariate distribution function can be split into its univariate marginal distribution functions and a copula function (Sklar's theorem (Sklar 1959)). In other words, copulas allow one to study the distributional dependence structure of random vectors irrespective of their marginal distributions and a copula function.

Definition 62.1 (Copula) Let $X = (X_1, \ldots X_d)'$, be a d-dimensional random vector with distribution function $F(x_1, \ldots, x_d) = P(X_1 \leq x_1, \ldots, X_d \leq x_d)$ and marginal distribution-functions $F_i(x_i) = P(X_i \leq x_i)$ for all i = 1, ... ,d. Then the distribution function C of the d-dimensional random vector $(F_1(X_1), \ldots, F_D(X_d))'$, is called the copula (or copula function) of X or F (Frees and Valdez 1998).

It can be shown that the copula function is uniquely determined by the multivariate distribution function F if all univariate marginal distribution functions are continuous and that

$$F\,(x_1, \ldots, x_\mathrm{d}) = C\,(F_1\,(X_1)\,, \ldots, F_\mathrm{d}\,(X_\mathrm{d})) \tag{62.1}$$

Thus, copulas can be utilized to build flexible multivariate distribution functions in two steps: First, model the distributional dependence via some copula(s), and second, plug in appropriate margins.

Copula functions represent standardized distributions in the sense that their one-dimensional margins are uniformly distributed on the interval [0,1]. An important property is that the copula of a random vector X stays the same regardless of any strictly increasing transformation of the margins X_j, $j = 1, \ldots, d$. This invariance property (also called "scale invariance") is a desired feature of dependence functions and dependence measures, as we understand dependence itself to represent the association between "large" and "small" realizations of random vectors irrespectively of their scale.

It is clear that the copula is a map from $[0, 1]^\mathrm{k}$ to [0, 1]. The copula is invariant under increasing and continuous transformations. This property is very useful as transformations are commonly used in economics. For example, no matter whether we are working with price series or with log price series, we have the same copula. To avoid confusions, we will use symbols x and y to denote the observations of random variables X and Y; and we will use symbols u and v to denote their marginal CDFs. So x and y could be any real numbers but u and v must be in [0, 1]. A simple example of a copula is an independent copula, which is defined as

$$C\,(u, v) = u \cdot v \tag{62.2}$$

Suppose one asks what is the probability that both returns in market A and in market B are in their lowest 20th percentiles? If these two markets are independent, using the independent copula, we have

$C(u,v) = C(0.2,0.2) = 0.2 \cdot 0.2 = 0.04$, which is simple and intuitive. More complicated copula functions usually contain one or more parameters, which are also called association parameters. If only one parameter presents in a copula function, this parameter usually reflects the strength of the dependence.

A proper dependence measure for multivariate distributions should be scale invariant (or invariant under change of the marginal distributions). All dependence measures derived from the copula are scale invariant, and so in line with our basic requirement. The most important scale-invariant dependence measure in financial applications is Kendall's τ.

Similar as a linear correlation coefficient, $\tau \in [-1, 1]$ and a positive τ implies positive dependence with the higher the value, the stronger the dependence. The relation between Kendall's τ and a copula is given in the following moment condition:

$$\tau = 4 \int_0^1 \int_0^1 C(u, v) dC(u, v) - 1 \tag{62.3}$$

Hence for a copula with single parameter, the parameter is determined if τ is fixed.

Another important notation is tail dependence, which measures the probability that both variables are in their extremes. The ability to capture tail dependence is an important property of a copula. A copula $C(u, v)$ is said to have left (lower) tail dependence if

$$\lim_{u \to 0} \frac{C(u, u)}{u} = \lambda_l > 0 \tag{62.4}$$

To define the right (upper) tail dependence, we need the notation of a joint survival function,

$$\overline{C}(u, v) = 1 - u - v + C(u, v) \tag{62.5}$$

A copula $C(u, v)$ measures the probability that its two arguments are below value u and v respectively, while its joint survival function measures the probability that these two arguments are above value u and v respectively. Now, a copula $C(u, v)$ is said to have right (upper) tail dependence if

$$\lim_{u \to 1} \frac{\overline{C}(u, u)}{1 - u} = \lambda_r > 0 \tag{62.6}$$

62.3 Data Description

This section explores the stock market volatility spillover issue caused by the 2011 Japanese earthquake using the bivariate copula model constructed above. The empirical investigation into regime switching of volatility spillover will be carried out by studying five stock market indices. These indices are Nikkei 225 index (N225), Financial Times 100 stock index (FTSE), Hong Kong Stock Exchange index(HKSE), the S&P 500 index (SP) and Shanghai Stock Exchange index(SSE). Since the earthquake originated in Japan, the N225 is assumed to be the volatility 'originator', while the other four indices are the volatility 'recipients'.

The daily range series of the five indices from December 27, 2010 to May 26, 2011 are retrieved. The full sample period was divided into two sub-periods before and after the earthquake (11, March 2011). The two sub-periods have the same sample size $N = 50$. In order to cope with the issue of trading in different time zones, the ranges of the N225 are matched with lagged ranges of the rest of the market indices.

Summary statistics for the daily returns of the five stock markets are presented in Table 62.1. The average returns in the stock markets range from −0.0723 % (in Japan) to 0.0577 % (in US), while the standard deviations range from 0.007 to 0.018 in full sample. It is evident that ranges for all markets become larger and more volatile in the post-earthquake period than in pre-earthquake from Table 62.1. Therefore, the earthquake disaster actually caused significant volatility increases in stock markets.

To obtain an idea for the extent of linkages between other four markets and Japan, Table 62.2 presents the Kendall's τ between different stock markets. For the whole sample period, the correlations between N225, FTSE, HKSE, SP and SSE range from 0.076 to 0.220. As expected, there is a positive correlation between the Japan and other stock markets. The highest correlation is between N225 and SSE and the lowest one is between N225 and FTFE. The intertemporal dependencies have strengthened after the Japan earthquake. This suggests that the globalization of the financial markets has substantially increased. On the other hand, the values of Kendall's τ between markets increase in the post-earthquake period in comparison with the pre-earthquake. This indicates that there might be changes in volatility spillover between markets during the disaster period, which points out the possible existence of regime switching due to the disaster.

62.4 Empirical Results

62.4.1 Copulas Choice

According to Deheuvels (1981), each multivariate distribution has at least one associated copula function. This copula function is unique when margins are continuous. Here, we assume that our bivariate vectors of N225 return and other stock returns exhibit continuous margins so that each copula is defined in a unique way.

Rather than specifying given margins and inducing then the related copula function, we prefer to specify a given copula function consistent with the studied empirical dependence structure. We consider the Archimedean copulas. To be able to fit conveniently estimated copulas, we choose to use mostly one parameter Archimedean copulas. Especially, our choice restricts to the following copulas: Gumbel, Frank and Clayton. As discussed before, the Gumbel features right tail dependence, the Clayton features left tail dependence, and the Frank is symmetric.

Table 62.1 Summary statistics for daily returns in five stock markets

		N225	FTFE	HKSE	SP	SSE
Pre-earthquake	Mean	0.00029969318	−0.00051042742	0.00085932992	0.00057334314	0.00054323450
	Std. deviation	0.009708443358	0.008545379003	0.010436004279	0.007871305778	0.012628067178
Post-earthquake	Mean	−0.00142668710	0.00014444237	−0.00047847755	0.00044864086	−0.00119414547
	Std. deviation	0.023672892407	0.009739882759	0.009645907987	0.007150234279	0.009419203634
Full sample	Mean	−0.00072319325	−0.00021287624	0.00003939320	0.00057706090	−0.00039275477
	Std. deviation	0.017942804178	0.009072821115	0.010097237394	0.007475708438	0.011103078426

Table 62.2 Unconditional Kendall T among stock markets

		N225	FTFE	HKSE	SP	SSE
N225	Correlation coefficient	1.000	0.076	0.162*	0.150*	0.220**
	Sig. (2-tailed)	–	0.263	0.017	0.027	0.001
FTFE	Correlation coefficient		1.000	0.096	0.242**	−0.010
	Sig. (2-tailed)		–	0.159	0.000	0.882
HKSE	Correlation coefficient				0.211**	0.040
	Sig. (2-tailed)				0.002	0.553
SP	Correlation coefficient				1.000	−0.090
	Sig. (2-tailed)				–	0.184
SSE	Correlation coefficient					1.000
	Sig. (2-tailed)					

*Correlation is significant at the 0.05 level (2-tailed)
**Correlation is significant at the 0.01 level (2-tailed)

62.4.2 Estimation of Copula Models

The estimation is semi-parametric and proceeds in two stages. In the first step we construct an empirical distribution for the margins. In the second stage we substitute the estimated margins into the copula and estimate the copula parameters by maximum likelihood.

62.4.3 Model Selection

In order to assess goodness of fit of the various copulas, we utilize information criteria. We use the Akaike Information Criterion (AIC), because of its optimality properties, and also since it can be used to compare nested or nonnested models, in contrast to likelihood ratio tests. We also use the Bayes Information Criterion (BIC). The BIC does not share the same optimality properties as AIC, but gives a stricter penalty for overfitting a model, which is sometimes a desirable property. The standard expressions for AIC and BIC are as follows. Consider a sample with size equal to T, and the number of estimated parameters equal to k. Then the AIC and BIC are defined as

$$\mathrm{AIC} = -2\log\ (\mathrm{max.likelihood}) + 2k \tag{62.7}$$

$$\mathrm{BIC} = -2\log\ (\mathrm{max.likelihood}) + k\log(T) \tag{62.8}$$

The best model is selected to be the one that minimizes AIC or BIC.

Table 62.3 Parameter estimates of the copulas

		Panel A: Full sample			Panel B: Pre-earthquake			Panel C: Post-earthquake		
		Clayton	Frank	Gumbel	Clayton	Frank	Gumbel	Clayton	Frank	Gumbel
N225-FTFE	θ	0.1453	0.5394	0.911	−0.012	−0.0143	0.7428	0.3397	1.2389	0.8808
	logL	0.6673	0.4119	3.6181	0.002	0.0016	12.2842	1.2809	0.0101	4.9849
	AIC	0.6654	1.1762	−5.2362	1.996	1.9968	−22.5684	−0.5618	1.9798	−7.9698
	BIC	0.36437	0.87517	−5.53723	1.69497	1.69577	−22.86943	−0.86283	1.67877	−8.27083
N225-HKSE	θ	0.5615	1.4855	0.9798	0.8472	2.3194	0.7668	0.273	0.5554	0.9702
	logL	7.0209	2.9352	0.3169	6.4561	3.4987	7.0939	0.945	0.2083	0.0258
	AIC	−12.0418	−3.8704	1.3662	−10.9122	−4.9974	−12.1878	0.11	1.5834	1.9484
	BIC	−12.3428	−4.17143	1.06517	−11.2132	−5.29843	−12.48883	−0.19103	1.28237	1.64737
N225-SP	θ	0.3778	1.3872	0.9903	0.4477	0.9578	0.8208	0.3623	1.6043	0.9977
	logL	3.6553	2.4094	0.09	2.2598	0.6056	8.4616	1.2838	1.5405	0.0071
	AIC	−5.3106	−2.8188	1.82	−2.5196	0.7888	−14.9232	−0.5676	−1.081	1.9858
	BIC	−5.61163	−3.11983	1.51897	−2.82063	0.48777	−15.22423	−0.86863	−1.38203	1.68477
N225-SSE	θ	0.4802	2.125	0.964	0.9572	3.937	0.9647	0.2189	0.6745	0.9692
	logL	5.4701	5.804	0.3382	6.9675	8.3735	0.3291	0.6025	0.3127	0.2596
	AIC	−8.9402	−9.608	1.3236	−11.935	−14.747	1.3418	0.795	1.3746	1.4808
	BIC	−9.24123	−9.90903	1.02257	−12.236	−15.048	1.04077	0.49397	1.07357	1.17977

62.4.4 Discussion of Results

Table 62.3 displays the parameter estimates and a comprehensive comparison of the various models examined, in three separate panels, based on the maximized log likelihood, the AIC and BIC, respectively. Panel A, Panel B and Panel C present the result on the full period, the pre-earthquake period and post-earthquake period, respectively.

For instance, the best model between Japan and London stock markets is Gumbel in spite of the full period and two sub-periods, but the dependence parameters in Gumbel function is 0.911, 0.7428 and 0.8808 on different period respectively, which means the strength of volatility spillover has been changed. At the same time, the best model is Clayton, Gumbel and Frank on the period of full sample, the pre-earthquake and post-earthquake between the Japan-US pair, which indicates the regime switching of volatility spillover has happened due to the 2011 Japan earthquake.

To sum up, strengthened cross-markets regime switching with significant evidence of volatility spillover are noticeable for Japan-Hong Kong, Japan-US and Japan-China pairs. Every national/regional stock market is found to suffer on the effect of the 2011 Japanese earthquake.

62.5 Conclusion

This chapter has investigated if the March 2011 Japanese earthquake influenced the regime switching of volatility spillover in international stock markets, which characterize the transmission of the major event which typically involve financial markets going from a period of low volatility to high volatility, and being followed by the other in the subsequent(s) period. The appealing feature of our specification is that it enables us to identify the dependence structure and the duration of the volatility transmission mechanism from one market to another. The potential applicability of the proposed procedure has been illustrated through an analysis of regime switching of volatility spillover between Japan and Britain, Hong Kong, US, Shanghai.

Acknowledgments This work is supported by the Fundamental Research Funds for the Central Universities (No. SWJTU12CX057, No. SWJTU12ZT14), the Youth Fund for Humanities and Social Sciences of Ministry of Education (No. 10YJCZH157) and College Students Innovation and Entrepreneurship Training Program (No. 111061354).

References

Ang A, Bekaert G (2002) International asset allocation under regime switching. Rev Financ Stud 15(11):37–87

Deheuvels P (1981) A non parametric test for independence. Publications de l'Institut de Statistique de l'Universite de Paris 26(1):29–50

Diebold FX, Yilmaz K (2009) Measuring financial asset return and volatility spillovers, with application to global equity markets. Econ J 119(534):158–171

Forbes K, Rigobon R (2002) No contagion, only interdependence: measuring stock market co-movements. J Financ 57(22):23–61

Frees EW, Valdez EA (1998) Understanding relationships using copulas. North American Actuar J 2(1):1–25

Hitt MA, Hoskisson RE (1994) A mid-range theory of the interactive effects of international and product diversification on innovation and performance. J Manag 20(2):297–326

King M, Sentana E, Wadhwani S (1994) Volatility and links between national stock markets. Econometrical 62(4):901–933

Li DX (2000) On default correlation: a copula function approach. J Fixed Income 9(4):43–54

Malik F, Ewing BT (2009) Volatility transmission between oil prices and equity sector returns. Int Rev Financ Anal 18(3):95–100

Milunovich G, Thorp S (2006) Valuing volatility spillovers. Glob Financ J 17(1):1–22

Sklar A (1959) Fonctions de repartition a n dimentional et leurs marges. Publications de l'Institut de Statistique de l'Universite de Paris 8(1):229–231 (in French)

Chapter 63
Comparative Study on the Performance of GEM Listed Companies Having or Not the Stock Right Incentive Mechanism

Gong-rong Chen, Cheng-yin Zheng, and Qi-shan Li

Abstract On the basis of GEM (Growth Enterprises Market) listed companies' financial data in 2011, this thesis builds a comprehensive assessment system for the company performance with application of factor analysis to examine whether there's distinct difference or not between the companies having or not the stock right incentive mechanism with the adoption of independent t-test. Through the empirical tests we concluded that: there's no distinct difference on the performance of these two kinds of companies and the implementation effect of the stock right incentive mechanism in GEM is still not very obvious. This thesis conducts detailed analysis on the current situation of stock right incentive mechanism in GEM and the managing structural characteristics of GEM listed companies and presents corresponding policy suggestion for the existing problems.

Keywords Company performance • Comparative study • GEM • Stock right incentive

63.1 Introduction

As a mid & long-term incentive tool, stock right incentive was adopted extensively in the Europe & America at the 1990s of twentieth century and enjoyed favorable effect. Its starting-point is to balance the stakeholder relationship between shareholders and management of the company for the purpose of avoiding damage to

---

This paper is financial supported by China Hunan Provincial Science & Technology Department (No. 2010FJ6033).

G. Chen • C. Zheng (✉) • Q. Li
Business School, Hunan University, Changsha, China
e-mail: zcy0830@gmail.com

E. Qi et al. (eds.), *The 19th International Conference on Industrial Engineering and Engineering Management*, DOI 10.1007/978-3-642-37270-4_63,

the company's interest because of the conflict of interest between these two parties hereof (Jensen and Meckling 1976). The development of stock right incentive mechanism is comparatively undeveloped in China (Wei Gang 2000). Limited by the lacking of internal experience and imperfection of assorting external environment, China has many obstacles and difficulties in developing this mechanism and it's hard for the listed companies to really actualize the stock right incentive mechanism and achieve a mutual development of both owner and management of one company (Zhou Jian-bo and Sun Ju-sheng 2003). For the purpose of regulating such stock right incentive system, establishing and perfecting a mid & long-term incentive mechanism with the combination of incentive and restraint for listed company, further completing the company's management structure, major modification towards the Securities Law of PRC and Company Law of PRC have been conducted and a series of policies PRC have been conducted and a series of policies such as Approach in Stock Right Incentive Management for Listed Company (trial) have been promulgated since 2005 to provide policy guidance for listed companies in designing and implementing the stock right incentive plan (Gu Bin and Zhou Li-ye 2007). Since the opening of GEM at October, 2009, the stock right incentive mechanism proved a big draw for GEM enterprises; by the end of May, 2012, 71 companies have publicized the stock right incentive plan taking 22.26 % of the total GEM listed companies. The market holds a positive attitude towards such mechanism in general and the mechanism also gained close attention by the government and active response from the market. In the meanwhile, although there're listed companies promulgating the stock right incentive plan constantly, a majority of GEM companies haven't chosen this mechanism. GEM is a security market with private high-tech enterprises as major listing companies, are there any differences on the performance of companies having or not such stock right incentive mechanism? What's the in-depth reason for the fact of having differences or no differences? This thesis would conduct research for this issue hereinafter.

63.2 Background of the Research

63.2.1 Current Stock Right Incentive Situation in Gem Listed Companies

The GEM mainly focuses on entrepreneur enterprises and was established for the purpose of adapting to the development of independent innovation enterprises and other growth entrepreneur enterprises. A majority of the companies listed in the GEM are engaged in high-tech area, having the characteristics of promising-growth potential and high-tech which make the enterprises in desperate needs of financing and expansion during their fast development. As the constant perfection of China's capital market, the promulgation of GEM in 2009 provided a direct financing platform for mega SME (small and medium enterprises) which remitted the financing-difficulty issue for SME to a certain extent and is hereby an important explorative way for resolving the development bottleneck of enterprises. Stock right incentive can remit the conflict of interest between owner and management of enterprises, restrain moral risk of management, reduce agent issue, decrease agent cost and hereby improve the company profit and value and actualize the sustainable development. GEM is a security market with promising-growth potential and high-tech features, hereby proper talent is the key factor for the development of innovation enterprises and other growth entrepreneur enterprises; for the purpose of attracting and retaining talents, many GEM companies adopted the stock right incentive mechanism. Therefore, the stock right incentive system can remit the current growth limit issue of GEM companies, adjust internal management structure of enterprises, breakthrough management bottleneck, contribute to the enterprise's re-entrepreneur development and as a result incentive the management and promote the enterprise's long-term development (Wu Chunshu and Zhu Xu 2012).

Whereas, some GEM companies are being questioned by the market because of their low criterion and broad coverage of the incentive plan. Currently 71 GEM companies have promulgated their stock right incentive plans but only 28 have executed such plan. From the viewpoint of amount, few companies have been approved by the CSRC (China Securities Regulatory Commission) and actualized the corresponding implementation and nine companies have already terminated their plan for several reasons. Therefore we can see that on the one hand China has strict laws and regulations in approving the stock right incentive plan and on the other hand the development of GEM companies' stock right incentive mechanism in China still remains immature (Liu Guo-liang and Wang Jia-sheng 2000). At the primary opening stage of GEM, over-assessment of the market caused the enterprises in setting over-high exercise price consciously; now that the stock market is at its downturn and the market returns to be rational, share price fell sharply and so did the exercise price and share price. When the share price comes near or breaks the exercise price of option and the grant price of restricted stock, the incentive mechanism would hereby loose the incentive effect it supposed to have and the affected companies would have to announce for withdrawing such stock right incentive plan. When the market efficiency is inadequate and the market is at its downturn, efforts of managers in boosting the share price would be futile and the market cannot present an objective evaluation for the company's profit and the manager's efforts. On the contrary, some GEM listed companies don't have any stock right incentive mechanism but still achieved favorable company performance.

63.2.2 Structural Characteristics of Gem Listed Companies' Management

From the viewpoint of the stock right structure, GEM listed companies are usually expanded and grew from family business with the stock rights extensively held by entrepreneur or family members thus has the feature of high-concentration of stock rights. Because of such phenomenon, there's no efficient mutual-restriction & balance between shareholders therefore the major shareholder occupies whole controlling towards the company and there's no efficient legal person management structure. The high-concentration of stock rights caused unreasonable wealth distribution system which leads to the wealth acquired through high IPO premium be extensively occupied by major shareholder and the wealth can only be distributed within certain group (Tian Cun-zhi and Wu Xin-chun 2010).

From the viewpoint of board of directors' structure, there's the phenomenon of highly-overlapping of the board of directors and managements in China's GEM listed companies with the entrepreneur highly controlling management authority of the company. Currently most of the GEM listed companies are established by entrepreneur, the company's scale is small and the main business decision-making person is the entrepreneur or his family members. For the purpose of improving decision-making efficiency and avoiding information transmission mistake, the owner usually is the top decision-maker. Such mechanism with chairman as general manager completely eliminated the supervision and binding function of board of directors; management staff can pursue mega benefit for themselves during the business operation hence interest of small and medium shareholders are damaged to a certain extend.

Judged from the current stock right incentive situation of GEM and the management structure characteristic of GEM listed companies, most owners of GEM enterprises is meanwhile the major manager, i.e. entrepreneur manager with the development of enterprises closely combined with personal interest of the entrepreneur; especially after the company being listed, the situation of company is directly related to the development and honor of the enterprise's actual controller and his family hereof, therefore the incentive function of stock right to entrepreneur can basically be ignored. Hence we consider that having or not having the stock right incentive mechanism is not very meaningful for the GEM listed companies (Chang-jiang et al. 2011).

63.3 Design and Sample of the Study

63.3.1 Selection and Source of the Data

This thesis selected all the GEM companies listed before 2011 as study sample, dislodging samples with incomplete financial data and anomalous data, finally confirmed 234 company samples. Cross-section data as of December 31, 2011 are selected to examine the performance of GEM listed companies. Financial data adopted in this thesis hereof are sourced from the CSMAR database and all the data processing and statistic analysis are conducted in the Excel2007 and SPSS20.0 statistic software.

63.3.2 Selection of Financial Indicators for Performance Assessment

For the purpose of assessing performance of the listed companies, financial indicators are required to be used as dimension to conduct the corresponding analysis. This thesis hereof selected 12 financial indicators listed hereinafter: operating margin ratio, rate of return on net assets, rate of return on assets, earnings per share; total asset turnover ratio, current assets turnover ratio; liquidity ratio, asset-liability ratio; rate of capital accumulation, growth rate of total assets, growth rate of net margin and the growth rate of operation revenue, covering the profitability, operating capacity, debt paying ability and development ability of the enterprises. The above mentioned financial rate index can dynamically combine the statement of assets and liabilities, cash flow statement and profit statement and eliminate influence brought by the scale of enterprises and difference of industries.

63.3.3 Method of the Empirical Study

This thesis adopted the factor analysis to conduct comprehensive grading for the performance of GEM listed companies and grouping the companies following the factor of having stock right incentive mechanism or not. We separately achieved two groups of data with the comprehensive score of companies having the incentive mechanism and not having such mechanism, then conduct independent sample T examination for these two groups of data to see if there's any difference between their mean values and then get the final conclusion.

63.4 Result of the Empirical Study and Corresponding Analysis

63.4.1 Descriptive Statistics Analysis and Examination Analysis on the Differences

Table 63.1 is a descriptive statistics for the 12 financial indicators of companies having or having not the stock right incentive mechanism.

From Table 63.1 we can see that only the mean value t examination of rate of capital accumulation and growth rate of operation revenue has distinct discrepancy with the number of 5 % which shows that these two types of companies has obvious difference at their rate of capital accumulation and growth rate of operation revenue. The rate of capital accumulation of company having stock right incentive mechanism is lower than that of don't have such mechanism; the higher the rate of

Table 63.1 Descriptive statistics

	Company having the stock right incentive mechanism		Company not having the stock right incentive mechanism		
	Mean value	Standard deviation	Mean value	Standard deviation	Mean value t-test
Operating margin ratio	0.1984	0.1257	0.1957	0.1166	0.1590
Rate of return on total asset	0.0773	0.0323	0.0791	0.0381	−0.3400
Rate of return on net assets	0.0953	0.0411	0.0979	0.0473	−0.4030
Earnings per share	0.5550	0.2780	0.5761	0.3505	−0.4350
Total asset turnover ratio	0.5038	0.2539	0.4997	0.2794	0.1030
Current assets turnover ratio	0.6722	0.3622	0.6500	0.3649	0.4190
Liquidity ratio	9.0557	15.3957	9.7996	9.8697	−0.4370
Asset-liability ratio	0.1805	0.1144	0.1601	0.1173	1.2000
Rate of capital accumulation	0.7930	1.4484	1.3217	1.7182	-2.370^{**}
Growth rate of total assets	0.5746	0.8777	0.8246	1.0919	−1.8180
Growth rate of net profit	0.7989	2.4800	0.5574	1.9151	0.793
Growth rate of operation revenue	0.4225	0.6241	0.2440	0.5841	2.0540^{**}

**Significant at the 5 % level

capital accumulation is, the more the company's capital accumulation is, its capital maintenance would be better and its ability in dealing with the risk and enjoying sustainable development is stronger. As for growth rate of operation revenue, company having stock right incentive mechanism has distinctly higher rate than that of don't have such mechanism. This indicator indicates the increasing situation of one company's major business shows the sustainable development trend and market expansion ability of one company and can reflect the growth of one company in its major development direction. And there's little discrepancy of these two types of companies in the rest ten financial indicators.

63.4.2 Factor Analysis

With the application of PCA (principal component analysis), we get the initial eigenvalue of common factor, contribution rate of variance and accumulative contribution rate of variance and confirm the number of initial common factors as 5. To conduct reasonable explanation for the actual implication of the extracted abstract factor, we adopt the maximum variance rotation method. From Table 63.2 we can see that the rotated factor loading matrix can demonstrate the actual economic meaning of the extracted common factor pretty clearly. Common factor F1 has the max loading

Table 63.2 Rotation ingredient matrix

	Components				
	1	2	3	4	5
Operation margin ratio	0.689	−0.613	0.012	0.018	−0.129
Rate of return on total asset	0.918	0.107	0.218	0.023	−0.124
Rate of return on net assets	0.881	0.297	0.256	0.044	0.036
Earnings per share	0.793	−0.045	0.298	0.077	0.026
Total asset turnover ratio	0.103	0.962	0.070	−0.038	0.039
Current asset turnover ratio	0.102	0.959	0.042	−0.047	0.078
Liquidity ratio	−0.068	−0.062	0.021	0.011	0.787
Asset-liability ratio	−0.001	0.186	−0.066	0.006	0.697
Rate of capital accumulation	0.288	0.086	0.934	0.058	−0.033
Growth rate of total asset	0.314	0.027	0.925	0.056	−0.033
Growth rate of net profit	0.012	0.001	0.057	0.906	−0.041
Growth rate of operation revenue	0.087	−0.077	0.037	0.895	0.062

at operation margin ratio, rate of return on total asset, rate of return on net asset and earnings per share which can reflect the income and profit status of GEM listed companies' financial performance from multiple angles; therefore we name F1 as profitability factor. Common factor F2 is determined by total asset turnover ratio and current asset turnover ratio to reflect the operation and turnover ability of capital and measure the asset management ability and operation situation of one company therefore is names as asset operation ability factor. Common factor F3 is mainly illustrated by rate of capital accumulation and growth rate of total asset for the purpose of measuring the future expansion of one company's asset and scale and is hereof named as expansion ability factor. Common factor F4 has the max loading at growth rate of net profit and operation revenue and is used to reflect the increase of operation revenue and profit and is hereby names as growth ability factor. Common factor F5 is mainly explained by liquidity ratio and asset-liability ratio to inspect one company's ability in resisting financial risks and conducting continuous operation and is thereof be named as debt paying capacity factor.

For the purpose of conducting a more comprehensive assessment towards the financial performance of the GEM listed companies, we calculate the comprehensive score for one company with the application of factor scoring method. The comprehensive score is achieved through the weighting sums of each factor variable score; the weight is the weight of each common factor's variance contribution rate in the accumulative variance contribution rate of all the extracted common factors. Hereinafter is the acquired formula:

$$F = 0.2921F1 + 0.2362F2 + 0.1936F3 + 0.1632F4 + 0.1149F5$$

Among which, F refers to the score of comprehensive performance; F1, F2, F3, F4 and F5 refers to the score of each factor.

Table 63.3 Independent samples test

		Levene's test for equality of variances		t-test for equality of means	
		F	Sig.	t	df
Comprehensive score	Equal variances assumed	1.313	0.253	0.035	232
	Equal variances not assumed			0.036	129.969

63.4.3 Mean Value Examination of Independent Samples

Mean value examination of independent samples is used to examine whether there's distinct difference between two samples. As shown in Table 63.3, the comprehensive score of companies having stock right incentive mechanism is 0.0023598 higher than companies having not such mechanism, the t statistic value is 0.035and it's not very obvious at the 5 % level, therefore there's no distinct differences of these two companies' performance. The results show that: performance of companies publicized the stock right incentive plan is slightly higher than companies didn't, but the difference is not obvious and their performances are almost equal.

63.5 Conclusion and Policy Suggestion

This thesis based on the characteristics of GEM enterprises and built an assessment indicator system for the financial performance of GEM listed companies with selection of 12 financial indicators in profitability, operation ability, debt paying ability and development ability; used the factor analysis method to extract 5 factors of the profitability, operation ability, debt paying ability, expansion ability and growth ability; calculated the comprehensive performance score of each company with the GEM listed companies as study samples; and finally get the conclusion that there's no difference on the performance of companies having and not having such stock right incentive mechanism with the application of independent sample t-test.

Above study conclusion shows that having or not the stock right incentive mechanism has any distinct affect for the GEM listed company's performance. China has just started the stock right incentive mechanism in our GEM listed companies which do require further exploration and perfection in the practice and to really resolve the problems and actualize the original intention of such stock right incentive mechanism, a reasonable stock right incentive plan shall be formulated and the corresponding management mechanism of companies shall be further perfected. Therefore, the writer presents several suggestions listed hereinafter:

First, design a reasonable stock right incentive plan which can not only incentive the managers sufficiently but also can protect the investor and other stakeholders from been damaged. (a) select appropriate stock right incentive method with the consideration of the actual situation and characteristics of the enterprise, e.g. as for high-tech GEM enterprises, the incentive subjects shall include not only the managements but also technical bones and even ordinary technical staff; (b) design a reasonable stock right performance assessment system. Use the multi-dimensional indicators to design an all-around assessment system with multiple-angles and the company can introduce several non-financial indicators such as industrial indicators, etc. including the market share, technical innovation, etc. in accordance to the situation of each industry; (c) strictly set the exercising indicator. Determination of exercise price shall reflect the actual current value of the company as real and the target shall be not only challenging but also feasible, i.e. the target hereof shall be able to achieve as of current HR ability but certain efforts must be conducted to achieve such target. Since the imperfection of current capital market and the easy to operate and short-term fluctuation features of the share price, we can require the listed companies to formulate the price in accordance to the average closing price within a longer period. To sum up, we should closely combine the self-characteristics of companies and flexibly apply the ordinary designing theory and thinking of stock right incentive; furthermore, the design of stock right incentive plan shall comprehensively take each and every aspects of the plan implementation into consideration, e.g. affect of the economic environment, dealing method when the share price deviates from the stock value, etc. (Yu Jian and He Wen-xi 2012).

Second, perfect the management structure of the company. (a) GEM listed companies shall strictly separate their chairmen and general managers; the system of separation shall be recommended. Strength the building of board of directors and board of supervisors, establish and perfect the independent director system; (b) develop and expand the institutional investor team. As professional agent institutions, institutional investors can supervise, bind and guide the GEM listed companies' managements with their information advantage and industrial experience. Institutional investors shall be encouraged to indirectly, authorized or directly participate in the formulation and examination process of the stock right incentive plan and participate in the process of market information creating for the purpose of identifying and reducing the false information operated and created by the enterprise's internal staff and guaranteeing the incentive mechanism in transmitting the factual market signal through share price variation and hereby promoting the long-term value increase of one company (Xu Ning and Xu Xiang-yi 2010).

References

Chang-jiang, Yan Ming-zhu, Zheng Hui-lian, Xu Jing-jing (2011) Why listing corporation choose equity incentive plan? Acc Res 1:68–75 (in Chinese)

Gu Bin, Zhou Li-ye (2007) Equity incentive effect of listed companies in China. Acc Res 2:79–84 (in Chinese)

Jensen M, Meckling W (1976) Theory of the firm: managerial behavior, agency costs and ownership structure. J Financ Econ 3(4):305–360

Liu Guo-liang, Wang Jia-sheng (2000) An empirical study in the ownership structure of listed companies, the incentive system and performance. Econ Theory Econ Manage 5:40–45 (in Chinese)

Tian Cun-zhi, Wu Xin-chun (2010) An empirical study of the impact of corporation equity and management incentive on information asymmetry. Nankai Bus Rev 3:28–34 (in Chinese)

Wei Gang (2000) Senior management incentive and operating performance of listed companies. Econ Res 3:32–39 (in Chinese)

Wu Chunshu, Zhu Xu (2012) Analysis of companies listed on GEM equity incentive. Collect Econ China 1:106–107 (in Chinese)

Xu Ning, Xu Xiang-yi (2010) Rationality and its binding factors of the stock option incentive contract. China Ind Econ 2:100–109 (in Chinese)

Yu Jian, He Wen-xi (2012) Analysis of issues related to companies listed on GEM equity incentive practice. Acc Commun 2:32–34 (in Chinese)

Zhou Jian-bo, Sun Ju-sheng (2003) The research on governance effect of top-management stock incentive – empirical evidences from Chinese listed companies. Econ Res J 5:85–94 (in Chinese)

Chapter 64
Benefit Embezzlement of Separation of Control Right and Cash Flow Right: Empirical Evidence from China Private Real Estate Listed Companies

Yan-hong Chen, Shi-jian Zhang, and Pei Zhu

Abstract Based on a 5 years' data from 2007 to 2011 of 185 China private real estate companies listed in shanghai and Shenzhen stock market , this paper makes an empirical study on the effects of the deviation between control right and cash flow right on benefit embezzlement. Furthermore, the function of outside director and institutional investor is analyzed. We document that benefit embezzlement is positively associated with the deviation between control right and cash flow right. Then, we find benefit embezzlement is negatively associated with the rate of outside director to directorate, and it is irrelevant to institutional holdings. Furthermore, our results suggest that the influence of deviation between control right and cash flow right on benefit embezzlement is weaker when the rate of outside director to directorate is high.

Keywords Benefit embezzlement • Institutional investor • Outside director • Separation of double rights

64.1 Introduction

A 1999 groundbreaking research about company ultimate ownership structure conducted by Porta L, Silanes LD and Shleifer A. shows that the relationship of controllers and minority shareholders gradually become the main content of the agency problem (Porta et al. 1999). After that many domestic and foreign

Y. Chen (✉) • S. Zhang
Business School, Hunan University, Changsha, China
e-mail: chyh5921@163.com

P. Zhu
Department of Accounting, Harbin Institute of Technology, Harbin, China

E. Qi et al. (eds.), *The 19th International Conference on Industrial Engineering and Engineering Management*, DOI 10.1007/978-3-642-37270-4_64,

scholars also made widely research about the effects of the separation of control right and cash flow right on the value of the company, dividend policy and corporate performance (Claessens et al. 2000; Lemmon and Lins 2003; Yong Ye et al. 2007; Shan-min Li et al. 2006; Peng Wang and Lian Zhou 2008; Qi Gu et al. 2006). The studies indicate that the interest conflicts between controllers and minority shareholders are the main problem of the corporate governance in different countries. The phenomena that control rights exceed cash flow rights are very common in emerging market economies of the eastern Asian. The more control rights exceeds cash flow rights, the smaller cost have the controllers to pay and the larger motivation they would have to seek gain for themselves. In theory, the outside directors and institutional investors existing as supervisors can effectively restrain the controllers' behavior to seek gain for themselves. So during the new transition period in China, dose the degree of the separation of control right and cash flow right and the supervision of outside directors and institutional investors have effects on the benefit embezzlement? Dose the benefit embezzlement caused by the separation of double rights exist differences when the supervision strength of outside directors or institutional investor is not same?

Known as a pillar industry of the Chinese economy, the real estate industry play a more and more important role in the development of the economy and the adjustment of economic structure. Due to the reasons of its history and system, the ownership structure of most private real estate listed companies is unreasonable. The ownership concentration of these companies is higher, and the ultimate controllers have strong motivation and ability to seek personal interests when their control rights exceed cash flow rights. So it is significant to research the problems of benefit embezzlement in private real estate listed companies from a perspective of the separation of double rights. We select a 5 year's data from 2007 to 2011 of 185 China private real estate companies listed in shanghai and Shenzhen stock market, and make an empirical study on the effects of the deviation between control right and cash flow right on benefit embezzlement. Furthermore, the function of outside director and institutional investor is analyzed. Our results provide further evidence for the research of the benefit embezzlement of the real estate industry.

The remainder of this paper is organized as follows. We develop our hypotheses in Sect. 64.2. In Sect. 64.3 we explain the sample and the research design. Then we present and discuss the empirical results in Sect. 64.4 and Sect. 64.5 concludes.

64.2 Literature Review and Hypotheses Development

Many ultimate controllers of the listed companies extend their resources through the use of pyramiding and management appointments, as well as through cross-ownership and the infrequent use of shares that have more votes. So their control rights exceed their cash flow rights (Porta et al. 1999). The ownership structure of a company not only affects the controllers' ability but also their motivation to occupy the interests of other shareholders (Gomes 2000; Shleifer and Wolfenzon 2002).

On the base of having certain quantity of cash flow rights, the ultimate controllers have more control rights and their ability to entrench upon other shareholders is stronger when the separation of double rights is serious. Similarly, on the base of having certain quantity of control rights, the ultimate controllers have less cash flow rights and their motivation to entrench upon other shareholders is stronger when the separation of double rights is serious. So the increasing of the degree of the separation of double rights amplifies the ability and the motivation of the controllers' to seek gains for themselves. Based on this discussion, our first hypothesis is:

H1 The degree of the separation of control right and cash flow right is positively associated with the level of ultimate controllers' benefit embezzlement.

Independent director system was formally introduced into China in 2001, whose purpose is to supervise and restrain major shareholders' expropriation behavior through the introduction of independent directors into director board. Current research shows that outside directors play a supervisory role to restrict the controllers' benefit expropriation behavior, which can reduce the size of the agency problems between the shareholders to some extent. Especially in the companies whose controllers' control rights exceed cash flow rights, the controllers' motivation to occupy other shareholders' interests is stronger, but the cost of benefit embezzlement increases under the supervision of the outside directors. This reasoning leads us to our second and third hypotheses:

H2 The strength of outside directors' supervision is negatively associated with the level of ultimate controllers' benefit embezzlement.

H3 The outside directors' supervision can restrain the impact of the degree of separation of double rights on the level of ultimate controllers' benefit embezzlement.

Along with the development of capital market, institutional investors have already become an important part of capital market. The results of existing literatures show that institutional investors can help ease the interest conflicts between major shareholders and minority shareholders (Kun Wang and Xing Xiao 2005), strengthen the independence of the directors board (Gallagher and Kochhar 1996; Xiao-hui Wu and Yan-fu Jiang 2006), enhance the earnings quality of the enterprise (Liu and Peng 2006; Xian-hui Bo and Lian-sheng Wu 2009), and improve the quality and performance of merger and acquisition (Xia Chen et al. 2007). But some other results of the research show that institutional investors are still short-term traders who "vote with their feet", and they don't play a positive role in corporate governance (Parrino et al. 2002). Taken together, the existing literatures imply that the institutional investors have some governance effect. Therefore, we think the level of the benefit embezzlement is lower when the supervision of institutional investors is stronger. In addition, they can effectively restrict the controllers' behavior to seek gain for themselves when their control rights exceed cash flow rights. Our next two hypotheses are:

H4 The strength of institutional investors' supervision is negatively associated with the level of ultimate controllers' benefit embezzlement.

H5 The institutional investors' supervision can restrain the impact of the degree of separation of double rights on the level of ultimate controllers' benefit embezzlement.

64.3 Research Design and Sample Selection

In this section, we discuss in detail our measures and of benefit embezzlement, the separation of double rights, the supervision of independent directors and the supervision of institutional investors. We then present our regression models, followed by a discussion of our choice of sample.

64.3.1 Definitions of Variables

64.3.1.1 Benefit Embezzlement (TUN)

There are many different forms of benefit embezzlement. Generally speaking, in the country which the legal protection for investors is better, the way of benefit embezzlement is more hidden. But the history of China's securities market is short and the environment of legal system is overall weak, which make the way of controller seeking gain for themselves more direct. And capital occupation is the direct form of the benefit embezzlement. Therefore in this paper we choose capital occupation as the measurement of the benefit expropriation. If there is capital occupation of controllers, it is an important part of accounts receivable-others, so "accounts receivable-others" can well reflect the degree of capital occupation of the listed companies by controllers and their affiliated parties (Jiang Guohua et al. 2009; Ke-min Wang et al. 2009). So in this paper we use "accounts receivable-others" divided by total assets to reflect the degree of the capital occupation of the listed companies.

64.3.1.2 The Separation of Double Rights (SR)

Control right is known as voting right, and cash flow right refers to the actual controllers' ownership of the company through pyramidal holdings, multiple tower holdings and cross-shareholdings of the listed companies. In this paper, we use the calculation method of previous researchers. In many cases, the ultimate controller has several control rights chains through which to control the votes in a company (Porta et al. 1999; Claessens et al. 2000). We trace those pyramidal and cross-holding chains individually, then describe control right as the sum of the weakest link in the individual chain of voting rights and describe cash flow right as the sum of the products of the ownership stakes along every chains. Then we take the ratio of voting rights to cash flow Rights as the separation of double rights (SR), the greater SR is, the bigger degree of deviations from one-share-one-vote is.

64.3.1.3 The Supervision of Outside Directors (OD)

Like Qi Gu et al. (2006), we use the proportion of the company's outside directors to the total number of directors board to describe the supervision on the ultimate controllers by outside directors.

64.3.1.4 The Supervision of Institutional Investors (IH)

Xia Chen et al. (2007) don't bring all institutional investors of the listed company into their research field, but focus on analyzing the first big or the top five institutional investors. The reason is that they expect institutional investors who have heavier warehouse holding can pay more attention to the company, which makes the supervisory effect more strong. Follow this approach, we sum up the proportion of institutional investors' holdings in top ten shareholders to be the proxy variable of the strength of their supervision.

In addition, in order to explain how explanatory variables affect the benefit embezzlement accurately, we control some other important variables, including the company size (SIZE), the asset-liability ratio (LEV), profitability (ROA) and macro annual influence (YEAR). SIZE is the natural logarithm of the total assets; LEV is the ratio of liabilities to assets; ROA is the ratio of profit to assets.

64.3.2 Regression Models

We use the following model to test the association between benefit embezzlement (TUN) and other variables. We make a regress of TUN on SR, OD and IH in our basic model to test our Hypothesis 1, Hypothesis 2 and Hypothesis 4:

$$\mathrm{TUN} = \alpha + \beta_1{}^{*}\mathrm{SR} + \beta_2{}^{*}\mathrm{OD} + \beta_3{}^{*}\mathrm{IH} + \lambda \sum \mathrm{CONTROLS} + \varepsilon$$

Then, we test our Hypothesis 3 (Hypothesis 5) using a multiple interaction effects model wherein we regress TUN on SR, OD(IH) and one two-way interaction terms OD*SR(IH*SR) to test for the individual and the joint effects of the explanatory variables on the benefit embezzlement.

$$TUN = \alpha + \beta_1{}^{*}SR + \beta_2{}^{*}OD^{*}SR + \lambda \sum CONTROLS + \varepsilon$$

$$TUN = \alpha + \beta_1{}^{*}SR + \beta_2{}^{*}IH^{*}SR + \lambda \sum CONTROLS + \varepsilon$$

Table 64.1 Distribution of the regression variables

	Mean	Median	MAX	MIN
TUN	0.030213	0.015345	0.493356	0.000453
SR	1.574895	1.25	7.810594	1
OD	0.3716	0.333333	0.555556	0.25
IH	2.227305	0.53	15.715	0
SIZE	21.84101	21.79244	24.34368	18.9892
LEV	0.555635	0.568527	1.810657	0.014459
ROA	0.060491	0.038001	2.142726	−0.06972
GROW	88.01212	0.125019	14883.06	−0.98143

64.3.3 Sample Selection Procedure

Our sample consists all shares A private real estate companies listed on Shenzhen Stock Exchange and Shanghai Stock Exchange from 2007 to 2011. We select the sample according to the following principles: (1) Eliminate ST, PT companies; (2) Eliminate companies whose data we could not precisely gain; (3) Eliminate companies whose data is extreme during our observation period. According to the above principles, finally we get 185 valid samples, including 26 samples in 2007, 30 samples in 2008, 38 samples in 2009, 45 samples in 2010 and 46 samples in 2011. We gain the main data from database CSMAR and then collected part of the missing data by hand through the listed company's annual reports on www.cninfo.com.cn.

64.4 Empirical Results

64.4.1 Descriptive Statistics

Table 64.1 presents the descriptive statistics for our samples. As shown in Table 64.1, the benefit embezzlement has a maximum, mean and median of 49.3, 3.0 and 1.5 % respectively, which show that there is benefit embezzlement phenomenon in China's private estate listed companies. In addition, the mean and median of SR is 1.575 and 1.250, which shows that the phenomena of the separation of control right and cash flow right are existent in the sample companies. On average, when the controllers have 1 unit of cash flow right, they actually have 1.575 units voting rights, so their motivation and capability to get private interest is stronger. And there is a big difference between the maximum and the minimum of OD and IH, which indicates the strength of outside directors' and institutional investors' supervision in different private estate listed companies is very different.

Table 64.2 Pearson correlation matrix

	TUN	SR	SD	IH	SIZE	LEV
TUN	1					
SR	0.1164***	1				
OD	−0.0851**	0.0626**	1			
IH	−0.0068	−0.0795*	0.0564	1		
SIZE	−0.1673*	−0.1496***	0.0505*	0.3828	1	
LEV	−0.2633**	−0.1147*	−0.0531	0.0126*	0.4070*	1
ROA	0.6946***	−0.0670**	0.0218***	−0.0396***	−0.1658***	0.4680***

***, ** and * denote significance at the 1, 5 and 10 % levels

64.4.2 Correlation Matrix

Table 64.2 presents Pearson correlations among our regression variables. Benefit embezzlement is significantly positively correlated with the separation of double rights, and company's outside directors to the total number of director board. So Hypothesis 1 and Hypothesis 2 are preliminarily verified. But the correlation between benefit embezzlement and institutional investors' holdings is not significant; we need further verify our Hypothesis 4. In addition, the size of the company is negatively correlated with benefit embezzlement, and property net profit rate is positively correlated with it. We find when a company's asset-liability ratio is higher, the level of its benefit expropriation is lower. Maybe the supervision of the creditors of the companies whose asset-liability ratio is higher is stronger, which can effectively restrain the benefit embezzlement of controllers.

64.4.3 Regression Results

Through the above analysis, some hypotheses have been preliminarily verified. In this part we will further verify our hypotheses by means of ordinary least squares regression (OLS) analysis.

As shown in Table 64.3, the results of models 1, 2 and 3 indicate that TUN is significantly associated with SR and OD but not significantly associated with IH, and they are consistent with the correlations in Table 64.2. We introduce all three variables in model 4 simultaneously. The coefficient of SR is positive and significant as expected based on previous published literature, so Hypothesis 1 can be verified. The focus of our study is OD and IH. Consistent with Hypothesis 2, OD is still negative and significant after controlling for SR, IH and other variables. But Hypothesis 4 can't be verified. Maybe institutional investors haven't really played a supervisory role in China.

Table 64.3 Tests of Hypothesis 1, Hypothesis 2 and Hypothesis 4

Independent variables	Predicted sign	Model 1	Model 2	Model 3	Model 4
Intercept	?	0.0207	0.0803	0.0785	0.0704
		(0.37)	(1.42)	(1.27)	(1.14)
SR	(+)	0.0076			0.008
		(3.04)***			(3.26)***
OD	(−)		−0.0965		−0.1073
			(−1.98)**		(−2.20)**
IH	(−)			0.0007	0.0008
				(0.82)	(0.95)
SIZE	?	−0.0002	−0.0005	−0.0024	−0.0008
		(−1.65)*	(−1.19)	(−1.71)*	(−1.24)
LEV	?	−0.0179	−0.0228	−0.0145	−0.0187
		(−1.98)**	(−2.01)**	(−1.76)*	(−1.97)**
ROA	?	0.2233	0.2223	0.2161	0.2249
		(11.09)***	(10.86)***	(10.40)***	(11.12)***
Year controls		Included	Included	Included	Included
F-statistic		23.79***	22.43***	21.67***	20.02***
Adjusted R2		49.80 %	48.20 %	47.30 %	50.80 %

The numbers in parentheses are the t-statistics. ***, ** and * denote significance at the 1, 5 and 10 % levels

We test for multiple interaction effects by performing OLS regression analysis on SR, OD*SR and IH*SR. Results are presented in Table 64.4. The OLS analysis reveals that the main effect of SR is positive, but the interaction effect of SR and OD is negative and significant and the interaction effect of SR and IH is positive but not significant. It suggests that there is a positive association between the variable SR and TUN, and the positive impact of SR on TUN decreased when OD is higher. So the Hypothesis 3 is verified. In private real estate listed companies, outside director has played a very important role to supervise the controllers and ease the impact of the separation of control right and cash flow right on the benefit embezzlement. The coefficient of the interaction between SR and IH is not significant. We do not find evidence to support Hypothesis 5 that the institutional investors' supervision can restrain the impact of the degree of separation of double rights on the level of ultimate controllers' benefit embezzlement. As an external supervisor, institutional investor didn't play the role to ease the benefit embezzlement in our private real estate listed companies.

64.5 Conclusions

By using the data from China private real estate listed during the period from 2007 to 2011, we make an empirical study on the effects of the deviation between control right and cash flow right on benefit embezzlement and analyze the function of

Table 64.4 Tests of Hypothesis 3 and Hypothesis 5

Independent variables	Predicted sign	Model 5	Model 6
Intercept	?	0.0048	0.0468
		−0.08	−0.75
SR	(+)	0.0204	0.0072
		(2.08)**	(2.84)***
OD*SR	(−)	−0.0348	
		(−1.67)*	
IH*SR	(−)		0.0005
			(0.96)
SIZE	?	0.0007	−0.0016
		(1.65)*	(−1.78)*
LEV	?	−0.0228	−0.0129
		(−1.99)**	(−1.69)*
ROA	?	0.2273	0.2198
		(11.20)***	(10.74)***
Year controls		Included	Included
F-statistic		21.45***	21.24***
Adjusted R^2		50.00 %	49.80 %

The numbers in parentheses are the t-statistics. ***, ** and * denote significance at the 1, 5 and 10 % levels

outside director and institutional investor. Following Conclusions can be made: (1) In China private estate listed companies, benefit embezzlement is larger when the separation of control right and cash flow right is higher; (2) If the proportion of the outside directors in the board of directors is high, the benefit embezzlement in the company can be decreased.(3) Outside director has played a very important role to supervise the controllers and ease the impact of the separation of control right and cash flow right on the benefit embezzlement;(4) As an external supervisor, institutional investor didn't play the role to ease the benefit embezzlement in our private real estate listed companies.

Based on the above conclusions, it indicates that the phenomena of the separation of double rights are ubiquitous in China private estate listed companies at present, which creates the conditions for the controllers to seek gain for themselves. Its chief consequence is that the value of the enterprise and the legal rights of minority shareholders' are hurt. And the outside directors play an important role to protect the interests of minority shareholders'. But the supervisory role of institutional investors' are not played now in China private real estate listed companies. Chinese institutional investors may be the short-term traders who are still "voting with their feet", and don't play a positive role in corporate governance.

In closing, our study didn't bring legal environment into the model, which might affect the conclusions and should be perfect in later work.

References

Claessens S, Djankov S, Fan JP (2000) The separation of ownership and control in East Asian Corporations. J Financ Econ 58(3):81–112
Gallagher PD, Kochhar R (1996) Barriers to corporate governance by institutional investors: implications for theory and practice. Eur Manage J 14(5):457–466
Gomes A (2000) Going public without governance: managerial reputation effects. J Financ 55(2):615–646
Jiang Guohua, Charles MC, Yue Heng (2009) Tunneling through inter-corporate loans: the China experience. J Financ Econ 6(1):78–83
Ke-min Wang, Mei-guang Ji, Wei Li (2009) Corporate transparency and expropriation by large shareholders. Nankai Bus Rev 12(4):83–91 (in Chinese)
Kun Wang, Xing Xiao (2005) The empirical study on the institutional ownership and resources expropriation by the related parties. Nankai Bus Rev 8(2):27–33 (in Chinese)
Lemmon M, Lins K (2003) Ownership structure, corporate governance, and firm value: evidence from the East Asian financial crisis. J Financ 58(4):1445–1468
Liu L, Peng E (2006) Institutional ownership composition and accruals quality. Working paper, California State University, 19–20
Parrino R, Sias R, Starks L (2002) Voting with their feet: institutional investors and CEO turnover. J Financ Econ 68:3–46
Peng Wang, Lian Zhou (2008) Large Shareholder's control rights, cash flow rights and corporate performance: study on family owned public firms in China. J Financ Res 2:88–98 (in Chinese)
Porta L, Silanes LD, Shleifer A (1999) Corporate ownership around the world. J Financ 54(2):471–517
Qi Gu, De-qiang Deng, Qian Lu (2006) Corporate value of separating cash flow rights from control: study on family owned public firms in China. Acc Res 4:30–36 (in Chinese)
Shan-min Li, De-you Wang, Tao Zhu (2006) The relationship between the separation of control right and cash flow right and the firms' performance. J Sun Yatsen Univ 6:83–91 (in Chinese)
Shleifer A, Wolfenzon D (2002) Investor protection and equity markets. J Financ Econ 66:3–27
Xia Chen, Jarrad Harford, Kai Li (2007) Monitoring: which institutions matter? J Financ Econ 86(2):279–305
Xian-hui Bo, Lian-sheng Wu (2009) The governance roles of state-owned controlling and institutional investors: a perspective of earnings management. Econ Res J 2:81–90 (in Chinese)
Xiao-hui Wu, Yan-fu Jiang (2006) Change of the independent director's governing efficiency under the influence of institutional investor: empirical evidence from China's listed firms. China Ind Econ 5:105–111 (in Chinese)
Yong Ye, Pei Hu, De-qing Tan (2007) Control rights deviate from cash flow rights, corporate value and corporate governance. J Eng Eng Manage 1:71–76 (in Chinese)

Chapter 65
Do Political Connections of Private Enterprises Impact Bank Loan Pricing

Gong-rong Chen, Qi-shan Li, and Cong-min Ding

Abstract Restricted by the factors of institutional system and information asymmetry, private firms are confronted with serious credit discrimination in the bank financing. Based on extant literatures studying the loan effect of political connections, this paper investigates the correlations between the political connections of private enterprises and bank loan pricing. The result of this paper indicates that political connections of private firms have the debt cost effect, thereby helping private firms finance with relatively low bank loan rates, which provides the new proof of the impact of political relationship on the credit discrimination that private corporates encounter from the bank loan pricing perspective for the extant literatures.

Keywords Bank loan rates • Credit discrimination • Debt financing costs • Political connections

65.1 Introduction

Along with the rapid development of Chinese economy, private firms have been the highlight and the important impetus that promote the economic growth of China. Nevertheless, in the background of transitional market economy, the deficiency of financial institutions makes private firms confronted with severe credit discrimination. Compared with state-owned corporates, private firms get lower loan scales

This paper is supported by the fund of China Hunan Provincial Science & Technology Department (No. 2010FJ6033).

G. Chen (✉) • Q. Li • C. Ding
Business School, Hunan University, Changsha, China
e-mail: Cgr62@126.com; L73001@126.com; Dcm2010@126.com

E. Qi et al. (eds.), *The 19th International Conference on Industrial Engineering and Engineering Management*, DOI 10.1007/978-3-642-37270-4_65,

but have to bear higher debt financing costs (Jiang Wei and Li Bin 2006; Brandt and Li 2003; Li Guangzi and Liu Li 2009). In the condition of transitional market economy, more and more literatures show that, establishing the relationship with government as private enterprises' strategic reflection to the particular institutional environments is favorable to private firms to obtain tax preference (Wu Wenfeng et al. 2009), government subsidy (Yu Minggui et al. 2010) and the allowance of entering regulated industries (Luo Danglun and Liu Xiaolong 2009). In the loan effect of political connections, the extant researches show that the political relations are conducive to private corporates to get more bank loans and longer loan maturity (Li et al. 2008; Yu minggui and Pan Hongbo 2008; Hao Xiangchao and Zhang Hongliang 2011). However, the prior documents only analyze the political connections' loan effect, but not the loan cost effect further. In reality, from the aspect of bank, when making the loan decisions, the bank will determine the loan scale and loan rate simultaneously. For the private firms, external institutional factors will affect their debt financing costs inevitably. So, what makes us curious is, in the present institutional circumstances, whether the political connections of private firms have the loan cost effect, thereby helping the political connected private firms finance with relatively low loan rates? If it is existed, how the loan cost effect of political relationship of private firms function?

For solving these questions, taking as the specimens the private direct listed A-share firms in China from 2007 to 2010, which disclose the information of data, sum, deadline and interest rate of every new bank loan simultaneously in the firms' annual financial reports, this paper aims to give a theoretical analysis and empirical test on the correlation between the political connections of private corporates and bank loan rates. And the main finding is that, in the background of transitional market economy, the political connections of private firms have the debt cost effect, which is conducive to the private corporates finance with relatively low bank loan rates, thereby decreasing firms' debt financing costs significantly. What the implication of this paper is that, it provides the new proof of the impact of political connections on the credit discrimination that private enterprises encounter in reality from the bank loan pricing perspective for the extant literatures, as well as furnishes the policy makers with direct and realistic evidence to take relevant measures to ameliorate the credit treatment of private corporates further.

65.2 Theory Analysis and Hypothesis

Based on the viewpoint of modern corporate finance theory, corporate debt costs usually depend on the corporate credit ratings, the correlation between of which is generally negative. Further, the company's credit rating is decided by a series of firms' characteristics, including the enterprise's profitability, solvency, operating risks and firm size. For private firms, influenced by the factors of institutional environment and asymmetric information, compared with state-owned enterprises, the credit rating evaluations of private companies that banks tend to give are usually a little lower, which leads to that private companies have to bear higher debt costs (Brandt and Li 2003; Li Guangzi and Liu Li 2009).

However, the situations that private corporates face may be meliorated by their political correlations. This is because that the political correlations of private companies, as the implicit relationship contracts between the private firms and the government or government officials in the specific institutional system of China, will facilitate the corporates to obtain competitive resources and enhance the adaptive ability of corporates to the external uncertain environment as well as improve the enterprises' performances undoubtedly (Zhou Xueguang 2005). In the background of transitional market economy, the incompleteness of market economy has made the government also has the ability of economic resources allocation, which means that the government can affect the convenience of the enterprise on the aspect of acquiring resources directly or indirectly through the right of control on the local economic resources (Chen Yunsen and Zhu Song 2009). Especially in relational society of China, establishing relationship with the government is beneficial to the corporate to enhance the ability and the opportunity of gaining essential resources through the nonmarket channels. The political background of corporate executives can be very important and useful tools to obtain relevant resources that the company needs in development, such as the policy resources that controlled by the government (Luo Danglun and Tang Qingquan 2009), which will decrease the risk of diversification and enhance operating performance of the firm (Zhang Min and Huang Jicheng 2009; Wu Wenfeng et al. 2008). Therefore, to private companies, the correlation between corporates and the government will transmit the information of better profitability about the political connected firms to the loan bank, which can decrease the loan bank's expectation of the uncertainty about the enterprise's performance and future cash flows. Further, when private entrepreneurs obtain the political identity through the way of being a deputy to the People's Congress or a member of the Chinese People's Political Consultative Conference, it means that the private company has owned considerable firm size, as well as the contributions the private corporate made to develop the local economy, which has been met with general acceptance of the government and local people. In that case, the political identity of a private firm itself has a very obvious function of information transmission about the corporate's economic strength (Luo Danglun and Zhen Liming 2008). Meanwhile, the government background of a executive of the private enterprise, who is a former or incumbent government official, also is beneficial to the firm to get the economic assistance and implicit guarantees from the government (Luo Danglun and Tang Qingquan 2009), which decreases the banks' expectation to the private corporate's loan defaults (Chen Yunsen and Zhu Song 2009), and then makes banks meliorate the evaluation of credit ratings as well as adjusts the loan interest rates downwards to private firms further. Therefore, building on the analysis above, this paper put forward the specific hypothesis as follows.

H1 Ceteris paribus, political correlations of private corporates have the debt costs effect, which helps to private firms finance with a low loan rate relatively.

65.3 Data and Methodology

65.3.1 Model and Tests

To examine the relation between the political connections of private corporates and the bank loan rates, this article designs the following multiple regression equation.

$$Rate_{it} = \alpha_0 + \lambda_1 PC_{it} + \lambda X_{it\ (-1)} + \varepsilon_{it} \tag{65.1}$$

where the dependent variable Rate is a firm's newly weighted average bank short-term or long-term loan rate spreads in a year. The variable is measured as the following formula, $\sum \text{rate}_{\text{it}} \times (\text{Newloan}_{\text{it}}/\text{Newtotalloan}_{\text{it}})$. Where rate_{it} is a firm's newly short-term or long-term loan rate of term i minus bank short-term or long-term benchmark lending rate of current period, $\text{Newloan}_{\text{it}}$ represents a firm's newly short-term or long-term loan rate of term i, and $\text{Newtotalloan}_{\text{it}}$ is a corporate's total sum of every newly short-term or long-term bank loan.

PL is an indicator variable, which is equal to one for political connected firms and zero otherwise. Building on Faccio et al. (2006), the article measures this variable based on the background of a firm's ultimate controlling shareholder, chairman or CEO. The firm is considered politically connected, if its ultimate controller, chairman or CEO is a former or incumbent government official, or a deputy to the People's Congress or a member of the Chinese People's Political Consultative Conference.

This paper tests the hypothesis by examining the correlation between Rate and PL. For all sample firms, we expect the coefficient of PL to be significant less than zero. To control the effect of other factors on the dependent variable, this article includes several control variables in Model (1) as follows: operating cash flow ratio (Cash), net fixed assets ratio (PPE), asset-debt ratio (Lev), return on assets (Roa), total assets turnover (Turnover), total assets growth rate (Growth) and the natural logarithm of total assets (Size), where the symbol in the bracket after every control variable is its sign of each. Considering that the loan decisions of banks are mainly influenced by enterprises' accounting information of previous year, so all of the control variables in this paper are lagged by 1 year.

65.3.2 Sample and Data

The sample of this paper is drawn from private direct listed A-share firms in China from 2007 to 2010, which disclose the information of data, sum, deadline and interest rate of every new bank loan simultaneously in the firms' annual financial report. To avoid the influence of outliers to the regression results, after eliminating the abnormal value of control variables, we obtain 67 samples of short-term bank loan rates and 229 samples of long-term bank loan rate respectively.

Particularly, the annual financial reports of corporates are from the websites of Shanghai Stock exchange and Shenzhen Stock exchange. The political backgrounds of firms' executives are abstracted form executives introduction in the financial reports of firms respectively. And the financial data of firms are from CSMAR database.

65.4 Empirical Results and Analysis

65.4.1 Summary Statistics

From the panel A of Table 65.1, we can see that the mean of short-term bank loan rates of private firms are 0.02, which shows that private firms' short-term bank loan rates are higher than the central bank benchmark loan rates in general. The mean of PL is 0.49. Moreover, the minimum of loan rate is −1.38, and the maximum is 1.71, which indicates the difference of loan rates in the sample is a little apparent.

As panel B of Table 65.1 shows, for the sample of long-term bank loan rate of private corporates, the mean of bank loan rate is −0.08, and the mean of PL is 0.54, which indicates that political connected private firms have significant debt financing costs effect on the whole, which drags down the overall mean value of the sample. In addition, the minimum and the maximum of loan rate is −1.98 and

Table 65.1 Variables

Variables	Mean	Median	Std.dev	Min	Max
Panel A Summary statistics of short-term rate model					
Short-rate	0.02	0.00	0.51	−1.38	1.71
PL	0.49	0.00	0.50	0.00	1.00
Cash	1.09	0.50	1.83	0.001	9.39
Lev	0.43	0.44	0.19	0.064	0.97
PPE	0.22	0.16	0.18	0.004	0.68
ROA	0.08	0.07	0.08	−0.19	0.46
Growth	0.53	0.29	0.60	−0.19	2.91
Turnover	0.79	0.64	0.64	0.03	3.78
Size	11.89	11.46	1.09	9.81	14.59
Panel B Summary statistics of long-term rate model					
Long-rate	−0.08	0.00	0.65	−1.98	2.88
PL	0.54	1.00	0.50	0.00	1.00
Cash	0.39	0.28	0.40	0.01	3.54
Lev	0.50	0.49	0.15	0.05	0.86
PPE	0.26	0.25	0.17	0.00	0.92
ROA	0.05	0.05	0.06	−0.14	0.46
Growth	0.24	0.15	0.42	−0.33	4.49
Turnover	0.71	0.63	0.46	0.02	3.24
Size	12.27	12.11	0.82	10.60	14.46

Table 65.2 Comparative analysis between groups

	The mean comparison			The median comparison		
Rate	*PL = 0*	*PL = 1*	*T stats*	*PL = 0*	*PL = 1*	*Z stats*
Short-rate	0.185	−0.14	2.81***	0.105	−0.045	3.54***
Long-rate	0.072	−0.22	3.53***	0.00	−0.058	3.92***

Note: *** stand for the significance at the 1 % level

2.88 respectively, indicating that the difference of long-term loan rate in the sample is distinct relatively. The summary statistics of control variables are listed in the Table 65.1.

65.4.2 Comparative Analysis Between Groups

This paper firstly divides the sample of bank loan rates into two groups according to whether the private firms are political connected, and then analyses the relevance between political connections of private firms and bank loan rates through the way of mean comparison (T test) and median comparison (Mann-Whitney test) between groups.

As Table 65.2 shows, for nonpolitical connected groups, the mean of short-term and long-term loan rate is 0.185 and 0.072 respectively; the median of short-term and long-term loan rate is 0.105 and 0 apart. For political connected groups, the mean of short-term and long-term loan rates is −0.1472 and −0.22 respectively, as well as the median of short-term and long-term loan rate apart is −0.045 and −0.058.

The result of comparative analysis shows that, no matter short-loan rates or long-term rates, the mean and the median of interest rates of political connected private firms are all lower than nonpolitical connected private firms distinctly. Moreover, the difference of loan rates between two groups has statistical significance. In brief, the analysis results above validate the hypothesis in this paper preliminary, which provides the empirical basis for the further multiple regression analysis.

65.4.3 Empirical Results

Table 65.3 gives a report of the regression results of model (1). In the model of short-term loan rates, the coefficient on PL is significantly negative, indicating a negative relation between political connections of private firms and short-term loan rates. For the control variables, Roa and Size are both negatively related to the short-term loan rates at the 5 % significant level. The value of adjusted R^2 is 0.252, and the value of F statistics is 3.778 at the significance of 1 %.

Table 65.3 The results of multiple regression

	Short-rate model		Long-rate model	
Variables	*Coefficient*	*T value*	*Coefficient*	*T value*
Cons	2.101**	2.642	1.437**	2.158
PL	−0.206*	−1.697	−0.193**	−2.212
Cash	−0.020	−0.403	−0.240*	−1.956
PPE	0.120	0.333	−0.678**	−2.57
Lev	0.605	1.114	0.282	0.757
Roa	−1.963**	−2.304	−0.858	−1.108
Turnover	0.127	1.221	−0.127	−1.363
Growth	0.169	1.285	0.158	1.532
Size	−0.192**	−2.302	−0.097*	−1.684
Adj. R^2	25.2 %		10.7 %	
F value	3.778***		4.426***	

Note: ***, **, * denote the significance at the 1 %, 5 % and 10 % level, respectively

In the model of long-term loan rates, PL is negatively related to long-term loan rates at the significance level of 5 %, which means that the long-term bank loan rates of political connected private firms are lower than that of nonpolitical private. For the control variables, the coefficient on Cash, PPE and Size are all negatively significant. The value of adjusted R^2 is 0.107, and the value of F statistics is 4.426 at the significance of 1 %. Moreover, the value of variance inflation factors in the bank loan rates model are all no more than 4, indicating that there is no multiple linear problem among the variables, which is in line with the requirement of multiple linear regression.

The empirical result above shows obviously that, for private firms, the debt costs effect of political connections exist in deed, which is beneficial to the private firms finance with a relatively low bank loan rates, and then decrease the political connected private firms' bank financing costs. All in all, the hypothesis of this paper is verified.

65.5 Conclusion and Discussion

Due to the structural contradictions existed in the financial system; private firms are confronted with serious credit discrimination from the bank (Brandt and Li 2003). However, with the reform of financial system marketization, there is not apparent tendency that credit discrimination on private firms has mitigated in recent years (Li Guangzi and Liu Li 2009). Based on extant literature studying the loan effect of private firm's political connection, this article emphasizes the research of the impact of political connections on credit discrimination that private firms encounter in reality from the debt financing cost perspective. As the result of this study show that, in the background of transitional market economy, the political connections of private firms formed by the way of private entrepreneurs' political participation and other means can influence bank's loan risk judgment and then the loan pricing to the private firms, which is conducive to the private firms to finance with a relatively low bank loan rate, and decrease the firms debt financing costs further.

Referring to the study of Li and Liu (2009), it is not difficult to conjecture that, in the conditions of transitional market economy, the reasons that political connections of private firms could play the role of debt financing costs effect lie in the fact that private firms are confronted with serious credit discrimination and have to pay higher loan costs for getting the bank loans. However, getting the support of external funding is one of the most important factors that spur the rapid development of private firms. But more importantly, the further development of the private economy has a critical significance in promoting social capital accumulation and improving social production efficiency. Therefore, the policy implications of this paper is that, for the policymakers, they should proceed with pushing forward the reform of financial system marketization, such as through the way of encouraging and supporting the development of local joint-stock commercial banks and other financial institutions vigorously to break the structural contradiction between the state-owned finance and private economy, and establishing relevant policies to conduct the commercial banks credit behaviors to provide the private firms with good credit environment and then the ideal credit treatment. Additionally, in the existing institutional environment, private firms should enhance and tamp the management level constantly, and by means of improving the sustainable profitability as well as cultivating the credit and reputation of enterprises themselves to improve corporates' financing qualification. Besides, in the backward financial institution of China, as an informal substitutable mechanism to the existing institutions, firm-bank relationship also has the bank debt contract function. Thereby, private firms can establish the firm-bank relationship by use of the long-term credit business with banks or employing the person with bank background as the corporate executive, to decrease the predicaments of credit discrimination purposefully.

References

Brandt L, Li H (2003) Bank discrimination in transition economies: ideology, information, or incentives? J Comp Econ 31:387–413

Chen Yunsen, Zhu Song (2009) Political connection, institutional environment and capital investment of listed companies. J Financ Econ 12:27–39 (in Chinese)

Faccio M, Masulis RW, McConnell JJ (2006) Political connections and corporate bailouts. J Financ 61:2597–2635

Hao Xiangchao, Zhang Hongliang (2011) Impacts of political status and background of government officials on access to bank loans for private firms in China. Financ Trade Econ 4:55–61 (in Chinese)

Jiang Wei, Li Bin (2006) Institutional environment, state ownership and leading discrimination. J Financ Res 11:116–126 (in Chinese)

Li Guangzi, Liu Li (2009) Debt financing cost and credit discrimination on private listed firms. J Financ Res 12:137–150 (in Chinese)

Li H, Meng L, Wang Q, Zhou L (2008) Political connections, financing and firm performance: evidence from Chinese private firms. J Dev Econ 87:283–299

Luo Danglun, Liu Xiaolong (2009) Political relationship, barriers to entry and enterprise performance: the evidences from the privately listed enterprises. Manage World 5:97–106 (in Chinese)

Luo Danglun, Tang Qingquan (2009) Political relationship, social capital and policy resource gains. J World Econ 7:84–96 (in Chinese)

Luo Danglun, Zhen Liming (2008) Private control, political relationship and financing constrain of private listed enterprises. J Financ Res 12:164–178 (in Chinese)

Wu Wenfeng, Wu Chongfeng, Liu Xiaowei (2008) Political connection and market valuation: evidence from China individual-controlled listed firms. Econ Res J 7:130–141 (in Chinese)

Wu Wenfeng, Wu Chongfeng, Rui Meng (2009) Between the special connections that high-ranking managers of some of China's listed companies have with the government and tax preference afforded to them. Manage World 3:134–142 (in Chinese)

Yu minggui, Pan Hongbo (2008) The relationship between politics, institutional environments and Private enterprises' access to bank loans. Manage World 8:9–21 (in Chinese)

Yu Minggui, Hui Yafu, Pan Hongbo (2010) Political connections, rent seeking, and the fiscal subsidy efficiency of local governments. Econ Res J 3:65–77 (in Chinese)

Zhang Min, Huang Jicheng (2009) Political connections, diversification and firm's risks. Manage World 7:156–164 (in Chinese)

Zhou Xueguang (2005) Property rights as a relational concept: a sociological approach. Sociol Res 2:1–31 (in Chinese)

Chapter 66
How the Largest Shareholder Control Affects the Company's Cash Value

Zhi-xiong Ling and Chun-ling Li

Abstract Starting from "incentive effects" and "encroachment effects" of the largest shareholders, this article analyzes how the largest shareholder's changes of stake and nature affect cash value. Empirical studies show that, when financing constraints, caused by the largest shareholder's moral hazard, do not exist, the relationship between stake and cash value is significant inverted U-shaped, and while the largest shareholder is government, it will weaken the positive effect of cash holdings on firm value; relative to non-government-controlled companies, the government-controlled company's cash value is lower. When the above-mentioned financing constraints exist, the largest shareholder's influences on cash value is not significant, but it can be seen the largest shareholder has negative effect on cash value in company with a high degree of financing constraints, which indicates that in the financing constrained company, largest shareholder's "encroachment effects" is on the performance.

Keywords Cash value • Encroachment effects • Incentive effects • The largest shareholder

66.1 Introduction

In recent years, scholars' researches about cash value focus on the static tradeoff theory, financing pecking order theory, free cash flow theory, investors legal protection, company characteristics, company governance mechanisms etc. In the domestic, a small amount of researches combined the controlling shareholder and cash holdings to consider the impact of this both on the value of the company, but these studies

Z. Ling • C. Li (✉)
School of Business and Management, Hunan University, Changsha, China
e-mail: lichunling1987@163.com

E. Qi et al. (eds.), *The 19th International Conference on Industrial Engineering and Engineering Management*, DOI 10.1007/978-3-642-37270-4_66,

only start from the perspective of the nature of the controlling shareholder, and no further study appeared to analyze the role of the largest shareholder's stake on cash holdings. This article analyzes how the largest shareholder's changes of stake and nature affect cash value, and provides empirical evidence for optimizing corporate cash holdings decision-making.

66.2 Literature Review and Research Hypotheses

There are two mainstream views in theoretical circles about the role of large shareholders on corporate governance and value creation. One view is "incentive effect" (also known as the "monitor effect"), Shleifer and Vishny show the largest shareholder will supervise managers and solve shareholders "free rider" problem (Shleife and Vishny 1997). When a considerable number of shares concentrated in the hands of a few large shareholders and the control right benefits outweigh the costs of monitoring, the largest shareholders will inspire managers, forcing the managers to act in accordance with the benefits of the shareholders, to maximize shareholder wealth. Another view is "encroachment effect" (also known as the "defense effect"), LaPorta et al. (2000) believes that the interests of the largest shareholders are not consistent with external minority shareholders', and there often appear serious conflicts between the two. While lacking of external oversight or the type of external shareholders diversified, the largest shareholders may use its controlling position to reap the private benefits, damaging other shareholders' interests.

The value of cash assets constitutes a large part of the enterprise value. Because of its strong liquidity, cash assets more likely to be a tool used by the largest shareholder to seek private benefits (Qi Luo and Guo-lou Qin 2009). So what the impacts of incentive effect and encroachment effect on the cash value are, foreign scholars have done the following research. First, these scholars believe that the largest shareholder's "incentive effect" reduces the agency costs between shareholders and managers, and the largest shareholders motivate managers to protect the interests of investors, thereby enhancing cash value. Faulkender (2002) found managers' ownership and corporate cash holdings significantly negatively correlated, and when increase proportion of the largest shareholder, cash holdings will decline, suggesting that the largest shareholders reduce agency problems, enhance corporate value. Second, the largest shareholder's "encroachment effect" increases the agency costs between the largest shareholders and minority shareholders, and harm interests of small and medium investors, thereby reducing value company's cash. Kalcheva and Lins (2007) have shown that controlling shareholder preferences in over-investment in liquid assets. According to theoretical predictions of hypothesis "encroachment effect", when the largest shareholder of listed companies has the motivation to obtain private benefits through the accumulation of large amounts of cash resources, rational investors will give cash lower value.

Based on the specific national conditions, China's economic and institutional environment are not yet ripe, and overall level of corporate governance is lower, as well as legal investor protection. The largest shareholder more often seizes minority shareholders' interests, and managers damage the interests of investors. At the same time, China's listed companies generally hold large amounts of cash; this provides an opportunity for interest transfer behaviors of the largest shareholder and management. On the one hand, due to the active participation of the largest shareholder, minority shareholders and the whole group will benefit from the increases of corporate value. That is to say, when the largest shareholders play a positive incentive effect, value cash is higher. On the other hand, in pursuit of maximizing their own interests, the major shareholder's behavior maybe deviate minority shareholders, even sacrificing minority shareholders interests, which lead to decline value of the company, forming "encroachment effect". When the largest shareholder's negative "encroachment effects" works, company's cash more tool to transfer resources by the largest shareholder. The company's cash value reduced. Based on the above analysis, we propose Hypothesis 1:

Hypothesis 1 The relationship between stake of the largest shareholder and cash value is significant inverted U-shaped.

Most Chinese listed companies restructure from state-owned enterprises. Government's motives and state-owned property nature play an important role on cash value. On one hand, government, as the largest shareholder of government-controlled companies supervise managers less efficient than non-government-controlled enterprises. In government-controlled company, the largest shareholder's "incentive effects" is weaker. Due to neither power nor capacity to supervise managers effectively, government-controlled companies' managers are more prone to acts of misuse of company cash resources to build a personal empire, low value of the cash (Chong-en et al. 2005). The other hand, relative to non-government-controlled companies, the government-controlled companies operate inefficiency, and the possibility of "encroachment effect" occurs. Zeng-quan Li et al. (2005) found that the largest shareholder and local governments in order to maintain the economic and social development of parent company, or maintain the continued existence of the local government, endowed with the motive to transfer resources from the listed companies. Some government officials may use the power in his hands, transfer the wealth of the state-owned enterprises, and encroach on the economic interests (Li-hui Tian 2005). The largest shareholder of most listed companies are acted by government; legally binding is difficult to limit the power of the government (Fang Yi-qiang and Li-jun Xia 2005). Meanwhile, the utility function of non-government – controlled property right is to maximize the pursuit of corporate profits, and less commitment to the social function and the policy burden, the company's behavior more follow the principle of market competition as opposed to the government-controlled company; the cash resources more are used to increase value of the company; cash value is also higher. Then, the second hypothesis can be expressed as follows:

Hypothesis 2 Compared to non-government-controlled companies, the largest shareholder's "incentive effects" of government-controlled companies is weaker; "encroachment effect" is higher, and cash value is lower.

66.3 Research Design

66.3.1 Sample Selection and Data Sources

Samples of this article are the A-share companies listed on Shanghai and Shenzhen. The observed time interval contains 9 years of panel data from 2003 to 2009. Because of using the relevant data of the previous year and the next year, the actual time interval is from 2002 to 2010. In the article we removed the following companies: financial sector and public utilities; Solvency abnormal financial status of the company; the company which cannot be traced to the nature of its largest shareholder; ST companies; IPO companies. After removing, 744 companies with 5,208 observations remain.

66.3.2 Model Design and Variable Definition

Literatures about company's cash value use two models. One is classic enterprise value regression model used in the study of Fama and French (1998), which checked effects of liabilities and dividends on corporate value. Another is the model used by Drobetza et al. (2010), the author instead the total cash holdings of excess cash holdings (defined as more than optimal level cash holdings). Scholars that use this model believe the existence of optimal level cash holdings. We use the first model, and make appropriate adjustments based on this model. Our model is designed as follows:

$$
\begin{aligned}
V_n = {} & \beta_0 + \beta_1 E + \beta_2 dE_1 + \beta_3 dE_2 + \beta_4 D + \beta_5 d D_1 \\
& + \beta_6 dD_2 + \beta_7 I + \beta_8 dI_1 + \beta_9 dI_2 + \beta_{10} NA + \beta_{11} dNA_2 \\
& + \beta_{12} dV_2 + \beta_{13} top1 + \beta_{14} cash + \beta_{15} top1 \times cash \\
& + \beta_{16} (top1 \times cash)^2 + \beta_{17} state + \beta_{18} state \times cash \qquad (66.1)
\end{aligned}
$$

Where V denotes the total market value of the firm (market value of equity plus book value of debt); E is earnings before interest and extraordinary items (after depreciation and taxes); dE_1, dE_2 expressed difference between the current (next issue) interest profit before tax and the previous (current) EBIT; D is total dividends paid; dD_1, dD_2 expressed the difference between the current (next) cash dividends and the previous (current) cash dividends; I is interest expense; dI_1, dI_2 expressed the difference between the current (next) interest expense and the previous (current) interest expense; NA is non-current assets; dNA_2 expressed the difference between the next issue of non-current assets and the current non-current assets; dV_2 expressed the difference between next phase of market value and the current market value; top1 is the stake of the largest shareholder; cash is cash holdings; state is a dummy variable expressing the nature of the company's largest shareholder; when the largest shareholder is government, the value of state is one, otherwise zero; the dependent variables are cash, state, and top1. The coefficient β16 and β18 show effects of the proportion and nature of the largest shareholders on cash value. The coefficient β14 on cash holdings is an estimate of the market value of cash. The differences between our model and the above two are as follows: first, because financial reports of listed companies in China don't list R & D expenditure; relevant data unable to get, so our model exclude R & D expenditures. Second, our model introduces the largest shareholding ratio variable, nature of the dummy variable, and these two variables interaction terms with the cash holdings. In order to test hypothesis 1, we also introduced in the square of interaction terms between the largest shareholder proportion and cash holdings.

66.4 Empirical Test Results

66.4.1 Descriptive Statistics

Xing-quan Yang and Zhao-nan Zhang (2010) show that cash holdings have a positive effect on firm value, the value of the company will increase with the increase of cash holdings. It can be seen from Table 66.1, the mean value of cash holdings in the two intervals of stake respectively are 0.1452 and 0.1516. Cash holdings show an increasing trend with the increase of the stake. However, the mean of companies' market value in the two intervals of stake respectively are 1.5721 and 1.3554. With the increase of the stake, market value is a downward trend. Our sample data don't show the positive effect of cash holdings on company value, indicating that the roles of cash holdings on company value are different with largest shareholding stake.

Yi-feng et al. (2008) showed that high level cash holdings is not conducive to maximize the company value, when referencing high level cash holdings, the state-owned controlling shareholders is on the motivation of interests occupation. Table 66.2 shows mean of cash holdings in government-controlled samples and non-governmental controlled samples, 0.1464 and 0.1518, showing little difference. But the mean 1.3261 of government-controlled corporate value is significantly lower than the mean 1.6904 of non-government-controlled companies'. Therefore, we predict different nature of the largest shareholder lead to different behaviors and different roles of cash holdings on firm value.

Table 66.1 Descriptive statistics of cash holdings and corporate value

Group		Mean	Median	Maximum	Minimum	S.D.
top1<median	Cash	0.1452	0.1250	0.7215	0.0012	0.0999
	v	1.5721	1.2449	14.3832	0.0000	0.9697
top1≥median	Cash	0.1516	0.1267	0.7768	0.0013	0.1050
	v	1.3554	1.1298	11.2219	0.0000	0.7244

Table 66.2 Descriptive statistics of cash holding and corporate value under different nature of the largest shareholder

Group		Mean	Median	Maximum	Minimum	S.D.
Government-controlled	Cash	0.1464	0.1247	0.7452	0.0013	0.1006
	v	1.3261	1.1276	8.4182	0.0000	0.6314
Non-government-controlled	Cash	0.1518	0.128	0.7768	0.0012	0.1056
	v	1.6904	1.3016	14.3832	0.0000	1.1094

Table 66.3 Variable regression results (The full sample)

	Model 1	Model 2	Model 3
Constant	2.3798*	3.0832*	2.9036*
cash	0.1521*	0.1670*	0.1401*
top1		−0.9411*	
top1*cash		0.0269	
$(\text{top1*cash})^2$		−3.8604	
state			−0.2089*
state*cash			−0.7869*
E	3.7351*	3.8872*	3.5865*
dE_1	−0.1417	−0.1539	−0.1151
dE_2	2.2001*	2.2875*	2.1420*
D	−7.1697*	−6.9110*	−6.2626**
dD_1	9.4804*	7.6611*	8.0526*
dD_2	−0.7934	−0.9925	−0.5595
I	−0.1331	−2.2464	−1.2545
dI_1	−18.0081*	−16.0632*	−15.7661*
dI2	−0.7475	−0.7039	−0.7765
NA	−1.6118*	−2.0346*	−1.9405*
dNA_2	0.1676*	0.1755*	0.1732*
dV_2	−0.1575*	−0.1647*	−0.1592*
n	5,208	5,208	5,208
R^2	0.1324	0.1671	0.1676
Adjusted R^2	0.1302	0.1646	0.1652

*, **Significance at the 1 %, 5 % level

66.4.2 Empirical Results

According to BP test and Hausman test, we rejected mixed regression and random effects regression and adopts a fixed effects regression. Table 66.3 reports regression results of the largest shareholder stake, nature and cash value.

Model 1 shows the regression results without considering the largest shareholder, the coefficient before cash is 0.1521 and significant at 1 % level, indicating that the company's cash value is 0.1521, and appears a significant discount. On the basis of model 1, model 2 introduces the largest shareholder proportion and the interaction terms between top1 and cash holdings top1*cash, as well as introduces the square of the interaction terms $(\text{top1*cash})^2$. The coefficient before cash is 0.1670, and before (top1 * cash)^2 is −3.8604, not significant. The results show that relationship between the largest shareholder ratio and cash value is inverted U-shaped, that is to say there both exist "incentive effects" and "encroachment effect". Model 3 introduces dummy variable of the nature of the largest shareholder and interaction terms between the nature and cash holdings state * cash. the coefficient before cash is 0.1401, and before state * cash is significantly negative, indicating that, comparing to non-government-controlled companies, the positive effect of cash holdings on firm value in government-controlled companies is weak, lower cash value.

Regression results of model 2 show that the inverted-U relationship between the largest shareholder and the cash value is not significant, indicating that there may be other factors that affect the two effects of the largest shareholder on cash value. In order to verify whether financing constraints caused by the largest shareholder's moral hazard lead to this insignificant relationship, we use the cash – cash flow sensitivity to measure the company's financing constraints; high cash – cash flow sensitivity leads to high degree of financing constraints; low cash-cash sensitivity lead to low degree of financing const-raint. As the financing constraints eventually caused by the asset preferences and motivation of interest occupations of the large shareholder, separation of control rights and cash flow rights can reflect the strength of the largest shareholder's motivation of interest occupation. Qi Luo and Zhi-qiang Hu (2011) shows that a high degree of separation of the two rights, a high cash-cash flow sensitivity; so we chose the degree of separation of the two rights to group samples. First, according to the degree of separation of the two rights, we divide the total sample into two sub-samples; the two rights undetached samples (non-existent financing constraints sample) names the ratio of control rights to cash flow right equal to 1, and the two rights detached samples (the existence of financing constraints samples) names the ratio of control rights to cash flow right is greater than 1. For the two rights detached samples, using the median as the cut-off point, we divide it into two sub-samples, the high degree of separation (high degree of financing constraints) and low degree of separation (low degree of financing constraints) groups. Specific groups are shown in Table 66.4.

Table 66.4 shows that in the non-existent financing constraints samples, the coefficient of (top1*cash)^2 is significantly negative, as well as the coefficient of state * cash, while in the financing constraints samples the coefficient of (top1*cash)^2 is insignificantly negative, and the coefficient of state * cash is significantly negative. In low degree of financing constraints samples the coefficients of top*cash and state*cash are respectively positive and negative; both are insignificantly. But this two coefficients in high degree of financing constraints samples are both negative. This shows that financing constraints caused by the largest shareholder moral hazard affect the cash value. When financing constraints does not exist, the relationship between the largest shareholder proportion and cash value is inverted U-shaped. Insignificant positive correlation between the stake and the cash value in the low degree of financing constraints samples indicates that "incentive effects" play a major role. The coefficient of interaction term between the largest shareholder proportion and cash holdings in the sample with a high degree of financing

Table 66.4 Variable regression results (The sub-samples)

	Non-existent financing constraints			Financing constraints		
	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9
Constant	2.9960*	4.1856*	3.4401*	2.1064*	2.5448*	2.5720*
cash	0.2661*	0.2881**	0.2401*	0.1019*	0.1299**	0.0930*
top1		−0.9852			−0.7282**	
top1*cash		0.0267			−0.0355	
$(\text{top1*cash})^2$		11.8567**				1.8434
state			−0.1759**			−0.2408*
state*cash			−0.9491**			−0.5794**
n	1,886	1,886	1,886	3,322	3,322	3,322
R^2	0.1317	0.1754	0.1597	0.1435	0.1733	0.1818
Adjusted R^2	0.1257	0.1684	0.153	0.1401	0.1693	0.178
	Low degree of financing constraints			High degree of financing constraints		
	Model 10	Model 11	Model 12	Model 13	Model 14	Model 15
Constant	2.2977*	2.7604*	2.4501*	1.9062*	3.9697*	2.7949*
cash	0.1502*	0.1185***	0.1375*	0.0521	0.0599	0.0491
top1		−1.2355*			0.6587**	
top1*cash		0.1233			−0.6017	
state			−0.3397*			−0.1281***
state*cash			−0.111			−1.1960*
n	1,658	1,658	1,658	1,664	1,664	1,664
R^2	0.1572	0.1929	0.2011	0.1368	0.1278	0.173
Adjusted R^2	0.1505	0.1855	0.1938	0.13	0.1204	0.1655

As there are too many variables, the table only lists the main variables of the regression results; *, **, *** Significance at the 1 %, 5 %, 10 % level

constraints is negative, indicating that with the increase of the degree of financing constraints "encroachment effects" play a major role. The coefficients of interaction term between nature of the largest shareholder and cash holdings in the four samples government-controlled companies, the positive effect of cash holdings on firm value in government-controlled companies is weak, lower cash value.

66.4.3 Robustness Test

Dittmar, Mahrt-Smith and Servaes believe, while examining the company's cash holdings, we use the company's sales income as the denominator deflating better than total assets, because the sales revenue less subject to accounting conservatism (Dittmar et al. 2003). In order to test the sensitivity of the conclusions to the variable definition of cash holdings, we also used the level of sales revenue to deflate the company's cash holding. However, when we made the above changes, the results of regression analysis had no substantive changes, showing the preceding conclusions of the study are robust.

66.5 Conclusion

The conclusion of empirical studies indicate that when financing constraints, caused by the largest shareholder's moral hazard, do not exist, the relationship between stake and cash value is significant inverted U-shaped. It means that the shareholder's moral hazard has both positive incentive effect and negative encroaching effect. While shareholder's moral hazard increases, cash holdings strengthen the positive effect on firm value and cash value increases. On the other hand, while the shareholder's moral hazard increase, cash hoard weakens the positive effect on firm value and cash value decreases. When the government acts as the largest shareholder, it will reduce the positive effect on firm value from cash hoard. Comparing to non-government-controlled companies, the government-controlled company's cash value is lower and equity restriction can increase cash value. When exists the above-mentioned financing constraints, the largest shareholder's influences on cash value is not significant but it can be seen in company with a high degree of financing constraints, the largest shareholder has negative effect on cash value, which indicates that in the financing constrained company, largest shareholder's "encroachment effects" is on the performance.

The significance of research conclusion lies on that decision-making on optimizing the cash holding is based on the largest shareholder's changes of stake and nature affect cash value. When the largest shareholder brings positive incentive effect, company chooses to hold more cash which is beneficial to increase firm value. On the contract, when the largest shareholder play a negative encroachment effects, the company can reduce the amount of cash holdings, and to avoid unnecessary waste of cash. The largest shareholder of the cash value is also affected by the nature of government control; to reduce the level of company's government-controlled, can also increase the value of the company's cash. On the same time, strengthen the prevention of the largest shareholder moral hazard also help to improve the value of the company's cash.

References

Chong-en Bai, Xiao Liu, Zhou Lu, Min Song, Jun-xi Zhang (2005) An empirical study on Chinese listed firms' corporate governance. Econ Res 51(2):81–91 (in Chinese)

Dittmar A, Mahrt-Smith J, Servaes H (2003) International corporate governance and corporate cash holdings. J Financ Quant Anal 38:111–133

Drobetz W, Grüningerb MC, Hirschvoglc S (2010) Information asymmetry and the value of cash. J Bank Financ 34:2168–2184

Fama EF, French KR (1998) Taxes, financing decisions, and firm value. J Financ 53:819–843

Fang Yi-qiang, Li-jun Xia (2005) Government control, environmental management and corporate value – empirical evidence from China's securities market. Econ Res 51(5):40–51 (in Chinese)

Faulkender M (2002) Cash holdings among small businesses (unpublished)

Kalcheva I, Lins KV (2007) International evidence on cash holdings and expected managerial agency problems. Rev Financ Stud 20:1087–1112

La Porta R, Lopez-de Silanes F, Shleifer A, Vishny RW (2000) Agency problems and dividend policy around the world. J Financ 55:1–33
Li-hui Tian (2005) Two hands of the government shareholder and the U curve between government shareholdings and corporate value. Econ Res 10:48–58 (in Chinese)
Qi Luo, Guo-lou Qin (2009) Investor protection and corporate cash holdings. Financ Res 10:162–178 (in Chinese)
Qi Luo, Zhi-qiang Hu (2011) Controlling shareholder moral hazard and the strategy of the company's cash. Econ Res 57(2):125–137 (in Chinese)
Shleife A, Vishny R (1997) A survey of corporate governance. J Financ 52:5–23
Xing-quan Yang, Zhao-nan Zhang (2010) Governance environment, the government controls and cash value. Econ Res 56(2):36–43 (in Chinese)
Yi-feng Shen, Xue-wen Kuang, Ya-juan Nie (2008) The impact of excess control of ultimate controlling shareholders on the value of cash holdings. Nan-kai Bus Rev 24(1):15–23 (in Chinese)
Zen-quan Li, Qian Yu, Xiao-kun Wang (2005) Tunneling, propping and M&A: evidence from Chinese listed companies. Econ Res 1:95–105 (in Chinese)

Chapter 67
Impact of the Corporate Income Tax Reform on Capital Structure Choices: Evidence from Data of Chinese Listed Firms

Qing-mei Tan

Abstract This paper investigates the influence of the corporate income tax reform in 2007 on capital structure choices of firms in China. The results show that corporate income tax reform decreases the effective tax rate of firms totally. The effective tax rate decreases from the year of the reform and then goes up with the increase of the average leverage during the year of 2009–2010. The results also show that the corporate income tax rate is a determinant of capital structure and the change of effective tax rate has positive impact on the change of leverage. There is no evidence that non-debt tax shields are determinants of capital structure and there is not a substitution effect between debt and non-debt tax shields in our sample.

Keywords Corporate income tax reform • Capital structure • Effective tax rate

67.1 Introduction

The corporate income tax reform is implemented in 2007 in China. On March 16, 2007, National People's Congress announced *Corporate Income Tax Law of China* (short for *CITLC*), which is officially implemented on January 1, 2008 in China. At the same time, *Income Tax Law of China for Foreign-invested Enterprise and Foreign Enterprise* which is adopted on April 9, 1991 and *Provisional Regulations of China on Corporate Income Tax* which are promulgated on December 13, 1993 are both repealed simultaneously. As a consequence, the statutory corporate tax rates for all resident taxpayers are 25 % except for some tax preference under *CITLC*, which means that the statutory corporate tax rate for most resident firms in China decrease by 8 % (from 33 to 25 %) from January 1, 2008. Although firms can avoid

Q. Tan (✉)
College of Management and Economics, Tianjin University, Tianjin, China
e-mail: tanqm@tju.edu.cn

E. Qi et al. (eds.), *The 19th International Conference on Industrial Engineering and Engineering Management*, DOI 10.1007/978-3-642-37270-4_67,

tax payments using alternative instruments, since interest expenses are deducted before corporate income tax, the corporate income tax reform will decrease the tax advantage of debt relative to equity and may have influence on firm's financing behavior, financing decision and capital structure choice, and then have impact on firm value. *CITLC* has been implementing for more than 4 years since January 1, 2008, if the capital structure has been changed since the corporate income tax reform? This paper aims at investigating whether firms adjust their capital structure in response to the corporate income tax reform based on the data of listed firms in China for the period from 2006 to 2010. Our paper sheds light on the understanding of the correlation between corporate income tax and capital structure in emerging markets.

67.2 Literature Review

MM theorem shows that a firm can't create value by choosing a specific capital structure in a world without taxes or other imperfections (Rajan and Zingales 1995). In 1963, Modigliani and Miller revise the MM theorem and argue that debt financing is associated with a high tax advantage since interest expenses are tax-deductible whereas dividend payments are not (Rajan and Zingales 1995). DeAngelo and Masulis (1980) argue that there is a firm-specific optimal capital structure by reducing the corporate income tax burden into account (DeAngelo and Masulis 1980). Brigham and Gapenski (1996) argue although that the MM theorem model may be true in theory, bankruptcy costs exist and they may increase when equity is traded off for liability in practice (Brigham and Gapenski 1996). Subsequently, they assert that firms adjust toward their target capital structure by offsetting the benefits of the tax shields that liabilities attract with the expected financial distress costs of firms. Rajan and Zingales (1995) find that the net tax advantages are positively related with leverage changes in most samples (Rajan and Zingales 1995). There are many empirical studies on corporate income tax and capital structure, but the results are mixed. Givoly et al. (1992) investigate the changes in enterprise financial policy following the 1986 US Tax Reform Act, and find both corporate income taxes and non-debt tax shields are determinants of capital structure, there is positive correlation between changes in leverage and changes in corporate income tax rates (Givoly et al. 1992).

Graham investigates the effect of taxes on corporate debt policy and provides evidence that high tax rate firms raise more debt, and marginal tax rates and leverage are simultaneously determined (Graham 1996). Gordon and Lee investigate the effects of corporate income tax rate change on debt policy of firms of different sizes, the results show that there is statistically significant relationship between taxes and debt levels (Gordon and Lee 2001). Overesch and Voeller analyze the effect of taxation on capital structure and find a higher tax benefit of debt has positive influence on capital structure (Overesch and Voeller 2008). Charalambakis, Espenlaub and Garrett find evidence of a positive relationship between taxes and leverage (Charalambakis et al. 2008). Wu and Yue investigate the effect of the increase of corporate tax rate in the year of 2002 on capital structure of listed firms in China, the result shows that the leverage of listed firms increases with the increase of corporate tax rate (Wu and Yue 2009).

According to the analysis of the preceding context, most empirical tests support that the tax deduction effect of debt interest payment has impact on capital structure decision. The statutory corporate tax rate of most listed firms decreases by 8 % after the corporate income tax reform, which will decrease the tax advantage of debt relative to equity. Although firms can avoid tax payments using alternative instruments, firms who intend to utilize tax advantage may decrease debt finance in response to the decrease of corporate income tax rate. However, since the financing sources in China are usually limited, even firms feel the benefits of adjusting capital structure, the adjustment may be constrained by the limited financing sources.

67.3 Research Design

67.3.1 Dependent Variables

We use two dependent variables, which are leverage and the change of leverage, respectively. Firstly, we investigate the impact of corporate income tax on leverage, and then we investigate the impact of the change of corporate tax rate on the change of leverage. Leverage (*LEV*) is defined as total debt-to-total assets ratio; the change of leverage (ΔLEV) is defined as the change in the level of leverage at year t minus the level of leverage at year $t-1$.

67.3.2 Independent Variables

Although most firms face the same statutory tax rate, capital structure decisions are not based on the statutory tax rate, but rather on the effective tax rate. In order to investigate the impact of corporate income tax reform on capital structure decision, we use effective tax rate and the change of effective tax rate as independent variables, respectively. The effective tax rate (*ETR*) can be calculated as follows: (1) If the taxable income is more than 0, the effective tax rate will be calculated as the ratio of income tax expenses to taxable income. (2) If the taxable income is less than or equal to 0, the effective tax rate will be 0. The change of effective tax rate (ΔETR) is defined as the change in the level of effective tax rate at year t minus the level of effective tax rate at year $t-1$.

67.3.3 Control Variables

We also choose some control variables, such as profitability, growth, firm size, non-debt tax shields and collateral value of assets to control for their impact on capital structure choice.

Profitability: According to tax-based models, since firms with high profitability have greater demands to shield income from corporate income tax, they should issue more debt compared with firms with poor profitability. The trade off theory argues that optimal capital structure lies at the point where incremental tax disadvantages and advantages from increased risk of financial distress are equal. As a consequence, firms with high profitability and lower risk of financial distress should have higher debt level. However, pecking order theory predicts that firms with high profitability can finance their growth using earned profit and tend to issue less debt, whereas firms with low profitability will be forced to use debt finance. Verschueren also find that there is negative relationship between profitability and capital structure (Verschueren 2001). In this study, return on equity (*ROE*) is used to measure firm's profitability.

Growth: According to Jensen and Meckling, too much debt can lead to asset substitution problem which may create financial distress and even bankruptcy. Therefore, firms with high growth opportunities are less likely to raise a loan (Jensen and Meckling 1976). Myers argues since higher leverage may cause the loss of growth opportunities, firms with more growth opportunities should use less debt (Myers 1977). However, the tradeoff theory points out that growth opportunity will reduce the leverage since firm growth increases financial distress cost and exacerbates debt-related agency problems. By contrast, pecking order theory argues that firms with high growth opportunities and big financing demands will increase leverage. We use sales growth rate as the proxy of growth (*GROW*), which is defined as sale revenues in year t divided by sale revenues in year t-1 minus 1.

Firm size: Larger firms may change their capital structure more readily than small firms (Givoly et al. 1992). Many previous studies find that there is positive relationship between capital structure and firm size (Rajan and Zingales 1995). However, Gordon and Lee suggest that larger firms rely less heavily on debt finance than small firms, independent of tax policy (Gordon and Lee 2001). Overesch and Voeller find that the capital structure of larger firms respond less heavily to the change in the tax benefit of debt (Overesch and Voeller 2008). Firm size is computed as the natural logarithm of total assets (Ln*A*) in this paper.

Non-debt tax shields: Some earlier studies, such as Titman and Wessles find no evidence that there is substitution effect between leverage and non-debt tax shields (Myers 1977). However, recent studies find that the non-debt tax shields and leverage are substituted (Overesch and Voeller 2008; Myers 1977). As a consequence, firms can choose to change leverage or change non-debt tax shields in response to the corporate tax reform. We use depreciation-to-total assets ratio as a proxy for non-debt tax shields (*NDTS*). In order to investigate the impact of non-debt tax shields on the change of leverage, we also use the change of non-debt tax shields ($\Delta NDTS$) as a control variable, which is defined as the non-debt tax shields in year t minus the non-debt tax shields in year $t-1$.

Collateral value of assets: Collateral value of firm's assets is a determinant of optimal capital structure (Myers 1977). A higher collateral value of assets is expected to be associated with high leverage (Titman and Wessels 1988). Myers and Majluf predict that the leverage is positively related with collateral value of assets due to the lower cost of debt (Myers and Majlfu 1984). Overesch and Voeller also indicate that there is positive effect of collateral value of assets on leverage (Overesch and Voeller 2008). In this paper, we use inventory and fixed assets-to-total assets ratio as indicator of collateral value of assets (*CVA*).

67.3.4 Data and Samples

The impact of the corporate income tax reform on capital structure choice is tested using financial data from a sample of listed firms which are selected by several criteria: (1) Since accounting earnings are more probably managed in ST firms and *ST firms, ST firms and *ST firms are both excluded; (2) Since the corporate tax rate for high tech firms is 15 % before and after the reform, high tech listed firms are removed; (3) Since there are differences between financial statements and capital structure of financial firms and non-financial firms, financial firms are excluded, also excluded are public services firms and firms with missing financial data or abnormal financial data. The requirement yields a sample of firms consisting of 4,546 firm-years, there are 993 observations in each year from 2006 to 2009 and there are 574 observations in the year of 2010, respectively. All the financial data are obtained from the CSMAR database.

67.3.5 Investigation Model

In order to investigate the impact of corporate tax reform on capital structure choice of firms in China, the following empirical estimations can be set up as (67.1) and (67.2) show. Since the corporate tax sensitivity may decrease with increasing non-debt tax shields (Overesch and Voeller 2008), in order to identify the effects of non-debt tax shields on leverage, we use an interaction term between the effective corporate tax rate and the non-debt tax shields. In order to investigate the change of capital structure caused by the corporate income tax reform, we investigate the impact of the change of effective tax rate and the impact of the change of non-debt tax shields on the change of leverage.

$$LEV = \beta_0 + \beta_1 ETR + \beta_2 ROE + \beta_3 GROW + \beta_4 \ln A + \beta_5 NDTS + \beta_6 CVA + \beta_7 ETR \times NDTS + \varepsilon \quad (67.1)$$

$$\Delta LEV = \beta_0 + \beta_1 \Delta ETR + \beta_2 ROE + \beta_3 GROW + \beta_4 \ln A + \beta_5 NDTS + \beta_6 CVA + \beta_7 \Delta NDTS + \varepsilon \quad (67.2)$$

Where *LEV* is financial leverage which is defined as total debt-to-total assets ratio, *ΔLEV* is the change of leverage, *ETR* is effective tax rate, *ΔETR* is the change of effective tax rate, *ROE* is return on equity to measure profitability, *GROW* is sales growth rate to measure firm growth, ln*A* is the natural logarithm of total assets to measure firm size, *NDTS* is non-debt tax shields which is defined as depreciation-to-total assets ratio, *ΔNDTS* is the change of non-debt tax shields, *CVA* is the collateral value of assets which is defined as inventory and fixed assets-to-assets ratio, β is the regression coefficient and ε is the error term.

67.4 Empirical Result

67.4.1 Descriptive Statistics

Table 67.1 presents descriptive statistics of all variables. The average change of leverage is positive with a mean of 0.33 %, which indicates that the leverage increased slightly for all firms of our sample during the period of 2006–2010. The average leverage decreases during the year of 2006–2007, then increases from 2008 and increases to 52.33 % in 2010. The average effective tax rate decreases from 21.38 to 17.77 % during the period of 2006–2008, and then increases to 18.84 % in 2010. Since the statutory tax rates for most resident taxpayers decreases from 33 to 25 % after the reform, the effective tax rate is lower than the statutory tax rate. The mean of *ΔNDTS* is 0.0143 which means the non-debt tax shields increase during the period of 2006–2010.

67.4.2 Correlation Analysis

Table 67.2 reports the Pearson correlation coefficients among variables. There is positive relationship between leverage and effective tax rate at 1 % significant level, while there is positive relationship between the change of leverage (*ΔLEV*) and the change of effective tax rate (*ΔETR*) only at 10 % significant level, which provides initial evidence that there exists relationships between firm's leverage and the effective tax rate in our sample. What's more, there are significant relationships between most of control variables and leverage, while only firm size and return on equity are related with the change of leverage (*ΔLEV*) in our sample.

Table 67.1 Descriptive statistics

	Year	N	*LEV*	*ΔLEV*	*ETR*	*ΔETR*	LnA	*CVA*	*ROE*	*GROW*	*NDTS*	*ΔNDTS*
Mean	2006	993	0.5087	0.0116	0.2138	−0.0013	21.5289	0.3303	0.0643	0.2808	0.0560	0.0115
	2007	993	0.5029	−0.0058	0.2112	−0.0026	21.7214	0.2921	0.0892	0.4110	0.0985	0.0224
	2008	993	0.5054	0.0025	0.1777	−0.0335	21.8059	0.2864	0.0512	0.1966	0.1109	0.0124
	2009	993	0.5101	0.0047	0.1876	0.0099	21.9506	0.2766	0.0705	0.4259	0.1191	0.0081
	2010	574	0.5233	0.0077	0.1884	0.0008	22.1724	0.2851	0.0645	0.3041	0.1035	−0.0103
	total	4546	0.5068	0.0033	0.1976	−0.0033	21.7517	0.2963	0.0688	0.3286	0.0961	0.0143
S.D.	2006	993	0.1730	0.1070	0.1663	0.1651	1.0133	0.1884	0.1777	0.8664	0.2271	0.4325
	2007	993	0.1709	0.0922	0.1544	0.1563	1.0748	0.1908	0.1649	1.3015	0.4788	0.3947
	2008	993	0.1789	0.0895	0.1558	0.1816	1.1210	0.1859	0.1689	1.1127	0.4839	0.6267
	2009	993	0.1820	0.0795	0.1529	0.1950	1.1780	0.1871	0.1345	1.3540	0.8244	0.8502
	2010	574	0.1846	0.0768	0.1301	0.1876	1.2415	0.1892	0.1652	1.2301	0.7056	0.6714
	total	4546	0.1762	0.0928	0.1581	0.19892	1.1086	0.1891	0.1629	1.1778	0.5468	0.6510

Coefficients marked with *, **, *** are significant at 10 %, 5 % and 1 %, respectively

Table 67.2 Pearson correlation coefficients

	LEV	*ΔLEV*	*ETR*	*ΔETR*	Ln*A*	*CVA*	*ROE*	*GROW*	*NDTS*	*ΔNDTS*
LEV	1	0.274**	0.054***	0.016	0.303***	0.009	−0.087***	0.078***	−0.081***	−0.018
ΔLEV		1	−0.030*	0.024*	0.098***	−0.024	−0.176***	0.015	−0.011	0.006
ETR			1	0.541***	−0.005	−0.050***	0.120***	0.072***	−0.027*	−0.035*
ΔETR				1	0.002	−0.022	0.094***	0.041***	−0.014	−0.016
Ln*A*					1	0.110***	0.150***	0.006	−0.152***	−0.039*
CVA						1	−0.037**	−0.178***	0.018	0.027
ROE							1	0.037**	−0.024	−0.017
GROW								1	0.003	−0.013
NDTS									1	0.747***
ΔNDTS										1

Coefficients marked with *, **, *** are significant at 10 %, 5 % and 1 %, respectively

Table 67.3 Regression results (*LEV*)

	Full sample	2006	2007	2008	2009	2010
(Constant)	−0.611	−0.504	−0.491	−0.687	−0.804	−0.614
	(0.000)***	(0.000)***	(0.000)***	(0.000)***	(0.000)***	(0.000)***
ETR	0.074	0.036	0.098	0.078	0.078	0.030
	(0.000)***	(0.071)*	(0.005)***	(0.019)**	(0.027)**	(0.012)**
Ln*A*	0.052	0.048	0.046	0.055	0.060	0.052
	(0.000)***	(0.000)***	(0.000)***	(0.000)***	(0.000)***	(0.000)***
CVA	−0.046	−0.087	−0.033	−0.049	−0.043	−0.036
	(0.001)***	(0.003)***	(0.228)	(0.019)**	(0.040)**	(0.014)**
ROE	−0.161	−0.111	−0.091	−0.244	−0.223	−.0025
	(0.000)***	(0.000)***	(0.005)***	(0.000)***	(0.000)***	(0.002)***
GROW	0.010	0.016	0.008	0.010	0.009	0.001
	(0.000)***	(0.011)**	(0.040)**	(0.048)***	(0.018)**	(0.034)**
NDTS	−0.009	−0.006	−0.023	−0.021	−0.004	−0.056
	(0.135)	(0.871)	(0.279)	(0.204)	(0.623)	(0.319)
ETF×*NDTS*	−0.013	0.023	0.043	−0.005	−0.053	−0.018
	(0.705)	(0.861)	(0.558)	(0.953)	(0.440)	(0.731)
F-statistics	80.409	14.461	15.2	28.55	27.943	18.351
Adjusted R^2	0.124	0.087	0.091	0.163	0.160	0.137
Mean Square	0.027	0.027	0.027	0.027	0.028	0.029

Coefficients marked with *, **, *** are significant at 10 %, 5 % and 1 %, respectively

67.4.3 Regression Results

The regression results are represented in Tables 67.3 and 67.4, respectively. There is significantly positive relationship between leverage and effective tax rate which suggests that the corporate income tax has positive effect on debt levels. The coefficients of both the non-debt tax shields and the interaction of non-debt tax shields and effective tax rate are not statistically significant. It suggests that the non-debt tax shields do not act as determinants of capital structure and there is not substitution effect between debt and non-debt tax shields in our sample. The coefficients of the natural logarithm of total assets, which are all positive and significant, mean that firm size has positive impact on debt level. Besides the sample of the year of 2007, the coefficients of collateral value of assets are all significant and strong which indicates the collateral value of assets has negative influence on the use of debt capital. The significant and negative coefficients of return on equity suggest highly profitable firms may use less liability, which is consistent with the pecking order theory. The coefficients between firm growth and leverage are all positive and significant which mean firms with high growth opportunities may reduce their debt level, which is consistent with the tradeoff theory.

There are significantly positive relationships between the change of leverage and the change of effective tax rate, which indicates that the corporate income tax rate is a determinant of capital structure of Chinese firms, and the change of effective tax rate has positive impact on the change of leverage. The coefficients of the change

Table 67.4 Regression results (ΔLEV)

	Full sample	2006	2007	2008	2009	2010
(Constant)	−0.239	−0.220	−0.146	−0.292	−0.246	−1.146
	(0.000)***	(0.003)***	(0.015)**	(0.000)***	(0.000)***	(0.000)***
ΔETR	0.005	0.016	0.016	0.016	0.010	0.026
	(0.035)**	(0.038)**	(0.034)**	(0.078)*	(0.022)**	(0.052)*
Ln*A*	0.011	0.011	0.007	0.014	0.012	0.053
	(0.000)***	(0.001)***	(0.016)**	(0.000)***	(0.000)***	(0.000)***
CVA	−0.005	−0.016	0.009	0.008	−0.033	−0.005
	(0.027)**	(0.077)*	(0.055)*	(0.074)*	(0.013)**	(0.000)***
ROE	−0.131	−0.069	−0.105	−0.151	−0.132	−0.023
	(0.000)***	(0.000)***	(0.000)***	(0.000)***	(0.000)***	(0.043)**
GROW	0.001	0.005	0.003	0.002	0.001	0.003
	(0.267)	(0.185)	(0.243)	(0.366)	(0.837)	(0.183)
$\Delta NDTS$	0.001	0.002	0.001	0.002	0.001	0.034
	(0.648)	(0.632)	(0.876)	(0.664)	(0.749)	(0.553)
F-statistics	35.354	4.362	6.483	19.306	11.837	8.852
Adjusted R^2	0.067	0.023	0.032	0.101	0.067	0.067
Mean Square	0.007	0.011	0.008	0.007	0.006	0.061

Coefficients marked with *, **, *** are significant at 10 %, 5 % and 1 %, respectively

of non-debt tax shields are not statistically significant, which means that non-debt tax shields are not substitutes of debt in our sample. There are significantly positive correlations between the change of leverage and the natural logarithm of total assets, which means that larger firms can change their leverage more readily than small firms. There are significantly negative relationships between the collateral value of assets and leverage which means the collateral value of assets has negative influence on the change of leverage. The coefficients of return on equity (*ROE*) are significant and negative, which indicates that profitability has negative influence on debt level in our sample. However, the coefficients of firm growth are positive but not significant.

67.5 Conclusion

We empirically investigate the effect of corporate tax reform on capital structure decisions using the panel data of Chinese listed firms during the period of 2006–2010. The empirical results suggest that the corporate income tax reform decreases the effective tax rate totally, the effective tax rate decreases from the year of the corporate income tax reform and then increase to 18.84 % in 2010 with the increase of the average leverage during the year of 2009–2010. The results also show that the corporate income tax rate is a determinant of capital structure choice and the change of effective tax rate has significantly positive impact on the change of leverage. The non-debt tax shields are not determinants of capital structure and there is not a substitution effect between debt and non-debt tax shields in our sample.

Acknowledgment We acknowledge financial support from National Natural Science Foundation of China (No. 71002104) and Tianjin University- Hainan University Innovation Funds.

References

Brigham EF, Gapenski LC (1996) Intermediate financial management. Dryden Press, Harcourt Brace College, Orlando

Charalambakis EC, Espenlaub SK, Garrett I (2008) Leverage dynamics, the endogeneity of corporate tax status and financial distress costs, and capital structure. http://ssrn.com/abstract=1102168

DeAngelo H, Masulis RW (1980) Optimal capital structure under corporate and personal taxation. J Financ Econ 8:3–29

Givoly D, Hahn C, Ofer A, Sarig OH (1992) Taxes and capital structure: evidence from firms' response to the tax reform act of 1986. Rev Financ Stud 5:331–355

Gordon RH, Lee Y (2001) Do taxes affect corporate debt policy? Evidence from U.S. corporate tax return data. J Public Econ 82:195–225

Graham JR (1996) Debt and the marginal tax rate. J Financ Econ 41:41–73

Jensen MC, Meckling WH (1976) Theory of the firm: managerial behavior, agency costs and ownership structure. J Financ Econ 3:305–360

Myers SC (1977) Determinants of corporate borrowing. J Financ Econ 5:147–175

Myers SC, Majlfu NS (1984) Corporate financing and investment decisions when firms have information that investors do not have. J Financ Econ 13:187–221

Overesch M, Voeller D (2008) The impact of personal and corporate taxation on capital structure choices. ftp://ftp.zew.de/pub/zew-docs/dp/dp08020.pdf

Rajan RG, Zingales L (1995) What do we know about capital structure? Some evidence from international data. J Finance 50:1421–1460

Titman S, Wessels R (1988) The determinants of capital structure choice. J Finance 43:1–19

Verschueren I (2001) Capital structure or debt tax shields ratio? An empirical investigation for Belgian firms. VUB Money & Finance working paper no. 01-03

Wu LS, Yue H (2009) Corporate tax, capital structure, and the accessibility of bank loans: evidence from China. J Bank Finance 33:30–38

Chapter 68
Optimization Model of Loan's Portfolio Based on Geometric Spectral Measure

Chong Duan and Xiu-min Jia

Abstract The loan's portfolio is a hot issue in bank's risk management. This paper puts forward an optimization model of loan's portfolio by using geometric spectral measure of risk to control extreme losses of portfolio. These innovations are as follows: firstly, the greater weight is distributed to greater extreme losses by the risk aversion function, which controls the risk of extreme losses. The risk aversion function fits investors' risk aversion characters. Secondly, an objective weight is given to extreme losses which avoids personal choices. Thirdly, the probability of disaster's risk occurrence is reduced while taking the geometric spectral measure minimum as an object function.

Keywords Extreme losses • Geometric spectral measure • Loan's portfolio • Risk aversion

68.1 Introduction

One important decision of bank is to optimize the loan's portfolio structure so as to hold a loan's portfolio with the highest yield rate and the lowest risk. One loan optimization does not mean loan portfolio optimization. Therefore, the research of optimization model of loan's portfolio in China can provide the decision support of risk control and optimization allocation of commercial banks, which has an important practical significance (Peter 2002).

The research are supported by the Inner Mongolia University of science and technology innovation fund project (the grantnumber is 2011NCL059).

C. Duan • X. Jia (✉)
Department of Mathematics Physics and Biology Engineering, Inner Mongolia University of Science and Technology, Baotou, China
e-mail: duanchong100@sina.com; jxm620168@126.com

E. Qi et al. (eds.), *The 19th International Conference on Industrial Engineering and Engineering Management*, DOI 10.1007/978-3-642-37270-4_68,

There are some researches of loan's portfolio model both at home and abroad. The American economist Markowitz proposed the mean-variance portfolios model (Markowitz 1952). Gollinger T.L., and Morgan J.B. proposed the efficient frontier model of loan portfolio, based on the hypothesis that portfolio yields are not less than objective yields (Gollinger and Morgan 1993). Altman E.I. proposed the loan portfolio analysis model, based on the same hypothesis (Altman 2000). Chi Guotai used VaRas the risk measure of loan portfolio taking the place of variance, proposed the mean-VaR portfolios model (Chi Guotai et al. 2000). According to the loan portfolio return and risk, Zhou Zongfang put forward a Multi-objective optimization model (Zhou Zongfang et al. 2002). Li Zhongfei found that calculation results existed large errors when taking VaR as the risk measure of loan portfolio, especially loan return distribution is fat tail (Li Zhong-fei and Yuan Zi-jia 2010).

To sum up, domestic and foreign researches have achieved considerable progresses, but problems have three points. Firstly, VaR is not a coherent measures of portfolio risk, which is contradiction with the fact that the diversified investment would reduce portfolio risk (Elliott et al. 2010; Andrew and Gah 2011), so VaR is not a good portfolio risk measure and can not measure portfolio risk accurately. Secondly, the probability of portfolio losses beyond value at risk is neglected, so extreme losses of portfolio is not controlled, which often results in huge losses for commercial banks. Thirdly, the risk preference is obtained subjectively, which lacks an objective basis (Yao Hai-xing and Li Zhong-fei 2009).

In view of above questions, this paper chose geometric spectral risk measure that is a coherent measures of portfolio risk to control the loan portfolio risk. On the premise of the target return rate, this paper considers the geometric spectral measure of loan portfolio risk as an objective function, takes the value at risk of loan portfolio as a constraint condition and sets up a new loan portfolio optimization model. The model is used to solve this problem that the value at risk can not measure portfolio risk accurately. This model makes up for the drawback of failing to control the extreme losses of portfolio.

68.2 Optimization Principle for Loan Portfolio Based on Geometric Spectral Measure

68.2.1 VaR Risk Controlled Principle

VaR (value at risk) is defined as the expected maximum losses of asset portfolio or one asset during holding periods in a certain confidence level (Altman 2000). It is shown as following (Yoshida 2009).

$$P(L < -\mathrm{VaR}) = \alpha \tag{68.1}$$

Where, $P(\cdot)$ is a kind of probability function; L- the loss of asset portfolio or one asset during holding periods; VaR-value at risk of a yield rate for asset portfolio or one asset, VaR > 0; $(1-\alpha)$-a certain confidence level.

VaR risk controlled principle is defined as controlling the expected maximum losses VaR within the scope of bank's capacity. It guarantees that the yield rate risk of loan portfolio can be limited within the scope of bank's capacity.

68.2.2 Geometric Spectral Risk Measure Principle

American mathematician Acerbi (2002) proposed Spectral Measure of Risk. Spectral Measure of Risk is a coherent measures of portfolio risk and has four good properties, such as Positively homogeneous and Subadditivity and Monotonicity and translational invariance (Acerbi 2002).

Suppose $f(p)$ is a real value function and integrable function, $\boldsymbol{p} \in (0, 1]$, and $f(p) \geq 0$.

If

$$\boldsymbol{p}_1 < \boldsymbol{p}_2 \Rightarrow f(\boldsymbol{p}_1) \geq \boldsymbol{f}(\boldsymbol{p}_2) \tag{68.2}$$

$$\int_0^1 \boldsymbol{f}(\boldsymbol{p})\, \boldsymbol{dp} = 1 \tag{68.3}$$

Thus (Acerbi 2002),

$$\boldsymbol{M} = \int_0^1 \boldsymbol{f}(\boldsymbol{p})\, \boldsymbol{VaR}_{\boldsymbol{p}} \boldsymbol{dp} \tag{68.4}$$

is a spectral measure of risk. M-spectral risk measure; $f(p)$-spectral density function, which reflects risk aversion degree of investors. VaRp-value at risk of loan portfolio in certain confidence level $1-p$; p-the probability when loan portfolio loss value equals to VaR$_P$.

This paper choices geometric spectral density function, and proposes geometric spectral risk measure.

The geometric spectral density function $f_G(p)$ is shown as following.

$$\boldsymbol{f}_{\boldsymbol{G}}(\boldsymbol{p}) = \frac{1}{\ln(1+\alpha)} \frac{1}{1+\boldsymbol{p}} \boldsymbol{I}_{\{0<\boldsymbol{p}\leq\alpha\}} \tag{68.5}$$

Where, $f_G(p)$-geometric spectral density function; $(1-\alpha)$-certain confidence level; α-the probability when loan portfolio loss has occurred; p-the loss probability, $p \in [0, 1]$. $I_{\{0<\mathrm{P}<\alpha\}}$-indicator function whose value is 0 or 1.

$$\boldsymbol{I}_{\{0<P\leq\alpha\}} = \begin{cases} 1 & 0 < \boldsymbol{P} \leq \alpha \\ 0 & \alpha < \boldsymbol{P} \leq 1 \end{cases} \tag{68.6}$$

The geometric spectral risk measure is shown as following.

$$\boldsymbol{GM} = \int_0^1 \boldsymbol{f}_{\boldsymbol{G}}(\boldsymbol{p})\boldsymbol{VaR}_{\boldsymbol{p}}\boldsymbol{dp} = \int_0^1 \frac{1}{\ln(1+\alpha)}\frac{1}{1+\boldsymbol{p}}\boldsymbol{I}_{\{0\leq P\leq\alpha\}}\boldsymbol{VaR}_{\boldsymbol{p}}\boldsymbol{dp} \quad (68.7)$$

geometric spectral risk measure GM reflects the weighted excess losses of loan's portfolio in certain confidence level.

There are two categories according different relations between the possible losses value (VaR_p) and the possible losses controlling value (VaR_α), which are as follows:

1. When the excess losses do not occur, in other words when the possible loss value (VaR_p) is less than the possible loss controlling value (VaR_α), considering $VaR_\alpha > 0$, so $-VaR_p > -VaR_\alpha$. Considering the definition of VaR, so the loss probability p is more than loss controlling probability α, thus the loss probability $p \in [\alpha, 1]$. Considering the definition of indicator function, thus $I_{\{0<p\leq\alpha\}} = 0$, and then geometric spectral density function $f_G(p) = 0$. In other words, the risk weight of VaR_p equals to 0 in the formula (68.7) that is the expression of geometric spectral risk measure.
2. When the excess losses occur, in other words when the possible losses value (VaR_p) is more than the possible losses controlling value (VaR_α), considering $VaR_\alpha > 0$, so $-VaR_p < -VaR_\alpha$. Considering the definition of VaR, so the losses probability p is less than losses controlling probability α, thus the losses probability $p \in [0,\alpha]$. Considering the definition of indicator function, thus $I_{\{0<p\leq\alpha\}} = 1$, and then geometric spectral density function $f_G(p) = \{1/\ln(1+\alpha)\}\{1/(1+p)\}$. In other words, the risk weight of VaR_p is $\{1/\ln(1+\alpha)\}\{1/(1+p)\}$ in the formula (68.7).

When the excess losses occur, and $p \in (0,\alpha]$, the geometric spectral density $f_G(p) = \{1/\ln(1+\alpha)\}\{1/(1+p)\}$ is a monotonically decreasing function. Therefore, with the losses probability p closer to 0, the possibility of excess losses occurrence is greater, and than the greater risk weight of VaR_p has been given by geometric spectral density function $f_G(p)$. With the losses probability p closer to α, the possibility of excess losses occurrence is more smaller, and than the smaller risk weight of VaR_p has been given by geometric spectral density function $f_G(p)$.

The greater excess losses are been given the greater risk weight with geometric spectral density function $f_G(p)$. In this way, the excess losses risk can be controlled, and the risk preference is obtained objectively.

68.2.3 Optimization Principle for Loan Portfolio Based on Geometric Spectral Risk Measure

Taking geometric spectral measure GM of loan portfolio risk as an objective function, taking the VaR of loan portfolio as a constraint condition, this paper sets up a new optimization model for loan portfolio, which solves the problem of risk controlling in some extreme cases. The excess loss of asset portfolio which comes from abnormal price changes can be avoided.

68.3 Optimization Model for Loan Portfolio Based on Geometric Spectral Measure

68.3.1 The Calculation of Yield Rate Expectation and Yield Rate Variance

1. the calculation of yield rate expectation for loan portfolio

Suppose: r_{ti}-yield rate of the t-th year for the ith loan, ($t = 1, 2, \ldots, m$; $i = 1, 2, \ldots, n$). The yield rate expectation of m years for the ith loan is as following (Shang Zhou et al. 2010).

$$\mu_i = \frac{1}{\boldsymbol{m}} \sum_{\boldsymbol{t}=1}^{\boldsymbol{m}} \boldsymbol{r}_{\boldsymbol{ti}} \tag{68.8}$$

Suppose: $\mu(X)$-the yield rate expectation of m years for loan portfolio; x_i-the proportion of the ith loan in the whole loan portfolio, ($i = 1, 2, \ldots, n$).

Thus (Shang Zhou et al. 2010),

$$\mu(\boldsymbol{X}) = \sum_{\boldsymbol{i}=1}^{\boldsymbol{n}} \mu_{\boldsymbol{i}} \boldsymbol{x}_{\boldsymbol{i}} \tag{68.9}$$

2. the calculation of yield rate variance for loan portfolio

Suppose: $\boldsymbol{\sigma}_{\boldsymbol{i}}^2$ – yield rate variance for the ith loan, ($i = 1, 2, \ldots, n$); σ_{ij}-yield rate covariance for loan portfolio, ($j = 1, 2, \ldots, n$).

Thus (Shang Zhou et al. 2010),

$$\sigma_i^2 = \frac{1}{\boldsymbol{m} - 1} \sum_{\boldsymbol{t}=1}^{\boldsymbol{m}} (\boldsymbol{r}_{\boldsymbol{t}i} - \mu_i)^2 \tag{68.10}$$

$$\sigma_{ij} = \frac{1}{\boldsymbol{m} - 1} \sum_{t=1}^{m} (\boldsymbol{r}_{ti} - \mu_i)\left(\boldsymbol{r}_{tj} - \mu_j\right) \tag{68.11}$$

The yield rate variance for loan portfolio is as following (Shang Zhou et al. 2010).

$$\sigma(\boldsymbol{X})^2 = \sum_{\boldsymbol{i}=1}^{\boldsymbol{n}} \sum_{\boldsymbol{j}=1}^{\boldsymbol{n}} \boldsymbol{x}_i \boldsymbol{x}_j \sigma_{ij} = \boldsymbol{X}^T \boldsymbol{U} \boldsymbol{X} \tag{68.12}$$

As shown in formula (68.12), x_i-the proportion of the ith loan in the whole loan portfolio, ($i = 1, 2, \ldots, n$). $X = (x_1, x_2, \ldots, x_n)^{\mathrm{T}}$, U-covariance matrix for loan portfolio, which is as following.

$$\boldsymbol{U} = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1\boldsymbol{n}} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2\boldsymbol{n}} \\ \vdots & \vdots & \vdots & \vdots \\ \sigma_{\boldsymbol{n}1} & \sigma_{\boldsymbol{n}2} & \cdots & \sigma_{\boldsymbol{nn}} \end{bmatrix} \tag{68.13}$$

68.3.2 The Establishment of Constraint Condition VaR

VaR risk controlled principle is defined as controlling the expected maximum loss VaR within the scope of bank's capacity. It guarantees that the yield rate risk of loan portfolio can be limited within the scope of bank's capacity. Therefore, this paper takes VaR of loan portfolio yield rate as a constraint condition.

Suppose: μ_0-the yield rate of loan portfolio; L-the loss of loan portfolio. When the loss of loan portfolio L is negative, that is to say loan portfolio have the positive yield rate, and than $\mu_0 = -L$. So formula (68.1) can be written as following.

$$P(\mu_0 < -\text{VaR}) = \alpha \tag{68.14}$$

Suppose the yield rate of loan portfolio μ_0 follows the normal distribution, so $[\mu_0 - \mu(X)]/\sigma(X)$ follows the standardized normal distribution according to the central limit theorem (Shang Zhou et al. 2010).

Thus (Shang Zhou et al. 2010),

$$\boldsymbol{P}\left(\frac{\mu_0 - \mu(\boldsymbol{X})}{\sigma(\boldsymbol{X})} < \frac{-\boldsymbol{VaR} - \mu(\boldsymbol{X})}{\sigma(\boldsymbol{X})}\right) = \boldsymbol{P}(\mu_0 < -\boldsymbol{VaR}) \tag{68.15}$$

Suppose: $\Phi(x)$-probability distribution function of the standardized normal distribution, then

$$\Phi\left(\frac{-\boldsymbol{VaR} - \mu(\boldsymbol{X})}{\sigma(\boldsymbol{X})}\right) = \boldsymbol{P}\left(\frac{\mu_0 - \mu(\boldsymbol{X})}{\sigma(\boldsymbol{X})} < \frac{-\boldsymbol{VaR} - \mu(\boldsymbol{X})}{\sigma(\boldsymbol{X})}\right) \tag{68.16}$$

From above formulas (68.14), (68.15), and (68.16), then

$$\Phi\left(\frac{-\boldsymbol{VaR} - \mu(\boldsymbol{X})}{\sigma(\boldsymbol{X})}\right) = \alpha \tag{68.17}$$

That is to say

$$\text{VaR} = \Phi^{-1}(1-\alpha)\sigma(X) - \mu(X) \tag{68.18}$$

Formula (68.18), $\Phi^{-1}(x)$-the inverse function of standardized normal distribution function $\Phi(x)$.

68.3.3 The Establishment of Objective Function

1. the calculation of parameter k_α

Above formula (68.7), the right integration expression of geometric spectral risk measure GM can be divided into two integration expressions with indicator function parameter α, which is as following.

$$\begin{aligned} \boldsymbol{GM} &= \int_0^\alpha \frac{1}{\ln(1+\alpha)} \frac{1}{1+\boldsymbol{p}} I_{\{0 \le P \le \alpha\}} \boldsymbol{VaR_p dp} \\ &+ \int_\alpha^1 \frac{1}{\ln(1+\alpha)} \frac{1}{1+\boldsymbol{p}} I_{\{0 \le P \le \alpha\}} \boldsymbol{VaR_p dp} \end{aligned} \tag{68.19}$$

According to the definition of indicator function $I_{\{0<\mathrm{P}<\alpha\}}$, the second integration equal to 0 above formula (68.20), so formula (68.19) can be written as following.

$$\boldsymbol{GM} = \int_0^\alpha \frac{1}{\ln(1+\alpha)} \frac{1}{1+\boldsymbol{p}} \boldsymbol{VaR_p dp} \tag{68.20}$$

Considering the expression of VaR above formula (68.18), so formula (68.20) can be written as following.

$$\boldsymbol{GM} = \int_0^\alpha \frac{1}{\ln(1+\alpha)} \frac{1}{1+\boldsymbol{p}} \left(-\Phi^{-1}(\alpha)\,\sigma(\boldsymbol{X}) - \mu(\boldsymbol{X})\right) \boldsymbol{dp} \tag{68.21}$$

Suppose:

$$\boldsymbol{k}_\alpha = \int_0^\alpha \frac{1}{\ln(1+\alpha)} \frac{1}{1+\boldsymbol{p}} \Phi^{-1}(\boldsymbol{p})\, \boldsymbol{dp} \tag{68.22}$$

Suppose: $p = \varPhi(x)$, so $x = \Phi^{-1}(p)$, $dp = \varphi(x)\, dx$, $\varphi(x)$ is the probability density function of standardized normal distribution. Formula (68.23) can be obtained by doing integral transformation for formula (68.22).

$$\boldsymbol{k}_\alpha = \frac{1}{\ln(1+\alpha)} \int_{\Phi^{-1}(0)}^{\Phi^{-1}(\alpha)} \frac{\boldsymbol{x}}{1+\Phi(\boldsymbol{x})} \phi(\boldsymbol{x})\, \boldsymbol{dx} \tag{68.23}$$

Where, $\varPhi^{-1}(\alpha)$-the α quantiles of standardized normal distribution; $\Phi(x)$-probability distribution function of the standardized normal distribution.

From above formula (68.23), k_α reflects the weighted average of yield rate in the confidence level $1-\alpha$ when the yield rate for loan portfolio is less than the α quantiles of standardized normal distribution $\varPhi^{-1}(\alpha)$.

And than, the Simplification of formula (68.23) is formula (68.24).

$$\boldsymbol{k}_{\alpha}=\frac{1}{\ln(1+\alpha)}\int_{\Phi^{-1}(0)}^{\Phi^{-1}(\alpha)}\boldsymbol{x}\boldsymbol{d}\left(\ln\left(1+\boldsymbol{\Phi}\left(\boldsymbol{x}\right)\right)\right) \tag{68.24}$$

With integration by part for formula (68.24), formula (68.25) can be obtained as following.

$$\begin{aligned}\boldsymbol{k}_{\alpha}&=\frac{1}{\ln(1+\alpha)}\left[x\ln\left(1+\boldsymbol{\Phi}\left(\boldsymbol{x}\right)\Big|_{\Phi^{-1}(0)}^{\Phi^{-1}(\alpha)}\right)-\int_{\Phi^{-1}(0)}^{\Phi^{-1}(\alpha)}\ln\left(1+\Phi\left(\boldsymbol{x}\right)\right)\boldsymbol{x}'\boldsymbol{dx}\right]\\&=\Phi^{-1}\left(\alpha\right)-\frac{1}{\ln(1+\alpha)}\int_{\Phi^{-1}(0)}^{\Phi^{-1}(\alpha)}\ln\left(1+\Phi\left(\boldsymbol{x}\right)\right)\boldsymbol{dx}\end{aligned} \tag{68.25}$$

2. the calculation of geometric spectral risk measure

According the k_{α} expression of formula (68.25), than

$$\boldsymbol{GM}=-\boldsymbol{k}_{\alpha}\sigma\left(\boldsymbol{X}\right)-\mu\left(\boldsymbol{X}\right)\int_{0}^{\alpha}\frac{1}{\ln\left(1+\alpha\right)}\frac{1}{1+\boldsymbol{p}}\boldsymbol{dp} \tag{68.26}$$

Because above formula (68.26) the integration equal to 1, so geometric spectral risk measure GM can be simplified as following.

$$GM=-k_{\alpha}\sigma(X)-\mu(X) \tag{68.27}$$

3. the calculation of objective function

This paper takes geometric spectral measure GM of loan portfolio risk as the objective function of optimization model to reduce the probability of catastrophic risk occurrence for commercial bank.

Thus

$$\text{obj}:\min\ GM=-k_{\alpha}\sigma(X)-\mu(X) \tag{68.28}$$

68.3.4 The Establishment of Optimization Model

Taking formula (68.28) as objective function, Taking formula (68.9) and formula (68.18) as constraint conditions, this paper sets up a new optimization model for loan portfolio as following.

$$\text{obj}:\min\ GM=-k_{\alpha}\sigma(X)-\mu(X) \tag{68.29}$$

$$\text{s.t.}\quad \mu\left(\boldsymbol{X}\right)=\sum_{i=1}^{n}\mu_i\boldsymbol{x}_i \tag{68.30}$$

$$\Phi^{-1}\,(1-\alpha)\,\sigma(X) - \mu(X) \le \mathrm{VaR} \tag{68.31}$$

$$\sum_{i=1}^{n} \boldsymbol{x}_i = 1 \tag{68.32}$$

68.3.5 The Establishment of the Range of Yield Rate for Loan Portfolio

We calculate a reasonable range of yield rate for loan portfolio, using the Lagrange multiplier method.

$\mu(r_m)$ is the yield rate for loan portfolio on the efficient frontier corresponds to the minimum of GM point.

Thus (Li Xiang et al. 2010; Rupak et al. 2009)

$$\mu\,(r_m) = \frac{\boldsymbol{A}}{\boldsymbol{C}} + \sqrt{\left(\frac{\boldsymbol{H}^2}{\boldsymbol{C}\boldsymbol{H}^2 - \boldsymbol{D}} - \frac{1}{\boldsymbol{C}}\right)\frac{\boldsymbol{D}}{\boldsymbol{C}}} \tag{68.33}$$

Where

$$A = I^{\mathrm{T}} U^{-1} R \tag{68.34}$$

$$B = R^{\mathrm{T}} U^{-1} R \tag{68.35}$$

$$C = I^{\mathrm{T}} U^{-1} I \tag{68.36}$$

$$D = BC - A^2 \tag{68.37}$$

$$I = (1, 1, \ldots, 1)^{\mathrm{T}} \tag{68.38}$$

$$H = -k_{\alpha} \tag{68.39}$$

$$R = (\mu_1, \mu_2, \ldots, \mu_n)^{\mathrm{T}} \tag{68.40}$$

Because the yield rate expectation for loan portfolio $\mu(X)$ is not more than the maximum rate of yield for a single loan, so

$$\mu\,(r_m) \le \mu(X) \le max\,\{\mu_i\} \tag{68.41}$$

Where, μ_i is the yield rate expectation for the ith loan.

This paper calculates an objective yield rate for loan portfolio, and makes up for the drawback of choosing an objective yield rate by practice experiences.

68.4 Practical Example Analysis

68.4.1 Basic Data

Suppose there are 15 years yield rate datas for these nine loan clients in the Bank Branch.

68.4.2 The Calculation of Yield Rate Expectation and Yield Rate Variance

1. the calculation of yield rate expectation for loan portfolio

Yield rate expectation $\mu_1$of the 1th loan can be got with the formula (68.8) and the first column of Table 68.1.

$$\mu_1 = (6.0 + 5.7 + \cdots + 5.4)/15 = 0.0532.$$

Similarly, yield rate expectation μ_i of the ith loan can be got with the formula (68.8) and the other columns of Table 68.1. The yield rate expectation vector for loan portfolio is as following.

$$\begin{aligned} R &= (\mu_1, \mu_2, \ldots, \mu_9)^{\mathrm{T}} \\ &= (0.0532, 0.0499, 0.0482, 0.0429, 0.0416, 0.0194, \\ &\quad 0.0005, 0.0497, 0.0545)^{\mathrm{T}} \end{aligned} \tag{68.42}$$

Table 68.1 The loan year yield (%)

	Yield rate datas for nine loan clients (%)								
Year	1	2	3	4	5	6	7	8	9
1	6.0	5.3	5.2	5.0	4.4	2.4	1.8	5.4	5.8
2	5.7	5.1	5.2	5.0	4.6	3.1	2.4	5.3	6.0
3	5.8	4.2	4.6	4.3	4.2	4.0	−3.4	5.2	5.9
4	5.4	4.6	4.4	4.4	3.7	3.0	−3.9	5.3	6.0
5	5.6	5.3	4.8	3.7	4.0	3.2	1.1	4.9	5.7
6	5.9	5.5	4.9	4.0	4.5	2.5	0.7	4.7	5.8
7	5.3	5.1	5.1	4.1	3.9	1.8	3.0	5.0	5.6
8	5.1	5.0	5.2	4.6	4.4	−2.6	1.5	5.1	5.9
9	4.5	4.9	4.4	3.8	3.8	−3.7	−4.7	5.0	4.9
10	4.8	4.8	4.6	4.2	4.0	1.4	−3.5	4.4	5.1
11	4.4	4.8	4.9	4.5	4.3	2.5	2.4	4.3	4.6
12	5.1	3.2	5.1	3.9	3.8	2.8	−1.4	4.5	5.0
13	5.5	4.7	5.2	4.3	4.1	3.6	1.0	5.7	5.3
14	5.3	5.3	4.0	4.1	4.5	3.1	2.0	5.1	5.2
15	5.4	5.1	4.7	4.4	4.2	2.1	1.7	4.6	5.0

2. the calculation of yield rate variance for loan portfolio

The yield rate covariance matrix U of nine clients for loan portfolio can be got with the formula (68.11) and the datas of Table 68.1 and yield rate expectation vector R, which are as Table 68.2.

The yield rate variance for loan portfolio can be got with the formula (68.12) and constant matrix U, and than standard deviation $\sigma(X)$ can be got with.

$$\sigma(X) = \sqrt{\boldsymbol{X}^T \boldsymbol{U} \boldsymbol{X}}$$

68.4.3 The Establishment of Constraint Condition VaR

Banks set the VaR value is 5 %, and confidence level $1-\alpha = 95\ \%$ according to risk-bearing capacity of banks itself.

The quantile $\Phi^{-1}(0.95)$ is 1.65 on the confidence level $1-\alpha = 95\ \%$ according to standardized normal distribution table. Thus

$$\Phi^{-1}(1-\alpha) = \Phi^{-1}(0.95) = 1.65.$$

The constraint condition VaR can be got with the formula (68.18) and $\Phi^{-1}(0.95) = 1.65$ and $\sigma(X) = \sqrt{\boldsymbol{X}^T \boldsymbol{U} \boldsymbol{X}}$, which is as following.

$$1.65\sqrt{\boldsymbol{X}^T \boldsymbol{U} \boldsymbol{X}} - \mu(X) \le 0.05 \tag{68.43}$$

68.4.4 The Establishment of Objective Function

1. the calculation of parameter k_α

Suppose

$$\Gamma_\alpha = \int_{\Phi^{-1}(0)}^{\Phi^{-1}(\alpha)} \ln(1 + \Phi(\boldsymbol{x}))\, \boldsymbol{dx} \tag{68.44}$$

The quantile $\Phi^{-1}(0.05)$ is -1.65 on the confidence level 95 % and the quantile $\Phi^{-1}(0)$ is $-\infty$ according to standardized normal distribution table. So, the expression of parameter Γ_α is as following on the confidence level 95 %.

$$\boldsymbol{T}(0.05) = \int_{-\infty}^{-1.65} \ln(1 + \Phi(\boldsymbol{x}))\, \boldsymbol{dx} \tag{68.45}$$

We calculate $\ln(1+\Phi(-1{,}000))$ by matlab software, and find $\ln(1+\Phi(-1{,}000)) < 10^{-30}$.

Table 68.2 The correlation matrix of nine clients' u

	u_1	u_2	u_3	u_4	u_5	u_6	u_7	u_8	u_9	Asset	Weight
u_1	0.2274	0.0315	0.0409	0.0483	0.0488	0.5392	0.2983	0.0960	0.1556	x_1	0.000
u_2	0.0315	0.1121	0.0208	−0.0141	0.0304	−0.0371	0.4604	−0.0249	−0.0016	x_2	0.412
u_3	0.0409	0.0208	0.1346	0.0549	0.0188	0.0265	0.4343	0.0187	0.0376	x_3	0.168
u_4	0.0483	−0.0141	0.0549	0.1484	0.0568	0.0925	0.3380	0.0502	0.0574	x_4	0.045
u_5	0.0488	0.0304	0.0188	0.0568	0.0826	0.0883	0.4388	0.0127	0.0295	x_5	0.000
u_6	0.5392	−0.0371	0.0265	0.0925	0.0883	4.7597	1.3079	0.0867	0.1772	x_6	0.000
u_7	0.2983	0.4604	0.4343	0.3380	0.4388	1.3079	7.0798	0.0478	0.0545	x_7	0.000
u_8	0.0960	−0.0249	0.0187	0.0502	0.0127	0.0867	0.0478	0.1595	0.1038	x_8	0.244
u_9	0.1556	−0.0016	0.0376	0.0574	0.0295	0.1772	0.0545	0.1038	0.2127	x_9	0.131

Therefore, above formula (68.45) the integration of the range of $(-\infty, -1.65]$ approximately equals to integration of the range of $(-1{,}000, -1.65]$, which hardly affects the accuracy of calculation result.

The numerical integration result of the range of $(-1{,}000, -1.65]$ can be easily got with matlab software. Then

$$\boldsymbol{T}\,(0.05) = \sum_{i=0}^{99835} \left[\ln\left(1 + \Phi\left(\boldsymbol{x}_{\boldsymbol{i}}\right)\right) \times 0.001\right] = 0.0482$$

The value of parameter k_α can be easily got with $\Gamma_{0.05} = 0.048\ 2$ and $\Phi^{-1}(0.05) = -1.65$ on the confidence level 95 %, which is as following.

$$k_\alpha = -1.65 - 0.0482/\ln\left(1 + 0.05\right) = -2.64$$

2. the calculation of objective function

The objective function of optimization model for loan portfolio can be easily got with $k_\alpha = -2.64$ and yield rate variance for loan portfolio $\sigma(X) = \sqrt{\boldsymbol{X}^{\boldsymbol{T}}\boldsymbol{U}\boldsymbol{X}}$, which is as following.

$$\text{obj: min} \quad GM = 2.64\sqrt{\boldsymbol{X}^{\boldsymbol{T}}\boldsymbol{U}\boldsymbol{X}} - \mu(X) \tag{68.46}$$

where, yield rate for loan portfolio $\mu(X)$ is got from E

68.4.5 The Establishment of Objective Yield Rate for Loan Portfolio

We can easily get the minimum yield rate expectation and the maximum yield rate expectation from the yield rate expectation vector $R = (\mu_1, \mu_2, \ldots, \mu_9)^{\mathrm{T}}$.

$$\min_i \mu_i = 0.0005 \tag{68.47}$$

$$\max_i \mu_i = 0.0545 \tag{68.48}$$

We can easily get $A = 0.267{\cdot}5 \times 10^{-3}$ with formula (68.34) and vector $I = (1,1,\ldots,1)^{\mathrm{T}}$ and yield rate expectation vector $R = (\mu_1,\mu_2, \ldots, \mu_9)^{\mathrm{T}}$ and constant matrix U.

We can easily get $B = 0.128{\cdot}9 \times 10^{-4}$ with formula (68.35) and yield rate expectation vector $R = (\mu_1, \mu_2, \ldots, \mu_9)^{\mathrm{T}}$ and constant matrix U.

We can easily get $C = 0.562{\cdot}0 \times 10^{-2}$ with formula (68.36) and vector $I = (1, 1, \ldots, 1)^{\mathrm{T}}$ and constant matrix U.

We can easily get $D = 0.091{\cdot}1 \times 10^{-4}$ with formula (68.37) and A =$0.267{\cdot}5 \times 10^{-3}$ and $B = 0.128\,9 \times 10^{-4}$ and $C = 0.562\,0 \times 10^{-2}$.

We can easily get $H = 0.562{\cdot}0 \times 10^{-2}$ with formula (68.39) and $k_\alpha = -2.64$.

We can easily get $\mu(r_m) = 0.0478$ that is the yield rate expectation corresponds to the minimum of GM point with formula (68.33) and those values of *A, B, C, D, H*.

The range of yield rate for loan portfolio can be got with formula (68.41) and $\mu(r_m) = 0.0478$ and the maximum yield rate expectation above formula (68.48), which is as following.

$$0.0478 \le \mu(X) \le 0.0545 \tag{68.49}$$

Therefore an objective yield rate $\mu(X)$ can be chosen in the range of [0.0478, 0.0545] from above formula (68.49).

We choice $\mu(X) = 0.05$ in this paper, and the constraint condition of yield rate expectation for loan portfolio can be got with formula (68.30) and $n = 9$ and the reasonable yield rate for loan portfolio $\mu(X) = 0.05$ and μ_i from above formula (68.42), which is as following.

$$0.0532x_1 + 0.0499x_2 + \cdots + 0.0545x_9 = 0.05 \tag{68.50}$$

68.4.6 The Establishment of Optimization Model for Loan Portfolio

The expressions of objective function and VaR constraint condition can be easily got with formula (68.46) and formula (68.43) and $\mu(X) = 0.05$.

We set up the optimization model based on geometric spectral measure GM, which is as following.

$$\text{obj: } \min \ GM = 2.64\sqrt{\boldsymbol{X}^{\boldsymbol{T}}\boldsymbol{UX}} - 0.05 \tag{68.51}$$

$$\text{s.t.} \quad 0.0532x_1 + 0.0499x_2 + \cdots + 0.0545x_9 = 0.05 \tag{68.52}$$

$$1.65\sqrt{\boldsymbol{X}^{\boldsymbol{T}}\boldsymbol{UX}} - 0.05 \le 0.05 \tag{68.53}$$

$$\sum_{i=1}^{n} \boldsymbol{x}_i = 1 \tag{68.54}$$

$$0 \le x_i \le 1, i = 1, 2, \ldots, n \tag{68.55}$$

68.4.7 Results of Optimization Model for Loan Portfolio

The optimal solution of loan distribution, which is used to control the excess loss risk for loan portfolio, is to drawn from the optimization model of formulas (68.51), (68.52), and (68.55).

It is shown in Table 68.2. At the same time, the geometric spectral measure GM of loan portfolio can be got. The GM value is equals to 0.024.

68.4.8 The Analysis of Optimization Model

We get three conclusion as following. Firstly, more credit funds have be assigned to the 2th loan and the 3th loan and the 8th loan for the higher yield rate expectation and the lower yield rate risk.

Secondly, the 5th loan and the 6th loan and the 7th loan have not get credit funds for the achievement of a suitable yield rate level.

Thirdly, credit funds do not have be assigned to the 1th loan which has the higher yield rate expectation and the higher yield rate risk for reducing the probability of excess loss occurrence for loan portfolio.

68.5 Conclusion

1. Taking geometric spectral measure GM of loan portfolio risk as an objective function, taking the VaR of loan portfolio as a constraint condition, this paper sets up a new optimization model for loan portfolio, which solves the risk controlling problem in extreme cases.
2. This optimization model takes geometric spectral measure GM as an objective function, which reducing the probability of loan portfolio loss beyond VaR. This model makes up for the drawback of existing researches failing to measure tail information of the yield rate distribution for loan's portfolio, and unable to distinguish the catastrophic loss from general loss for loan's portfolio.

Acknowledgment This research is supported by the Inner Mongolia University of science and technology innovation fund project (the grantnumber is 2011NCL059). We thank Inner Mongolia University of science and technology for its financial support.

References

Acerbi C (2002) Spectral measures of risk: a coherent represent action of subjective risk aversion [J]. J Bank Finance 26(7):505–518

Altman EI (2000) Predicting finance distress of companies: revisit the Z-score and ZETA models [J]. J Finance 55(7):18–20

Andrew EB, Gah YV (2011) Conditional value-at-risk in portfolio optimization: coherent but fragile [J]. Oper Res Lett 39(4):163–171

Chi Guotai, Xu Cheng, Li Yanxi (2000) Optimal decision making model of asset liability portfolio for banks. Proceedings of 2000 international conference on china joins the world trade organization [C]. Shanghai Scientific and Technological Literature Publishing House, Shanghai, pp 67–73

Elliott RJ, Siu TK, Alex B (2010) On mean variance portfolio selection under a hidden Markovian regime switching model [J]. Econ Model 10(6):678–686

Gollinger TL, Morgan JB (1993) Calculation of an efficient frontier for a commercial loan portfolio [J]. J Portfolio Manag 19(2):39–49

Li Xiang, Qin Zhong-feng, Kar S (2010) Mean variance skewness model for portfolio selection with fuzzy returns [J]. Eur J Oper Res 202(1):239–247

Li Zhong-fei, Yuan Zi-jia (2010) A dynamic mean-variance model of portfolio selection under parameter uncertainty [J]. J Manag Sci 13(12):1–9

Markowitz H (1952) Portfolio selection: efficient diversification of investments [J]. J Finance 7(1):77–91

Peter SR (2002) Commercial bank management [M], 5th edn. The McGraw-Hill Companies, New York

Rupak B, Samarjit K, Dwijesh DM (2009) Fuzzy mean variance skewness portfolio selection models by interval analysis [J]. Comput Math Appl 32(7):126–137

Shang Zhou, Xian Shi-quant, Pan Chen (2010) Improbability and mathematical statistics [M], 4th edn. Higher Education Press, Bee Jing, pp 239–247

Yao Hai-xing, Li Zhong-fei (2009) Portfolio selection with different borrowing-lending rates: utility maximization model based on mean and VaR [J]. Syst Eng Theory Pract 29(1):22–29

Yoshida YJ (2009) An estimation model of value at risk portfolio under uncertainty [J]. Fuzzy Set Syst 45(7):3250–3262

Zhou Zongfang, Jiang Dazhi, Xi Yang, Lin Jianhua (2002) Decision making model of loan's portfolio optimization based on the yield of VaR [J]. Chin J Manag Sci 10(6):1–7

Chapter 69
The Effect of Market Selection on Lagged Currency Risk Exposure

Xiao Wu and Long-wen Zhou

Abstract This paper studies the exchange rate exposure of wood furniture companies of China by using the Shanghai and Shenzhen A-share market listed companies as a basis for our sample. We use the enterprise unique exchange rate index to improve the model and analyze the sensitivity of the stock returns of listed companies to different lag period changes in exchange rates by using distributed-lag model. We also study the influence of marketing strategy in the choice of the market on lagged exchange rate risk. The results show that the foreign market share is proportional to the risk exposure and the market dispersion is in contrast with risk exposure (MOE Research Project Fund of Humanities and Social Sciences (No. 09YJA630031) and Hunan University Research Projects (No. 09HDSK092; No. 11HDSK118)).

Keywords Exchange rate risk exposure • Lag period • Market selection • Wood furniture industry

69.1 Introduction

July 21, 2005, China began a managed floating exchange rate system based on market demand and supply, with reference to a basket of currencies. RMB had no longer been pegged to the US dollar. Hence the volatility of its exchange

X. Wu
Department of Management, University of Hunan University, Changsha, China

Department of Economics and Business, University of Xinjiang Agricultural University, Urumqi, China

L. Zhou (✉)
Department of Management, University of Hunan University, Changsha, China
e-mail: zlw088210@163.com

E. Qi et al. (eds.), *The 19th International Conference on Industrial Engineering and Engineering Management*, DOI 10.1007/978-3-642-37270-4_69,

rates increases. But the overall trend has been appreciation. In the context of globalization, regardless of whether a company has import and export business in its daily operations or not, it more or less likely to encounter exchange rate risk. Therefore, no matter what type of business, it is necessary to control and prevent exchange rate risk. The company's marketing strategy including market selection, product strategy, price strategy and promotional strategy. We only study the influence of market selection on exchange rate risk here. The choice of the market includes two aspects: one is the domestic and foreign market share, the other is the number of fragmented markets. We believe that the share of foreign markets is positively related to exchange rate risk exposure, the dispersion of the market is negatively related to exchange rate risk exposure.

China is a big furniture producing country, this huge market provides a broad space for furniture industry's development. The United States is the most important countries as Chinese furniture exports regions, nearly half of Chinese furniture exports to the United States. In addition, Europe, Japan, Hong Kong is also the major countries or regions. Furniture's raw materials in China rely on import and furniture export has a relatively large proportion in furniture industry. The RMB's appreciation has a large impact on the imports of raw materials and furniture exports. Since 2005, the RMB have appreciated a lot, and this trend will continue to maintain, will greatly affect China's furniture exports and imports of foreign furniture products. According to 2010 statistics, the furniture industry's export is $15,777,930,000,000, an increase of 31.3 % compared to 2009, import amounted to $1,394,830,000,000, an increase of 38.7 % compared to 2009. As a larger proportion of export and import in this industry, the existence of significant exchange rate risk exposure is the focus of our research. We choose wood furniture A-share companies listed in Shanghai and Shenzhen Stock Exchange for the study.

We concentrate on this because since Alder and Dumas defined currency exposure as the firm value which may be subject to exchange rate risk on the basis of the relationship between a future cash flow of a firm and its firm value (Adler and Dumas 1984), much progress has been achieved in this research field, but our investigation goal of this paper might be one of the issues which have not been fully explored. Jorion amended the original Alder-Dumas model, added market yield rat to the model and formed the augmented market model (Jorion 1990, 1991). This is regarded as a classic model of measuring currency exposure. This method is widely used in literature, such as (Khoo 1994; Bodnar and Gentry 1993; Doukas et al. 1999; Chow et al. 1997). The researchers had different views in using Jorion's augmented market model about the definition of the exchange rate changes. Some used a single exchange rate, e.g., (Doukas et al. 1999; Williamson 2001), while some others employed trade-weighted exchange rate index (Muller and Verschoor 2007; Bartov and Bodnar 1994).

Some earlier literature argued that there is little link between the current exchange rate changes and stock price, while the lagged exchange rate changes are much more significantly relevant. In other words, exchange rate fluctuation will take some time to impact firm cash flow, and in turn to influence its stock price. This time period of influence process is called lag period. For example, Bodnar and Wong found that the full impact of exchange rate changes is not instantaneous reflected in the stock returns. There is no evidence that exchange rate changes have a significant impact on the spot price of the company. When the lagged exchange rate change was introduced into the model, they found strong evidence to support the existence of currency exposure (Bodnar and Wong 2003). Amihud used monthly data and the real trade-weighted exchange rates, and found that when the exchange rate lagged 1 and 3 months, 32 US exporters faced significant currency exposure (Amihud 1994). Donnelly and Sheehy studied 39 British exporters and found that the exchange rate fluctuation needs to go through a few months to be reflected in the stock prices (Donnelly and Sheehy 1996). These studies showed that the corporate value is not only influenced by the current change in exchange rate, but also by the lagged exchange rate changes. Among some others, Masry studied the UK financial industry, and showed that the UK non-financial enterprises were significantly exposed to the lagged exchange rate risk, more significant than the current currency exposure (El-Masry 2003). He and Ng found that the lagged exchange rate fluctuation had no significant effect on Japanese multinational companies (He and Ng 1998).

On the basis of the literature survey, we found lagged exchange rate fluctuation was reported to have no significant impact on Japanese multinational companies, but most of the literature showed the lagged exchange rate changes had influence on firms. However, these methods in the literature are often simply adding a fixed lag period length in the capital market model, and used OLS analysis or cash flow method. We think that each firm is different in specific circumstance so that its production and business process are differently affected by exchange rate fluctuation. We attempt to combine capital market model with distributed lag model, the length of the lagged period of each company is determined after several calculations.

69.2 Methodology

We use the classic Jorion model to estimate currency exposure:

$$R_{it} = \alpha + \beta_m R_{mt} + \beta_x X_t + \xi_t \tag{69.1}$$

Where R_{it} is yield of company *i* at time *t*, R_{mt} is the market portfolio yield at time *t*, X_t is change of exchange rate at time *t*, β_x measures the extent of effect of exchange rate change on corporate earnings.

In order to estimate the lagged impact of exchange rate volatility, we add lagged variables to Eq. (69.1), namely:

$$\begin{aligned} R_{it} = \alpha + \beta_m R_{mt} + \beta_x X_t + \beta_{x-1} X_{t-1} \\ + \beta_{x-2} X_{t-2} + \cdots + \beta_{x-s} X_{t-s} + \xi \end{aligned} \tag{69.2}$$

Where s is the lag length interval, X_{t-s} is the sth lag period of the explanatory variable X, the other variables are identically defined as in Eq. (69.1).

The skewness of the return distribution is one of the important features of the security price. To concentrate on the impact of lag periods, as in literature, we do not consider the relationship between the skewness and the coefficient of risk premium. The introduction of lagged variables may make the model produce multicollinearity. When there are a lot of lag variables, direct estimation will cost a lot degree of freedom. And the current fluctuation and lagged values of exchange rate change may lead high degree of collinearity. This may cause the parameter estimates meaningless.

Foreign market share equal to the income of foreign sales divided by the total sales revenue. We use the kinds of company's assets (including the selection and marketing of the assets of foreign currency and local currency assets) on behalf of the market dispersion. The impact of the selection of market share and market dispersion on exchange rate risk exposure can be measured with the following formula:

$$\beta_{sun} = c + \eta_1 MR + \eta_2 NUM + \zeta \tag{69.3}$$

Where β_{sum} is the total exchange rate risk, MR is the proportion of the foreign market share of the total market, NUM is market dispersion.

There are a total of 11 A-share listed companies listed on the wood furniture industry, there are five companies listed before 2006, respectively, Yihua Timber, Sichuan Guodong Construction, Guangming Group Furniture Corporation, Daya technology and Babby rabbit. As Guangming is ST firm, we removed it. The company has foreign sales which is reported in its annual report has been identified as exporter. On the contrary, compared with domestic enterprise (where the definition of domestic enterprise is the corporation's sale market all based on the domestic market, and all the income comes from sales in the domestic market). Through reading the company's annual report, which is from the Shanghai and Shenzhen Stock Exchange, we can know:

1. Yihua Timber sets up a wholly owned subsidiary in the United States, it also has operations in Europe and Australia. Its foreign currency is U.S. dollars, Hong Kong dollars, Australian dollars, Euro. Its export business is stronger than domestic business;
2. Sichuan Guodong Construction doesn't publish foreign Operation income in its annual report, so we divide it into the domestic market enterprise;
3. Daya technology has an associated company belong to its subsidiary in the United States. The main business is to sell products to the United States customers. According to annual report, Foreign currency include the dollar and the Euro, and compared to the dollar, Euro only a small part;
4. Babby rabbit does not have foreign subsidiaries, but has foreign currency that is dollar and the Euro, the ratio is about 8:2.

We calculate weighted exchange rate mainly by considering the foreign currency in the proportion of the total foreign currency. The method is calculated from the average proportion of foreign currency from 2006 to 2010, the results reserve four

Table 69.1 Exchange rate risk exposure

Company name	Yihua Timber	Guodong construction	Daya	Babby rabbit
Lag period	2	1	3	3
Adjusted R^2	0.5758	0.6216	0.6691	0.7022
Total exposure coefficient	6.0153	1.1594	2.0580	−2.8170

decimal. Yihua Timber's exchange rate is the weighted exchange rate of the dollar and the Euro, after analysis, weights are 0.9995 and 0.0005 (From the 2007 financial reports began to have Australian dollars, but the Australian dollar converted into RMB, ranging from 20 to30 yuan. The calculation is still 0 when retain six decimal. The proportion of the HK is 0.000030, considering Hong Kong dollars and U.S. dollar-linked exchange rate system and the proportion is not large, so not included in the weighted exchange rate). Guodong Construction doesn't have foreign currency, as the dollar is the main currency used by the import and export, it only take the U.S. dollar into account. The Daya technology has a weight of 0.9965 and 0.0035 relate to dollar and Euro. Baby rabbit can also use the weighted exchange rate of the dollar and the Euro, the weights are 0.8685 and 0.1315, respectively. Monthly exchange rate data is from the Chinese Administration of Foreign Exchange. Consideration of cash dividends reinvested monthly returns and market month yields are from the GTA database.

69.3 Results

We consider the lag periods of exchange rate fluctuations up to 4, that is, consider the 4-month lag periods, and select the largest adjusted R2 to determine the optimal lag period. We sum the current coefficient and the lag coefficient as total exchange rate exposure coefficient, retaining the results of four decimal. We use the VIF test to check multi-collinearity. The results showed that VIF are less than 2, far less than 10, the model does not exist multi-collinearity (Table 69.1).

69.4 Discussion

Through empirical research, it is found that Yihua Timber, Guodong Construction, and Daya Technology have been negatively affected by RMB's appreciation, Babby rabbit benefited from the appreciation of the RMB. Through browsing the company's annual report, we can know: Yihua Timber's export sale accounts for the main part, and mainly sells to North America. The main currency used is dollars, and export is the main mode of operation of the enterprise. The labor cost advantage allowing companies to export products at competitive prices. As RMB appreciation against the dollar, the superiority will gradually disappear. Sichuan Construction has no foreign sales, that means there is no export, but it will face the competition of imported goods, because of the appreciation of the RMB, imported goods have more competitiveness, so we think Construction exists exchange rate risk, the empirical results show that exchange rate risk indeed exist. The Daya's domestic income accounts for the majority of total sales revenue. So, its exchange rate risk exposure is lower than Yihua. The share of euro of Babby rabbit is larger than the other three companies. From the entire range, RMB against the euro appreciated a little, there was even devaluation in some period. RMB depreciate against the euro to enhance product's competitiveness in European market, Babby rabbit selects this type of market to expand business sales. Therefore, companies can consider to expanding the European market or other markets where the RMB depreciation against its currency.

We first plus every year's sales revenue, then use the sum of foreign sales income to divide to total sales income to obtain the proportion of the foreign market share of the total market ,which is represented by MR, the results reserve four decimal places. Yihua is 0.9737, Guodong Construction is 0, the Daya is 0.1090, Babby rabbit is 0.2695. We can not know exactly the business partners to each firm, through read the annual reports, we know that Yihua Timber has the greatest variety of assets, which has five kinds of assets, market dispersion is regarded as 1. Daya and Babby rabbit have three kinds of assets, market dispersion is 0.6 ($3/5 = 0.6$), Guodong Construction only has local currency asset, the market dispersion is 0.2 ($1/5 = 0.2$). We choose MR and NUM as the independent variables, the total exchange rate risk exposure as the dependent variable, you can get MR's standardized coefficient is 0.940, the NUM's standardized coefficient is -0.306. MR coefficient is positive, show the relation between the proportion of the foreign market share of the total market and the total exchange rate risk exposure is positive, the greater the share of foreign market, the greater the total exchange rate risk exposure. NUM coefficient is negative, indicating that market dispersion and exchange rate risk exposure is negatively correlated, the greater the dispersion, the smaller the total exchange rate risk exposure. This is consistent with our hypothesis (Table 69.2).

69.5 Conclusion

The augmented market model is used in this paper, and adds the lagged values of explanatory variables. We choose the wood furniture companies listed on the Shanghai and Shenzhen A-share market, and carried out empirical research about the exchange rate risk exposure from 2006 to 2010. Scholars on whether use of a single exchange rate or weighted exchange rate has not formed a unified answer, we believe that different enterprise has different exchange rate, some use the weighted exchange rate, others use a single exchange rate. By empirical study, we found that the greater the share of foreign market, the greater the total exchange rate

Table 69.2 Foreign market share

		Sum	MR
Yihua	FS	9,296,345,802.43	0.9737
	DS	251,371,068.36	
Construction	FS	0	0
	DS	11,356,529,627.66	
Daya	FS	3,109,127,109.67	0.1090
	DS	25,427,879,716.50	
Baby rabbit	FS	1,021,540,620.88	0.2695
	DS	2,769,665,135.23	

risk exposure, and the greater the market dispersion, the smaller the total exchange rate risk exposure. So when the foreign market share is the greater, the need to hedge exchange rate risk exposure is more important, companies can hedge through diversifying the market, when the enterprises expand overseas market share, you can consider investment and sell in different markets, it not only achieve the purpose of expanding the market, but also reduce exchange rate risk exposure. At the same time, it can consider expanding the European market or other market where RMB depreciation against its currency.

The inadequacies of this article is unable to obtain the life of product cycle of each enterprise, can not set a specific lag period, and only generally set the lag period up to four. This is the direction of future research.

Acknowledgment We are grateful to Chi Xie, Zhijian Zeng, and Xiaoli Li for helpful discussions.

References

Adler M, Dumas B (1984) Exposure to currency risk: definition and measurement. Financ Manag 13:41–50

Amihud Y (1994) Exchange rates and corporate performance. Irwin Professional Publishing, New York

Bartov E, Bodnar GM (1994) Firm valuation, earnings expectations, and the exchange-rate exposure effect. J Finance 49:1755–1785

Bodnar G, Gentry WM (1993) Exchange rate exposure and industry characteristics: evidence from Canada, Japan, and the USA. J Int Money Finance 12:29–45

Bodnar GM, Wong MF (2003) Estimating exchange rate exposures: issues in model structure. Financ Manag 32:35–67

Chow EH, Lee WY, Solt ME (1997) The economic exposure of US multinational firms. J Financ Res 20:191–210

Donnelly R, Sheehy E (1996) The share price reaction of U.K. exporters to exchange rate movements: an empirical study. J Int Bus Stud 27:157–165
Doukas J, Hall PH, Lang LHP (1999) The pricing of currency risk in Japan. J Bank Finance 23: 1–20
El-Masry AA (2003) The exchange rate exposure of UK nonfinancial companies industry-level analysis. Working Paper, Manchester Business School, Manchester, UK
He J, Ng LK (1998) The foreign exchange exposure of Japanese multinational corporations. J Finance 53:733–754
Jorion P (1990) The exchange rate exposure of U.S. multinationals. J Bus 63(3):331–345
Jorion P (1991) The pricing of exchange rate risk in the stock market. J Financ Quant Anal 26: 363–376
Khoo A (1994) Estimation of foreign exchange exposure: an application to mining companies in Australia. J Int Money Finance 13:342–363
Muller A, Verschoor WFC (2007) Asian foreign exchange risk exposure. J Japan Int Econ 21: 16–37
Williamson RG (2001) Exchange rate exposure and competition: evidence from the automotive industry. J Financ Econ 59:441–475

Chapter 70
Chairman Overconfidence and Dividend Payout Policy: *An Analysis Based on China's Listed Companies*

Xiu-qun Hu, Rong-shen Lv, and Guo-liu Hu

Abstract Considering the individual characteristics of chairman and organizational environment, this paper constructs a comprehensive evaluation index as proxy measure of chairman overconfidence and develops a panel regression model to empirically examine the impact of chairman overconfidence on corporate dividend policy. The results show that there is a notable negative correlation relationship between overconfidence behavior of the chairman and the dividend payout policy in China's listed companies. Relative to the rational chairman, overconfident chairman will be less likely to pay cash dividend.

Keywords Behavioral corporate finance • Dividend • Overconfidence • Payout policy

---

Foundation item:
1. Humanities and Social Science Research Funds of the Ministry of Education (11YJC630067).
2. National Natural Science Foundation of China (NO. 71062009).

X. Hu (✉)
Department of Management and Economic, Tianjin University, Tianjin, China

School of Economics and Management, Hainan University, Hainan, Haikou, China
e-mail: hxqun7377@yahoo.con

R. Lv
Department of Management and Economic, Tianjin University, Tianjin, China

School of Management, Tianjin University of technology, Tianjin, China

G. Hu
School of Economics and Management, Hainan University, Hainan, Haikou, China

E. Qi et al. (eds.), *The 19th International Conference on Industrial Engineering and Engineering Management*, DOI 10.1007/978-3-642-37270-4_70,

70.1 Introduction

Psychology studies have shown that people are not entirely rational, one of the most important characteristics of which is that people would be overconfident. Many Studies conducted by western scholars have proved that leaders and managers in company commonly behave more overconfident than the ordinary people, overconfidence is one of the most important features of managers' non-rational behavior. In their studies on entrepreneur, Cooper et al. (1988) found entrepreneurs are generally overconfident. Langer (1975), Camerer and Lovallo (1999), Gervais et al. (2003) and Truong and Heaney (2007) have proven that business executives typically exhibit greater overconfidence than ordinary employees.

To traditional Chinese culture and present economic environment, the overconfident behavior of company executives is more serious in China (Yates et al. 1996; Fu-xiu Jiang et al. 2009). In recent years, many state-owned enterprises leaders and private enterprise founders are described as "godfather", their overconfidence are overmuch strengthened by their meritorious contribution and continued success. Thus, studying the effect of overconfident behaviors of entrepreneurs and company executives to the company's financial decision-making has become an important part of behavioral corporate finance research.

Dividend policy is one of the core financial policies in listed companies. Mature capital market experience has shown that a reasonable dividend policy is a crucial factor to optimize capital structure, improve capital efficiency and maximize company's value. Therefore, how to develop a reasonable dividend policy becomes a vital issue for corporate decision makers. According to "Company Law" in China, the company's dividend distribution project is ultimately determined by the board. And chairman of the board has great influence on the board to make financial decisions, whose psychological deviation would inevitably generate significant impact on the company's dividend policy-making. Meanwhile, the psychological research shows that the formation of psychological deviation is often relevant to the individual characteristics like age, sex, education, professional skills, and to the organizational environment as well, such as financial structure of companies and corporate governance structure. Therefore, based on individual characteristics of chairman and organizational environment, this paper builds a comprehensive evaluation indicator as proxy measure of chairman overconfidence to investigate the influence of chairman overconfidence on listed company's dividends policies in China.

70.2 Literature Review and Research Assumptions

In recent years, with the development of psychology and behavioral science, human behavior factors are introduced into the dividend policy research. In the existing literature, scholars mainly studied the effect of human behavior deviation on dividend policies from two aspects. First, assume that managers are rational and investors are irrational. Its main findings include prospect theory, regret aversion theory, self-control theory, rational expectations hypothesis, catering theory of dividends and so on. Second, assume managers are non-rational and investors are rational. To study the effect of managers' non-rational behavior, especially managerial overconfidence on company's dividend policy is becoming a front subject. Based on data of foreign markets, Deshmukh et al. (2008) established a model considering the relationship between dividend policies and CEO overconfidence and conducted empirical tests on this model, then drew the conclusion: relative to rational CEOs, over-confident CEOs prefer a lower level of dividend payment. Leonardo Cordeiro (2009) reached the same conclusion from a managerial perspective. But, these studies were based on the developed capital market, in the traditional culture background and Special and imperfect capital market of china, especially after the reform of non-tradable shares, whether can we get the same conclusion?

Unlike some countries such as American and the United Kingdom where different stakeholders maintain a balanced force, ownership and management right of most Chinese enterprise is not or cannot be completely separated due to China's cultural characteristics and Special economic system, additionally, the members of the board and managers of the state-owned enterprise are generally appointed by the government, and the chairman of private enterprise has a strong control power as a result of the typical family management, So a mutual balancing relationship between board and managers can not fundamentally exist in the Chinese listed companies. That is to say, as the highest representative of the interests of shareholders, chairman has very important influence on the company's financial decision-making, and lead to the existence of chairman overconfidence. Therefore, specific to listed companies in China, this paper proposes the following hypothesis:

Hypothesis Relative to the rational chairman, overconfident chairman will be less likely to pay cash dividend.

One point to note is that the key point and difficulty of studying overconfidence behavior lie to the measure of overconfidence. In recent years, in the field of behavioral finance, some scholars innovatively tried to put forward a number of measurement methods, such as CEO share holding conditions (Malmendier and Tate 2003, 2005), Evaluation of the related mainstream media on CEO (Malmendier and Tate 2003), frequency of Merger & Acquisition implemented by CEO (Malmendier and Tate 2003), relative remuneration of the CEO (Hayward and Hambrick 1997), current performance of the enterprise (Hayward and Hambrick 1997), The business climate index (Ming-gui Yu et al. 2006). However, using the above mentioned single alternative indicators to characterize the degree of overconfidence of managers is relatively rough and may generate large noise, its reliability and validity is doubtful.

Bertrand and Schoar (2003) suggest that different company's managers are not homogeneous, manager's personal characteristics will affect the company's financial policy. Further, some research in Forbes (2005), Beyer (1990) and Lundeberg et al. (1994), Landier and Thesmar (2009), Heath and Tversky (1991), Ben-David et al. (2007), Paredes (2005) find that human overconfidence behavior will be affected by individual characteristics like age, sex, education level, professional skills. At the same time, the research of Bondt and Thaler (1990) show that organizational environment such as financial structure, governance structure, size, past performance of companies can also impact overconfidence psychology of entrepreneurs and managers. For the above reasons, to make up for the deficiency of the previous single alternative indexes, this paper applies a comprehensive evaluation indicator related to individual characteristics of chairman and organizational environment as proxy measure of chairman overconfidence to empirically test the influence of chairman overconfidence on listed company's dividends policies in China.

70.3 Research Design

70.3.1 Data

We gather A-share listed companies in China as an initial sample, then select the listed companies that consecutively launched a cash dividend payout scheme and in which the chairman continuously held the post in the period of 2007–2010, and the following companies are excluded: (1)ST companies. (2)Three types of monopoly industries such as communication and cultural industry, production of electricity, gas and water supplying industry, finance and insurance industry. (3) Companies whose information of corporate finance and individual characteristics of chairman is incomplete. Finally, we obtain 1,072 samples observations. All the data are quoted from the CSMAR Chinese listed companies database of Shenzhen GTA Information Technology Co, Ltd. and RESSET database, except part of them are obtained by hand.

70.3.2 Variable Selection

(a) Explained variable. Since the rationality of listed companies' cash dividend distribution is affected by the scale and profitability of the company, it can not be simply measured by the absolute total amount of cash dividends or unit dividend. Here the research experience of LLSV and Todd Mitton used for reference, we select cash dividend payout rate as the explained variable of cash dividend. One thing needs to be noted is that this research focuses on the influencing factors of companies' annual cash dividend allocation decisions, as for companies that distribute cash dividends more than one time a year, the accumulated cash dividend in one year is regarded as an annual cash dividend distribution.

(b) Explanatory variable. A comprehensive evaluation index of the chairman overconfidence. Considering individual characteristics of chairman and organizational environment of his company, employing the Generalized principal components analysis and an objective weight method, we construct a comprehensive evaluation index of the chairman overconfidence and put the score of this comprehensive

Table 70.1 Variable symbols and metrics

Variable type	Variable name	Symbols	Metrics
Explained variable	Dividend-payout ratio	DR	Cash dividend per share/earnings per share
Explanatory variable	Comprehensive evaluation indicator about chairman overconfidence	F	Generalized principal components analysis
Control variables	The first major stockholder's holdings	H1	The number of the first major stockholder shares/the total number of shares
	The ratio of The first major shareholder's holdings to the second major shareholder's holdings	Z	The first major stockholder's holdings/the second major stockholder's holdings
	Controlling Shareholder	ECD	The first major stockholder's holdings are greater than 30 % of the total number of shares, take 1, otherwise take 0.
	The ratio of the independent director in the board	IDP	The number of independent director/the total number of the board
	Company size	SIZE	Ln(The book value of the total assets at the year-end)
	Asset-liability ratio	DEBT	Total debt at the year-end/Total assets at the year-end
	Return on equity	ROE	Net Profit after tax/Net asset *100 %

index as the criteria to show the degree of the chairman overconfidence (Owing to the space forbids, the calculation of the score of the comprehensive index is not stated in detail).

(c) Controlled variables. A company's dividend policy is relevant to corporate governance structure, corporate characteristics, performance of the company and other factors. In order to ensure the reliability of the results, several key variables that reflect company governance structure and company characteristics and performance of the company would be controlled. These variables mainly include the first major stockholder's holdings, the ratio of the first major shareholder's holdings to the second major shareholder's holdings, the existence of a controlling shareholder, the ratio of the independent director in the board, company scale, asset-liability ratio and net return on equity. Variable symbols and the metric are shown in Table 70.1.

70.3.3 Model and Estimation Results

Taking into account the individual differences among the variables, the following variable intercept panel data model is established to study the relationship between chairman overconfidence and dividend payout policy:

Table 70.2 Hausman test results

Test summary	Chi-Sq. statistic	Chi-Sq. d.f.	Prob.
Cross-section random	15.21127	8	0.0552

$$DR_{it} = \alpha_i + \beta_1 F_{it} + \lambda_1 H1_{it} + \lambda_2 Z_{it} + \lambda_3 ECD_{it} + \lambda_4 IDP_{it} + \lambda_5 SIZE_{it} + \lambda_6 DEBT_{it} + \lambda_7 ROE_{it} + \varepsilon_{it} \quad (70.1)$$

The definitions of the variables in model (1) are shown in Table 70.1.

In general, a variable intercept model can be a fixed effects model or a random effects model, we need to test the selected data before deciding which model is suitable for the analysis. Now Hausman tests on the selected data through Eviews6.0 are carried out. Hypothesis H0: the model should be set as a random effects model. H1: The model should be set as a fixed effects model. Test results are shown in Table 70.2.

From the results in Table 70.2, $P = 0.0552 < 0.1$, thus it is significant at the 0.1 level and the null hypothesis should be rejected, we shall construct a fixed effects model. However, since the selected sample data in this article are from a short panel, from the statistical point of view, using a fixed effects model would cause a great lose in freedom, while using a random effects model can avoid the loss of degrees of freedom. Thus, the following table lists the results of both the fixed effects model and random effects model for comparison.

The regression analysis results in Table 70.3 show that, whether it is a fixed effects model or a random effects model, there are significant negative correlation between listed companies' dividend distribution policies and chairman overconfidence at the 0.01 level, which is accordant to above hypothesis. The result indicates that whether the chairman is overconfident or not does generate a strong influence on the level of cash dividends payout, those overconfidence chairmen tend to retain cash in the company and distribute less cash dividends.

In addition, the analysis results of the controlled variables in fixed effects model (1) show the following: There is a significant negative correlation relationship between the ratio of the first major shareholder's holdings to the second major shareholder's holdings (Z) and the cash dividend payout level, which is accordant to our expectation, indicating that the more centralized the equity, the weaker the constraints of the second largest shareholder to the largest shareholder, and the lower the level of listed companies' dividend payout is. There is a notable negative correlation relationship between the ratio of the independent director in the board (IDP) and the cash dividend payout level, which is contrary to our expectation. It may be the reasons that the independent directors of listed companies in China have long been lack of independence, the definitions of legal rights and responsibilities are not clear, the relationship with the Board of Supervisors is chaotic, thus independent directors did not effectively play their monitoring role in enhancing the transparency of the Board of directors, ensuring the fairness of the Board of directors, protecting the interests of small investors, etc. Contrary to the expectation,

Table 70.3 Panel data regression analysis results (Dependent variable: Dr)

		Model(1)	Model(2)
Independent variables	Expected symbols	Fixed effects	Random effects
F	–	−0.004***	−0.004***
		(0.0000)	(0.0066)
H1	–	−0.003	0.006**
		(0.2380)	(0.0444)
Z	–	−0.001***	−0.001
		(0.0000)	(0.1424)
ECD	–	0.025	−0.274***
		(0.6095)	(0.0073)
IDP	+	−0.627***	−0.975
		(0.0003)	(0.1068)
SIZE	+	0.011	−0.069**
		(0.6181)	(0.0208)
DEBT	–	0.554***	0.056
		(0.0000)	(0.8005)
ROE	+	−1.805***	−3.240***
		(0.0000)	(0.0000)
C	?	0.568	2.739***
		(0.2417)	(0.0000)
Sample observations		1,054	1,054
R2 (W.Stat)		0.660	0.104
Adj-R2 (W.Stat)		0.540	0.097
F Test		5.498	15.133
		(0.0000)	(0.0000)
D.W Test (W. Stat)		2.590	2.067

***Significant at the 0.01 level
**Significant at the 0.05 level

there is a notable positive correlation relationship between the level of asset-liability ratio (DEBT) and the cash dividend payout level, implying the phenomenon of violating the interests of creditors through paying cash dividends to shareholders are still serious in China's listed companies. There is a notable negative correlation relationship between return on equity (ROE) and the cash dividend payout level, which is contrary to our expectations, the reason might be a vast number of tradable shareholders are lack of "sense of dividends", they value profits much higher than dividends, high profits raise the expectation of the majority shareholders on the company's future development prospect, push them more willing to hold stocks but not cash dividends. meanwhile, since listed companies' behavior of hiding profits have reduced after China's non-tradable shares reform, the release of previously "hided" profits of listed companies substantially boost the performance growth of the companies, which eventually leads to the anomalous result that listed companies with high profitability pay low cash dividends.

70.4 Conclusions

Through investigating the individual characteristics of chairman and organizational environment that influence the level of chairman overconfidence, this paper has constructed a composite evaluation index of chairman overconfidence using the generalized principal component analysis method. And then has built a panel regression model of the effect of chairman overconfidence on cash dividend policies to empirically analyze the relationship between the two. The results indicate: there is a significant negative correlation between chairman overconfidence and listed company's cash dividend payout policy, in comparison to rational chairmen, overconfident chairmen tend to retain cash in the company and pay less cash dividends. Through a new way of thinking to solve the problem of measuring chairman overconfidence, this paper will further provide an empirical evidence for the explanation of listed companies' dividend policies in China.

References

Ben-David I, Graham JR, Harvey CR (2007) Managerial overconfidence and corporate policies. Working paper, Ohio State University, Columbus, pp 1–55

Bertrand M, Schoar A (2003) Managing with style: the effect of managers on firm policies. Quart J Econ 118(4):1169–1208

Beyer S (1990) Gender differences in the accuracy of self-evaluation of performance. J Pers Soc Psychol 59(5):960–970

Camerer C, Lovallo D (1999) Overconfidence and excess entry: an experimental approach. Am Econ Rev 89(1):306–318

Cooper AC, Woo CY, Dunkelberg WC (1988) Entrepreneurs' perceived chances for success. J Bus Ventur 3(2):97–108

Cordeiro L (2009) Managerial overconfidence and dividend policy. London Business School working paper, London Business School, London, pp 1–36

De Bondt WFM, Thaler RH (1990) Do security analysts overreact? Am Econ Rev 80(2):52–57

Deshmukh S, Goel AM, Howe KM (2008) CEO overconfidence and dividend policy: theory and evidence. Report, DePaul University, Chicago, USA, pp 1–39

Forbes DP (2005) Are some entrepreneurs more overconfident than others? J Bus Ventur 20(5):623–640

Fu-xiu Jiang, Min Zhang, Zhen-fei Lu (2009) Managers overconfidence, enterprise expansion and financial distress. Econ Res J 44(1):131–143 (in Chinese)

Gervais S, Heaton JB, Odean T (2003) Overconfidence, investment policy, and executive stock options. Rodney L. White Center for Financial Research working paper no 15–02, Pennsylvania, pp 1–50

Hayward MLA, Hambrick DC (1997) Explaining the premiums paid for large acquisitions: evidence of CEO hubris. Adm Sci Q 42(1):103–127

Heath C, Tversky A (1991) Preference and belief – ambiguity and competence in choice under uncertainty. J Risk Uncertain 4(1):5–28

Landier A, Thesmar D (2009) Financial contracting with optimistic entrepreneurs. Rev Financ Stud 22(1):117–150

Langer EJ (1975) The illusion of control. J Pers Soc Psychol 32(2):311

Lundeberg MA, Fox PW, Puncochar J (1994) Highly confident but wrong-gender differences and similarities in confidence judgments. J Educ Psychol 86(1):114–121
Malmendier U, Tate G (2003) Who makes acquisitions? CEO overconfidence and the market reaction. Mimeo Stanford University, Graduate School of Business (Draft)
Malmendier U, Tate G (2005) Does overconfidence affect corporate investment? CEO overconfidence measures revisited. Eur Financ Manag 11(5):649–659
Ming-gui Yu, Xin-ping Xia, Zhen-song Zou (2006) Managers overconfidence and enterprise radical liabilities behavior. Manage World 21(8):104–112 (in Chinese)
Paredes TA (2005) Too much pay, too much deference: behavioral corporate finance, CEOs, and corporate governance. Working paper, Florida, pp 1252–1267
Truong T, Heaney R (2007) Largest shareholder and dividend policy around the world. Q Rev Econ Finance 47(5):667–687
Yates JF, Price PC, Lee JW, Ramirez J (1996) Good probabilistic forecasters: the 'consumer's' perspective. Int J Forecast 12(1):41–56

Chapter 71
Shared Knowledge and Shared Understanding Between CIO and Top Management Team: A Literature Review

Rong-jie Lv, Zhao Zhao, Xiao-chun Chen, and Li-jun He

Abstract CIO and top management team, called as information system leadership team (ISLT), are the key leaders during the development of information systems (IS). Without the coordination among the members of ISLT, IS success will be impossible. The quality of the partnership is one of the key successful factors of IS. However, the measurement of the relationship remains unsolved. In prior studies, two concepts (i.e. shared knowledge and shared understanding) are often found in literatures. Even so, some authors think they are the same construct and use them alternatively. And others distinguish one from the other and measure them with different scales. In this way, this paper will reconstruct the two concepts and put forward the propositions to describe the relationships between them.

Keywords CIO • Shared knowledge (SK) • Shared understanding (SU) • Top management team (TMT)

71.1 Introduction

Given the essential characteristic of information system (IS) implementation, the Chief Information Officer (CIO) and top management team (TMT) (named as IS leadership team, abbreviated as ISLT) have to involve themselves into the whole progress and cooperate closely from beginning to end to ensure the success of IS application. As the experts of IS or business, they are not similar to the other

This work is supported by Natural Science Foundation of Hebei Province (G2011202064, G2011202154) and Social Science Foundation of Hebei Province (HB12GL050).

R. Lv • Z. Zhao • X. Chen • L. He (✉)
School of Economics and Management, Hebei University of Technology, Tianjin, China
e-mail: sm_hebut@163.com

E. Qi et al. (eds.), *The 19th International Conference on Industrial Engineering and Engineering Management*, DOI 10.1007/978-3-642-37270-4_71,

domain knowledge. The value of IS relies on collaboration among ISLT to align IT and business seamlessly. From this perspective, the quality of the relationship between CIO and TMT is one of the key successful factors of IS implementation. Although some concepts in literatures emerge to describe the relationship among ISLT, there are no consistent constructs. Shared knowledge (abbreviated as *SK*) and shared understanding (abbreviated as *SU*) is in the pool. Some researchers use them alternatively (Nelson and Cooprider 1996; Ray et al. 2005), while others suggest that *SK* and *SU* have essential difference (Preston and Karahanna 2009; Bassellier and Benbasat 2007). Even so, they measure the two constructs with various scales.

So then, the current study will clear the definitions, dimensions of *SK* and *SU*, propose the appropriate scales, and put forward the propositions to explore the possible relationships between *SK* and *SU*, to lay the foundation for future research. The paper proceeds as follows: it first reviews the literatures in *SK* and *SU*, and then constructs the model to explain the causal relationships. The paper ends with the conclusions.

71.2 The Definition and Measurement of *SK*

71.2.1 The Definition of SK

Boynton et al. (1994) use *managerial IT knowledge* to describe the common knowledge, which reflects the overlap of the knowledge of IT and business managers, especially the business and strategy knowledge of IT managers and the potential application of IT in the domain of business managers. Following this way, other researchers propose similar definitions. They state *SK* refers to as the inter-departmental knowledge that is exchanged, shared between the two parties.

Nelson and Cooprider (1996) define *SK* as the understanding and appreciation about the technologies and processes which influence the common performance of CIO and TMT. According to their definition, *SK* can be regarded as synergy, consisting of agreement and mutual respect. Mutual understanding and appreciation are the key components of *SK*. They suggest that deep level knowledge sharing is the necessary condition for the development of *SU*, and *SU* induces knowledge sharing in turn. Similarly, Reich and Benbasat (2000) regard *SK* as the ability of IS and business executives to understand and participate into the key process in the other domain and respect the unique contribution and challenge, which represents their relevant cognition and experience. However, Reich and Benbasat construct the two variables into the same model to explore whether *SK* influences communication, which induces *SU* in turn. Using this definition, *SK* is defined as objective knowledge (Armstrong and Sambamurthy 1999), including business knowledge of CIO and IT knowledge of TMT, by Preston and Karahanna (2009).

Ray et al. (2005) expand prior definitions and propose that *SK* means the knowledge of IT managers about customer service process and that of customer service mangers about IT using opportunity, and their common understanding about improving customer service performance with IT support. In nature, *SK* is a kind of important competence, which facilitates the change of the organization structure, business process and the improvement of process performance. So the studies following this way don't differentiate *SK* and *SU*, but also combine them into one construct. Despite all this, Ray et al. (2005) operationalize *SK* as shared domain knowledge only.

The review of the definition of *SK* indicates that although the researchers propose different description based on their research design and situation, most of the definitions focus on the knowledge and understanding of IT and business professionals about the other domain. In the case of ISLT, *SK* is composed of strategic IT knowledge and business knowledge (Armstrong and Sambamurthy 1999). Strategic IT knowledge refers to as the potential of IT infrastructure in organization, strategic IT action of the competitors, and the opportunity of emerging new IT. The IT knowledge is defined not from tactical but from strategic perspective, for the level and responsibility of CIO and TMT in the organization. Strategic business knowledge means that of business strategy, business process, the recipes for success, the strengths and weakness of the competitors. Business knowledge reflects the relationship between stakeholders and the organization, the competitive policy and the tendency of the competitors. Business knowledge of CIO represents his understanding and appraisal of the organization's competiveness, strategy and process. Those CIOs with high business knowledge are easy to understand the priority, opportunities and needs of business activities to facilitate the alignment of IT application with key business process.

71.2.2 The Dimensions of SK

Though most authors acknowledge *SK* consists of IT professional's business knowledge and business professional's IT knowledge, there are no agreement on the construct structure. Some insist that it's single dimensional (i.e., the score is the sum or multiplication of two parts), and others emphasize the two components are independent dimensions.

Boynton et al. (1994) suppose that both of the two kinds of knowledge are high level, which meet "push-pull theory". Therefore, the multiplication of the two scales which measure IT and business knowledge is regarded as the score of *SK*. Reich and Benbasat (2000) suggest that the amount of IT experience of business executives and business experience of IT executives represent the level of *SK* among ISLT. Hence, they measure IT knowledge and business knowledge of senior leadership and the results fall into three categories: high, middle and low. If both kinds of knowledge are graded as high, the level of *SK* is high; if none of them is high, the level of *SK* is low; otherwise, *SK* is middle. Ranganathan and Sethi (2002) admit the two-component structure of *SK*; however, the theoretical inference and data analysis indicate *SK* has only one dimension. Using the scales developed by

Boynton et al. (1994) to measure the knowledge of the two domains, Ray et al. (2005) get the same results according to the sum or multiplication. Obviously, they regard *SK* as a single-dimensional construct also.

Unlike the above authors, other researchers (such as Bassellier and Benbasat, Preston and Karahanna) divide *SK* into two dimensions based on its content. Bassellier (2003) suggests *SK* refers to the knowledge beyond the domain, especially that of the other domain, including IT knowledge of business professionals and business knowledge of IT professionals. In that study, Bassellier measure the two dimensions independently with the scales self-developed. IT knowledge in business people consists of IT-related knowledge and experience; business knowledge in IT people is defined as business and interpersonal knowledge and skills. Bassellier and Benbasat (2007) argue that though many studies pay attention on the effects of *SK*, few of them distinguish different type knowledge. Hence, besides of two component of *SK*, they test the influence of their own domain knowledge of IT and business professionals. Following the definition and measurement of Armstrong and Sambamurthy (1999), Preston and Karahanna (2009) treat *SK* as a two-dimensional construct, and confirm that the strategic business of CIO and strategic IT knowledge of TMT have important role during the development of *SU*.

The structure of a construct relies on its definition and analysis results of the data. In other words, clear meaning and the results of factor analysis determines the dimensions of a construct. If the items load on a few of factors which have non-crossed meaning, then we can draw conclusion that it is multidimensional; and if all the items load on only one factor, it should be regarded as a single-dimensional construct.

71.3 The Definition and Measurement of *SU*

71.3.1 The Definition of SU

In prior studies, researchers use different term to describe *SU*, including *mutual understanding*, *convergence*, *shared cognition* and *social alignment*. Generally, *SU* represent the extent of the consistency about the understanding on the role of IS in the organization between IT and business professionals. For example, Preston and Karahanna (2009) define shared knowledge among ISLT as mutual understanding about the role of IS between CIO and TMT. Lind and Zumd (1991) regard *SU* (in term of convergence) as shared understanding about the importance of business process and the support of technology for business between technology providers and users.

During the long-time cooperation, the assumption, belief and expectation of CIO and TMT change, and shared understanding develops with time going on. They receive information from the other side, and adjust their own cognition, eliminate the differences and obstacles (Lind and Zumd 1991), and reach a consensus on the application and planning of IS/IT.

As is known from the above review, *SU* describes the cognition difference between the two parties. That is to say, the smaller the difference is, the higher the consensus is; conversely, the lower is. Therefore, to some extent, *SU* is a continuum of state, not an "either/or" question (i.e., there are only two possible state: reaching consensus or not). Consensus between the two sides means that they have the same understanding of the importance of some problem. The key of the definition of *SU* should focus on the following two questions. The first is that the subjects reaching consensus involves whom; and the second is what the two parties have common understanding on. We suggest that future research may follow the definition proposed by Preston et al. and Johnson and Lederer (2005), and adapt it according to the situation.

71.3.2 The Measurement of SU

Generally, although the authors reach agreement on the definition of *SU*, they measure it with different scales. Some researchers (such as Preston and Karahanna 2009) regard it as a single-dimensional construct. Others (such as Reich and Benbasat 2000; Johnson and Lederer 2005) measure the construct with multiple independent scales.

Reich and Benbasat (2000) suggest that *SU* is composed of that on the current role and that on the future role from time framework perspective. Short-term *SU* means the understanding and commitment of business and IT executives on the short-term (1–2 year) plans and objectives of each other. Long-term *SU* is defined as the state in which business and IT executives shared visions of the ways in which IT contributes to the success of business strategy. There another study has proved that the two dimensions are distinct, for they find that, in some organizations, when one of them is high, the other dimension is low.

Referring to other literatures, Johnson and Lederer (2005) indicate that the current and future role of IS in the organization are different obviously. The former focuses on the day-to-day operation, and the latter focuses on the future strategy. The current role can't determine the future role. In other words, the support of IS on business operation wouldn't ensure the support for strategy realization. Indeed, the function and planning of current IS may hinder the objectives of business strategy. The current role has only one dimension which means the extent to which the recent dependence of the organization on IT. However, the future role has three dimensions, including managerial support, differentiation, and system enhance. Hence, they divide *SU* into *SU* on the current role and *SU* on the future role.

In a study about *SU* between business and IT section, Stoel (2006) proposes that *SU* consists of that on operation and that on strategy. *SU* on operation level refers to as the proper appraisal on business process, resource requirement, and challenge of business and IT department. *SU* on strategy level is defined as mutual knowing about the improvement objectives of business process, the strategic principles which guide investment and resource obtaining, and the ways in which IT is used to support the above goals. The former focuses on daily behavior, and the latter stress long-time orientation.

Although Preston and Karahanna (2009) regard *SU* as a single dimensional construct, Preston points out in his Ph. D dissertation that besides of shared language and shared understanding, future research may test whether shared vision and shared values are the components of shared mental model (Preston 2004). Therefore, we can draw a conclusion that most authors admit *SU* is a multidimensional construct, including short-term *SU* and long-term *SU*, which are independent.

71.4 The Relationship Between *SK* and *SU*

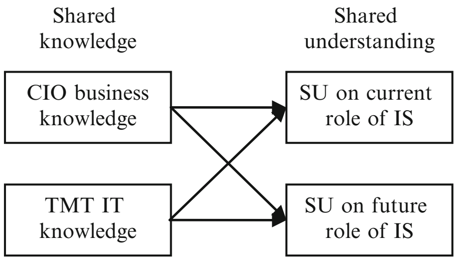
Even if some authors (Nelson and Cooprider 1996; Ray et al. 2005) don't distinguish *SK* and *SU*, other recent studies (Preston and Karahanna 2009; Bassellier and Benbasat 2007; Stoel 2006) treat them as two different constructs, measure them by diverse scales, and discuss the relationship between them.

The asymmetry of knowledge between CIO and TMT causes the cultural gap. Keen (1991) argues that the organization can't tolerate IT professionals without business knowledge, and business professionals who know few of (information) technology. If senior leaders have no rich strategic IT knowledge, they have to transfer initiative to CIO and IT department, which weakens the effectiveness of IT application. In other words, top management team should master some degree of IT knowledge, which facilitates their cognition on how to use IT to support business operation and strategy.

Explicit information resources don't necessarily bring competitive advantage to the organization; they improve organizational performance through *SU* between IT and business people. Top business managers should understand IS capability and commit to develop IS to meet business need. Matching of IT and business knowledge is the necessary condition for the organization to carry out organizational IT-based changes. *SU* between CIO and TMT correlates positively with the level of IT application and the absorbing and usage of IT in the organization (Boynton et al. 1994; Armstrong and Sambamurthy 1999).

The precondition of *SU* is that some knowledge in the overlapped knowledge is grasped by each other. The knowledge structures of the individuals are not the same extremely, but compatible for each other and helpful for the development of common expectation (Cannon-Bowers and Salas 2001). The knowledge of the individuals is the basement for *SU*. However, *SU* is not the sum of the knowledge of the individuals, but the outcome of adjustment, creation and integration. Hence, the strategic business knowledge of CIO and the strategic IT knowledge of TMT are the essential condition for the development of *SU* on the role of IS in the organization. CIO with high level of business knowledge can understand the strategy and objectives of the organization, and put forward IS-related advices to top managers. On the other side, TMT with sufficient IT knowledge can understand CIO's proposal and the role of IS in the organization.

Fig. 71.1 The relationships between *SK* and *SU*



Henderson (1990) suggests that the cooperation between IT and business department doesn't depend on *translation*, but relies on the work experience, and the understanding and appraisal of the knowledge in the other domain; *SK* plays important role during the establishment of cooperation and joint decision making. IT experiences of CEO and business orientation of CIO are the key factors that influence the cooperation. Accumulated knowledge among ISLT determines the extent of alignment of business and IT. Bassellier (2003) finds that IT knowledge of business professionals affects their cognition, involvement and enthusiasm about IT; IT professionals with business knowledge know how to apply IT to ensure the business objectives of the organization. Stoel (2006) indicates that *SK* between IT and business department has significant effect on shared operational understanding, and IT knowledge of business section is more important. Bassellier (2003) propose that, besides the knowledge of their own domain, the knowledge of the other domain (i.e., *SK*) influences the development of *SU* between IT and business professionals. *SK* encourages CIO and TMT to cooperate, act coherently, and share objectives, risk and responsibility. *SK* induces *SU*, which facilitates alignment of business and IT in turn. The relationships between *SK* and *SU* are illustrated in Fig. 71.1.

Based on the above, we put forward the following propositions:

P1 Business knowledge of CIO affects *SU* on the current role of IS positively.
P2 Business knowledge of CIO affects *SU* on the future role of IS positively.
P3 IT knowledge of TMT influences *SU* on the current role of IS positively.
P4 IT knowledge of TMT influences *SU* on the future role of IS positively.

71.5 Conclusions

Based on the above review and analysis, it's easy to be found that a few of authors confuse *SK* and *SU* because of their unclear and inconsistent definition and measurement. On the other side, as some researchers suggest, the two concepts are distinct constructs, which have different meaning and content, and *SK* affects *SU* positively in logic. Based on the structure and dimensions analysis, this paper put forward four propositions about the effects of *SK* on *SU* at the dimension level, which will guide the approaching theoretical and empirical research. The future studies should collect the data in various organizations in China and other countries with different culture and development level, to test the reliability and validity of the model.

Acknowledgment R.J. Lv, Z. Zhao, X.C. Chen and L.J. He expressed sincere thanks to Professor Y.F. Wang (Hebei University of Technology) for her research advices and all the authors in reference list.

References

Armstrong CP, Sambamurthy V (1999) Information technology assimilation in firms: the influence of senior leadership and IT infrastructure. Inf Syst Res 10(4):304–342

Bassellier G (2003) Knowledge and partnership for information technology and business people. Ph.D. dissertation, The University of British Columbia, Ottawa

Bassellier G, Benbasat I (2007) Assessing the contribution of business and IT knowledge to the development of IT/Business partnership. In: Proceedings of the 2007 ACM special interest group on management information system conference on computer personal research, Missouri, pp 203–206

Boynton AC, Zmud RW, Jacobs GC (1994) The influence of IT management practice on IT use in large organizations. MIS Quart 18(3):299–318

Cannon-Bowers JA, Salas E (2001) Reflections on shared cognition. J Organ Behav 22(2):195–202

Henderson JC (Spring 1990) Plugging into strategic partnership: the critical IS connection. Sloan Manag Rev 31(3):7–18

Johnson AM, Lederer AL (Fall 2005) The effect of communication frequency and channel richness on the convergence between Chief Executive and Chief Information Officers. J Manag Inf Syst 22(2):227–252

Keen PG (1991) Shaping the future: business design through information technology. Harvard Business School Press, Boston

Lind M, Zumd RW (1991) The influence of a convergence in understanding between technology providers and users on information technology innovativeness. Organ Sci 2(2):195–217

Nelson KM, Cooprider JG (1996) The contribution of shared knowledge to IS group performance. MIS Quart 20(4):409–432

Preston DS (2004) Shared mental models between the Chief Information Officer and Top Management Team: towards information systems strategic alignment. Ph.D. dissertation, The University of Georgia, Athens

Preston DS, Karahanna E (2009) Antecedents of IS strategic alignment: a nomological network. Inf Syst Res 20(2):159–179

Ranganathan C, Sethi V (2002) Rationality in strategic information technology decisions: the impact of shared domain knowledge and IT unit structure. Decis Sci 33(1):59–86

Ray G, Muhanna WA, Barney JB (2005) Information technology and the performance of the customer service process: a resource-based investigation. MIS Quart 29(4):625–651

Reich BH, Benbasat I (2000) Factors that influence the social dimension of alignment between business and information technology objectives. MIS Quart 24(1):81–113

Stoel MD (2006) The antecedents and consequences of shared business-IT understanding: an empirical investigation. Ph.D. dissertation, The Ohio State University, Columbus

Chapter 72
The Effect of Environmental Awareness on Consumers' Green Purchasing: Mediating Role of Green Perceived Value

Jian-xin Li, Dan Liu, and Li Liu

Abstract According to the existing academic research, it appears that consumer's environmental awareness do not keep in line with green purchasing behavior. In this paper, the author tried to demonstrate the relationship between environmental awareness impacts and green purchasing behavior. Green perceived value was added to the regression model, finding that the impact of environmental awareness on green purchasing behavior was mediated by green perceived value. Meanwhile, it is known that green perceived value consists of healthy value, emotional value, image value and environmental value, which is different from traditional products' perceived value.

Keywords Environmental awareness • Green perceived value • Green product • Green purchasing behavior

72.1 Introduction

Green purchasing is a process of consideration on product's environmental attributes or characteristics referring to the environmental friendly or green product during the purchase activity (Mainieri et al. 1997). Green purchasing is an important way for consumer to practice environmental responsibility. Getting aware of the factors which affect consumer's green purchasing behavior would help government to formulate appropriate policy to guide green consumption, help enterprises to launch effective green marketing.

In the research of green consuming, environmental awareness has been regarded as a significant factor which influences consumer's environmental behavior (green

J. Li (✉) • D. Liu • L. Liu
School of Management and Economics, Changsha University of Science and Technology, Changsha, China
e-mail: jxlee01@163.com; ldld0602@163.com; 664172132@qq.com

E. Qi et al. (eds.), *The 19th International Conference on Industrial Engineering and Engineering Management*, DOI 10.1007/978-3-642-37270-4_72,

purchase behavior) (Bamberg and Moser 2007). However from a practical perspective, consumer's environmental awareness does not keep in line with environmental behavior. From a theoretical perspective, the relationship between environmental awareness and green purchasing is not as simple as what the theory assumes. The relationship between environmental awareness and green purchasing could be roughly calculated into three categories. Firstly, environmental awareness is a predictor to green purchasing, and its explanatory power to green purchasing behavior is much high (Schlegelmilch et al. 1996). Secondly, environmental awareness is of low explanatory power to green purchasing, in Roberts' study (1996), environmental awareness can only explain green consumer behavior (mainly measuring the green product purchases) by 5 % of the variance. Later, Straughan & Roberts found that environmental awareness can only explain green consumer behavior by 1.5 % of the variance. Thirdly, environmental awareness and green purchase behavior is lack of correlation (Minton and Rose 1997). Within the understanding of low correlation between environmental awareness and green purchasing, scholars have proposed two solutions: First one, don't use environmental awareness (as a general attitude) but specific attitude to predict specific behavior directly (Bamberg 2003). This method is just in line with matching principle proposed by Ajzen and Fishbein. Another approach is to consider other variables and (or) situational factors which would adjust the relationship (Stern 1992).

Perceived value is an important driver to consumer purchasing behavior. Green perceived value, on which some scholars have made preliminary explorations, is lack of depth in diversity and content. Based on the general theory of perceived value, this study is about to explore the specific content of green perceived value and its impact on green purchasing, besides, to research whether the green perceived value has an effect on the relationship between environmental awareness and green purchasing behavior or not.

72.2 Literature Review and Research Hypothesis

72.2.1 Environmental Awareness and Green Purchasing Behavior

The concept of environmental awareness first appeared in 1968, in the U.S., the scholar Roth proposed "environmental literacy", which is the prototype of environmental awareness. In later studies, due to each scholar's personal background, the concept of environmental awareness gradually expanded. Generalized environmental awareness has a broad coverage, including relevant aspects such as geography, ecology, philosophy, environics, ethics, law, political science and other disciplines. Some scholars believe that the basic meaning of environmental awareness includes natural concepts and values, environmental science and technology, environmental law and policy, environmental ethics and environmental psychology, there are significant interaction within these five territories; Hong claims environmental awareness should include environmental knowledge, environmental value, environmental attitude and environmental behavior. What these definitions have in common is that people's environmental awareness consists of not only cognitive independence of environment, but also the conscious of protecting the environment. However, with further research, some scholars proposed that environment behavior should act as an independent variable distinguished from environmental awareness, or the influence of environmental awareness on environmental behavior should have been settled in advance, therefore it would be difficult to explore the extent, to which people would transform environmental awareness into specific environmental actions.

From the narrow sense, environmental awareness refers to cognition about nature and value of environment-related human behavior, it's a new independent ideology, a progressive value, and a sum of social idea, theory, emotion, willingness, consciousness and other concepts which reflects relationship between human and environment. The definition specifies and refines environmental awareness to environmental value. It's easy to control in the exploration. This article uses the narrow definition of environmental awareness, which is mainly about perceived environmental values or "supportive behavioral tendencies towards environmental protection and environmental matters."

The purpose of studying environmental awareness is mainly focusing on predicting and guiding the people's environmental behavior. Hines et al. indicated that the average correlation coefficient of environmental awareness and environmental behavior is 0.35(Hines et al. 1986). Consumer's environmental awareness directly affects his consideration about products and significantly predicts his environmental behavior. Consumer who has strong environmental awareness will consider environment-friendly product or service in priority. Environmental awareness can affect consumer's green purchasing behavior has been proved by Peattie and Goose (2001). Therefore, this article made the following assumption:

H1 Consumer's environmental awareness has a positive effect on green purchasing behavior.

72.2.2 Environmental Awareness and Green Perceived Value

Zeithaml's definition of customer perceived value(CPV) is generally accepted, he proposed that customer perceived value is consumer's overall assessment based on balancing perceived interests in obtaining products or services against the cost of products or services(Zeithaml 1988). As for dimensions of CPV, Sheth, Newman and Gross proposed the theory of market value and consumer choice, in which CPV includes five dimensions: social value, emotional value, functional value, cognitive value and context value, customer's choice is a function of multi-dimensional perceived value; these values work in differential degrees in different purchasing situation (Sheth et al. 1991). Based on Sheth, Sweeny and Soutar designed a set of scale for durable goods to measure CPV, in the research; CPV consists of the following dimensions: quality value, price value, emotion value and social value (Sweeney and Soutar 2001). Afterwards, dimensions of CPV has been enriching, "perceived payment", "potential risks", "perceived happiness", "green perceived value" were gradually raised. With growing popularity of green product, Ottman suggested green product would include the following five kinds of perceived value: efficiency and cost effectiveness, health and safety, performance, symbolic and status, convenient (Ottman et al. 2006). These dimensions make green products different from ordinary ones and show more green attributes, they represent CPV of green products exactly. His notion has not been empirically testified yet, but it gives us enlightenment. According to Externality Theory of Economics, green products can be divided into self-interest products and altruistic products, which mean green perceived value, may consist of self-interest value and altruistic value. On that base, combining the thoughts of Sheth and Ottman, this study tries to summarize a more comprehensive content of green products' green perceived value.

Environmental awareness based on the values is a general attitude, however, green perceived value is a special, specific psychological perception for green products, is a positive attitude towards specific green product. Value affects individual's inner mechanism, and influences how specific attitude and behavior form. Ecological value affects the formation of Ecology-Center environmental attitude directly or indirectly. Fulton indicated in their cognitive hierarchy model of human behavior that the value in the bottom was the basic of the upper variables (Fulton et al. 1996). Therefore, this article made the following assumption:

H2 Environmental awareness has a positive effect on green perceived value.

72.2.3 Environmental Awareness, Green Perceived Value and Green Purchasing

Zeithaml found that high quality do not lead to purchasing behavior directly, CPV would affect relationship between quality and customer behavior as an intermediary variable. This conclusion has been supported by Pan Yu, who claimed that CPV would impact on consumer behavior directly, and serve as intermediary variable to explain the relationship between traditional Chinese value and consumer behavior. Environmental awareness has a subtle influence on consumers' green consuming tendency, this influence is restricted by consumer's perceived actual efficiency, which means if his own words or actions can help to protect environment effectively. CPV is subjective value-estimated activity in consumer's mind, and green perceived value is a positive judgment for green product. Based on values – attitude – behavior theory, green perceived value, as an intermediary variable, can stimulate environmental awareness transform into green purchase motivation. Therefore, this article made the following assumption:

H3 Environmental awareness has a positive effect on green purchasing behavior, mediated by green perceived value.

72.3 Research Methods

Adopting 7-point Likert-type scale, this study involves three variables: environmental awareness, green perceived value and green purchasing behavior. Eight items of environmental awareness refers to Hong's questionnaire; the measurement of green purchasing behavior refers to Roberts and Bacon (1997), including four items; the measurement of green perceived value is based on Sheth's model, referred to the point of Ottman, combining qualitative with quantitative methods to form suitable items as many as possible from self-interest and altruistic aspects, as a consequence, adjusted by professors of business administration, finally consists of 15 items. In this study, convenient sampling method is adopted, the researcher had published 400 questionnaires online, and got 349 valid back. SPSS16.0 was used for statistical analysis.

72.4 Data Analysis

72.4.1 Reliability and Validity Analysis

The Cronbach's Alpha scores of environmental awareness, green perceived value and green purchasing behavior were 0.880, 0.954, and 0.804. In factor analysis, the KMO values of these variables were 0.876, 0.932, 0.772, indicating that the questionnaire is of good reliability and validity. Table 72.1 shows the statements' factor loadings on each of the four factors of green perceived value. The four factors "explained" 82.538 % of the variance, and separately named "healthy value", "emotional value", "image value" and "environmental value". "Healthy value" is defined as perceived benefits on consumer's health and safety when they choose or buy green product, "emotional value" is defined as that consumers would feel very pleased and proud of themselves psychologically because their purchasing behaviors not only meet their own demands, but also contribute to the environmental protection. "Image value" means that consumers look forward to establish a positive or socially-responsible image and get recognition and praise when they purchase or use green products. These three factors are self-interest type perceived value. "Environmental value" implies that consumers wish they would contribute to society and environmental protection directly and effectively through purchasing green products. This is an altruistic-type value.

Table 72.1 Factor analysis of green perceived value

Items	Factor loadings Healthy value	Emotional value	Image value	Environmental value
Using green products can help to improve ecological environment				0.883
Using green products can help to reduce the pollution to environment				0.837
Using green products can drive the others doing as me				0.552
Using green products can help me own a good image			0.842	
Using green products can help me earn lots of praise			0.874	
Using green products can help me build a pro-environment self-image			0.745	
Using green products makes me relaxed		0.814		
Using green products makes me feel good		0.811		
Using green products lets me in a good mood		0.803		
Using green products gives me a feeling of harmony with nature		0.716		
Green products contains less ingredients harmful to human	0.648			
Using green products can secure our healthy and safety	0.738			
Using green products is a guarantee of the high quality of life	0.803			
Green products are much healthier compared to traditional products	0.757			
To buy green products is in order to rest assured	0.733			
Standard Cronbach Alpha coefficients	0.918	0.954	0.902	0.876
Each factor explained variance %	24.008	22.313	18.873	17.344
Explained variance accumulation %	24.008	46.321	65.194	82.538

72.4.2 Correlation Analysis of Main Variables

Correlation analysis is used to examine the relationships among environmental awareness, green perceived value and green purchasing behavior. Pearson correlation values of environmental awareness and green perceived value, environmental awareness and green purchasing behavior, green perceived value and green purchasing behavior were 0.530, 0.545, and 0.638, which means the three variables

Table 72.2 Correlation analysis of main variables (N = 349)

	Environmental awareness	Healthy value	Emotional value	Image value	Environmental value
Environmental awareness	1				
Healthy value	0.499**	1			
Emotional value	0.458**	0.749**	1		
Image value	0.347**	0.613**	0.645**	1	
Environmental value	0.520**	0.699**	0.681**	0.553**	1
Green purchasing	0.545**	0.618**	0.603**	0.444**	0.504**

**Correlation is significant at the 0.01 level (2-tailed)

positively correlated. Table 72.2 also shows that every dimension of green perceived value positively correlates to environmental awareness and green purchasing behavior.

72.4.3 Regression Analysis of Environmental Awareness, Green Perceived Value, Green Purchasing Behavior

In order to explore further relationship among environmental awareness, green perceived value, green purchasing behavior, this paper takes environmental awareness, four kinds of green perceived values (healthy value, emotional value, image value, environmental value) as independent variables, and takes the green purchasing behavior as the dependent variable to conduct a multiple regression analysis. In Table 72.3, environmental awareness significantly impacts on green perceived value, and explains green perceived value by 27.9 % variance; environmental awareness significantly impacts on green purchasing behavior, and explains green purchasing behavior by 29.5 % variance. Four kinds of green perceived value significantly impact on green purchasing, and these could explain green purchasing by 42.1 % variance, healthy value mainly affects green purchasing behavior, its β coefficient is 0.358, followed by emotional value, β coefficients of image value and environmental value are not significant. In conclusion, environmental awareness would predict green perceived value, green purchasing behavior significantly, otherwise green perceived value can also significantly predict green purchasing behavior.

72.4.4 Analysis of Mediate Effect of Green Perceived Value

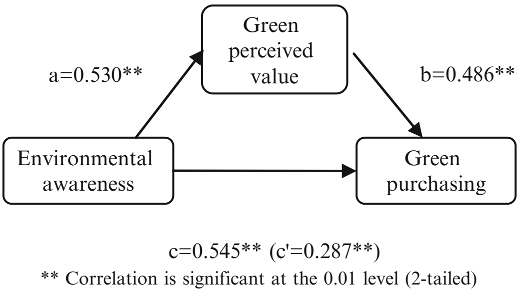
Hierarchical regression analysis was used to examine the intermediary role of green perceived value. Firstly, independent variable environmental awareness and dependent variable green purchasing behavior were put into a regression model, $\beta = 0.545$, $t = 12.108$, $p < 0.01$, the main effect is $c = 0.545^{**}$, Secondly,

Table 72.3 Regression analysis (N = 349)

Variables	Green perceived value (β)	Green purchasing (β)	Variables	Green purchasing (β)
Environmental awareness	0.530**	0.545**	Healthy value	0.358**
			Emotional value	0.300**
			Image value	0.006
			Environmental value	0.046
F	135.86**	140.61**	F	64.325**
R^2	0.281	0.297	R^2	0.428
Adjusted R^2	0.279	0.295	Adjusted R^2	0.421

**Correlation is significant at the 0.01 level (2-tailed)

Fig. 72.1 Causality



independent variable environmental awareness and intermediary variable perceived green value were put into a regression model, $\beta = 0.530$, $t = 11.656$, $p < 0.01$, the effect from independent variable to intermediary variable is $a = 0.530^{**}$. Thirdly, independent variable environmental awareness, intermediary variables green perceived value and dependent variable green purchasing behavior were put into a regression model, the effect from intermediate variables to dependent variable is $b = 0.486^{**}$. Adding intermediate, the effect from independent variable to dependent variable is $c' = 0.287^{**}$. All the results implied that the relationship between environmental awareness and green purchase behavior has significantly changed after adding green perceived value into the model. We can draw the causality diagram like Fig. 72.1, according to Wen Zhonglin's point, the mediating effect of green perceived value is partial mediation, $ab = 0.258$, 47.26 % of total effect C.

72.5 Summary

72.5.1 Research Conclusions

The result shows that: (1) Consumer's environmental awareness has directly significant impacted on green perceived value and green purchasing behavior. Consumers with strong environmental awareness would consider highly of green products, easily perceive green products differing from the ordinary ones, as a follow, generating a buying preference. (2) Green products' dimensions of perceived value are different from traditional products. It is not accurate to simply add "green value" into traditional model. Individuals pay more attention to green products' healthy value, emotional value and image value which can benefit themselves and environmental value which is altruistic. Healthy value and emotional value are the most important to consumers. They are of predictive ability towards green purchasing behavior. (3) Green perceived value is the mediating variable of the relationship between environmental awareness and green purchasing behavior. Environmental awareness not only has impact on green purchasing behavior directly, but also affects the green purchasing behavior through green perceived value. This is in line with consumers' "values – attitude – behavior" model.

There are two paths from environmental awareness to green purchasing behavior. Environmental awareness can partially promote consumers' green purchasing, if the characteristics of green products are not distinct, consumers' green perceived value would not be obvious, therefore, the ordinary environmental awareness in general level is not deep enough to drive consumers towards green purchasing behavior.

72.5.2 Practice Enlightenment

The above conclusions obtain the following enlightenments: Firstly, if government and business want consumers to choose green products, it is essential to help consumers increase awareness about ecological environment, and to guide consumers to form a sustainable development and responsible environmental values. Secondly, the government should improve the publicity. We can neither copy perceived value of traditional products nor emphasize too much on environmental effectiveness of green products. We should address both self-interest and altruistic values of green products.

72.5.3 Limitations and Future Research Directions

This article has theoretical and practical significances, but there are still some shortcomings. Firstly, owing to the limited budget, this study did not fully authorize professional questionnaire website to collect data, which may affect the representative of the conclusions. Secondly, this article does not focus on green perceived value: the exploration is not deep enough. In the future, we should increase the objectivity and representative of sample in order to ensure the quality of data. Otherwise, we could focus on the green perceived value and the mechanism of it. In addition, according to the study purpose and need, we could add new variables to enrich the research framework.

Acknowledgment The authors conveyed their thanks to a grant from the Open Fund Project of Key Research Institute of Philosophies and Social Sciences in Hunan Universities, the Fund Project of Humanities and Social Science from Ministry of Education, P.R.C (09YJA630141), and the Fund Project of Hunan Soft Science (10ZK033).

References

Bamberg S (2003) How does environmental concern influence specific environmentally related behaviors? A new answer to an old question. J Environ Psychol 23:21–32

Bamberg S, Moser G (2007) Twenty years after Hines, Hungerford, and Tomera: a new meta-analysis of psycho-social determinants of pro-environmental behavior. J Environ Psychol 27(1):14–25

Fulton DC et al (1996) Wildlife value orientation: a conceptual and measurement approach. Hum Dimens Wildl 1(2):34–47

Hines JM, Hungerford HR, Tomera AN (1986) Analysis and synthesis of research on responsible environmental behavior: a meta-analysis. J Environ Educ 18:1–8

Mainieri T et al (1997) Green buying: the influence of environmental concern on consumer behavior. J Soc Psychol 137(2):189–204

Minton AP, Rose R (1997) The effects of environmental concern on environmentally friendly consumer behavior: an exploratory. J Bus Res 40:37–48

Ottman JA, Stafford ER, Hartman CL (2006) Avoiding green marketing myopia. Environment 48(5):103–115

Peattie K, Goose GOGW (2001) The hunt for the green consumer. Bus Strat Environ 10(4):18–71

Roberts JA (1996) Green consumers in the 1990s: profile and implications for advertising. J Bus Res 36(3):217–231

Roberts JA, Bacon DR (1997) Exploring the subtle relationships between environmental concern and ecologically conscious consumer behavior. J Bus Res 40:79–89

Schlegelmilch BB, Bohlen GM, Diamantopoulos A (1996) The link between green purchasing decisions and measures of environmental consciousness. Eur J Mark 30(5):35–55

Sheth JN, Newman BI, Gross BL (199l) Why we buy what we buy: a theory of consumption values. J Bus Res 22(2):54–72

Stern PC (1992) What psychology knows about energy conservation. Am Psychol 47(10):1224–1232

Sweeney C, Soutar GN (2001) Consumer perceived value: the development of a multiple item scale. J Retail 77:203–220

Zeithaml VA (1988) Consumer perceptions of price, quality, and value: a means—end model and synthesis of evidence, J Mark 52:23–46

Chapter 73
Developing an Analysis Model for Public Participation in Sustainable Construction

Qiu-bo Yang, Xue-qing Wang, and Yuan Ren

Abstract Sustainable construction is generally regarded as the specific application of the thought of sustainable development in construction business. This study specifically demonstrates the importance of public participation towards sustainable construction, and establishes an analysis framework of public participation in sustainable construction. The goal of this research is to better understand the relationship between public participation and sustainable construction. Some index like public participation index (PPI) and sustainable development Ability (SDA) is introduced to measure the sustainability and public participation in construction project.

Keywords Behavioral framework • Public participation • Sustainable construction • Sustainable development

73.1 Introduction

Awareness of the importance of sustainable development has been growing around the globe for the last few decades, especially after the Copenhagen Climate Conference in 2009. Sustainable construction is one of the applications of sustainable development practices in construction industry. Construction industry consumes approximately 40 % of total energy, and should be responsible for some 30 % of CO_2 emissions. This industry also produce vast amounts of waste, for example, in

Q. Yang (✉)
School of Education, Tianjin University, Tianjin, China

College of Management and Economics, Tianjin University, Tianjin, China
e-mail: yqb@tju.edu.cn

X. Wang • Y. Ren
College of Management and Economics, Tianjin University, Tianjin, China

E. Qi et al. (eds.), *The 19th International Conference on Industrial Engineering and Engineering Management*, DOI 10.1007/978-3-642-37270-4_73,

European Union, it generates approximately 40 % of all man-made waste (Bourdeau 1999). The new discipline of 'sustainable construction' is addressed in the First International Conference on Sustainable Construction in Tampa, Florida, in 1994, which has been dubbed as 'green construction'. Sustainable construction adopted the principle of sustainability and relates it to building and construction activities. Originally the term 'sustainable construction' was proposed to describe the responsibility of the construction industry in attaining 'sustainability'. Sustainable construction was tentatively defined as 'creating a healthy built environment using resource-efficient, ecologically based principles' by Kibert in 1994. Generally it was used to describe a process which started well before construction in the planning and design stages, and last to the moment that the construction team left the site. It also included managing the serviceability of a building during its lifetime and extended to its eventual deconstruction and the recycling of resources to reduce the waste stream usually associated with demolition (Wyatt 1994). The emphasis was on how to reduce their impact to natural environment from the beginning. A decade ago, the emphasis was focused on more technical issues in construction such as materials, technologies, building components and on energy related to design concepts. Nowadays, however, the emphasis has been shifted onto economic, social, cultural issues, etc. (Bourdeau 1999). Sustainable construction has different approaches and priorities in different countries.

Sustainable construction addresses the responsibility of construction industry in attaining sustainable development, which is based on the mutual coordination and joint progress of society, economics, human beings, resources and environment. Its principle of justice, sustainability and commonality requires active public participation in public issues. The World Bank Participation Sourcebook follows the definition of participation adopted by the Bank's Learning Group on Participatory Development, participation is a process through which stakeholders influence and share control over development initiatives and the decisions and resources which affect them (The World Bank 1996). James L. Creighton (2005) argues that, public participation is the process by which public concerns, needs, and values are incorporated into government and corporate decision making.

This study specifically demonstrates the importance of public participation towards sustainable construction, and an analysis framework of public participation in sustainable construction is addressed. The goal of this research is to better understand the relationship between public participation and sustainable construction.

73.2 Towards Sustainable Construction Through Public Participation

Public participation consists of different mechanisms. Public can participate in the following ways: (1) before authorities making decisions, giving advice, participating in consultations, and promoting projects; (2) after decisions being made, acting in or controlling the execution of them. Public participation makes government actions transparent, effectively avoids corrupt behaviors, and clarifies responsibilities and

Table 73.1 SWOT analysis of public participation in sustainable construction

Strength	Weakness
1. To acquire the technical experience from the public and other channels; 2. To attain the local experience that experts do not know; 3. To encourage the multi-angled observation to study on problems rather than make decisions recklessly 4. To promote the public to better understand the system; 5. To utilize the passion and enthusiasm of public; 6. To ensure a better evaluation of problems.	1. The lack of resources such as time, funds, staff, etc.; 2. The insufficiency of the legal framework; 3. The lack of participation consciousness and Experience; 5. The difficulty of getting the related information; 6. The lack of technical support; 7. The results may not been fully valued; 8. The lack of enough public participation (the scope of public participation is not enough).
Opportunity	**Threats**
1. To establish trust and ability; 2. To improve the environment form groups and avoid the waste of resources; 3. To empower people through conducting dialogues and promoting development; 4. To promote understanding through solving problems together; 5. To avoid conflict through the early public intervention; 6. To save the time of the whole decision-making process by reducing the objection.	1. The public possibly mistake participation process for merely a form, rather than a substantive link; 2. Only a few strong groups dominate the public meetings; 3. Not enough time for discussing and deciding the scheme; 4. If the short-term interests are superior to the long-term interests, the long-term interests will not be recognised; 5. Environmental impacting assessment report is not in a high quality or does not covers all the problems.

facilitates the eventual application of sanctions as well, in case of reproachable conducts. The public participation is commonly recognized to implement sustainability within projects including the promotion of equity and fairness; distribution of power; empowerment and capacity building; integration of stakeholder knowledge; better understanding of contextual issues; greater commitment to project goals; as well as enhanced process legitimacy through transparency and credibility of the decision-making.

Participation itself cannot guarantee sustainability or sustainable outcomes. Yet, it does provide a basis to understand the expectations of different stakeholders, which is a necessary precursor to reconciling the diverse range of ideas and expectations that surround projects. Ultimately, all developments involving multiple stakeholders are compromises. Participation can provide legitimacy to those compromises by making the decision-making process transparent, particularly in the areas of equity and fairness.

Based on the analysis frame addressed at the Aarhus Convention implementation forum, we would like to illustrate the necessity of public participation in sustainable construction by SWOT analysis (As in Table 73.1).

The function of public participation in sustainable construction is discussed as follow:

1. Public participation is the continuous motivation of sustainable construction. The utmost goal of sustainable construction is to improve living condition, accelerate comprehensive development of human beings through improving the building environment. Public interest is the starting point and the ending result.

The transformation from material-oriented to humanity-oriented idea in construction industry asks for public participation. Traditional economic theories take material-oriented economics as their theoretical frame, and use the relationship of substantial resources and commodity to explain the economic phenomenon and the law of production and reproduction of substantial resources. The new frame emphasizes the promises to humanity, and values human as the most precious resources. The human-oriented principle demands that relevant companies have the right to learn about the facts, participate, supervise and make decisions, in order to guarantee their interests, especially in terms of environmental aspect. Environmental protection, special public goods, has extreme externality. The initial motive of international environmental protection is public participation. The public welfare litigation system, generally used in law of environment in western countries, bestows the right of litigation to national judicial department upon any citizen, social organization, and national department for the sake of social public welfare.

2. Public participation could improve the skill of decision and execution in construction effectively. In the design stage of environment influence, public participation can identify more comprehensive stakeholders and make costs and profits of the project clear. Also, it's good for clarifying the targets and requirements of the project or plan and absorbing the experience and knowledge of local public, which can cover a lot of experts' knowledge, thus improving the decision quality. The public participation can establish consensus to reduce or avoid the conflicts among different stakeholders, offering the public a sense of participation and provide the implementation of project with appropriate resources, thus improving the execution skill.

3. Public participation in sustainable construction can increase the public satisfaction through bestowed rights. By the means of public participation in all stages of construction, including decision of plan, implementation of design, supervision and assessment of whole process and benign interaction, public can understand and respect the end product, thus increasing the satisfaction. Moreover, it improves the awareness of democracy and participation.

4. It is the best practices in sustainable construction that demands for public participation. A series of conventions about best practices emerged for sake of promoting the application of sustainable construction. These conventions demand for public participation, mainly covering Ceres principles, Hannover Principles and Bellagio Principles.

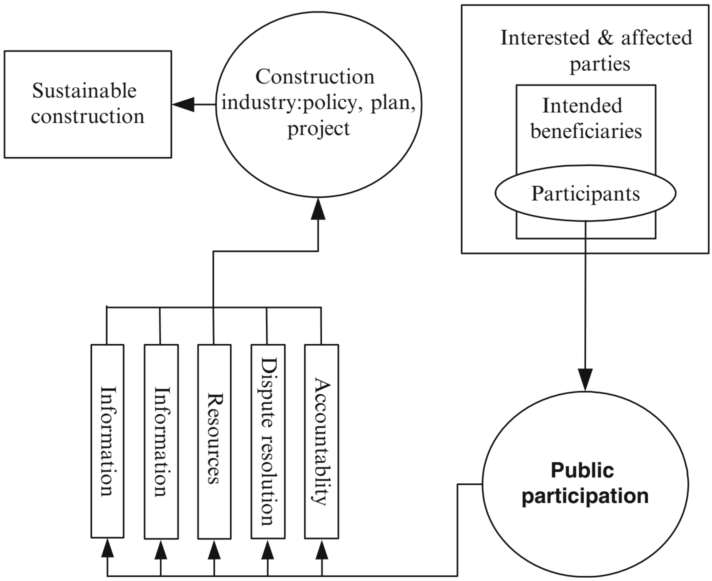


Fig. 73.1 An analysis framework of public participation in sustainable construction

Based on the above discussion, referring to the benefits of public participation from Jan Johannes Perold (2006), an analysis framework of public participation in sustainable construction could be designed as Fig. 73.1.

73.3 Identification of the Public in Sustainable Construction

The subject of public participation means who should be involved in sustainable construction. It belongs to the scope of stakeholders. Freeman (1984) analyzed stakeholders from three dimensions—ownership, economic dependence and social interest. Currently, the most widely used definition of stakeholders is score based approach by Mitchell in 1997, in which stakeholders is divided into three categories—determined, expected, and potential (Mitchell et al. 1997). Rietbergen-Mc Cracken and Narayan-Parker (1998) discussed five key questions to confine the stakeholders. Integrated former studies on public participation, the subject of public participation in sustainable construction should be one of the organizations, groups or individuals that: (1) directly or indirectly affected by construction project; (2) providing important knowledge or information for construction project; (3) affecting the implementation of the construction project; (4) interested in the construction project. Thus, the definition and relationship of public in sustainable construction is proposed as in Fig. 73.2.

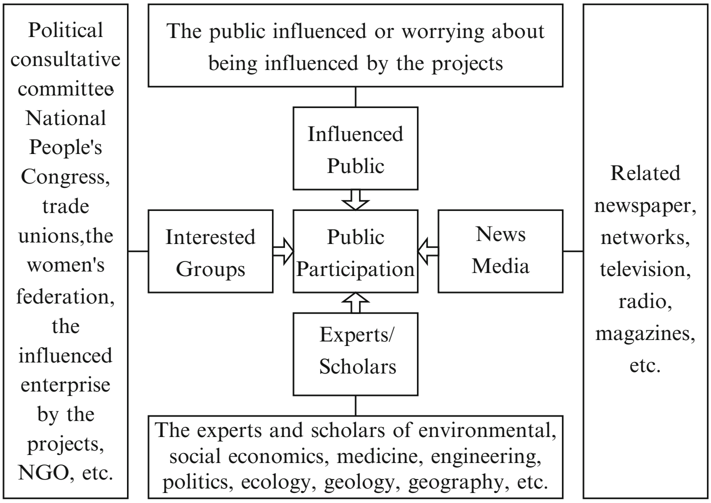


Fig. 73.2 Definition and relationship of the public in sustainable construction

73.4 Measurement of Sustainability in Construction Project

According to the definition of sustainable development by World Committee of environment (WCED) in 1987, the sustainable development ability of construction project involves economic development ability, social development ability and environmental development ability. Referring to L. Y. Shen, M. Wu, and J. Y. Wang (2002), the value of sustainable development of construction project can be illustrated as following:

$$SDV = f\ (E, S, E_n) \tag{73.1}$$

Where: SDV (Sustainable Development Value) = the value of construction project which is sustainable development oriented;
E = the economic value of construction project which is sustainable development oriented;
S = the social value of construction project which is sustainable development oriented;
En = the environmental value of construction project which is sustainable development oriented.

For the reason that E, S, En is all the function of time (t), the Formula (73.1) could be transformed into:

$$SDV = f\ [E(t), S(t), E_n(t)] \tag{73.2}$$

In individual phases of the whole life-circle of the construction project, *SDV* may be negative; however, it is absolutely positive in the whole cycle. Suppose the sustainable development ability of whole life-circle of the project as *SDA* (Sustainable Development Ability), then:

$$SDA = \oint_{\Omega} SDV \, d\delta = \int_{0}^{Tend} f\,[E(t), S(t), E_n(t)]\, dt \tag{73.3}$$

Where: Ω indicates the whole life cycle of the project; ***Tend*** indicates the end of life cycle or the rejection phase of the project. As to the measurement of *E, S, Sn*, it's possible to use Analytical Hierarchy Process to establish indicator system, and set up each weight in order to get their values. Commodities and service efficiency, allocation of resources consumption and some other indicators can be taken into the consideration of economic development ability of construction project. Demands, preferences, culture, population, politics, justice, life quality and other indicators can be taken into the consideration of social development of construction project. The consideration of environmental development involves renewable resources, biodiversity, assimilation capacity, ecological resilience capacity, etc. (Gibberd 2003).

73.5 Measurement of Public Participation Index in Construction Project

PPI (*Public Participation Index*) means the scope and depth of public participation in all stages of construction project. The scope of public participation indicates whether the affected and interested public participates or not, which is noted as *PPIs*; the depth of public participation means the empowerment to the public, which is noted as PPI_d. Therefore,

$$PPI = f\,(PPI_s, PPI_d) \tag{73.4}$$

73.5.1 The Measurement of PPIs

PPIs encompass the individual scope of public participation (PPI_{sd}) and organizational scope of public participation (PPI_{so}). The weight each are noted as K_d and K_o. Then:

$$K_d + K_o = 1 \tag{73.5}$$

Referring to the definition method of *Gini coefficient* in economics, PPI_{sd} could be illustrated as:

$$PPI_{sd} = \sum_{i=1}^{n} W_i Y_i + 2\sum_{i=1}^{n-1} W_i\,(1 - V_i) - 1 \tag{73.6}$$

Where: W_i = the ratio of the number of the affected and interested individual to the whole population;

Y_i = the ratio of the number of participants in the whole the affected and interested individual;

$$V_i = \sum_{i=1}^{i} Y_i \tag{73.7}$$

Suppose the number of the affected and interested organization as Q_i and the actual number of it as Q_a, then:

$$PPI_{so} = Q_a/Q_i{}^{*}100\,\% \tag{73.8}$$

Therefore:

$$\begin{aligned} PPI_s &= K_d \cdot PPI_{sd} + K_o \cdot PPI_{so} \\ &= K_d \left[\sum_{i=1}^{n} W_i Y_i + 2\sum_{i=1}^{n-1} W_i\,(1 - V_i) - 1 \right] + K_o Q_a/Q_i{}^{*}100\,\% \end{aligned} \tag{73.9}$$

73.5.2 The Measurement of PPI_d

Referring to Pretty (1995), divide the depth of public participation into seven categories, such as Passive participation, participation in information- giving, Participation by consultation, Participation for material incentives, Functional participation, Interactive participation, and Self-mobilization, and noted each weight as $W_i, i = (1, 2, 3, \ldots 7)$ is the evaluation to the model j as Pretty, $0 \leq W_i \leq 1, \sum W_i = 1$. Suppose the weight of public participation in the stage of decision, design and construction, trial-operation, operation, and rejection as K_j $(j = 1, 2, \ldots 5)$, $\sum K_j = 1$; then:

$$PPI_d = \sum W_i \lambda_j \tag{73.10}$$

73.6 Conclusion

It is essential to apply public participation in sustainable construction. The framework is addressed to discuss the institutional arrangement of public participation in sustainable construction. Empirical studies are needed to test such hypotheses.

It's a fact that the attitudes of the public and government towards participation become mature, from rivalry to seek dialogue and consultation based on the legal framework. As Arnstein (1969) stated in her famous article on citizen participation: There is a critical difference between going through the empty ritual of participation and having the real power needed to affect the outcome of the process. In other words, meaningful participation is not a matter of process, but of deliberate intention. Conflicts between the government, media, enterprises and the public are not bad things. More important is how the decision-making process can become more scientific, democratic, open and transparent.

References

Arnstein SR (1969) A ladder of citizen participation. J Am Plann Assoc 35:216–224

Bourdeau L (1999) Agenda 21 on sustainable construction. CIB report publication no. 237, Rotterdam

Creighton JL (2005) The public participation handbook: making better decisions through citizen involvement. A Wiley Imprint, San Francisco

Freeman RE (1984) Strategic management: a stakeholder approach. Pitman, Boston

Gibberd J (2003) Building systems to support sustainable development in developing countries. Facilities Planning and Management, CSIR Building and Construction Technology, Pretoria

Kibert CJ (1994) Establishing principles and a model for sustainable construction. In: Proceedings of the first international conference of CIB task group 16 on sustainable construction, Tampa, FL, 6–9 Nov 1994, pp 3–12

Mitchell RK, Agle BR, Wood DJ (1997) Toward a theory of stakeholder identification and salience: defining the principle of who and what really counts. Acad Manag Rev 22:853–886

Perold JJ (2006) The psychosocial dynamics of public participation: a systemic analysis. PhD thesis of University of Pretoria, South Africa

Pretty J (1995) Participatory learning for sustainable agriculture. World Dev 23:1247–1263

Rietbergen-Mc Cracken J, Narayan-Parker D (1998) Participation and social assessment: tools and techniques. International Bank for Reconstruction and Development, Washington, DC

Shen LY, Wu M, Wang JY (2002) A model for assessing the feasibility of construction project in contributing to the attainment of sustainable development. J Constr Res 3:255–269

The World Bank (1996) The World Bank participation sourcebook. Environmentally Sustainable Development, Washington, DC

Wyatt DP (1994) Recycling and serviceability: the twin approach to securing sustainable construction. In: Proceedings of the first international conference of CIB Task Group 16 on sustainable construction, Tampa, FL, 6–9 Nov 1994, pp 113–122

Chapter 74
The Cooperative and Competitive Complexity Model of Emerging Strategic Marine Biological Medicine Virtual Industry Cluster Based on Value Net Theory

Peng Du

Abstract Emerging strategic marine biological medicine virtual industry cluster (ESMBMVIC) is a new kind of cluster mode in knowledge economy field. Since ESMBMVIC has the unique characteristics of high technicalization and cluster virtualization, cooperative and competitive activities of ESMBMVIC have characteristics of complexity. Therefore, it is an important issue for ESMBMVIC to describe the characteristics of complexity and measure the complexity. In this paper, the basic cooperative and competitive value net model and DNA structure imitation cooperative and competitive value net model of ESMBMVIC have been established to make concept extraction on cooperative and competitive activities. On the basis of those value net models, the characteristics of complexity of ESMBMVIC have been further analyzed, which have also been measured combined generalized entropy theory.

Keywords Cooperation • Competition • Emerging strategic marine biological medicine virtual industry cluster (ESMBMVIC) • Value net

74.1 Introduction

Since the late twentieth century, with the development of marine technology, a number of marine resources types have been gradually discovered, utilized and explored commercially, which have strategic importance to human survival and

Supported by Teaching Reform Research Project of Dalian Ocean ***University*** **(*JG2011YB35*)**

P. Du (✉)
College of Humanities and Law, Dalian Ocean University, Dalian, China
e-mail: dupeng0470@163.com

E. Qi et al. (eds.), *The 19th International Conference on Industrial Engineering and Engineering Management*, DOI 10.1007/978-3-642-37270-4_74,

development. Emerging strategic marine industry is a kind of industry that can effectively enhance the total national ocean development capabilities. And this industry is based on marine high-tech development, whose core content is marine high-tech achievement industrialization.

It is noteworthy that emerging strategic marine biological medicine industry is a significant kind of emerging strategic marine industry, which takes marine biological resources as R&D objects, marine biological technology as leading technology, and marine medicine as leading products. In China, however, the R&D on marine biological medicine industry started late, is relatively weak in technology, and is still in its gestation period. Although China is the largest raw material producer of marine biological products, there is still a big gap between China and developed countries on deep processing of marine biological products. In addition, the main R&D institutions engaged on marine biological products are small and medium enterprises. Resources like the state-owned research institutes, colleges and universities have not been fully utilized. The traditional geographic centralized form of marine biological medicine industrial clusters cannot completely meet the needs of marine biological medicine enterprise. Therefore, a new type of clusters called emerging strategic marine biological medicine virtual industry cluster (ESMBMVIC) is rising. Relying on advanced network technology methods, ESMBMVIC integrates organizations from all around the world according to marine biological and medicine high technology industrial value chain, which has the characteristics of organization approach replacing the geographical proximity(Molina and Bremer 2001; Porter 1998).

74.2 The Cooperative and Competitive Value Net Models of ESMBMVIC

Value net model explains the relationship between participants of all commercial activities, and emphasizes that value is created through the cooperation between participants (Porter 1998; Passiante and Seeundo 2002). Since the main cooperative and competitive activities of ESMBMVIC are about process of value-added related with high technology of marine biological medicine industry, numbers of basic value nets are formed, whose core are high-tech enterprises of marine biological medicine industry, which form the whole ESMBMVIC cooperative and competitive system (Han-mi 2007; Von Hippel 1994). It is for value net that is formed by members not only to create value, but also achieve the flows of logistics, information and energy, which thereby promote the evolution of ESMBMVIC. In this paper, value net model of cooperation and competition of ESMBMVIC are constructed based on value net theory to analyze cooperative and competitive complexity of ESMBMVIC. The value net models of ESMBMVIC cooperative and competitive activities contain the cooperative and competitive basic value net model and DNA structure imitation cooperative and competitive value net model.

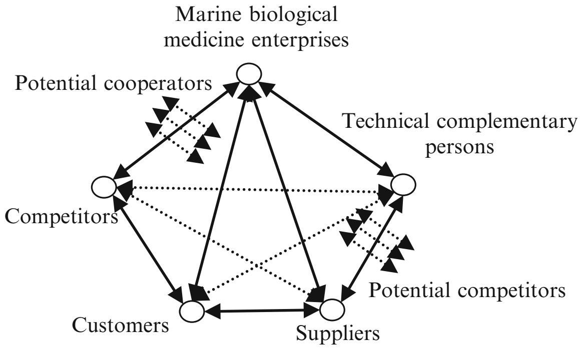


Fig. 74.1 The seven factors cooperative and competitive basic value net model

74.2.1 Cooperative and Competitive Basic Value Net Model

In this paper, the cooperative and competitive basic value net model is set up to explain the basic cooperative and competitive relationship of members of ESMBMVIC (Blau 1985). As is shown in Fig. 74.1, since the marine biological medicine products have characteristics of high risk, high input and high degree of difficulty in development, it is very common to make technical cooperation during the development process of the marine biological medicine products. Marine biological medicine enterprises cooperate to develop a product, and each enterprise focuses on its core technology (Bretz and Judge 1994; Coombs 1999). The cooperative and competitive basic value net model contains marine biological medicine enterprises, technical complementary persons, competitors, suppliers, potential competitors and potential cooperators.

The cooperative and competitive basic value net of ESMBMVIC can reflect the basic relationship between cooperation and competition and the value add process. Under the cooperation of marine biological medicine enterprises, technical complementary persons and suppliers, the marine biological medicine products are created, and the marine biological medicine value is implemented through the transaction with customers.

In order to adapt to the characteristics of high risk, high input and short life cycle of high-tech products, marine biological medicine enterprises in ESMBMVIC continuously integrate, split, and reorganize the basic value net model to achieve a temporary balance status. In this process, the role of the original basic value net model may be changed by the potential competitors and potential cooperators to meet the demands of marine biological medicine market.

74.2.2 The DNA Structure Imitation Value Net Model

ESMBMVIC cooperative and competitive basic value net model explains the basic cooperative and competitive relationship between the participants. However, viewed from the perspective of cluster, cooperative and competitive activities of

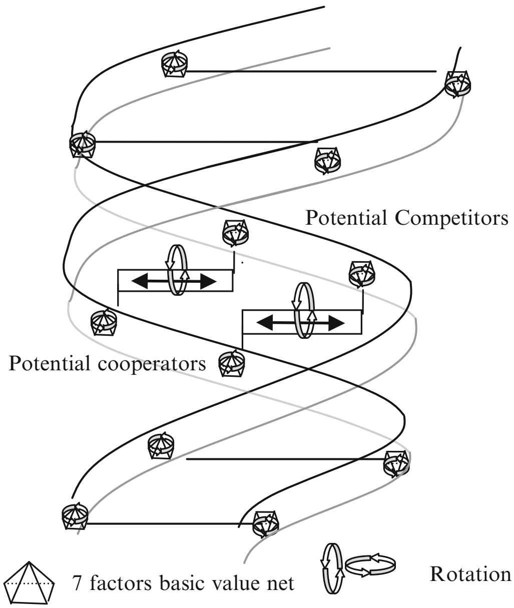


Fig. 74.2 DNA structure imitation value net stereo model

ESMBMVIC members are very complex. In the virtual space, any members can have cooperative and competitive activities, and some members may have more than two basic value nets, which put value net in a cross-status. In order to clearly explain the cooperative and competitive activities of members of ESMBMVIC, DNA structure imitation value net model is established based on basic value net model with the idea of DNA double helix structure.

As is presented in Fig. 74.2, DNA structure imitation value net model is not only on the structure but also on features similar with DNA model:

Firstly, the member's cooperative and competitive activities can achieve information transmission and knowledge sharing, while DNA can transfer genetic information.

Secondary, since the members of ESMBMVIC are adaptive subjects who continuously adjust cooperative and competitive activities by studying others' successful experience, the DNA structure imitation value net of ESMBMVIC has the process of self-adjustment, repair, split and reorganization, while the DNA has the function of copy, restructuring, and self-repair.

74.3 The Complexity Characteristics of ESMBMVIC Cooperation and Competition

In this paper, according to the cooperative and competitive value net model of ESMBMVIC, the complexity characteristics of ESMBMVIC cooperation and competition are analyzed as follows:

74.3.1 The Characteristics of Open and Dynamic

In ESMBMVIC, since the information, energy and logistics can be exchanged through members competing and cooperating with each other or with external enterprises, cooperative and competitive activities of members of ESMBMVIC are completely open, and the external enterprises can compete and cooperate with ESMBMVIC members freely, which indicates that the ESMBMVIC cooperative and competitive system is an open system.

The dynamic characteristic of cooperation and competition of ESMBMVIC is as follows:

Firstly, at the embryonic stage of ESMBMVIC, the members mainly compete and cooperate with local enterprises. With the further development of ESMBMVIC, members gradually break the geographical constraints to implement cooperative and competitive activities in virtual clusters.

Secondary, Members can choose to join or exit the cluster. Then according to market demand, the value net, which is composed by the members, constantly repairs, splits and reorganizes.

74.3.2 Adaptive Subjects

The members of ESMBMVIC are all adaptive subjects who have the ability of perception and response, and can learn to adjust their status automatically to acclimatization. If the members make the wrong decision, they can independently repair and restructure the value net (Tracey and Gordon Clark 2003).

74.3.3 Nonlinear Characteristics

ESMBMVIC cooperative and competitive system is a nonlinear system which does not meet the principle of superposition. In the system, a change of a small part will cause the change of the other parts till the whole system is changed. Since the ESMBMVIC cooperation and competition system is composed by numbers of subsystem whose composition are value nets, it has nonlinear feedback not only between subsystems but also between the subsystem and the whole system, which generates the synergistic effect and multiplier effect. Based on Lotka-Volterra model, this paper establishes the following model to show the nonlinear characteristics of ESMBMVIC cooperation and competition system (Jian Zhuang 2006).

$$\frac{dX_i}{dt} = \frac{\sum_{r=1}^{m_i} \frac{dx_r}{dt}}{m_i} = \frac{1}{m_i} \left\{ \sum_{r=1}^{m_i} \left[\left(A_i - B_i x_r + C_i + D_i - \frac{x_r}{K} \right) x_r + \theta_i \sum_{\substack{j=1 \\ j \neq i}}^{N} (E_{ijr} X_i Y_j) \right] \right\} \tag{74.1}$$

As is showed in formula 74.1, x_r is the core technology r of enterprise X_i. E_{ijr} is the interaction coefficient which presents that enterprise X_i has influence on enterprise Y_j on the technology x_r. m_i presents the number of core marine biological medicine enterprise that the enterprise i possesses. N presents the total number of enterprises in ESMBMVIC. $\frac{dx}{dt}$ presents the change rate of market share of the enterprise X. $\frac{dy}{dt}$ presents the change rate of market share of the enterprise Y. t presents time. A_x and A_y respectively present the natural growth rate of market share of enterprise X and Y. B_x and B_y respectively present influence coefficient that marine biological medicine products life cycle has effect on the growth rate of market share of enterprise X and Y, since the marine biological medicine products have the characteristic of short cycle, marine biological medicine products life cycle is very important to the growth rate of market share. C_x and C_y respectively presents the research capability of marine biological medicine enterprises of X and Y. K presents the demand of the marine biological medicine products of α, while K_1 presents the demand of the high-tech products of α and β; D_x and D_y respectively presents that brand and knowledge infiltration have influence on the on the growth rate of market share of enterprise X and Y. θ_x and θ_y respectively presents coefficient that the ESMBMVIC have effect on enterprise X and Y.

As is shown in formula 74.1, there are Nonlinear characteristics in the cooperative and competitive activities of ESMBMVIC, which emerges the whole advantage of the ESMBMVIC.

74.4 The Measurement of Cooperative and Competitive Complexity of ESMBMVIC

At the natural and social sciences fields, generalized entropy is used to measure complexity of the random events sets in different levels and types. The definition of the generalized entropy is as follows:

$$H(x) = -\sum_{i=1}^{q} p(x_i) \log p(x_i) \tag{74.2}$$

q presents the total number of events in random events sets; $p(x_i)$ presents case x_i random probability; $H(x)$ presents the entropy of the random events x.

In this paper, the complexity of the ESMBMVIC can be measured through calculating the entropy of the various random events with the tool of generalized entropy. Since ESMBMVIC is a dynamic network organization and the cooperation and competition value net of ESMBMVIC also have the characteristic of dynamic, in this paper, the complexity of the ESMBMVIC is measured for a moment.

There are various random events sets in ESMBMVIC which will generate entropy because of uncertainty and randomness. By analyzing DNA structure imitation value net model, this paper argues that the complexity of ESMBMVIC, which includes the following entropy:

Firstly, the entropy is generated through the information interaction of ESMB-MVIC which is included by the intra value nets and inter value net.

Secondary, the entropy is generated by the uncertainty of structure, which mainly is generated by the potential competitors and potential cooperators.

Thirdly, the entropy is generated by the uncertainty of the state of the value net.

74.4.1 The Entropy Is Generated by Information Interaction of ESMBMVIC

In the basic value net of ESMBMVIC, if m_{ij} is the amount of information interaction of subject i and j, and $p_1(ij)$ is probability of information implementation of subject i and j, then:

$$p_1(ij) = \frac{m_{ij}}{\sum_{i=1}^{n} \sum_{j=1, j \neq i}^{n} m_{ij}} \tag{74.3}$$

n is the total number of subject in the basic value net of ESMBMVIC, while $\sum_{i=1}^{n} \sum_{j=1, j \neq i}^{n} m_{ij}$ is the total amount of information interaction in the basic value net.

$H_1(ij)$ is entropy which is generated by the information interaction between subject i and j.

$$H_1(ij) = -p_1(ij)\log_2(P_1(ij)) \tag{74.4}$$

Then in the ESMBMVIC, there has a basic value net called k whose entropy generated by the information interaction.

$$H_k = \sum_{i=1}^{n} \sum_{j=1, j \neq i}^{n} (H_1(ij)) \tag{74.5}$$

If G is the total number of basic value nets of ESMBMVIC, then H_1 is the total entropy generated by the information interaction of all the intra basic value nets.

$$H_1 = -\sum_{k=1}^{G} \sum_{i=1}^{n} \sum_{j=1, j \neq i}^{n} (p_1(ij)\log_2(P_1(ij))) \tag{74.6}$$

In the ESMBMVIC, if m_{ij} is the amount of information interaction of basic value net i and j, and G is the total number of basic value nets of ESMBMVIC, $p_2(ij)$ is probability of information implementation of basic value net i and j, then

$$p_2(ij) = \frac{r_{ij}}{\sum_{i=1}^{G} \sum_{j=1, j \neq i}^{G} r_{ij}} \tag{74.7}$$

$\sum_{i=1}^{n}\sum_{j=1,j\neq i}^{n} r_{ij}$ is the total amount of information interaction during the inter basic value nets.

$H_2(ij)$ is entropy which is generated by the information interaction between the basic value net i and j.

$$H_2(ij) = -p_2(ij)\log_2(P_2(ij)) \tag{74.8}$$

H_2 is the total entropy generated by the information interaction of all the inter basic value nets.

$$H_2 = -\sum_{i=1}^{G}\sum_{j=1,j\neq i}^{G} p_2(ij)\log_2(P_2(ij)) \tag{74.9}$$

74.4.2 The Entropy Is Generated by the Uncertainty of Structure

At ESMBMVIC, the structure of the basic value nets is always in the uncertain state because of the existence of potential competitors and potential cooperators, which will generate the entropy and increase the complexity of the ESMBMVIC.

In the basic value net called i, w_i is the total number of potential competitors and potential cooperators. $p_3(i)$ is probability of structure change of the basic value net i.

$$p_3(i) = \frac{w_i}{\sum_{i=1}^{G} w_i} \tag{74.10}$$

G is the total number of basic value net of ESMBMVIC, and $\sum_{i=1}^{n} w_i$ is the total number of all the potential competitors and potential cooperators in ESMBMVIC.

$H_3(i)$ is the entropy of structure change of the basic value net i.

$$H_3(i) = -p_3(i)\log_2(P_3(i)) \tag{74.11}$$

H_3 is the entropy of structure change of the ESMBMVIC.

$$H_3 = -\sum_{i=1}^{G} p_3(i)\log_2(P_3(i)) \tag{74.12}$$

74.4.3 The Entropy Is Generated by the Uncertainty of the Status of the Basic Value Net

The basic value nets appear to different status, such as steady-status, easily dissolution-status, easily absorbing the new individual's status, which will generate entropy and increase the complexity of ESMBMVIC.

There is v status in the basic value net i, and $H_4(i)$ is the entropy of the uncertainty of the status of the basic value net i.

$$H_4(i) = \left(-\sum_{k=1}^{v} p_4(i_k)(\log_2(p_4(i_k))) \right) \tag{74.13}$$

$p(i_k)$ is the probability of the uncertainty of the status k of the basic value net i.

The formula 74.13 only shows the entropy generated by the uncertainty of the status of the basic value net itself. But the entropy can also be generated by transferring process between the inter value nets. The basic value nets may be influenced by other basic value nets to be changed the status. Therefore, $H_4(i, j)$ is conditional entropy which is generated by the status changes between the basic value nets.

If there have some random basic value net i and j, then,

$$H_4(i, j) = -\sum_{k=1}^{v} \sum_{f=1}^{v} p_4(i_k, j_f)\log_2 p_4(i_k, j_f) \tag{74.14}$$

$p_4(i_k, j_f)$ is the joint probability with the status k of the basic value net i and the status f of the basic value net j.

$$H_4(i, j) = H_4(i) + H_4\left(\frac{j}{i}\right) \tag{74.15}$$

$$H_4(i, j) = H_4(j) + H_4\left(\frac{i}{j}\right) \tag{74.16}$$

$H_4\left(\frac{i}{j}\right)$ is the conditional entropy of the basic value net i under the conditions of the certain status of the basic value net j. It means that the joint entropy of basic value net i and j equals to the structure entropy of the basic value net i pulse the entropy that the basic value net j transfers to the basic value net i.

G is the total number of basic value net of ESMBMVIC, and H_4 is entropy that is generated by the uncertainty of the status of the basic value net in ESMBMVIC.

$$H_4 = -\sum_{i=1}^{G} \sum_{j=1, j \neq i}^{G} \sum_{k=1}^{v} \sum_{f=1}^{v} p_4(i_k, j_f)\log_2 p_4(i_k, j_f) \tag{74.17}$$

74.4.4 The Complexity Measurement of the ESMBMVIC Cooperation and Competition

H_{ESMBMVIC} is the complexity measurement of the ESMBMVIC cooperation and competition, which are the sums of the three above entropy.

$$H_{ESMBMVIC} = H_1 + H_2 + H_3 + H_4 \tag{74.18}$$

74.5 Conclusion

In this paper, the basic value net model and DNA structure imitation value net model are established to make concept extraction on the cooperative and competitive activities of ESMBMVIC, which can be used to analyze the characteristics of complexity much more clearly. Combined the value net model and extended entropy theory, this paper preliminarily measures the cooperative and competitive complexity of ESMBMVIC. However, when measuring the cooperative and competitive complexity of ESMBMVIC, this paper only considers some major factors, and discusses not in depth on some other factors which in the future will be further exploration and research.

References

Blau G (1985) The measurement and prediction of career commitment. J Occup Psychol 58:277–288

Bretz RD Jr, Judge TA (1994) The role of human resource systems in job applicant decision processes. J Manag 20:531–551

Coombs R (1999) Human resource management and the motivation of technical professionals. Int J Innov Manag 3:17–26

Han-mi L (2007) Research on organization virtualization of traditional industry cluster. Sci Technol Prog Policy 24:76–77 (in Chinese)

Jian Zhuang Z (2006) Research on the risk and organization virtualization of the traditional industry cluster. Technol Econ 25:44–45 (in Chinese)

Molina A, Bremer CF (2001) Achieving critical mass: a global research network in system engineering. Foresight 3:59–64

Passiante G, Seeundo G (2002) From geographical innovation clusters towards virtual innovation clusters: the innovation virtual system. Presented at the 42th ERSA congress, Dortmund

Porter ME (1998) Clusters and the new economics of competition. Harv Bus Rev 76:36–38

Tracey P, Gordon Clark L (2003) Perspective-alliances, networks and competitive strategy, rethinking cluster of innovation. Growth Change 34:1–16

Von Hippel E (1994) Sticky information and the locus of problem solving: implications for innovation. Manag Sci 40:429–439

Chapter 75
Research on Functional Demands of Urban Integrated Transportation Hub Project Based on the Stakeholder Theory: *A Case Study of Tianjin Station*

Xiao-li Zhang, Yi-lin Yin, Gong-Jie Ao, and Xiao-Dong Zhang

Abstract The orientation and realization of urban integrated transportation hub project is established on the basis of the satisfaction of stakeholders' interest demands. It is crucial to recognize the parties related to the project and to analyze the needs of every part for realizing the value of urban integrated transportation hub project. Based on the stakeholder theory, taking Tianjin station as an example, to recognize and define the stakeholders of integrated transportation hub project of Tianjin station, then confirm the core stakeholders and analyze the interests demands, finally determine the project functional demands which provides effective evidence for construction of urban integrated transportation hub project.

Keywords Interest demands • Project value • Stakeholders • Urban integrated transportation hub project

75.1 Introduction

Urban integrated transportation hub is an important part of the transportation system, the intersection of the different modes of transport network and the integrity of fixed and mobile devices connected by several transport, sharing the direct

X. Zhang (✉) • Y. Yin
School of Management, Tianjin University of Technology, Tianjin, China
e-mail: utopiaxiaoli@163.com

G. Ao
School of Management, University Du Québec A Chicoutimi (UQAC), Chicoutimi, Canada

X. Zhang
Tianjin Design Institute, China Petroleum Pipeline Bureau, Tianjin, China

E. Qi et al. (eds.), *The 19th International Conference on Industrial Engineering and Engineering Management*, DOI 10.1007/978-3-642-37270-4_75,

operations, transit operations, hub local operations and urban external transport related operations in the hub of the region (Yuan Hong and Lu Huapu 2001).

The characteristics of urban integrated transportation hub are huge scale, numerous functions, complex internal structure and external environment, long construction period and extended rang of stakeholders involved. Therefore, according to the realization of the project value, if they intend to obtain the goal of integrity value maximization, urban integrated transportation hub must pay attention to the needs of the related stakeholders. As the project progresses, however, the property of the stakeholders will change and the impact of the change will appear to be horizontal of dynamic development. So, it becomes an eager problem that how to identify stakeholders and define and analyze their interest demands during the project construction period of urban integrated transportation hub.

75.2 The Related Literature Review

75.2.1 Review of Urban Integrated Transportation Hub Project Functional Demands

According to the analysis of the project demands, the existing researches, such as Yi-lin Yin has pointed out that the core value of public projects is the results obtained from the common consult and compromise of the project stakeholder, measure the realization of project value from the realization of demand degree (Yin Yi-lin and Hu Jie 2006). Ding-rong Yu regards the related demands of all project stakeholders as a key standard to measure the success of the project (Ding Rong yu 2005). Sha Wang considers the stockholders' demands of urban integrated transportation hub preliminary design according to its later period evaluation (Wang Sha 2009). Moreover, in their researches of the key project stakeholders, Xiao and Jing have demonstrated that how to identify the key stakeholders and their related relationships is the most important basis to determine whether the project could success (Wang Xiaojin and Huang Jing 2006). Jing-fu Yang says we should not only stand on the organizational point to analyze the recognition of the project stakeholders, but also accurately divide the extent of stakeholders' impacts when researching them (Yang Jing-fu 2002). Shi-Bo Dong, etc would think the public engineering project as an incomplete contractual relationship of effectively allocated limited public resources by stakeholders and dynamically identify the stakeholders in the different stages of the project life cycle (Dong Shi-bo 2009). Some researches have proposed a combined method of questionnaire and expert advice for how to identify the stakeholders (Wang Wen-xue and Yin Yi-lin 2008). However, the researches are only subjective levels with no identification and definition of the stakeholders from the different dimension.

75.2.2 Application of Stakeholder Theory

The theory about the development of the stakeholders has been existed for several years, but the definition problem of its concept has never been generally recognized so far. In 1983, Freeman divided the stakeholders into general stakeholders and special stakeholders. This expanded the connotation of stakeholders, which brought the social organizations, the public, government departments, and environmentalists groups into the stakeholder theory's category and became a standard paradigm about the definition of stakeholders (Freeman and Evan 1990). Based on this, this paper defines the stakeholders of urban integrated transportation hub as the concept that the benefited or damaged construction of the project activities can affect the project objectives or can affect the individuals or groups of project goals. The stakeholders' theory applied to the project is a new trend of domestic and international project management in recent years and has become a very popular analytical tool at the International. Cleland and Kerzner provide a theoretical basis for the introduction of the project stakeholders (Cleland and Kerzner 1985). He believes that the project is a combination of monetary resources, human resources, management resources and other resources into a temporary organization of special groups in order to achieve specific objectives. A series, systemic contract relationship and interest demands constitute the project itself demands. The project has the characteristics of the organization and the premise of its existence is to solve a specific goal from Williamson's theory. Therefore, the project organization can also use the method of the stakeholders to analyze (Williamson 1995).

In China, the stakeholders' theory focuses on corporate governance research to solve the problem of value and objective of corporate governance optimization (Yang Rui long and Zhou Ye an 2000). Along with the development of the stakeholder theory gradually, the scope and application gradually expanded, such as south to north water transfer project, the World Bank loan programs and the social assessment of the project research. For the application in the project stakeholders, more and more domestic scholars thought that the large part of the success of the project depends on the project stakeholders' satisfaction to the project (Turner 2004). Conclusively, the success of the stakeholder management would largely determine the value maximization realization of urban integrated transportation hub project.

75.3 Analysis of Demands of Core Stakeholders in Urban Integrated Transportation Hub Project of Tianjin Station

75.3.1 Recognition of Project Stakeholders and Determination of Core Stakeholders

Since it has been involved the huge range of stakeholders, long period and frequent varieties during the construction process of Tianjin Railway Station transportation hub project, this paper starts mainly from the identification of the project life cycle to the project and considers the stakeholders' interests demands influenced by the project construction activities as far as possible (Chen Hong-hui and Jia Sheng-hua 2004). Planning phase is a starting stage of project life cycle and the main content of work is project proposal, feasibility study report and design plan descriptions (Deng Pingjun and Wang Yan 2007), so the main related stakeholders included the construction company, the administrative department in charge of construction, project investors as well as the public. Design phase is completed by design company, which involves the main stakeholders such as the construction company, consulting company and the investigation company, government related administrative departments and so on. The Construction and delivery stage mainly refers to the construction of the project phase and the completion acceptance stage work, the main stakeholders included design company, the construction company, supervision company, suppliers, etc. Operation phase is the stage after the completion of the project and put the normal function into use. So it is the stage embodied the value of the project. The main stakeholders contain related investors, operation management company, project users and the public. Based on this, the paper combined with the practice of Tianjin station, the recognition of stakeholders is shown in Table 75.1.

It can be seen from the analysis above that the various stakeholders involved in the entire life cycle of Tianjin Railway Station integrated transportation hub and that different stakeholders in different stages as well as different interests demands of construction projects. Because some stakeholders who are closely related to the project have largely influenced the project and the others are opposite, therefore, we should determine the main stakeholders also called the core stakeholders who mainly influence Tianjin Railway Station integrated transportation engineering and then analyze the demand of that engineering based on the interest demands.

This paper first adopts the survey questionnaire to identify and determine the core stakeholders, then obtains the information by investigating the influence degree of stakeholders on Tianjin Railway Station integrated transportation hub, and finally combines with the government departments, the City Investment Group, related investors and the users to determine the core stakeholders of Urban integrated transportation hub while designing the questionnaire to analyze the individual interests and understand the real demands of the core stakeholders.

75.3.2 Interests Demands of Core Stakeholders

75.3.2.1 Interests Demands of the Government Departments

The Government is a provider of the public project. The start point of the project construction includes macro and micro level. From macro level, the goal of the project construction is to meet the demands of national and regional transport

Table 75.1 Recognition of stakeholders in Tianjin station

Planning phase	Design phase	Construction phase	Deliver phase	Operation phase
Tianjin Municipal Construction Committee	Tianjin Municipal Construction Committee	Tianjin Municipal Construction Committee	Tianjin Municipal Construction Committee	Tianjin Municipal Government
Tianjin Municipal Government	Tianjin Municipal Government	Tianjin Municipal Government	Tianjin Municipal Government	Tianjin Municipal Government
Tianjin organizing committee	Tianjin organizing committee	Government executive department	Tianjin organizing committee	Tianjin city cast construction Co., LTD
Planning department	Planning department	Tianjin City Investment Group	Government executive department	Related investors
Government executive department	Government executive department	Tianjin city cast construction Co., LTD	Tianjin City Investment Group	Users
Tianjin City Investment Group	Tianjin City Investment Group	Design organization	Tianjin city cast construction Co., LTD	The surrounding communities and shops
Related investors	Related investors	Construction organization	Design organization	
Tianjin city cast construction Co., LTD	Tianjin city cast construction Co., LTD	Supervision organization	Construction organization	
Users	Design organization	Consulting organization	Supervision organization	
	Consulting organization	Suppliers		
	Users			

development, to promote the use of Tianjin urban public transport and to guide the layout of urban space. From micro level, its goal is to improve the traffic chaos of Tianjin Railway Station integrated transportation hub region.

75.3.2.2 Interests Demands of Tianjin City Investment Group

As the shareholder of Construction Company, Tianjin City Investment Group contributes to participate in the construction of Tianjin Railway Station integrated transportation hub project. The subsidiary of the Group, called Tianjin City Investment Construction, is responsible for the construction of hub and the operation and management of the facility. As one of the investors, the Group needs to ensure that its investments can be well compensated. Therefore, it is necessary to develop some business facilities for the operation and management and take these revenues for the compensation of construction investment as well as normal operation. In addition, Tianjin City Investment Group can acquire large construction management experience by participating in the large-scale construction project. As a result, it will contribute for the Group to participate in the similar construction projects in the future and enhance its visibility and influence in the industry.

75.3.2.3 Interests' Demands of Investors

The interest demands of related investors such as the Ministry of Railways, Tianjin Metro Company, Jinbin Light Rail Company, the Haihe River Park and Pipeline Company have two aspects in summary: the one is to provide better transport facilities to meet with the requirements of the users; the other is to be able to obtain Investment incomes such as commercial advertising income, parking spaces income and commercial rental income, etc.

75.3.2.4 Interests' Demands of Users

The users are the "end users" of Tianjin Railway Station integrated transportation hub project. It is a key standard to measure the success of the transportation hub project whether the services provided by the users could meet with the requirements of users. As the users of project, they contain several public demands of city integrated transportation hub shown in Table 75.2.

In summary, the interests demand of government stressed on the problems of urban traffic, such as solving the un-equilibrium between public transportation and railway traffic and so on. The city investment group most belong to the state-owned property, their interests demand are consistent with the government, besides they

Table 75.2 Interests demands analysis of Tianjin station

Number	Interests demands	Specific objectives
1	Transport functions	Reasonable transport cost; arrive on time; safe and reliable; no obstacle; travel convenience; high operating efficiency
2	Service functions	Ticketing points set up user-friendly (automatic ticket vending machines); signs set clear and obvious; prevailing set of related facilities; staff, enthusiastic, and thoughtful; rail transportation within the external space is clean, neat and orderly; waiting, clean place to rest clean and tidy; set of automated teller machines; traffic news and information.
3	Commercial functions	Provide food and beverage, accommodation services and facilities; commercial outlets: supermarkets, gift shops, etc.; leisure facilities
4	Landscape functions	Laying green spaces; planting trees; fresh air; decorating murals

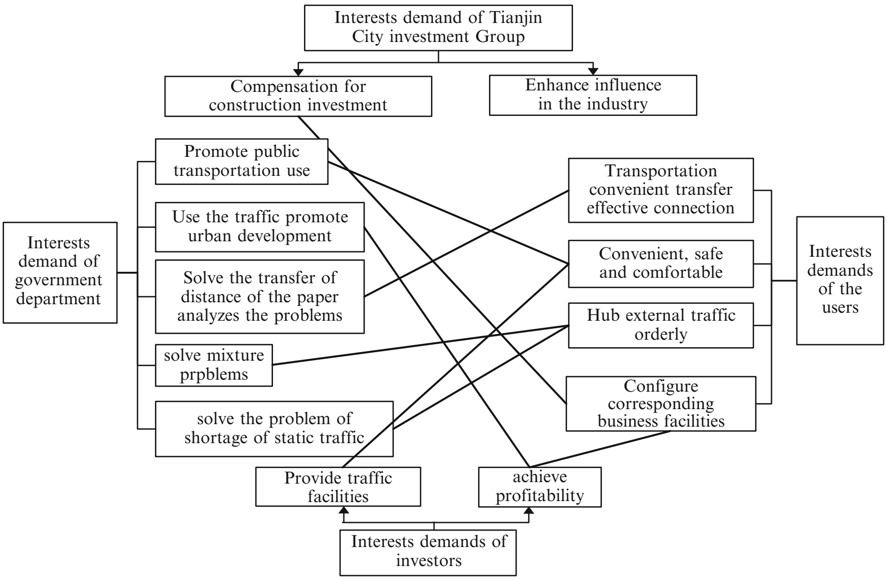


Fig. 75.1 The relationship of core stakeholder demands of Tianjin station

hope to obtain corresponding development profit on the basis of the demand. The users stress on the requirement of the traffic hub of auxiliary facilities such as the convenient transfer condition, neat internal environment and external environment. Therefore, the relationship of interest demands among the government departments, investors, Tianjin City investment and the users is shown in Fig. 75.1.

75.4 Functional Demands of Tianjin Station Urban Integrated Transportation Hub Project

Based on the analysis above, this paper first concludes the overall functional demands of Tianjin Railway Station integrity transportation hub project from the point view of the demands of the stakeholders and then summarizes four aspects to get the functional demands of Tianjin Railway Station integrity transportation hub project, as shown in Fig. 75.2.

75.4.1 Drive Regional Development

The construction of the transportation hub project can bring a lot of flow and the certain construction of commercial and property facilities can enrich the region's commercial activities as well as attract more people. The location of Tianjin Railway Station is the geometric center of Tianjin city. The station closes to the Haihe River and echoes the plan of Tianjin Haihe River development to create a new Haihe River. The development of the commercial and property formats would drive the economic development of the region in which Tianjin Station transportation hub project is and would plays a great role in the goal of urban development strategies about promoting Tianjin Haihe coast.

75.4.2 Promote the Use of Public Transport

To promote the development of public transport is an effective measure to solve urban traffic problems and is an improved performance level of urban traffic

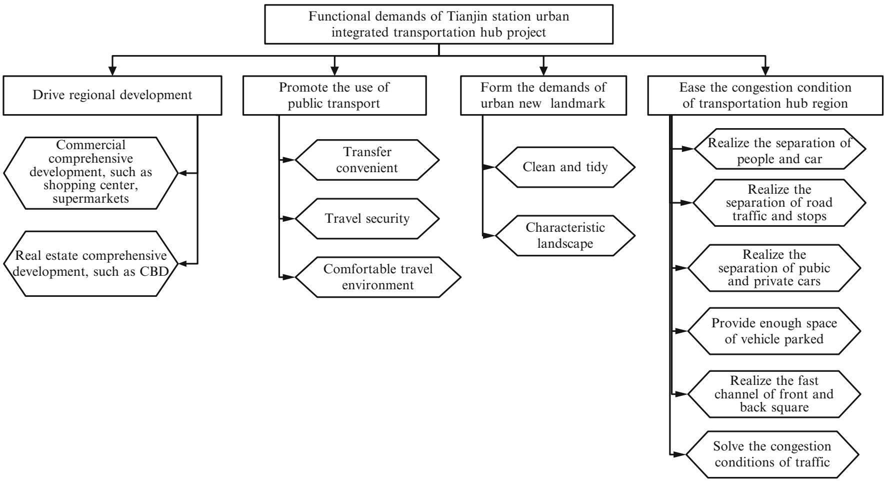


Fig. 75.2 Functional demands of Tianjin station

development. Since the improvement of urban traffic conditions has effectively improved the travel conditions and reduced environmental pollution, the government departments and users are the main interest demands. The main demands of the users are convenient transfer, safety travel and comfortable travel environment. The new construction of transport facilities such as Orbit transfer center, City Liner and he Plaza contact channels could provide the better travel services for the users. At the same time, the set of metro line 2, 3, and 9 also provide faster speed, more convenient transport services and more selective transfers.

75.4.3 Form the Demands of Urban New Landmark

Tianjin Railway Station integrated transportation hub is the door of Tianjin external transportation. Because of its good environment and characteristics of the landscape construction, it could shape Tianjin Railway Station integrated transportation hub into new landmark in Tianjin, giving a deep impression to foreign visitors and improving the city overall image.

75.4.4 Ease the Congestion Condition of Transportation Hub Region

The original transportation facilities can not meet with the growing demands of urban transportation and the traffic chaos of the hub region include the human and car mixed, the crossing traffic and arrival traffic mixed and the serious traffic congestion. These conditions lead to the bad travel conditions for the travelers, can not meet with the travel demands and are not good for the city to further development. Therefore, the contraction such as Haihe East Road and the main square underground engineering building, comprehensive building, five longitude road tunnel project and Li Gong House overpass reconstruction works of the building can effectively solve the traffic problems in the hub region and meet with the users' requirements of smooth and order traffic.

75.5 Conclusion

Tianjin Railway Station integrity transportation hub project involves many stakeholders. This paper starts from the stakeholders' theory to identify the stakeholders of the project. Based on the actual construction, we finally determine the core stakeholders including government departments, investors, Tianjin City investment and users by designing questionnaire and make a further analysis to ensure integrated functional demands of Tianjin Railway Station project. It provides a reliable basis for project planning and design engineering, leading the value function maximization of integrated transportation hub project under the condition with the requirements of all individuals.

References

Chen Hong-hui, Jia Sheng-hua (2004) Research on stakeholders 3-dimension classification in Chinese enterprises. Econ Res J 4:80–90

Cleland DI, Kerzner HA (1985) Performance and project success: perceptions of Chinese construction supervising engineers. Int J Proj Manag 3:35–37

Deng Pingjun, Wang Yan (2007) The planning of risk prevent of planning stage in construction. Shaanxi Archit 9:52–53

Ding Rong yu (2005) Research on success factors of construction project under DB mode, Central South University

Dong Shi-bo (2009) Study on the contracts of public projects—a perspective based on the stakeholders theory. J North China Inst Sci Technol 6:93–98

Freeman RE, Evan WM (1990) Corporate governance's stakeholder interpretation. J Behav Econ 19:337

Turner JR (2004) Five necessary conditions for project success. Int J Proj Manag 22:349

Wang Sha (2009) Social impact post-assessment of transport infrastructure project based on core value. Tianjin University of Technology, Tianjin

Wang Wen-xue, Yin Yi-lin (2008) Research on management of the stakeholders of Tianjin transportation hub project. Urban Mass Transit 9:4–6

Wang Xiaojin, Huang Jing (2006) The relationships between key stakeholders' project performance and project success: perceptions of Chinese construction supervising engineers. Int J Proj Manag 3:35–37

Williamson OE (1995) Organization theory: form Chester Barnard to the present and beyond. Yale University Press, New York, pp 58–60

Yang Jing-fu (2002) Analysis of investment project stakeholders. Hohai University, Nanjing

Yang Rui long, Zhou Ye an (2000) The application of stakeholders theory in enterprise. Economic Science Press, Beijing

Yin Yi-lin, Hu Jie (2006) Study on the success criteria of public project based on the core value of the stakeholders. China Soft Sci 5:149–155

Yuan Hong, Lu Huapu (2001) Study on model and method of comprehensive transportation terminal planning. J Highw Transp Res Dev 6:101–105

Chapter 76
Environmental Policy and Enterprise Hiring Decisions

Jin-ying Zhang

Abstract This study analyzes the mechanism that environmental policy affects enterprise-hiring decisions. By decomposing the substitution effect and cost effect by mathematics and graphs, it shows that the impact of environmental policy on enterprise hiring decisions is uncertain. The net effect depends on the substitutability between labor and other inputs, the selected environmental policy instruments and the harmonization of these instruments.

Keywords Cost effect • Environmental policy • Hiring decisions • Labor • Substitution effect

76.1 Introduction

Environmental policy refers any action for preventing or reducing the harmful effects of human activities on natural resources and environment, and for ensuring that man-made environmental changes do not endanger human health. Because the allocation of natural resources and environmental products by a free market is not efficient, proper government intervention is necessary. Environmental protection and employment were regarded contradictory in the 1970s and 1980s. In 1991, Pearce argued that the financing of the public expenditures through environmental taxation can thus provide a double dividend if the revenue is utilized to cut back other taxes that bring about large excess burden (Pearce 1991), which triggered a large debate about the real effectiveness of the notion of a double dividend. There was a surge of papers on double dividend in 1990s. Some papers have shown that an energy tax and carbon tax does not necessarily lead to job losses,

J. Zhang (✉)
School of Economics, Shandong University of Finance and Economics, Jinan, China
e-mail: jinying0011@hotmail.com

E. Qi et al. (eds.), *The 19th International Conference on Industrial Engineering and Engineering Management*, DOI 10.1007/978-3-642-37270-4_76,

many research results have shown that the environment can be improved while the employment be increased (Bossier and Brechet 1995; Bovenberg and van der Ploeg 1994, 1996; Bovenberg 1998; Schneider 1997; Crowley 1999). Some scholars emphasized the limitations on double dividend and argued that green jobs measures can not fundamentally solve the employment problem. If the job is really increased, environmental pollution will increase, therefore, the double dividend is difficult to achieve (Nielsen et al. 1995; Carraro et al. 1996; Brunello 1996; Ackerman 1997). The research has become less controversial since 2000. More and more researches proved that environment policy gain employment and environment (Hinterberger et al. 2002; Bezdek et al. 2008; Marx 2010).

Government's environmental policy instruments become increasingly robust. Environmental policy may cause massive changes to the enterprise hiring decisions. Recent research has focused on estimating how many jobs environmental tax reform created, and exploring methods to improve the employment creativity of environmental policy. This paper focuses on analyzing the effect and mechanism that environmental policy affects enterprise hiring decisions from the microscopic point of view. Section 76.2 decompose the effect of environmental policy on enterprise's hiring decisions by mathematic method. Sections 76.3, 76.4, and 76.5 analyses the substitution effect and cost effect by graphs. Conclusions are presented in Sect. 76.6.

76.2 Decomposing the Effect of Environmental Policy on Enterprise Hiring Decisions

Environmental policy instrument comprises two categories, direct means and indirect means. Direct means is implemented directly through public investment by government departments, such as the infrastructure construction focus on environmental protection, low-carbon cities construction, monitoring and control carbon emissions from production, urban transport and living, establishing information database of energy use and carbon emissions, and widely conducting environmental publicity work, gradually popularizing environmental education. Indirect means is implemented through controlling or incentive methods to urge polluters protect environment, including regulatory and economic measures. Regulation is administrative legislation that constitutes or constrains emission rights and allocates pollution control responsibilities. Such as to urge power plants use certain technologies to reduce greenhouse gas emissions, to urge the surface coating companies use some type of ventilation system to re-process the emitted steam, to lay down auto emissions criterion and building energy efficiency standard. One can consider regulation measures as actions of conduct imposing sanctions, such as a fine. Economic measures take use of taxes, fees, capital subsidy on cleaner production and environmental protection innovation, low-interest or interest-free loans, tax breaks and emissions trading to urge or motivate polluters to control pollution.

Government controls or stimulates enterprises to achieve environmental protection objectives through a variety of environmental policy instruments. Some enterprises may suffer cost increase because of a variety of taxes, fees, and all the work in order to achieve government regulation standards. Some enterprises may enjoy cost reduction due to the Government capital subsidies, low-interest or interest-free loans or tax breaks. The cost change caused by environmental policy will lead to changes in the number of enterprise employment. This paper will simply analyze the hiring change in the enterprise level. Suppose hiring decisions are not affected by the change in product price and demand resulted from cost changes.

Suppose that the production function as follows:

$$Y = f(L, X) \tag{76.1}$$

Where Y is the production level, L is hiring cost, and X represents cost of other inputs, including the cost of pollution control (E). The cost of pollution control is the burden on enterprises caused by the environmental policy.

Suppose that the production cost function is:

$$TC = W \cdot L + \gamma \cdot X \tag{76.2}$$

Where TC represents the total cost, W is the wage rate, γ is the price of other inputs (X).

$$L = \frac{1}{W}\sigma_L \cdot TC \tag{76.3}$$

Where σ_L indicates the proportion of hiring cost to total cost.

Then, the hiring change resulted from environmental policy can be written:

$$\frac{\partial L}{\partial E} = \frac{TC}{W}\frac{\partial \sigma_L}{\partial E} + \frac{\sigma_L}{W}\frac{\partial TC}{\partial E} \tag{76.4}$$

The first part on the right side of Eq. (76.4) indicates substitution effect of environmental policy on enterprise employment. It represents changes in the proportion of hiring cost to total cost induced by environmental policy, reflecting substitution relationship between labor and other inputs. In order to achieve energy saving and emission reduction targets, enterprise may substitute labor for energy sources or other production factors, which will enhance hiring, and the hiring cost share will increase in the proportion of the total cost. In order to achieve cleaner production, enterprise may increase the level of automated production, which will reduce hiring, and the proportion of hiring cost to total cost will decline. The second part represents changes in total cost caused by environmental policy. It shows the cost effect of environmental policy on enterprise hiring decisions. If the environmental policy leads to total cost added, the enterprise will increase all the factors of production in order to maintain original production, including labor input;

conversely, companies will reduce labor and other production inputs. There are differences in the conditions and modalities in implementing environmental policy in different enterprises, as a result, the direction and degree of the substitution effect and cost effect are not quite the same, and the net effect of environmental policy on the hiring decisions may be different.

76.3 The Effect of Restrictive Environmental Policy

In order to achieve environmental policy objectives, the government will adopt regulatory instruments and restrictive economic instruments, such as setting emission standards, levying taxes, fees, and fines on carbon emissions and pollution behavior. These instruments lead to direct impact on enterprise, resulting in added production cost. Suppose the price of labor keeps unchanged, incremental production cost caused by environmental policy is marked by the rise of other inputs' price (γ). Then, the absolute value of the slope of iso-cost line, input price ratio (γ/W), increases. The horizontal intercept of iso-cost line decreases. As shown in Figs. 76.1 and 76.2, the iso-cost line moves down from AB to AC, and the profit-maximum inputs combination adjust from N_1 (X_1, L_1) to N_2 (X_2, L_2).

Iso-cost line *DE* is drew to decompose the substitution effect and cost effect of environmental policy on hiring decisions, which is parallel to iso-cost line *AC* and tangent to isoquant I_1 at N_3. Profit-maximum inputs combination change from N_1 to N_3, which reflect the level of the total cost and production keep unchanged, while the factor price ratio changes. Combination changing from N_1 to N_3 shows the substitution effect. Because the price of other inputs rises, the price of labor becomes relatively declined, resulting in substitution of labor for other inputs. The amount of hiring increases from L_1 to L_3, and other inputs reduces from X_1 to X_3. The change of inputs combination from N_3 to N_2 shows the result of cost effect, which reflects that the factor price ratio keeps unchanged. Because total inputs decreases, hiring decreases from L_3 to L_2, other inputs also decreases.

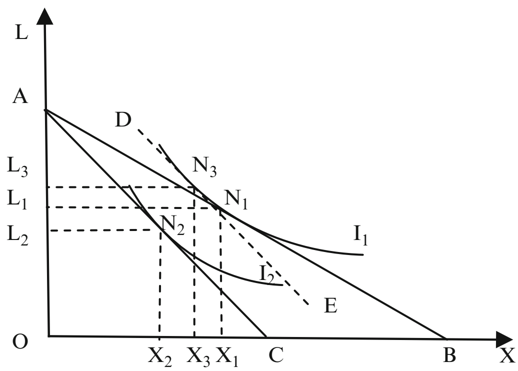
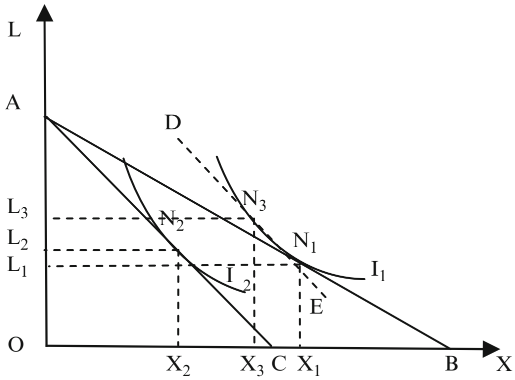


Fig. 76.1 The effect of restrictive environmental policy (a)

Fig. 76.2 The effect of restrictive environmental policy (b)



Changes of other inputs caused by substitution effect and cost effect take the same direction, so the net effect is the summation. The net effect shows other inputs ultimately reduced from X_1 to X_2. But the situation is different in the changes of hiring. In the example showed in Fig. 76.1, hiring decreases from L_1 to L_2 however, in the example in Fig. 76.2, hiring increases from L_1 to L_2. The reason lies in that hiring changes caused by substitution effect and cost effect take opposite direction. In Fig. 76.1 substitution effect is less than cost effect, implying that the substitutability of labor for other inputs is small. In Fig. 76.2 substitution effect is greater than cost effect, implying that the substitutability of labor for other inputs is large.

76.4 The Effect of Incentive Environmental Policy

Government may take incentive environmental policy instruments to motivate polluters to control pollution, such as production and research subsidies, low interest or interest-free loans, tax breaks, and rewards for clean production enterprises. These instruments will cut down the production cost, so the iso-cost line moves down. Suppose the price of labor keeps constant, production cost reduction caused by environmental policy is reflected by the decline of γ, so the absolute value of the slope of iso-cost line drops as shown in Figs. 76.3 and 76.4. The iso-cost line AB moves to AC, and the profit maximize inputs combination adjust from N_1 (X_1, L_1) to N_2 (X_2, L_2).

Iso-cost line DE is drew to decompose the substitution effect and cost effect of environmental policy on hiring, which is parallel to iso-cost line AC and tangent to isoquant I_1 at N_3. Profit-maximum inputs combination changes from N_1 to N_3, which reflect the level of the total cost and production keep unchanged, while the factor price ratio changes. Combination changing from N_1 to N_3 shows the substitution effect. Because the price of other inputs decreases, the price of labor becomes relatively increased, resulting in substitution of other inputs for labor. The amount of

Fig. 76.3 The effect of incentive environmental policy (a)

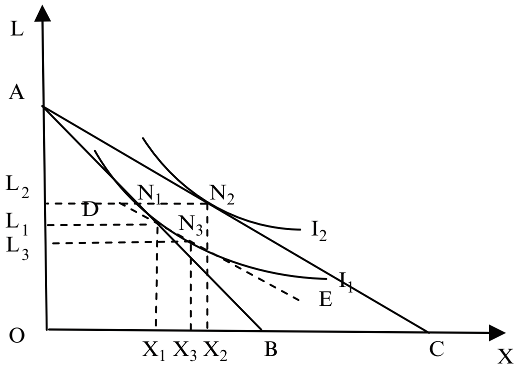
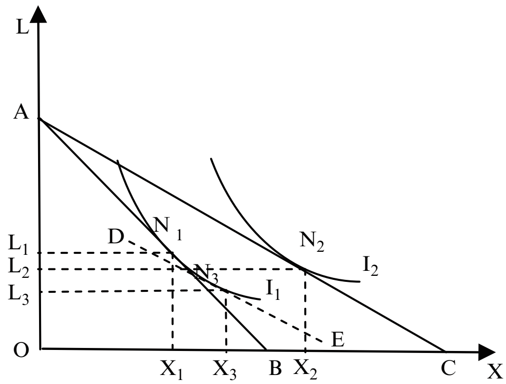


Fig. 76.4 The effect of incentive environmental policy (b)



hiring decreases from L_1 to L_3, and other inputs increases from X_1 to X_3. The change of inputs combination from N_3 to N_2 shows the result of cost effect, which reflect that the input price ratio keeps unchanged while total cost decreases. Because total inputs increases, hiring increases from L_3 to L_2, other inputs also increases to X_2.

Changes of other inputs caused by substitution effect and cost effect take the same direction, so the net effect is the summation. The net effect shows other inputs increase in both examples. But the situation is different in the changes of hiring. In the example showed in Fig. 76.3, hiring increases from L_1 to L_2, however, in the example in Fig. 76.4, hiring decreases from L_1 to L_2. The reason is that hiring changes caused by substitution effect and cost effect take opposite direction. In Fig. 76.3 substitution effect is less than cost effect, implying that the substitutability of labor for other inputs is small. In Fig. 76.4 substitution effect is greater than cost effect, implying that the substitutability of labor for other inputs is large.

The rationale of the effect of restrictive environmental policy and incentive environmental policy on enterprise hiring decisions is the same, but the result has difference.

76.5 The Effect of Green Fiscal Reform

Sections 76.3 and 76.4 are limited to the assumption that environmental policy do not change labor tax and wage rate. In recent years, some developed countries initiated the green fiscal reform. The core content of green fiscal reform is to increase tax on resources, energy and carbon emissions, while reducing labor tax. Whether green fiscal reform obtains double dividend of environmental and employment? This section will analyze from the enterprise level.

If the labor supply curve is perfectly inelastic, the personal income tax does not affect wage rate and hiring. If the labor supply curve is elastic as shown in Fig. 76.5, pre-tax market wage rate increases and enterprise hiring falls (McConnell et al. 2002). Curve D indicates pre-tax enterprise hiring curve. D_{t1} indicates after-tax labor demand curve. The vertical distance between D and D_{t1} indicates personal income tax. The vertical distance between D and D_{t1} increases as the wage rate increases, implying that the tax rate is progressive. Income tax causes pre-tax wage rate increase from W_1 to W_2 and enterprise hiring reduces from L_1 to L_2. Green fiscal reform reduces individual labor income tax rate. The after-tax labor demand curve is D_{t2}. This will result in direct and indirect effect on enterprise hiring. The direct effect shows that hiring declines compared with the hiring in D_{t1}. The indirect effect shows that wage increment declines, add the influence of the price change of other inputs caused by increased taxation, the price ratio of labor and other inputs will be affected, meaning that the substitutability of labor and other inputs changes.

In Fig. 76.6, line AB shows the iso-cost line before green fiscal reform. if the reform causes the factor price ratio (γ/W) to increase, other inputs become relatively expensive, and the absolute value of the slope of iso-cost line increases, so the iso-cost line move to the CD direction. If the factor price ratio declines, labor becomes relatively expensive, and the absolute value of the slope of the iso-cost line declines, so the iso-cost line moves to EF direction. Enterprise substitutes other inputs for labor, hence enterprise hiring decreases. So the effect of green fiscal reform on enterprise hiring is uncertain. Figure 76.6 describes the substitution effect caused by the green fiscal reform, without considering the impact of cost change.

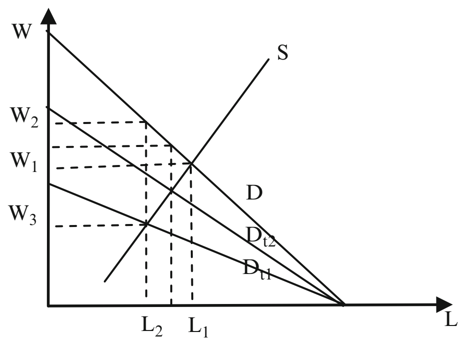


Fig. 76.5 The impact of income tax on wage and hiring

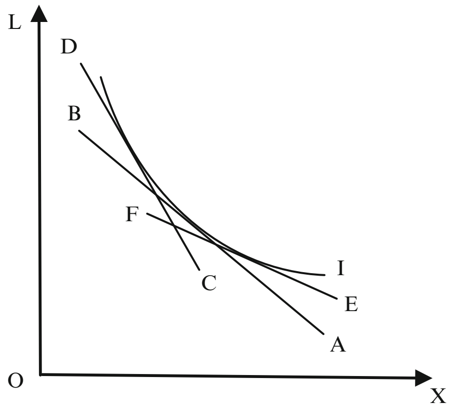


Fig. 76.6 The effect of green fiscal reform

76.6 Conclusion

Through the above analysis we can see that the impact of environmental policy on enterprise hiring is uncertain. It depends on the substitutability of different inputs, what environmental policy instruments are selected and whether these instruments are harmonized. In order to achieve double objects of environmental protection and employment growth, government intervention should be established in full market information and scientific regulatory standards. To figure out the acceptable pollution control degree and emission standard in the current technical conditions and select appropriate environmental policy instruments is of special importance. Otherwise, it will lead to inadequate intervention or excessive intervention, which may cause new environment and employment problems and lead to government failure. This paper does not make empirical analysis. It will be completed in following research.

References

Ackerman F (1997) Why do we recycle? Island Press, Washington, DC

Bezdek RH, Wendling RM, DiPerna P (2008) Environmental protection, the economy and jobs: national and regional analyses. J Environ Manage 86:63–79

Bossier F, Brechet T (1995) A fiscal reform for increasing employment and mitigating CO2 emissions in Europe. Energy Policy 23:789–798

Bovenberg AL (1998) Tax reform, structural unemployment and the environment. Scand J Econ 100:593–610

Bovenberg AL, van der Ploeg F (1994) Environmental policy, public finance and the labour market in a second-best world. J Public Econ 55:349–390

Bovenberg AL, van der Ploeg F (1996) Optimal taxation, public goods and environmental policy with involuntary unemployment. J Public Econ 62:59–83

Brunello G (1996) Labour market institutions and the double dividend hypothesis. In: Carraro C, Siniscalco D (eds) Environmental fiscal reform and unemployment. Kluwer Academic Publishers, Dordrecht

Carraro C, Galeotti M, Gallo M (1996) Environmental taxation and unemployment: some evidence on the double dividend hypothesis in Europe. J Public Econ 62:141–181

Crowley K (1999) Jobs and environment, "the 'double dividend' of ecological modernisation?". Int J Soc Econ 26:1013–1026

Hinterberger F, Omann I, Stocker A (2002) Employment and environment in a sustainable Europe. Empirica 29:113–130

Marx A (2010) Ecological modernization, environmental policy and employment. Can environmental protection and employment be reconciled? Innovation 13:311–325

McConnell CR, Brue SL, Macpherson DA (2002) Contemporary labor economics, 6th edn. McGraw Hill Education, New York, pp 396–397

Nielsen SB, Pedersen LH, Sørensen PB (1995) Environmental policy, pollution, unemployment, and endogenous growth. Int Tax Public Finance 2:185–205

Pearce DW (1991) The role of carbon taxes in adjusting to global warming. Econ J 101:938–948

Schneider K (1997) Involuntary unemployment and environmental policy: the double dividend hypothesis. Scand J Econ 99:45–59

Chapter 77
Internal Mechanism Research on Synergy of Enterprise Network Organizations Based on Value Structure

Jing-wen Han and Zheng-yin Peng

Abstract Based on the view of value structure, the mechanism of the synergy of enterprise network organizations has been studied in use of the analysis method of transaction costs. The findings suggest that correct cooperative strategy, appropriate distribution of relationship rents, mutual trust and commitment help the organizations achieve synergy. Moreover, the perfect cooperative strategy is the key condition of the achievement of high-order synergy, meanwhile the commitment of organizations is the key condition of the achievement of low-order synergy.

Keywords Commitment • Relationship rent • Synergy • Transaction costs • Trust

77.1 Introduction

Cooperative competition in the unity of competition and cooperation has become the inevitable trend of development of competitive strategies of the enterprise network organizations. In recent years, under the intricate tasks' situations, most of strategic alliances, entrepreneurship and enterprise network organizations lie in the problem of low overall efficiency and the instable development. One of the main reasons of the above phenomenon is due to the mismatching of cooperation and competition,

The paper is supported by the National Natural Science Foundation of China (No: 71172018, 71002007, 61004056), National Social Science Foundation (No: 11&ZD153) and the research fund of Tianjin University of Finance and Economics (No: Y1107).

J. Han (✉) • Z. Peng
School of Management, Tianjin University of Finance and Economics, Tianjin, China
e-mail: Jingwenhan@126.com; Pengzhy2002@yahoo.com.cn

E. Qi et al. (eds.), *The 19th International Conference on Industrial Engineering and Engineering Management*, DOI 10.1007/978-3-642-37270-4_77,

and thus, the development of our corporate network needs urgently to set up a synergistic-led strategy to reduce operational risks. Meanwhile, the synergy state of the enterprise network organizations means through the integration and optimization of value chain, the efficiency of the allocation of organizations' resource achieves optimization. Therefore, the study of the internal mechanism of the enterprise network organizations' synergy from the perspective of value structure can provide the theoretical basis to the implementation of enterprise network organizations' synergy, and have the guiding significance to promote the effective operation and the development of our business network.

Synergy research of the enterprise network organizations both at home and abroad mainly come from the evolution of cooperative and competitive relations of the organizations. Firstly, the relationship of cooperation and competition has the symbiotic characteristics. Barry and Adam first proposed the formulation of cooperation and competition, and used game theory to describe the phenomenon of containing coexistence of competition and cooperation (Barry and Adam 1996). Secondly, the effect of cooperation and competition is to gain competitive advantage. Cravens, etc constructed a conceptual framework, defining and classified the relationship of cooperation between organizations, emphasized the characteristics and differences between them, and analyzed strategic Union, thus pointed out that through this form the organization will gain competitive advantage (Cravens and Shipp 1993). Thirdly, the goal of cooperation and competition is to generate synergy. Tsai studied the effect of knowledge sharing in cooperative and competitive organizations and pointed out the possible ways to achieve synergy (Tsai 2002). Wu and Choi build a framework for co-produced by the transaction cost theory and social relations theory which based on the Chinese business network scenarios integrated, meanwhile collaborative generation refers to the exchange of useful resources by the enterprise with other partners, and has stronger competitiveness and higher efficiency after uniting (Wu Wei-Ping and Choi 2004). The research of organization and coordination by some domestic and abroad scholars embodies in the areas of reference and content development. Gu etc studied the configuration of the shared resources within the enterprise group from the perspective of the cooperation theory, and explored the issues relating to the enterprise group synergy (Gu Baoguo and Fang Xiaojun 2004). Xu etc constructed an organizational system of learning model based on synergy mechanisms (Xu Xueguo et al. 2006) while Yun etc concluded that synergy is the main reason to cause the evolution of enterprise group (Yun Jiang et al. 2006).

The above researches are all belong to results-oriented research, and there are few researches neither studied the synergy of enterprise network organizations from the behavior-oriented nor explained the source of behavior. This paper studied the microcosmic enterprises in the enterprise network organization with the theory of transaction cost economics to investigate the source of the synergy of organizations and promote the synergistic factors of the enterprise network organizations, so as to provide some theoretical basis for enterprise network organizations to achieving synergy.

77.2 The Realization of Enterprise Network Organization

By the virtue of differences of study object and perspective, there is no uniform definition of "relationship assets" in academic fields, the relationship assets mentioned in this article refers to the interests of embodiment based on the corporate relations which contribute to creating value for them. Therefore, it is the function of the degree of cooperation". From the perspective of the influence of relationship assets to relationship rent and transaction costs, the relationship assets have the following properties:

Property 1 The relationship rent is an increasing function of the relationship assets and presents the rule of marginal returns increasing, as shown in Fig. 77.1.

With the increasing of the relationship assets, enterprise cooperation has increased from the depth, which can be reflected from the shallow sharing of a technology to the deep sharing of it or from a product of a single technical cooperation to expanding cooperation for multi-technology or from the increasing of co-proprietary assets investment. The results will enable the enterprise to obtain more relationship rents. Furthermore, with the increasing depth of cooperation and the strength of cooperative relations, cooperation will become increasingly close to the company's core technologies or core shared, which will undoubtedly lead to the rule of marginal income increasing. The limitation is that the two sides fully shared core technologies or core product, which leads to the maximum relationship rent. Meanwhile, the competitive advantage as an independent enterprise will disappear. It is the merger border for cooperative enterprise on cooperative projects, as shown in Fig. 77.1.

Property 2 The transaction cost is an increasing function of relationship assets, and presents the rule of marginal returns diminishing, as shown in Fig. 77.2.

The reason is that with the increasing of relational assets, cooperation will increase which leads to the increasing of the contracted number. Therefore, the preparation costs (searching information costs), decisive contract costs (negotiation and decision-making costs), and contract performance costs will increase, which means the increasing of the marketable transaction costs (Furubotn and Richter 1997). Meanwhile, the proprietary investments to maintain or improve

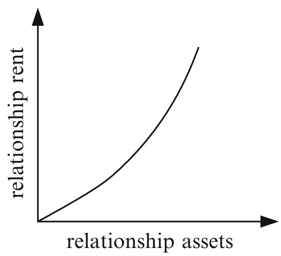


Fig. 77.1 Relationship rent Relationship assets curve

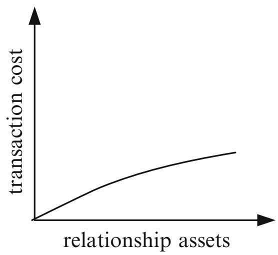


Fig. 77.2 Relationship cost-relationship assets curve

the partnership, and the costs based on cooperation such as information management will increase, which indicating that design costs of maintaining or changing a running organization and organizational costs will also increase, that is to say, the management of transaction costs increase. Therefore, the transaction cost is an increasing function of the relationship assets. We also noticed that with the increasing of relational assets, cooperation will get deeper and deeper, and accumulate more and more information on cooperation, which will lead to the searching and information costs decrease marginally, marginal decrease of negotiations and decision-making costs; at the same time, many times of collaborating make the cooperative enterprise mutual trust become stronger than before, supervisory costs and contractual obligations fulfillment costs also presents the rule of marginal diminishing. Therefore, the marketable transaction costs on the relationship assets decline marginally. At the same time, with the increasing of cooperation, the skillful administrative procedures make the management of transaction costs on the relationship assets also presents the rule of marginal diminishing. In summary, the transaction cost is an increasing function of relationship assets, and presents the rule of marginal returns diminishing.

"Synergy" means working together, the simplest formula of synergy philosophy is: $2 + 2 > 4$ (Yang Botang et al. 2004). The target of business cooperation in the enterprise network organization is making the relationship rent maximum. Therefore, from the view of the single enterprise's value structure, the realization of synergy not only referring to the "$2 + 2 > 4$", the more exact expression should be "$2 + 2 = \mathrm{m} + \mathrm{n} > 4$, $\mathrm{m} > 2$, $\mathrm{n} > 2$". That is, on the one hand, the gross income created by the collaborate enterprise is greater than the sum of individual companies; on the other hand, the participation and cooperation of each individual enterprise are to obtain greater benefits than before. In a viewpoint of single enterprise, the relationship rent from cooperation should be greater than the transaction costs paid by the cooperation.

Based on the influential of relationship assets on relationship rent and transaction costs, collaborative enterprise network organization includes the following two kinds of situations, as shown in Figs. 77.3 and 77.4. Figure 77.3 is the optimal co-state, it has the high requirements of the cooperative property, which is called high-order synergy; the synergy state in Fig. 77.4 arise after some states that have higher requirements of cooperative attitude, and we call it low-order synergy. In fact,

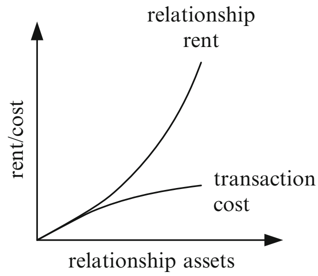


Fig. 77.3 High-order synergy

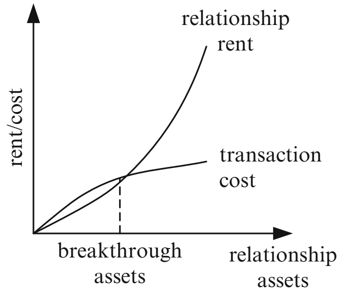


Fig. 77.4 Low-order synergy

no matter what kind of synergy realizations, they are required to have a possible higher relationship rents and the possible lower transaction costs. Therefore, in the following, we analyze the internal synergy mechanism of enterprise network organization from the determinants of the relationship rents and transaction costs.

77.3 The Internal Mechanism Analysis Results of the Enterprise Network Organization's Synergy

1. From the perspective of relationship rents.

In the same level of relationship assets, the larger the parties of cooperative network organization get relationship rents, the more likely it is to achieve synergy. Relationship assets are the key assets that can bring monopoly rents and innovative rents for the enterprise (Wu Miao 2002). So the scarcity of the relationship assets and the degree of innovation affects the number of the relationship rents. In addition, the relationship assets with the complementary property can also bring complementary benefits to the enterprise in use of the resources of other enterprises. Meanwhile, the appropriate relationship assets even have the function of integration the resource of the enterprise which brings conformity rents. In summary, the property of relationship assets' degree of scarcity and innovation, complementary extent and degree of integration decided the relationship assets. Therefore, considering of these four aspects, establishing appropriate cooperation strategy is the guarantee of getting enough relationship rents and achieving high-order synergy. In accordance with the cooperation of apportionment of the total relationship rents, the reasonable relationship rent distribution coefficient is also the key condition to achieve synergy.

2. From the perspective of transaction cost.

In the same relationship assets, the less payment of transaction cost, the organization can be more likely to achieve synergy. Therefore, the effective savings in transaction cost is an important index of the development of collaborative strategy. Williamson, from the perspective of the property of the transactions, described three important factors affecting the transaction cost, that is, asset specificity, transaction uncertainty and transaction frequency (Williamson 1985).

The stronger of the asset specificity, the larger of transaction cost (Fei Fangyu 1996). The larger transaction cost may make the cooperation cannot achieve synergy. Building up the mutual trust between the partners is the main solution to this problem. Here trust means all parties believe that each will not make use of other's weaknesses to obtain benefits. Trust can be divided into the credible trust and trust of goodwill. The former is the rational part of the trust, and it means that the cooperative enterprises believe each other has goodwill and ability to assume their alliance obligations and to fulfill the commitments they made for the union. The latter is the emotional part of the trust, it means that the confidence of the partners who will exchange good faith and fair to get along well with each other (Johnson et al. 1996). Once the network members established trust of goodwill, the transaction cost generated by the asset specificity is almost negligible, the enterprise network organization is easy to achieve synergy.

The higher uncertainty transaction, the greater the transaction cost (Fei Fangyu 1996). Williamson divided the uncertainty into three categories: the main uncertainty, secondary uncertainty and behavioral uncertainty. The main uncertainties vary depending on the circumstances; Secondary uncertainties engender because of lacking of communication which make decision-makers cannot grasp the situation; behavioral uncertainty refers to the traders' confidentiality and distortion for strategy. From the perspective of the three uncertainty categories, as long as members of the corporate network organization maintain the smooth flow of information alternation, transaction cost caused by the secondary uncertainty can be minimized. If the members of the network organization can build mutual trust, their confidentiality and distortion for strategy will be greatly reduced or even disappear, which makes the transaction cost caused by behavior uncertainty minimize. Apparently, the smooth exchange of information and mutual trust is an effective strategy to reduce the transaction costs generated by the uncertainties and help to achieve synergy for network organizations. The frequency of the transaction does not affect the absolute value of the transaction cost, so does the synergy of enterprise network organizations.

From the above analysis, the correct strategy for cooperation ensure that relationship assets have the higher degree of scarcity, degree of innovation, degree of complementing and degree of integration, which would enable the relationship assets produce the relationship rents as much as possible; for the cooperation in accordance with the proportion of the total relationship rent, reasonable relationship rent distribution coefficient can guarantee the individual firm to obtain a sufficient number of relationship rent. What mentioned above are the basis of realizing high-order energy, Meanwhile, to establish trust between the partners, especially goodwill trust, and to maintain the smooth exchange of information allow the payment of transaction cost as less as possible, which has great significance for whether high-order or low-order synergy.

Furthermore, the excellent cooperation strategy, a reasonable relationship rent distribution coefficient and mutual trust (especially goodwill trust) enable the joint venture's "relationship rent – transaction cost" curve similar to Fig. 77.3 so as to realize the high-order energy, in which the cooperation strategy plays an important role.

If cooperation strategy can only make the enterprises' "relationship rent – transaction cost" curve similar to Fig. 77.4, breaking through the critical bottleneck is the key for the enterprise network organizations to achieve synergy. The bottleneck breakthrough is closely related to the organizational "commitment". Here the "commitment" refers to the purpose of continued cooperation and goodwill of the partners, which is to say the partners are willing to maintain cooperative relations and make efforts to make cooperation successful. "Commitment" can be divided into two parts of rationality and emotion. The rational part based on the countable enterprises' benefit, also known as the measurable commitment. It refers to that the cooperative enterprise requires contact which can bring tangible income. Once the partners achieve its strategic objective, commitment will be enhanced (Cullen et al. 2000). Clearly, it is easy to achieve measurable commitments. The emotional part of the commitment is also called the attitude commitment, which refers to a psychological identity and a pride for the members of the network organization. It implies that the partners who have commitment are willing to offer resources and extra efforts, and have courage to face the risk of the alliance operation (Johnson et al. 1996; Lu Qian 2006). Therefore, although the realization of low-order energy of enterprise network organizations have a lower demand for the relationship assets, it put forward a higher demand for the members of organization, not only the mutual trust but also the attitude commitment.

From the above analysis, bounded rationality and opportunism are the background of discussing a synergy mechanism of the enterprise network organization, which is the for some network organizations who cannot achieve synergy. Accurate cooperation strategy (relationship assets with a high degree of scarcity, innovation, complementing and integration), the relational rent partition coefficient, mutual trust and commitment of members of the organization are the four major factors to eliminate the adverse effects for synergy. The correct strategy for cooperation and a reasonable relationship rent distribution coefficient ensure synergy from the point view of the "relationship rent", and excellent cooperation strategy (relational assets with a high degree of scarcity, innovation, complementing and integration) ensure the network organization to achieve high-order synergy; the mutual trust among the members ensures synergy from the perspective of transaction cost, the internal

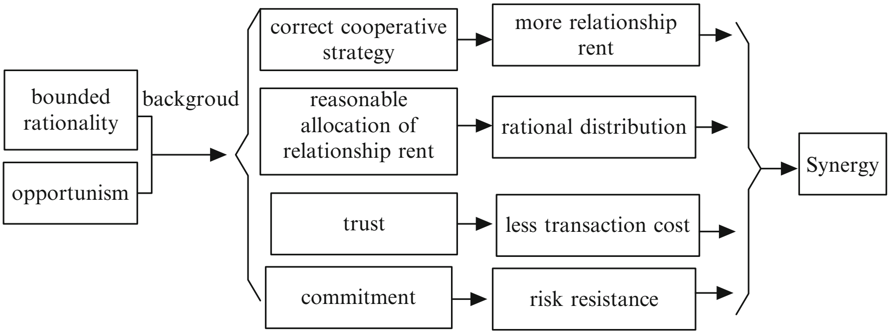


Fig. 77.5 Internal synergy mechanism of enterprise network organization

organization's commitment enhances the organization's ability of resisting risks and ensures the comprehensive achievement of low-order coordination in reality, as shown in Fig. 77.5.

77.4 Conclusion

From the perspective of the value structure, this paper studies the internal synergy mechanism of enterprise network organization in use of transaction cost economics theory. The study shows that the realization of synergy requires more revenue got for each member of the network organization than ever. From the perspective of transaction cost economics theory, it requires each cooperating member of the network organization to obtain more relationship rent than the transaction cost paid. Because of the two properties of relationship rent and transaction cost, the enterprise network organizations will present us low-order synergy and high-order synergy. A correct cooperative strategy, a reasonable allocation of relationship rent, mutual trust and commitment of the organization members can promote the network organization achieve synergy, in which excellent co-operation strategy is the key point to achieve high-order energy, and inter-organizational commitment is the key point to realize low-order energy.

References

Barry N, Adam B (1996) Co-opetition [M]. Random House Audio, Oskarshamn

Cravens DW, Shipp SH (1993) Analysis of co-operative interorganizational relationships, strategic alliance formation and strategic alliance effectiveness [J]. J Strateg Mark 1(1):55–70

Cullen JB, Johnson JL, Sakano T (2000) Success through commitment and trust: the soft side of strategic alliance management [J]. J World Bus 35:223–240

Fei Fangyu (1996) Determinants of assumed contract person and transaction costs – one of the economics comments of Williamson transaction cost [J]. Foreign Econ Manag 5:26–29 (in Chinese)

Furubotn EG, Richter R (1997) Institutions and economic theory: the contribution of the new institutional economics [M]. University of Michigan Press, Ann Arbor

Gu Baoguo, Fang Xiaojun (2004) Analysis based on the synergy of shared resources of the enterprise group [J]. Focus Facts 7:20–23

Johnson JL, Cullen JB, Sakano T, Takenouchi H (1996) Setting the stage for trust and strategic integration in Japanese-U.S. cooperative alliances [J]. J Int Bus Stud 27(5):981–1004

Lu Qian (2006) On the formation and management of strategic alliance relational capital [J]. Econ Manag 20(7):50–52

Tsai W (2002) Social structure of "coopetition" within a multiunit organization coordination competition and intraorganizational knowledge sharing [J]. Organ Sci 13(2):179–190

Williamson OE (1985) The economic institute of capitalism [M]. Free Press, New York

Wu Miao (2002) The creation of the relationship assets and corporate earnings – also on income changes of the different social forms of enterprise [J]. Zhongnan Univ Econ Law 2:103–107 (in Chinese)

Wu Wei-Ping, Choi WL (2004) Transaction cost social capital and firms' synergy creation in Chinese business networks: an integrative approach [J]. Asia Pac J Manag 21(3):325–343

Xu Xueguo, You Jianxin, Peng Zhenglong (2006) Study of learning model of organizational collaboration [J]. Ind Eng Manag 1:14–17 (in Chinese)

Yang Botang, Yang Yegong, Jiangying (2004) Analysis based on the collaborative resources of the enterprise's core competitiveness [J]. Enterp Econ 2:8–10 (in Chinese)

Yun Jiang, Liu Li, Gao Jie (2006) Value-creation and strategic improvement of business group: a perspective of synergy [J]. Res Financ Econ Issues 4:79–86 (in Chinese)

Chapter 78
The Application of Fuzzy Analytic Hierarchy Process and Grey Relational Analysis in the Taxi Passenger Satisfaction Evaluation

Ting Liu

Abstract Taxis are an important supplementary and complementary tool of public transport and analysis of passenger satisfaction aims to identify existing problems in the process of taxi service, then to propose improvements. By fuzzy analytic hierarchy process and grey relational analysis, to calculate the weight and correlation degree of the impact factors, then to obtain the key factors. The data analysis shows that the range of between dissatisfaction and the general, it is in the lower level. The waiting time, refusing to take passengers, service attitude and complaints handling are the most dissatisfied factors. The government should improve the overall satisfaction of the taxi passengers from three strategies about raising the price of the taxi, developing the public transport, strengthening management.

Keywords Evaluation • F-AHP • GRA • Satisfaction

78.1 Introduction

78.1.1 Research Background

The existing taxi number is 14120 in Wuhan City, this number was far from satisfying the needs of passengers, strong demand is an important reason, in addition to the increase of passengers, mainly taxi driving is difficult, the root cause is that road congestion increased year by year. The survey shows a substantial increase

T. Liu (✉)
School of Transportation, Wuhan University of Technology, Wuhan, P.R. China
e-mail: 515501708@qq.com

E. Qi et al. (eds.), *The 19th International Conference on Industrial Engineering and Engineering Management*, DOI 10.1007/978-3-642-37270-4_78,

in car ownership in Wuhan City. In urban areas, the vehicle speed is showing a declining trend, lower than the average speed in some of the congested road, and road congestion caused the decline in turnover efficiency of taxi operations.

The passengers pay attention to the time efficiency, when public transport which is urban transport aorta were blocked, it prompted the explosive development of private transport, so road blocking was harder than before. For taxi drivers, they will consider cost accounting, some sections which the efficiency is not high will not travel, so that makes congestion sections of very few taxi trips, which is exacerbated by a “take a taxi is hard” phenomenon.

78.1.2 Research Contents

From the perspective of the passengers, this paper analyzes the taxi passenger satisfaction evaluation factors firstly, through questionnaires, randomly selected passengers to score the 17 indicators of the impact on passenger satisfaction evaluation, and then calculated the satisfaction evaluation data, to analyze the problems of the taxi service process, then to propose improvements for increasing passenger satisfaction and easing “take a taxi is hard” phenomenon.

78.1.3 Research Methods

In this paper, fuzzy analytic hierarchy process (F-AHP) and grey relational analysis (GRA) combined to increase the reliability and accuracy of the satisfaction evaluation.

F-AHP combines the advantages of the AHP and fuzzy comprehensive evaluation method, in a fuzzy environment, to take into account many factors, and use the fuzzy membership theory quantitative to quantify the qualitative indicators reasonably. This method solves the question that evaluation process is too subjective, and to effectively compensate for the lack of establishing the weight. Come to the scientific, objective results of evaluation (Zheng Xin et al. 2010; Xu Ge-ning and Jiang Fan 2010).

GRA is an important method in grey system theory that the objective things have a wide range of grey, according to similar or different degree of the development trend between factors, to measure the correlation degree, which reflect the merit order of evaluation indicators (Wu Qi-bing et al. 2008). GRA has obvious advantage of the theoretical analysis for inaccurate information and the uncertainty of small sample system (Hu Da-li 2003).

78.2 F-AHP and GRA Model of Taxi Passenger Satisfaction Evaluation

78.2.1 Index System

Index system of taxi passenger satisfaction evaluation is the premise and foundation to build satisfaction evaluation model. Whether index system is reasonable and accurate, directly impact on the scientific, reliability and accuracy of the evaluation results. The evaluation index system can be divided into three layers: target layer, rule layer and indicator layer (Wang En-xu and Wu Chun-you 2009) (Fig. 78.1, Li Hong-ji 2005).

78.2.2 Determining the Weights of the Index System

To establish weight layer $W=\{W_1,W_2,\ldots,W_n\}$, W_i is the weight of the i-th indicator in the rule layer, $W_i=\{W_{i1},W_{i2},\ldots,W_{ij}\}$, W_{ij} is the weight of the j-th indicator in the i-th indicator (Lu Yue-jin 2002).

A number of experts judge the importance of n indicators on the same layer by Delphi method (Table 78.1), then to establish the judgment matrix, and check the consistency. Consistency index is $\mathrm{CI}=\frac{\lambda_{\max}-n}{n-1}$, according the mean random consistency index RI (Table 78.2), obtains relative consistency index CR, $\mathrm{CR}=\frac{\mathrm{CI}}{\mathrm{RI}}$, if $\mathrm{CI}<0.1$, the weight order meets the requirements, the results are satisfied, or to adjust the judgment matrix, and to obtain the weights again by the above steps (Morash 2001).

1. *Judgment matrix $O-U$:*

$$A = \begin{pmatrix} 1 & 3 & 1/2 & 2 & 1/3 \\ 1/3 & 1 & 1/4 & 1/5 & 1/5 \\ 2 & 4 & 1 & 1/2 & 1/2 \\ 1/2 & 5 & 2 & 1 & 1/2 \\ 3 & 5 & 2 & 2 & 1 \end{pmatrix}, \ \lambda_{\max}=5.374,\ \mathrm{CI}=0.094,\ \mathrm{CR}=0.084$$

<0.1, the judgment matrix $O-U$ has consistency. $W=(0.168, 0.055, 0.193, 0.202, 0.382)^{\mathrm{T}}$.

2. *Judgment matrix U_1-U_{1j}:*

$$A_1 = \begin{pmatrix} 1 & 2 & 3 \\ 1/2 & 1 & 3 \\ 1/3 & 1/3 & 1 \end{pmatrix}, \ \lambda_{\max}=3.054,\ \mathrm{CI}=0.027,\ \mathrm{CR}=0.046<0.1, \text{ the}$$

judgment matrix U_1-U_{1j} has consistency. $W_1=(0.527, 0.333, 0.140)^{\mathrm{T}}$.

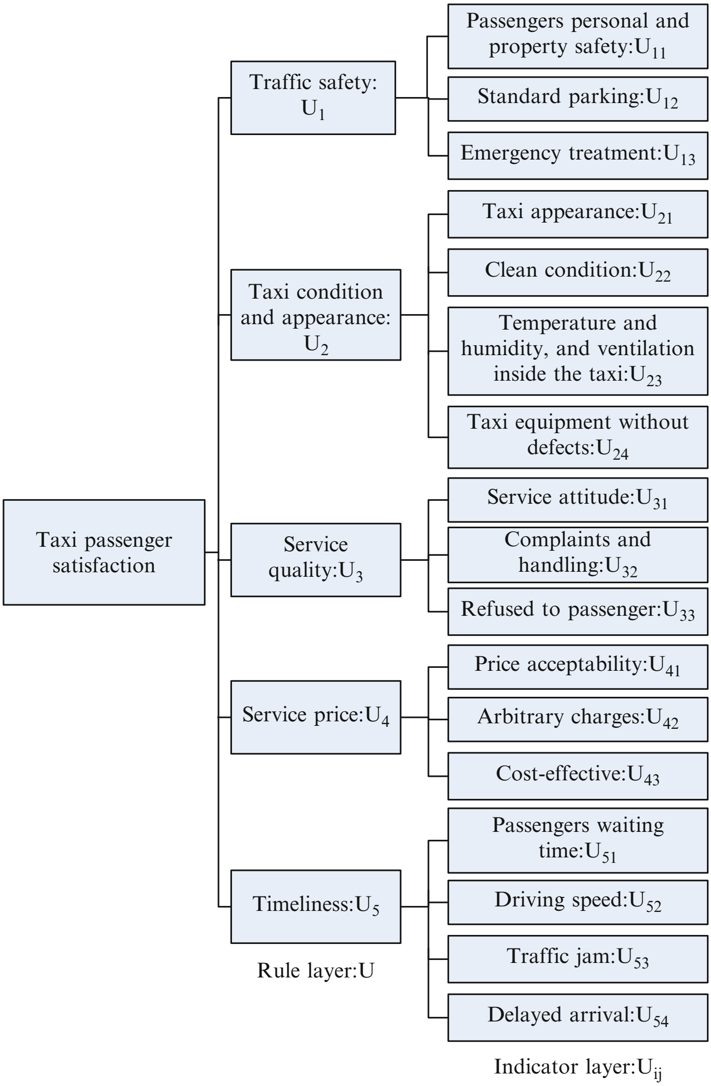


Fig. 78.1 Index system of taxi passenger satisfaction evaluation

Table 78.1 RI value

Size(n)	1	2	3	4	5	6	7
RI	0	0	0.58	0.9	1.12	1.24	1.32

Table 78.2 Evaluation results of various indicators

Rule layer	Indicator layer	Very satisfactory	Satisfactory	Genearl	Dissatisfactory	Very dissatisfactoy
R_1	R_{11}	0.1	0.3	0.4	0.2	0
	R_{12}	0.2	0.4	0.3	0.1	0
	R_{13}	0.1	0.3	0.4	0.2	0.1
R_2	R_{21}	0.2	0.4	0.3	0.1	0
	R_{22}	0.3	0.4	0.3	0	0
	R_{23}	0.1	0.3	0.4	0.2	0
	R_{24}	0.3	0.4	0.3	0	0
R_3	R_{31}	0	0.1	0.4	0.4	0.1
	R_{32}	0	0.1	0.5	0.3	0.1
	R_{33}	0	0.1	0.4	0.4	0.1
R_4	R_{41}	0.3	0.4	0.2	0.1	0
	R_{42}	0.1	0.2	0.5	0.1	0.1
	R_{43}	0.1	0.2	0.4	0.2	0.1
R_5	R_{51}	0	0.1	0.4	0.4	0.1
	R_{52}	0.2	0.4	0.3	0.1	0
	R_{53}	0.1	0.2	0.4	0.2	0.1
	R_{54}	0.1	0.2	0.3	0.3	0.1

3. *Judgment matrix U_2—U_{2j}:*

$$A_2 = \begin{pmatrix} 1 & 1/2 & 1/3 & 1 \\ 2 & 1 & 1 & 2 \\ 3 & 1 & 1 & 4 \\ 1 & 1/2 & 1/4 & 1 \end{pmatrix}, \lambda_{\max} = 4.046, \mathrm{CI} = 0.015, \mathrm{CR} = 0.017 < 0.1, \text{the}$$

judgment matrix U_2—U_{2j} has consistency. $W_2 = (0.141, 0.314, 0.413, 0.132)^{\mathrm{T}}$.

4. *Judgment matrix U_3—U_{3j}:*

$$A_3 = \begin{pmatrix} 1 & 5 & 3 \\ 1/5 & 1 & 1/3 \\ 1/3 & 3 & 1 \end{pmatrix}, \lambda_{\max} = 3.039, \mathrm{CI} = 0.019, \mathrm{CR} = 0.033 < 0.1, \text{the}$$

judgment matrix U_3—U_{3j} has consistency. $W_3 = (0.637, 0.105, 0.258)^{\mathrm{T}}$.

5. *Judgment matrix U_4—U_{4j}:*

$$A_4 = \begin{pmatrix} 1 & 1/4 & 2 \\ 4 & 1 & 3 \\ 1/2 & 1/3 & 1 \end{pmatrix}, \lambda_{\max} = 3.108, \mathrm{CI} = 0.054, \mathrm{CR} = 0.093 < 0.1, \text{the}$$

judgment matrix U_4—U_{4j} has consistency. $W_4 = (0.218, 0.630, 0.152)^{\mathrm{T}}$.

6. *Judgment matrix U_5—U_{5j}:*

$$A_5 = \begin{pmatrix} 1 & 5 & 2 & 2 \\ 1/5 & 1 & 1/5 & 1/3 \\ 1/2 & 5 & 1 & 1/2 \\ 1/2 & 3 & 2 & 1 \end{pmatrix}, \lambda_{\max} = 4.155, \mathrm{CI} = 0.052, \mathrm{CR} = 0.057 < 0.1,$$

the judgment matrix U_5—U_{5j} has consistency. $W_5 = (0.438, 0.070, 0.219, 0.273)^{\mathrm{T}}$.

78.2.3 Establising the Evalutation Set and Comprehensive Evaluation Matrix

The data were collected by the questionnaires. Issued 500 copies of the questionnaire, the recovered questionnaires were 387, effective rate was 77.4 %. The age of investigators divided into 4 ranges, under 18, 18–30, 30–40 and over 40 years old. Each indicator divided into 5 levels of evaluation, to establish the evaluation set V, $V = \{V_1, V_2, V_3, V_4, V_5\}$ = {very satisfactory, quite satisfactory, general, dissatisfactory, very dissatisfactory}, score range is from 0 to 100 points, and score range for each grade are: very satisfactory (over 90 points), quite satisfactory (70–90 points), general (50–70 points), dissatisfactory (30–50 points), very dissatisfactory (under 30 points.). Except "very satisfactory" grade rating, each grade ratings can choose the lower limit as parameters, namely: $C = (1, 0.7, 0.5, 0.3, 0)$, then obtained the results of comprehensive evaluation (Table 78.2), finally to build the total evaluation matrix (Miao Zhi-guo and Zheng Quan-cheng 2010; Niu Hui-yong 2005).

Then to combine two fuzzy subset W_i and R_i to the total evaluation matrix, namely $B_i = W_i{*}R_i$ (Wang Gui-cheng 2008).

$$B_1 = W_1 R_1 = \begin{pmatrix} 0.527 & 0.333 & 0.140 \end{pmatrix} \begin{pmatrix} 0.1 & 0.3 & 0.4 & 0.2 & 0 \\ 0.2 & 0.4 & 0.3 & 0.1 & 0 \\ 0.1 & 0.3 & 0.4 & 0.2 & 0.1 \end{pmatrix}$$
$$= (0.133,\ 0.333,\ 0.367,\ 0.167,\ 0.014)$$

Similarly calculate the B_2, B_3, B_4 and B_5, obtained a total evaluation matrix R. Then,

$$B = W^* R = \begin{pmatrix} 0.168 & 0.055 & 0.193 & 0.202 & 0.382 \end{pmatrix}^* \begin{pmatrix} 0.133 & 0.333 & 0.367 & 0.167 & 0.014 \\ 0.203 & 0.359 & 0.341 & 0.097 & 0 \\ 0 & 0.100 & 0.411 & 0.390 & 0.100 \\ 0.144 & 0.244 & 0.419 & 0.115 & 0.078 \\ 0.063 & 0.170 & 0.366 & 0.308 & 0.093 \end{pmatrix}$$
$$= (0.087,\ 0.209,\ 0.384,\ 0.250, 0.073)$$

Grade calculation result is as follows:

$$F = B^* C = (0.087,\ 0.209,\ 0.384,\ 0.250,\ 0.073)^* \,(1,\ 0.7,\ 0.5,\ 0.3,\ 0)$$
$$= 0.5003.$$

According to the previous evaluation set shows that the satisfaction of the taxi passengers is between general and dissatisfaction. Then using GRA to get the

Table 78.3 The average score of each indicator

Indicator	Passengers (ages)				Comparative sequence X_{ij}
	Under 18	18–30	30–40	Over 40	
U_{11}	65.5	65.2	61.7	72.0	72.0
U_{12}	72.9	65.9	76.3	62.5	76.3
U_{13}	61.1	56.3	57.4	64.0	64.0
U_{21}	64.2	74.4	60.8	71.0	74.4
U_{22}	71.4	82.7	67.7	79.0	82.7
U_{23}	73.4	70.4	66.0	70.0	73.4
U_{24}	74.3	72.1	79.7	80.5	80.5
U_{31}	50.9	48.3	45.7	51.3	55.9
U_{32}	53.6	52.0	50.8	55.2	62.0
U_{33}	50.1	46.7	44.3	51.7	54.1
U_{41}	71.9	79.5	83.3	68.1	83.3
U_{42}	65.0	58.8	68.1	55.7	68.1
U_{43}	57.6	66.7	63.6	54.5	66.7
U_{51}	55.8	48.4	45.0	48.4	56.0
U_{52}	77.6	67.0	74.0	63.5	77.6
U_{53}	58.2	52.6	49.9	60.9	60.9
U_{54}	60.3	52.1	49.3	57.5	60.3

satisfaction evaluation value of each factor and identify the key factors of the taxi passenger satisfaction (Li Juan-fang and Liu Xing 2011).

78.2.4 Getting the Satisfaction Evaluation Value of Each Factor by GRA

1. *Determining the analysis sequence:*

GRA must be sure the reference sequence X_0(the evaluation indicator of n), $X_0 = \{X_0(1), X_0(2), \ldots, X_0(n)\}$, then select the maximum value of satisfaction evaluation of the m-class passengers record as the comparative sequence X_i, $X_i = \{X_i(1), X_i(2), \ldots, X_i(n)\}$ (Zhou Yan-fang and Zhou Leishan 2007; Yang Ya and Qi Xiao-yu 2011). The average score of each indicator are shown in Table 78.3.

2. *Calculating the correlation coefficient* $\xi_i(k)$ (Dai Yi et al. 2008):

$$\xi_i(k) = \frac{\min\limits_i \min\limits_k |X_0(k) - X_1(k)| + \zeta \max\limits_i \max\limits_k |X_0(k) - X_1(k)|}{|X_0(k) - X_1(k)| + \zeta \max\limits_i \max\limits_k |X_0(k) - X_1(k)|}$$

$$(i = 1, 2, \ldots, \mathrm{m}; k = 1, 2, \ldots, n); \tag{78.1}$$

$\xi_i(k)$: the relative difference between X_i and X_0 of the kth factor;

ζ: distinguishing coefficient, $\zeta \in [0, 1]$, in order to reduce the influence of extreme values on the calculation (Miller 1992).

According to Table 78.3, $\min_i \min_k |X_0(k) - X_1(k)| = 0$, $\max_i \max_k |X_0(k) - X_1(k)| = 15.1$, take $\zeta = 0.5$, coefficient matrix E as follows:

$$E = \xi_i(k)_{m*n} = \begin{pmatrix} \xi_1(1) & \xi_1(2) & \cdots & \xi_1(n) \\ \xi_2(1) & \xi_2(2) & \cdots & \xi_2(n) \\ \vdots & \vdots & \ddots & \vdots \\ \xi_m(1) & \xi_m(2) & \cdots & \xi_m(n) \end{pmatrix}$$

$$= \begin{pmatrix} 0.54 & 0.53 & 0.42 & 1.00 \\ 0.69 & 0.42 & 1.00 & 0.35 \\ 0.72 & 0.49 & 0.53 & 1.00 \\ 0.43 & 1.00 & 0.36 & 0.69 \\ 0.40 & 1.00 & 0.33 & 0.67 \\ 1.00 & 0.72 & 0.51 & 0.69 \\ 0.55 & 0.47 & 0.90 & 1.00 \\ 0.60 & 0.50 & 0.43 & 0.62 \\ 0.47 & 0.43 & 0.40 & 0.53 \\ 0.65 & 0.51 & 0.43 & 0.75 \\ 0.40 & 0.67 & 1.00 & 0.33 \\ 0.71 & 0.45 & 1.00 & 0.38 \\ 0.45 & 1.00 & 0.71 & 0.38 \\ 0.98 & 0.50 & 0.41 & 0.50 \\ 1.00 & 0.42 & 0.68 & 0.35 \\ 0.73 & 0.48 & 0.41 & 1.00 \\ 1.00 & 0.48 & 0.41 & 0.73 \end{pmatrix}$$

3. *Calculating the correlation degree r_i:*

$$r_i = \sum_{k=1}^{n} W(k)\xi_\iota(k) \tag{78.2}$$

then $r_{11} = 0.55$, $r_{12} = 0.64$, $r_{13} = 0.60$, $r_{21} = 0.68$, $r_{22} = 0.66$, $r_{23} = 0.67$, $r_{24} = 0.70$, $r_{31} = 0.50$, $r_{32} = 0.44$, $r_{33} = 0.53$, $r_{41} = 0.71$, $r_{42} = 0.66$, $r_{43} = 0.76$, $r_{51} = 0.52$, $r_{52} = 0.56$, $r_{53} = 0.55$, $r_{54} = 0.55$; the order is $r_{43} > r_{41} > r_{24} > r_{21} > r_{23} > r_{22} > r_{42} > r_{12} > r_{13} > r_{52} > r_{11} > r_{53} > r_{54} > r_{33} > r_{51} > r_{31} > r_{32}$ (Qin Peng et al. 2011).

According to the sort of correlation degree r_i, in all the factors, taxi passengers satisfied with taxi cost-price, price and taxi equipment without defects. They dissatisfied with taxi refusing to take passengers, waiting time, the attitude of taxi drivers, complains handling (Li Li 2012). To find out the factors that taxi passengers are less satisfied, in order to propose the measures to improve the overall satisfaction (Lin Xiao-yan et al. 2005).

78.3 Improvements

78.3.1 Raising the Price

Bus and subway are the most important public transport, taxis as a supplementary mode of transport service for the relatively small number of people and special population. The start-price of Wuhan taxi is 6 yuan, 1.4 yuan per kilometer after 2 km, the lower price and underdeveloped public transport make a large number of passengers to choose a taxi trip, to take up the limited resources of the community. If the taxi price are raised, not only can improve the ability of taxi operators, increase drivers' income, but also balance taxis demand by the price leverage.

78.3.2 Developping Public Transport

Public transport is underdeveloped, causing passengers to travel difficulties, and part of the bus passengers diverted to the taxi. Government needs to increase public transport investment to improve the public transport grade, reduce price, raise the service quality of public transport, and avoid the loss of bus passengers on the taxi.

78.3.3 Strengthening Government Management

To install the GPS system for taxi, and establish a taxi dispatch center, form a comprehensive supervision; to increase the on-call service for improving the utilization of the taxi; strengthen the service of handling the passengers complaints for increasing satisfaction rate; severely crack down on a small number of taxi refuse to take passengers; increase the intensity of traffic congestion management; speed up the road widening maintenance and new construction, to fundamentally solve taxi passenger waiting time, traffic jam and raising the turnover rate of a taxi.

78.4 Conclusion

Through F-AHP and GRA, establish evaluation model of taxi passenger satisfaction, then analyze the main factors of the impact on passenger satisfaction (Chen Kai et al. 2011). And propose improvements for increasing satisfaction. Combination of two methods, simple and easy to operate practical, and the conclusion was reasonable and scientific (Yan Zhi-heng et al. 2010).

References

Chen Kai, Xu De-qian, Wang Mei-ting (2011) A study of comprehensive evaluation of urban ecological river bank based on fuzzy hierarchy method. J Hefei Univ Technol (Nat Sci) 34(6):56–58

Dai Yi, Huo Jia-zhen, Zhang Qian (2008) Fuzzy comprehensive evaluation of enterprise internal risk. J Tongji Univ (Nat Sci) 36(6):864–868

Hu Da-li (2003) Applying the grey system theory evaluating the competitive competence of enterprises. Sci Technol Prog Policy 20(1):159–161

Li Hong-ji (2005) The basis of fuzzy mathematics and the use of algorithms. Science Press, Beijing

Li Juan-fang, Liu Xing (2011) Satisfaction evaluation of the construction project stakeholders based on grey clustering method. Technoecon Manag Res 1(7):12–16

Li Li (2012) Fuzzy AHP – based quality evaluation of health examination service. Chinese Health Qual Manag 19(1):53–55

Lin Xiao-yan, Rong Chao-he, Chen You-xiao (2005) Study on comprehensive fuzzy hierarchy analysis of investment effects of the land development railway. J China Railw Soc 27(1):107–113

Lu Yue-jin (2002) Weigh calculation method of fuzzy analytical hierarchy process. Fuzzy Syst Math 16(2):79–85

Miao Zhi-guo, Zheng Quan-cheng (2010) Grey comprehensive evaluation for customer satisfaction index in packaging. Packag Food Mach 28(1):22–24

Miller KD (1992) A framework for integrated risk management in international business. J Int Bus Stud 23(2):311–331

Morash EA (2001) Supply chain strategies, capabilities, and performance. Transportation 3(4):124–126

Niu Hui-yong (2005) Study on safety evaluation of urban road traffic based on gray theory. China Saf Sci J 15(9):92–95

Qin Peng, WangYin-hua, Wang Wei-han, Li Meng (2011) Integrated model of fuzzy analytical hierarchy process and variable fuzzy set model on evaluating river health system. J Zhejiang Univ (Eng Sci) 45(12):2172–2174

Wang En-xu, Wu Chun-you (2009) Study on the satisfaction index of tourist based on gray correlation analysis. Commer Res 392(12):133–135

Wang Gui-cheng (2008) The synthesize evaluation of the enterprise management effect based on the grey statistics method. Hunan Nonferr Metals 24(4):71–73

Wu Qi-bing, Liu Zu-de, Zhao Yun-sheng (2008) Safety assessment model for ironworks based on fuzzy AHP and its application. China Saf Sci J 18(10):136–140

Xu Ge-ning, Jiang Fan (2010) Safety assessment on the crane based on FAHP. J Saf Environ 10(2):196–200

Yan Zhi-heng, Yuan Peng, Huang Yan, Qian Xiao-yan (2010) Study on hierarchical fuzzy recognition model for evaluation of water saving society establishment. Water Resour Power 28(4):35–39

Yang Ya, Qi Xiao-yu (2011) Evaluation of passenger's satisfaction about No1. Subway of Chengdu metro. Technol Econ Areas Commun 13(6):48–50

Zheng Xin, Zhao Zhi-min, Xu Jian (2010) Study on degree of passenger satisfaction to a certain subway station based on fuzzy comprehensive evaluation method. Technol Econ Areas Commun 12(6):41–43

Zhou Yan-fang, Zhou Leishan (2007) Establishing assessment system about passenger satisfaction indices in urban rail transit. Urban Rapid Rail Transit 20(5):25–28

Chapter 79
The Impact of Consumption Emotions on Service Quality and Customer Satisfaction

Guo-li Pu and Qin Su

Abstract Although the important role of service quality and customer satisfaction has been generally recognized, almost all the studies have been based on the cognition. Recent researches suggest that consumption emotions have important impact on both of them. Based on a framework of "Attribution—Emotion—Action" come up by Weiner, this article develops a model including emotion, cognition, service quality and customer satisfaction. The model and hypotheses are empirically tested, the results indicate that consumption emotions are significantly related to consumer cognitions, both of which have significant influence on total service quality and customer satisfaction. Positive emotions have positively significant influence while negative emotions of self-attribution and external attribution are significantly and negatively related to total service quality and customer satisfaction.

Keywords Attribution theory • Consumption emotions • Customer satisfaction • Service quality

---

This research is supported by the Special Foundation of Key Disciplines of Shaanxi Province (Grant: 00X901) and the Soft Science Foundation of Shaanxi Province (Grant: 2012KRM118)

G. Pu (✉)
School of Economics and Management, Xi'an University of Technology, Xi'an, PR China

School of Management, Xi'an Jiaotong University, Xi'an, PR China
e-mail: pgl@xaut.edu.cn

Q. Su
School of Management, Xi'an Jiaotong University, Xi'an, PR China
e-mail: qinsu@mail.xjtu.edu.cn

E. Qi et al. (eds.), *The 19th International Conference on Industrial Engineering and Engineering Management*, DOI 10.1007/978-3-642-37270-4_79,

79.1 Introduction

There is an increasing interest in consumption emotion and customer behavior in recent years, and great progress has been made in this field based on the theory of psychology and marketing. Compared with the traditional manufacturing industry, customers require more energy and emotional devotion in the service process so as to cope with the interpersonal interactions commonly existing in the service industry. Once understanding the consumption emotions well and taking proper assessments, customer service staff can improve the service efficiency and service quality effectively. With the service quality getting more attentions (such as banks queuing events), it has greater practical significance in improving the service quality through more exploration in the field of consumption emotions.

In this paper, the authors analyze the mechanisms how consumption emotion affects customer satisfaction and evaluation of customer service quality. Following is the structure of this paper. The authors first review literatures in the field of service quality, customer satisfaction and consumption emotion. The conceptual model (as shown in Fig. 79.1) and hypotheses are then presented. Third, the authors describe the study method and test the model on the basis of data from questionnaire survey. Finally, the authors summarize the major findings and the implications, as well as directions for future research.

79.2 Theory Background

79.2.1 Service Quality

Grönroos (1982) put forward the concept of perceived service quality, namely that the service quality depends on the gap between expected quality and the experienced

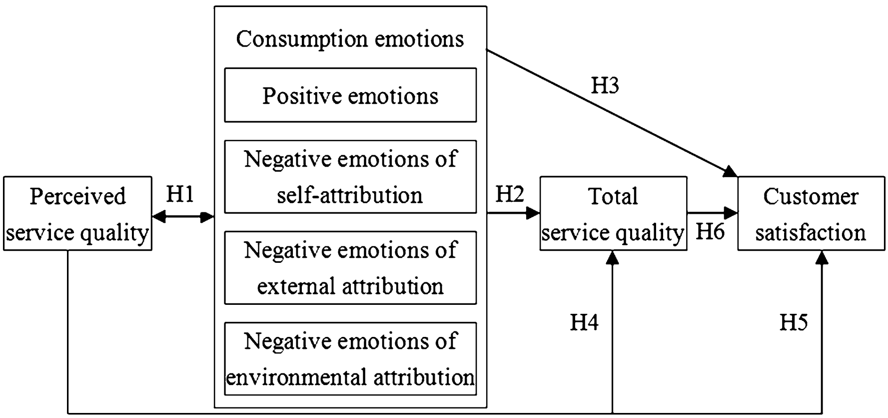


Fig. 79.1 Conceptual model

quality of customers, which was still the most important theoretical basis in the field of service quality. Based on the gap model of service quality, Parasuraman, Zeithamal and Berry (referred to as PZB) (1991) developed an important scale called “SERVQUAL” to measure service quality. However, both the widely used “perceived service quality” concept and “SERVQUAL” scales are based on the customers’ cognition and assessment. Recent studies have noted that customer’s assessment of products and service not only attributes to rational cognition, but also their subjective emotions and feelings. Edvardssen (2005) suggests that cognitions and emotions together construct the perceived service quality. Since the purpose of this paper is to compare the role of consumer cognitions and consumption emotions that used to explain the total service quality, the concepts of perceived service quality and total service quality are differently defined in order to assess the service quality more accurately. SERVQUAL are employed to assess perceived service quality while cognitions and emotions are included into the total service quality.

79.2.2 Customer Satisfaction

It’s broad consensus that customer satisfaction is the model of gap between customer expectations and actual perceptions. The model was developed by Oliver (1980) and tested and supported by many scholars on basis of a large number of empirical studies. The impact of cognition process on customer satisfaction is emphasized but the rules of emotional factors are ignored. PZB (1985) argued that the most significant factor that affects the customer satisfaction is not the gap but consumption emotion. Westbrook (1987) agreed with that emotional factors and cognitive factors have equal impact on customer satisfaction. Oliver (2000) also suggested that the role of emotion variable be investigated. However, traditional model is still generally employed in the field of customer satisfaction; little attention has been given to the impact of consumption emotion on customer satisfaction.

79.2.3 Consumption Emotions

Westbrook and Oliver (1991) defined consumption emotion as “a specific set of emotional response that triggered during the use of product or consumption experience”. Menon and Dube (2000) argued that “consumption emotions are the consumers’ emotional response to product attributes and value of the consumption”. It’s generally accepted that consumption emotions are divided into positive emotions and negative emotions, Westbrook (1987) pointed out that the positive and negative emotions are independent, that is, the customer may experience the excitement, pleasure and other positive emotions as well as worries, anger and other negative emotions during the same service process. This conclusion suggests that human emotion is a complicated psychological process and the different emotions have different impacts on customer satisfaction and assessment of service quality. It's generally supported that the positive emotions is positively related to customer satisfaction and assess of service quality. But there has no general conclusion about the relationship between negative emotions and customer satisfaction. In order to solve the problem, Oliver (2000) suggested negative emotions need to be classified and he divided negative emotions into three categories: external attribution, self-attribution and environment attribution on basis of attribution theory, which is helpful to in-depth and comprehensively interpret the relationship between the negative emotions and customer satisfaction.

Although there are lots of literatures existing in the field of service quality and customer satisfaction, most of which are based on cognition model. Little attention has been paid to relationship between consumption emotions and service quality and customer satisfaction. Understanding about consumption emotions are quite different, the studies about the impact of emotion attribution on service quality and customer satisfaction are limited. This paper tries to resolve the following issues: (1) taking account the cognitive and emotional variables into the impact model of service quality and customer satisfaction so as to make a more comprehensive understanding and more accurate interpretation of them; (2) exploring the impact of consumption emotions on service quality and customer satisfaction employing the empirical method.

79.3 Model and Hypothesis

79.3.1 Conceptual Model

The conceptual model is constructed based on the Weiner's (1985) theory of "attribution-emotion-action" (as shown in Fig.79.1) (Weiner and Handel 1985). According the Weiner's action chain, customers will attribute the service process and effect, and different consumption emotions attribute to different effects, which will affect the assessment of total service quality and customer satisfaction. Cognitions and emotional behaviors are interactive when customers assess the cognitions and emotions simultaneously.

79.3.2 Hypothesis

That there is a close relationship between the cognitions and emotions has been generally recognized by psychologists. It's inevitable that emotions involve in the process of cognition, and emotions have significant impact on activities of information processing, attention, memory and decision-making, etc. Additionally positive and negative emotions have contrary effect on cognitions. At the same time emotions are regulated by the cognition process. Thus, we propose the hypothesis H1 as follows.

H1 consumption emotions will be interactive with assessment of cognitions.

Izard's (1977) emotional scale (DES II) is generally used to assess the emotions in the field of psychology, which divides the emotion into positive and negative emotions. According to the theory of Smith and Ellsworth (1985), negative emotions are divided into external attribution, self-attribution and environmental attribution of negative emotions in this paper. It has been discussed that the consumption emotions and customer cognitions are the basic properties of service quality and customer satisfaction and have a significant impact on both of them. Additionally different negative emotions have different impact on customer satisfaction and service quality. For example, Dube and Menon (2000) found that only negative emotions of external attribution will lead to customer dissatisfaction, while the others are positively related to post-purchase satisfaction. We argue that negative emotions of environmental attribution have significant impact on customer satisfaction, but there is not an absolute positive or negative relationship between them. Therefore, the following hypotheses are proposed.

H2a positive emotions will positively influence total service quality directly;
H2b negative emotions of self-attribution will positively influence total service quality directly;
H2c negative emotions of external attribution will negatively influence total service quality directly;
H2d negative emotions of environmental attribution will be significantly related to total service quality;
H3a positive emotions will positively influence customer satisfaction directly;
H3b negative emotions of self-attribution will positively influence customer satisfaction directly;
H3c negative emotions of external attribution will negatively influence customer satisfaction directly;
H3d negative emotions of environmental attribution will be significantly related to customer satisfaction.

Traditional studies on service quality and customer satisfaction are based on evaluation of customer assessment of cognitions, and conclude that cognitive processes significantly and positively affect customer service quality and satisfaction. At the same time the relationship between customer satisfaction and service quality is the focus in the field of service marketing, and that there is a significant correlation between them is generally agreed.. Accordingly, this gives:

H4 assessment of cognition will positively influence total service quality;
H5 assessment of cognition will positively influence customer satisfaction;
H6 total service quality will positively influence customer satisfaction.

79.4 Methodology

79.4.1 Survey Instrument

To test the hypotheses, a questionnaire survey method was used. All measurement items were picked up from existing literatures, and the authors modified them according to interviews and investigations. Reliability and validity of items were analyzed and guaranteed. All multi-item scales were measured on a seven-point Likert scales (1 = "strongly disagree" to 7 = "strongly agree").

79.4.2 Sample

Undergraduate and graduate students in the school of management were selected as convenient simple and all of them have the experience of banking services. One reason was that respondents could better understand and complete the questionnaire, which would reflect their true desires. The other is that it has been judged using a sample with a high degree of similarity can make more accurate forecasts and more efficient testing of theories. Three hundred responses were received, of which 245 were complete and usable (the effective rate is 82.7). Fifty-three percent was male, 88.3 % was 25–30 years old, 24 % was undergraduates, and more than 50 % of responses had the experience of banking services for 3 years and surplus, which were representative.

79.5 Analysis and Findings

79.5.1 Data Analysis

The internal consistency of the measures of the questionnaire is between the Cronbach's alpha 0.75 and 0.9, which indicates all scales demonstrate good reliability. In order to test the efficacy of sample data, LISERL 8.53 software was employed to take confirmatory factor analysis. Analysis results indicated that factor loadings of all the indicators is highly significant, indicating that the sample data has good convergence validity. Meanwhile, the variances of the latent variable indicators are greater than the covariance, indicating that the sample data has good discriminant validity. In addition, Oliver and Smith's classification of emotions were confirmed.

79.5.2 Test of Model

The structural model based on a path analysis was estimated after achieving a satisfactory goodness in the measurement model, and path analysis produced the

Table 79.1 Path analysis of structural model

Hypotheses	Standardized loading	Path	Result
H1	0.39**	PE → PSQ	Support
H1	−0.12	SE → PSQ	Reject
H1	−0.36**	OE → PSQ	Support
H1	0.11	EE → PSQ	Reject
H2a	0.41**	PE → TQ	Support
H2b	−0.24**	SE → TQ	Reject
H2c	−0.27**	OE → TQ	Support
H2d	0.07	EE → TQ	Reject
H3a	0.25**	PE → CS	Support
H3b	−0.15**	SE → CS	Reject
H3c	−0.41**	OE → CS	Support
H3d	0.18	EE → CS	Reject
H4	0.36**	PSQ → TQ	Support
H5	0.22**	PSQ → CS	Support
H6	0.26**	TQ → CS	Support

Note: (1) *PE* passive emotions, *SE* negative emotions of self-attribution, *OE* negative emotions of external attribution, *EE* negative emotions of environmental attribution, *PSQ* perceived service quality, *TQ* total service quality, *CS* customer satisfaction. (2) **means significance at level 0.05

following fit statistics for structural model: χ^2/df = 2.85, NFI = 0.94, NNFI = 0.98, CFI = 0.98, IFI = 0.97, GFI = 0.91, RFI = 0.96, SRMR = 0.056(less than 0.08), RMSEA = 0.069 (less than 0.1), all the fit indexes are well accepted.

79.5.3 Test of Hypothesis

Standardized loadings are given in Table 79.1. The result indicates that: (1) positive emotions have significant positive impact on perceived service quality, negative emotions of external attribution have significant negative impact on perceived service quality, H1 are partially supported; (2) positive emotions have significant positive impact on total service quality and customer satisfaction, while negative emotions of external attribution have significant negative impact on total service quality and customer satisfaction, H2a, H3a, H2c, H3c are supported. (3) H2b, H3b are negative significant and opposite to the hypothesis, H2d, H3d are not significant, all of which are rejected; (4) customer perceived service quality have significant positive impact on total service quality and customer satisfaction, H4 and H5 are supported; (5) the total service quality have significant positive impact on customer satisfaction, H6 is supported.

79.6 Discussion

The results validate the Oliver's classification of negative emotions, which can be divided into positive emotions, negative emotions of external contribution, self-contribution and environmental contribution. And different emotions have different impact on service quality and customer satisfaction.

The results suggest that positive emotions are positively and significantly related to consumer cognitions, negative emotions of self-attribution and external attribution are negatively and significantly related to consumer cognitions, which agrees with the relationship between emotions and cognitions in the field of psychology. It also suggests that customer emotions are complicated psychological phenomenon, positive and negative emotions may exist simultaneously.

The results indicate that emotions and cognitions play the important role in the process of customers' assessment of service quality and formulation of satisfaction. Consumption emotions not only directly affect the customer satisfaction and total service quality but also indirectly affect them through cognitions. Additionally Dube and Menon (2000) suggests that negative emotions of environmental contribution are positively related to customer satisfaction, which is not supported in this study. This may attribute to background of banking services which has a high standard in environment. And customers will have little emotion wave when environment is changed. There is no inevitable relationship between negative emotions of environmental contribution and customer satisfaction, which will depend on the type of service industry.

79.7 Conclusions

Most of the studies about service quality and satisfaction are based on customer assessment of cognitions; the role of consumption emotions has not been well recognized. This study employs empirical method to investigate the relationship between consumption emotions, customer cognitions and service quality and customer satisfaction. It is believed that the findings are to enrich the achievements in the field of emotions, providing theoretical and practical insights to help understand the service quality and customer satisfaction.

However, there are also some limitations existing in this study. First, although the samples are representative, the sample are limited to college students and needs to be enriched; Secondly, this study just focuses on the background of banking services, but different background may affect the results. Future studies should expand the model to other types of service industries and test the validity and generality.

References

Dube L, Menon K (2000) Multiple roles of consumption emotions in post-purchase satisfaction with extended service transactions. Int J Ser Ind Manag 11:287–304

Edvardssen B (2005) Service quality: beyond cognitive assessment. Manag Ser Qual 15:127–131

Grönroos C (1982) An applied service marketing theory. Eur J Mark 16:30–41

Izard E (1977) Human emotion. Plenum Press, New York

Menon K, Dube L (2000) Ensuring greater satisfaction by engineering salesperson response to customer emotions. J Retail 76:285–307

Oliver R (1980) A cognitive model of the antecedents and consequences of satisfaction decisions. J Mark Res 17:460–469

Oliver R (2000) Customer satisfaction with service. In: Handbook of service marketing and management. Sage Publications, Thousand Oaks

Parasuraman A, Zeithaml V, Berry L (1985) A conceptual model of service quality and its implications for future research. J Mark 49:41–50

Parasuraman A, Zeithaml V, Berry L (1991) Refinement and reassessment of the servqual scale. J Retail 67:420–450

Smith CA, Ellsworth PC (1985) Patterns of cognitive appraisal in emotion. J Pers Soc Psychol 48:813–838

Weiner B, Handel SJ (1985) A cognition-emotion-action sequence: anticipated emotional consequences of causal attributions and reposed communication strategy. Dev Psychol 21:102–107

Westbrook R (1987) Product/consumer-based affective responses and postpurchase processes. J Mark Res 24:258–270

Westbrook R, Oliver R (1991) The dimensionality of consumption emotion patterns and consumer satisfactions. J Consum Res 18:84–91

Chapter 80
Exploratory Study on the Healthy Organization in Manufacturing Enterprise

She-hong Liang, Lei Xing, Kan Shi, and Peng Gao

Abstract This study explores the relationship between the healthy organization and three structural dimensions and organizational performance by 188 subjects in a manufacturing company. The results show that three structural dimensions of healthy organization that include organization's physical and mental harmony, competence and efficiency, innovation and development have positive impact on organizational performance by controlling for three demographic variables: age, seniority and education. Healthy organization also has positive impact on organizational performance.

Keywords Competence and efficiency • Healthy organization • Innovation and development • Organizational performance • Physical and mental harmony

80.1 Introduction

As one of the world's fastest growing and most competitive markets, Chinese entrepreneurs and employees are facing increasing pressure. More and more entrepreneurs and employees frequently suffer from psychological and emotional problems. Psychological tension, depression, anxiety, pain, guilt and loss of confidence have become psychological defects of new generation of professional

Fund Project: NSFC China (No. 90924007), China 973 Project (No. 2010CB731406)

S. Liang (✉) • L. Xing • K. Shi • P. Gao
School of Management, Graduate University of Chinese Academy of Sciences, Beijing, China
e-mail: LLLSH@126.com

E. Qi et al. (eds.), *The 19th International Conference on Industrial Engineering and Engineering Management*, DOI 10.1007/978-3-642-37270-4_80,

groups. Psychological crisis of young people not only bring pain and suffering to themselves, but also cause great resistance to the business goals. Therefore, attention to mental and behavioral health of employee, promoting "Organizational Health", building "Healthy Organization" has been becoming a focus for researchers and business managers (Argyris 1958).

World Health Organization clearly articulated in its founding charter in 1948: "Health is a kind of physical, psychological and social perfection, not merely the absence of disease and weakness of the state"(World Health Organization 1946). That is, health is unity of three aspects which include physical, psychological and social function. Mental health is an integral part of overall health, which includes positive mental health status, life stress and effective recovery, outstanding and effective work, loose and innovative organizational culture. It can contribute to society (World Health Organization 2003). Siu OL proposed from the point of individual resilience that individual health should be divided into three levels, which include body, mind and spirit (Siu 2006). Body refers healthy lifestyle, balanced diet and regular exercise. Mind refers to develop good psychological quality, avoid low self-esteem and negative inertia ideas. Spirit involves higher level, including values and outlook on life, knowing how to face failure and avoid over-emphasis on the material pursuit.

Research on Organizational Health rose in the late 1990s. An organization, community and society are also good or bad as human health. People agree that health is a state worth to get because healthy organization and its structure, culture and management processes would help to achieve a high level of performance (McHugh and Brotherton 2000). Cooper and Cartwright pointed out that healthy organization is not only successful and financial, but also having physical, mental health staff (Cooper and Cartwright 1994). It can maintain a longer period of health and satisfying work environment and organizational culture, especially during the period of market turmoil and change. In new economic era, some scholars defined Organizational Health as a competitive organization that can operate properly and effectively, both enhancing internal capabilities and adapting to changes of external environment, in which employee have satisfactory performance and health body, organization maintain sustainable development and have good social and economic benefits (Ruibin et al. 2006).

Kan Shi and Rui Zheng firstly proposed the concept of Healthy Organization from the point of body, mind and spirit of organization, as Fig. 80.1 shown. The normal state of mind refers to "physical and mental harmony and health of the organization. Successful competence refers to "competence and efficiency". Innovative organizational culture advocates rational concept of happiness and building a harmonious, competent, innovative organizational culture (Kan Shi and Rui Zheng 2007). Based on this theory, Dr. Shi group developed nine dimension of Healthy Organization Questionnaire which includes organizational culture, social responsibility, labor relations, psychological feelings, coping with stress, staff counseling, transformational leadership, behavioral health, organizational learning.

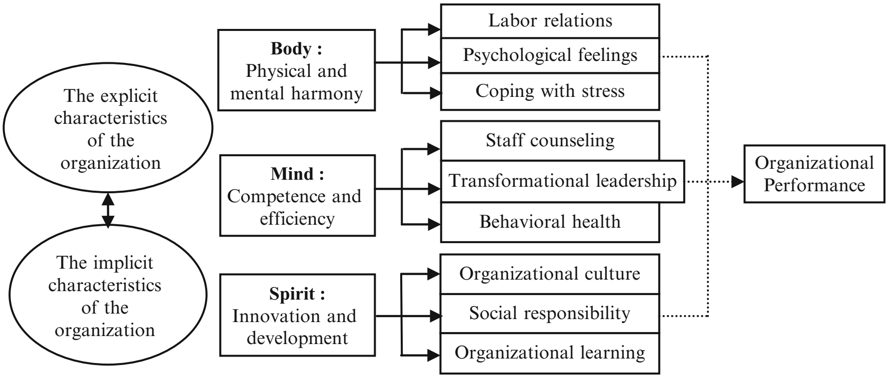


Fig. 80.1 Body, mind and spirit model of healthy organization

The questionnaire is used to diagnose organizational health and its findings will be used to provide professional advice for some enterprises which want to build healthy organization.

The current research paradigm of organizational health is divided into two categories (Xinqiong Wang and Weizheng Chen 2008). On the one hand, it is divided into individual level and organizational level. On the other hand, it is divided into independent variables and dependent variables. The current research of healthy organization is used as dependent variable, the major study concentrated on the relationship between leadership and organizational health. Quick et al. proposed that healthy leader is center of healthy organization and can promote individual health and organizational health (Quick et al. 2007). Keller RC and Stang PE (2006) proposed that healthy leadership model would have a high degree of individual and organizational health, which include vitality, productivity, flexibility and so on. Luthans (2002) have shown healthy leader and their attitudes about health will affect the health of other people in the organization and organizational health outcomes. In addition, some scholars study organizational culture, leader style, trust, decision-making involved and other factors that affect healthy organization (Avolio et al. 2004; Bruhn 2001). Organizational conditions on health as independent variables, such as Edwards GP and Gill R pointed out that when individuals and organizations act in a healthy way they will get the results which have a positive impact on a large number of stakeholders in the organization (Edwards and Gill 2006). Therefore, this study predicts Healthy Organization will have a positive impact on organizational performance.

In summary, this study has two basic hypotheses. First, three dimension of healthy organization which include physical and mental harmony, competence and efficiency, innovation and development have positive impact on organizational performance. Second, healthy organization has positive impact on organizational performance.

80.2 Methods

80.2.1 Participants

Hypotheses were tested using data from a survey of front-line managers in a manufacturing enterprise in Yunnan. The response rate was 90 % and the final analysis was based on a sample of 180 participants. The sample contained 19 % females and 81 % males. The average age was 39, and the average tenure period was 19 years.

80.2.2 Measure

Physical and mental harmony was measured with three-item ($\alpha = 0.778$) which include psychological experience of employee, corporate labor relations and coping with stress. Front-line managers were asked to response scale ranges from 1 = "very poor", to 7 = "very good".

Competence and efficiency was measured with three-item ($\alpha = 0.706$) which include transformational leadership, staff-caring and behavioral health. Using from 1 = "very poor" to 7 = "very good" Likert scale measurement.

Innovation and development was measured with three-item ($\alpha = 0.753$) which include innovative culture, social responsibility and organizational learning. Using from 1 = "very poor" to 7 = "very good" Likert scale measurement.

Healthy Organization was measured with Physical and mental harmony, Competence and efficiency, Innovative culture three scales ($\alpha = 0.872$). Using from 1 = "very poor" to 7 = "very good" Likert scale measurement.

Organizational performance was measured with eight-item ($\alpha = 0.856$). Using from 1 = "very poor", to 7 = "very good" Likert scale measurement.

80.3 Results

80.3.1 Questionnaire Internal Reliability and Correlation Analysis Between the Main Variables

Table 80.1 lists the mean, standard deviation and correlation coefficient of various research variables. The internal reliability coefficient of the questionnaire used in this study is 0.70 above. The mean of physical and mental efficiency, competence and efficiency, innovation and development, and healthy organization are 3.88 (SD = 1.07), 3.62 (SD = 1.19), 3.98 (SD = 1.09), 3.84 (SD = 1.00). It shows organizational health level are in above average. Physical and mental harmony, competence and efficiency, innovation and development of organization, healthy organization has positive impact on organizational performance.

Table 80.1 Mean, standard deviation and correlation coefficient of study variables

Variables	Mean	Standard deviation	1	2	3	4	5	6	7
Physical and mental harmony	3.88	1.07							
Competency and efficiency	3.62	1.19	0.73**						
Innovation and development	3.98	1.09	0.69**	0.67**					
Health organization	3.84	1.00	0.90**	0.90**	0.88**				
Organizational performance	4.40	1.28	0.54**	0.55**	0.50**	0.60**			
Age	2.32	0.83	0.00	0.01	−0.04	−0.037	0.01		
Seniority	2.37	0.88	0.00	0.00	−0.03	−0.025	−0.02	0.89**	
Education	1.96	0.94	−0.17	−0.13	−0.21*	−0.20*	−0.11	−0.50*	−0.41**

Note: $^{*}p < 0.05$, $^{**}p < 0.01$

Table 80.2 Regression results for Hypotheses 1

	Organizational performance			
Variable	β	R2	ΔR2	F
Step1: Control variable				
Age	0.11			
Seniority	−0.02			
Education	0.00			
		0.07		5.89
Step2: Main effect				
Physical and mental harmony	0.26*			
Competency and efficiency	0.30**			
Innovation and development	0.27*			
		0.33	0.26**	25.96**

Note: *p < 0.05, **p < 0.01

Table 80.3 Regression results for Hypotheses 2

	Organizational performance			
Variable	β	R2	ΔR2	F
Step1: Control variable				
Age	0.11			
Seniority	−0.10			
Education	0.01			
		0.01		0.55
Step2: Main effect				
Health organization	0.55**			
		0.30	0.29**	61.97**

Note: **p < 0.01

80.3.2 Each Variable on the Dependent Variable Regression Analysis

Table 80.2 lists the results of regression analysis that three dimensions of healthy organization effect on organizational performance. Table 80.2 shows, after controlling for three demographic variables regression effect of age, seniority and educational background, organization's physical and mental harmony has positive impact on organizational performance ($B = 0.26$, $p < 0.05$), competence and efficiency of organization has positive impact on organizational performance ($B = 0.30$, $p < 0.01$), organizational innovation and development has a positive impact on organizational performance ($B = 0.27$, $p < 0.05$). Therefore, this study verified the hypothesis one.

Table 80.3 lists the results of regression analysis that healthy organization effects on organizational performance. As shown in Table 80.3, after controlling for demographic variables of age, seniority and educational background, healthy organization has a positive effect on organizational performance ($B = 0.55$, $p < 0.01$). Therefore, this study validated the hypothesis two.

80.4 Discussion

With rapid changes in organizational environment, such as global competition, corporate mergers and acquisitions, organizational flat, flexible working, outsourcing and other new trends, all kinds of workers face various pressures and challenges. Work stress effects on individual health and it ultimately affects organization, which has been becoming the focus of study in the past few years. Practitioners make efforts to seek work environment which help to implement organizational change and enhance employee health, so that employees and organizations can avoid negative consequences. With a deep understanding of the concept of organizational health and initial exploration of building health-based organizations, we can further improve the structural elements of Healthy Organization and measurement tools. So we can implement more effective strategies and measures which improve the health status of employees and job competence in workplace. It is very significant to theoretical and practical research.

This study verified the hypothesis that three level of Healthy Organization, which is physical and mental harmony, competence and efficiency, innovation and development of healthy organization, effect organizational performance. This is an innovation of theory of Healthy Organization in Chinese culture. In this study, the main effect of healthy organization and its three structural dimensions are significant.

From practical aspect, we can perform a general survey of Chinese enterprises and provide more targeted measures of organizational change by exploring the structural dimension and measurement tools of Healthy organization. First, organization should encourage innovation and learning, support staff development and bear social responsibility. Second, employee should have healthy behavior and core competence. Business managers have to have capabilities of psychological counseling and transformational leadership. Last, organization should have harmonious labor relations which employee in the organization can effectively deal with pressure to achieve harmonious development of body and mind. This requires business executives to establish strategic concept of building healthy organization and combine actual development stage of company to build healthy organization in trying new methods, so as to promote Chinese enterprises to a new management level.

This study has some shortcomings. First, the study sample focused on a business, and insufficient sample size, which to some extent, limit the application of the results to other companies, industries or cultural background. Second, the evaluation of organizational performance is with a certain degree of subjectivity which may not fully reflect actual performance of organization. Third, the classification and measurement of Healthy Organization in academic has not been agreed. Therefore, this study proposes three dimensional structure of Healthy Organization which needs further empirical exploration in the future.

80.5 Conclusion

The results show that Healthy Organization and its three structural dimensions that include organization's physical and mental harmony, competence and efficiency, innovation and development have positive impact on organizational performance. In practice of management, a comprehensive understanding the characteristics and structure of Healthy Organization, effectively diagnose organizational health and build healthy organization will be a new ways that meet the fierce competition in global market, rationally develop and effectively use of corporate human resources.

Acknowledgment This paper is supported by NSFC China No. 90924007 and China 973 Project No. 2010CB731406, to the third author.

References

Argyris C (1958) The organization: what makes it healthy? Harv Bus Rev 36(6):107–116

Avolio BJ, Gardner WL et al (2004) Unlooking the mask a look at the process by which authentic leaders impact follower attitudes and behaviors. Leadersh Quart 15(6):801–823

Bruhn JG (2001) Trust and the health of organizations. Kluwer Academic/Plenum Publishers, New York

Cooper CL, Cartwright S (1994) Healthy mind, healthy organization – a proactive approach to occupational stress. Hum Relat 47(4):455–461, New York

Edwards GP, Gill R (2006) Self-awareness of full range leadership behavior and followers. The British Academy of Management

Kan Shi, Rui Zheng (2007) The thinking of building health organization. J Cap Univ Econ 9(1):12–19 (in Chinese)

Kessler RC, Stang PE (2006) Health & work productivity. University of Chicago Press, Chicago

Luthans F (2002) The need for and meaning of positive organizational behavior. J Organ Behav 23(6):695–706

McHugh M, Brotherton C (2000) Health is wealth-Organizational Utopia or Myopia. J Manag Psychol 1(8):744–770

Quick JC, MacikFrey M, Cooper G (2007) Managerial dimensions of organizational health: the healthy leaderwork. J Manag Stud 44(2):189–205

Ruibin Hon, Zhihong Li, Zhaoming Liu, Zhiming Zhou (2006) From the perspective of organizational culture to explore the meaning of healthy organization. In: Fifth Hong Kong business psychology academic and practical seminars (in Chinese)

Siu OL (2006) Resilience. Peking Bus Rev 1(4):72–74

World Health Organization (1946) Preamble to the constitution of the World Health Organization. International Health Conference, New York

World Health Organization (2003) Investing in mental health. World Health Organization, Geneva

Xinqiong Wang, Weizheng Chen (2008) Review of organization health research. East China Econ Manag 22(5):134–139 (in Chinese)

Chapter 81
Study on Customer Satisfaction of Express Enterprises Based on AHP-FCE

Hai-yan Qi and Jiang-bo Zheng

Abstract Express industry, a fast-developing sector, is facing increasingly fierce competition. As a service industry, the aim of express enterprises is to satisfy customers. Therefore, it's necessary to apply quantitative analysis on customer satisfaction of express companies. For this reason, this article established the express enterprise customer satisfaction evaluation model and the customer satisfaction index system. Then this article applied the model to an express enterprise and evaluated its customer satisfaction using Analytic Hierarchy Process and Fuzzy Comprehensive Evaluation. In this way, this article provides an effective method and tool for express enterprises to identify customers' demand precisely as well as enhancing their competitiveness.

Keywords AHP • Customer satisfaction • Express • Fuzzy comprehensive evaluation

81.1 Introduction

The express industry started late in China. It's not until 1980 that China Post set up Express Mail Service. In 1993, SF Express and STO Express set up in Pearl River Delta and Yangtze River Delta respectively. In early 1994, ZJS Express was

Foundation: This paper is sponsored by the Research & Development Fund of Shenzhen Science and Technology (JC201105201194A).

H. Qi (✉)
School of Management, Shenzhen polytechnic, Shenzhen, China
e-mail: hyqi@szpt.edu.cn

J. Zheng
School of Management, Jinan University, Guangzhou, China
e-mail: zhengjbjnu@126.com

E. Qi et al. (eds.), *The 19th International Conference on Industrial Engineering and Engineering Management*, DOI 10.1007/978-3-642-37270-4_81,

founded in Beijing. Since the 1990s, with the booming of the economy in our country, the social economic activities are increasing. Against this background, people are becoming more and more demanding on delivery time. Not only documents and parcels, but items of high-value, low volume and customized are being delivered by express companies now. It's reported that there are more than 10,000 express enterprises in our country, which can be divided into three categories, stated-own companies, foreign companies and private companies. At present, the category of private express enterprise is the largest for there are more than 2,000 business entities and about 7,000 branches in this category. In recent years, the thriving rise of e-commerce brings dramatically growth to the express industry. While the express industry is growing rapidly and keeps bringing benefits to consumers and the society, express companies are encountering bottlenecks and defections. Customers are more demanding on express service and competition among express enterprises are intensifying. Therefore, understanding the needs of customers accurately, enhancing service continually and improving customer satisfaction are essential for an enterprise to march toward success in today's competitive market place. For this reason, this research established the express industry customer satisfaction evaluation model and the index system and applied Fuzzy Comprehensive Evaluation to assess customer satisfaction, which enables express enterprises to understand customers' need as well as their own limitations. In this way, this article provides an effective method and tool for express enterprises to enhance their competitiveness.

81.2 Express Enterprise Customer Satisfaction Indexes Selection

Scholars at home and abroad have undertaken a lot of research into customer satisfaction of the express industry. The theory of Logistics Service Quality was first proposed by 7Rs theory (Perrault and Russ 1974). JACKSON et al. maintained that the protection of packaging and the cooperation between logistics companies and customers were included in the field of logistics service. LALONDE maintained that the goal of logistics quality is to improve customer satisfaction (Lalonde and Zinszer 1976). MENTZER et al. pointed out that physical distribution service quality could be measured by a three dimension construct composed of availability, timeliness, and quality from the customers' perspective. In 1989 (Mentzer et al. 1989), PARASURAMAN et al. developed SERVQUAL, which was a watershed for the study on logistics service quality (Parasuraman et al. 1985). BRENSINGER et al. applied SERVQUAL to industrial purchasing of motor carrier transportation services and developed a four-factor structure service quality scale. On the basis of their own work (1999), MENTZER et al. put forward the customer-oriented model LSQ (Mentzer et al. 1999, 2001; Bienstock et al. 1997), which is conceptualize in temporal order. Ronald H. Ballou divided the components of customer service into three types, pre-transaction, transaction, and post-transaction factors, and selected important components from these three categories to establish logistics service quality evaluation indicators (Huiskonen and Pirttila 1998). The latest complete description of logistics service quality was given by Tennessee University in 2001(Chiu 2002). Upon the in-depth survey of large third party logistics enterprises and consumers, they summarize nine indicators for the assessment of logistics service quality from customers' perspective, including personnel contact quality, order release quantities, information quality, ordering procedures, order accuracy, order condition, order quality, order discrepancy handling and timeliness.

Based on the general model of customer value realization, Li Wenli decomposed the evaluation of customer satisfaction into five secondary indicators, which are customer expectation, customer perception of service quality, customer complaints, and customer retention. With further analysis and refinement of five secondary indicators, Li Wenli established customer satisfaction evaluation index model of logistics system (Xu Jincan et al. 2002). Wang Jing and Li Jin-fei (2006) pointed that modern logistics quality management had three characteristics, including total involvement, full control and overall management. The evaluation index model of logistics operation system consists of operation indicators, warehouse service indicators, information indicators, and customer satisfaction degree. In addition, many other scholars have also done a lot of research in this field (Sun Hua-li et al. 2006).

Summarizing the aforementioned research of logistics service, we can find that there are various kinds of evaluation index system of logistics services with different emphasis. However, most of the literature available on logistics service research has focused on third-party logistics at macro level, only a few academic studies have focused directly on customer satisfaction of express industry. There's some difference between the express industries and logistic in general sense, for express companies serve individuals who attach high importance to timeliness. Based on the existing research and the interview information from the personnel in express industry, and combined with ideas from service management experts, this article proposes the initial dimensions structure of evaluation indexes for express service. Then this article revises the initial structure according to advice from the experts and customers, and put forward the final model of express service evaluation index (Table 81.1).

81.3 An Empirical Study on Customer Satisfaction of Express Enterprise

This article analyzed a large private express company in Shenzhen, which has set up departments of data collection, marketing, logistics distribution and express service and has established a nationwide service network. Besides, the company is expanding its business to international express service.

Table 81.1 Express enterprise customer satisfaction evaluation index

The first level	The second level (criteria)	The third level (alternatives)
Customer Satisfaction of Express Enterprise A	Corporate Strength B_1	Corporate recognition C_{11}
		Corporate reputation C_{12}
		Possession of physical facilities C_{13}
		Coverage of delivery network C_{14}
	Service Personnel B_2	Attitude and Behavior of service personnel C_{21}
		Image of service personnel C_{22}
		Professionalism and skills of service personnel C_{23}
	Timeliness B_3	Response time on pickups C_{31}
		Committed delivery time C_{32}
		The ratio of on time delivery C_{33}
		Flexibility of service time C_{34}
	Reliability B_4	Security and accuracy of the cargo C_{41}
		Availability of delivery status information C_{42}
		Stability of Service C_{43}
		The standardization of service process C_{44}
		The ability to resolve complaints C_{45}
	Economics B_5	Satisfaction degree on price C_{51}
		Insurance and compensation policy C_{52}

81.3.1 The Quantification of Evaluation Indexes

This article utilizes a 5-point Likert scale to evaluate customer satisfaction, which resulted in five rating categories, namely very satisfied, satisfied, basically satisfied, dissatisfied, and very dissatisfied. Customer satisfaction degree could be classified into multiple levels, or converted to scores, showed in the form of interval. Its value is limited to [0,5], as [5,4] refers to very satisfied, [4,3] to satisfied, [3,2] to basically satisfied, [2,1] to dissatisfied, and [1,0] to very dissatisfied. According to the extent of satisfaction, the scores are converted to (4.5, 3.5, 2.5, 1.5, 0.5).

This research conducted a survey on customers who have used the service of the selected express company. We handed out 200 questionnaires and 197 questionnaires were returned, of which 190 questionnaires were available. The statistical result of the survey is shown in Table 81.2.

81.3.2 Determine the Weight of the Indexes with the Application of AHP

The key of AHP is the construction of the judgment matrix, before which we need a scale that indicates the priority of the indexes. A commonly used measurement

Table 81.2 Statistical result of empirical study on customer satisfaction

Index	Satisfaction degree				
	Very satisfied ratio (%)	Satisfied ratio (%)	Basically satisfied ratio (%)	Dissatisfied ratio (%)	Very dissatisfied ratio (%)
Corporate recognition	0.19	0.27	0.33	0.13	0.07
Corporate reputation	0.20	0.26	0.29	0.17	0.08
Possession of physical facilities	0.11	0.19	0.33	0.25	0.11
Coverage of delivery network	0.20	0.32	0.26	0.15	0.08
Attitude of Behavior of service personnel	0.12	0.20	0.32	0.26	0.11
Image of service personnel	0.10	0.23	0.28	0.24	0.15
Professionalism and skills of service personnel	0.18	0.25	0.33	0.17	0.06
Response time on pickups	0.11	0.19	0.28	0.25	0.17
Committed delivery time	0.18	0.25	0.27	0.19	0.11
The ratio of on time delivery	0.15	0.17	0.31	0.22	0.16
Flexibility of service time	0.21	0.36	0.27	0.15	0.02
Security and accuracy of the cargo	0.11	0.17	0.31	0.36	0.05
Availability of delivery status information	0.21	0.25	0.29	0.20	0.05
Stability of Service Level	0.05	0.16	0.28	0.35	0.16
The standardization of service process	0.20	0.26	0.31	0.20	0.04
The ability to resolve complaints	0.11	0.20	0.22	0.32	0.15
Satisfaction degree on price	0.13	0.19	0.31	0.24	0.13
Insurance and compensation policy	0.05	0.14	0.21	0.34	0.26

scale in the AHP is the 9-point scale, which is shown in Table 81.3 (Zhang Zhen et al. 2006).

Once the scale has been determined, we can go on to build the judgment matrix. In the process of building the judgement matrix, the decision maker has to implement pair-wise comparison between the indexes repeatedly, in other words, compares A_i with A_j and determines the relative importance with respect to the elements in the level immediately above it, then presents the result by numerical values. Given n elements $A_1, A_2, \ldots, A_n$, we thus compare the relative importance of one element with respect to a second element, using the 9-point scale shown in Table 81.3, and generate the judgement matrix $A = (a_{ij})_{n \times n}$. Of which, a_{ij} is the relative importance ratio for factor i and factor j, as shown in Table 81.4.

Table 81.3 Scale of importance

Relative importance	Score	Explanation
Equal importance	1	In pairwise comparison, two elements contribute equally to the objective
Moderate importance	3	In pairwise comparison, one element is a little more important than another
Strong importance	5	In pairwise comparison, one element is obviously more important than another
Demonstrated importance	7	In pairwise comparison, one element is demonstrate more important than another
Extreme importance	9	In pairwise comparison, one element is extremely more important than another
2, 4, 6, 8 are intermediate values between the two adjacent judgments		

Table 81.4 Judgement matrix of the criteria

Judgment Matrix B	B_1	B_2	B_3	B_4	B_5
Corporate Strength B_1	1	1/5	1/7	1/9	1/5
Service Personnel B_2	5	1	1/5	1/5	1/3
Timeliness B_2	7	5	1	1/3	3
Reliability B_4	9	5	3	1	5
Economics B_5	5	3	1/3	1/5	1

Table 81.4 is the judgement matrix of the criteria; similarly, we can yield the judgement matrix of the alternatives. The relative weights of the alternatives under a single criterion could be formulated as following:

1. Normalize each column of matrix A:

$$b_{ij} = \frac{a_{ij}}{\sum\limits_{j-1}^{n} a_{ij}} (i, j = 1, 2, \ldots, n) \tag{81.1}$$

2. Sum the rows of the normalized matrix:

$$W_i = \sum_{j=1}^{n} b_{ij} \tag{81.2}$$

3. Normalize the row sums and yield the relative weight vector:

$$W_i = \frac{W_i}{\sum\limits_{i=1}^{n} W_i} (i = 1, 2, \ldots, n) \tag{81.3}$$

Substitute the judgement matrix to Eqs. (81.1), (81.2), and (81.3), and then we can yield the relative weight vector of the criteria:

$$\overline{W_A} = (0.031, 0.080, 0.268, 0.489, 0.133)$$

Table 81.5 Random index

n	1	2	3	4	5	6	7	8	9
R.I.	0.00	0.00	0.52	0.90	1.12	1.24	1.35	1.42	1.46

Remark: n is the number of factors

Estimate the consistency of the pairwise comparison matrix $A = (a_{ij})_{n\times n}$:

1. Compute the maximum eigenvalues of matrix A

$$\lambda_{\max} = \sum_{i=1}^{n} \frac{[AW_i]}{n(W_i)} = 5.3828 \tag{81.4}$$

2. Calculate the Consistency Index

$$C.I. = \frac{\lambda_{\max} - n}{n-1} = 0.0975 \tag{81.5}$$

3. Generate $R.I.$ (the average value of CI for random matrices) from Table 81.5.
4. Calculate the Constituency Ratio (*C.R.*)

$$C.R. = \frac{C.I.}{R.I.} = 0.085 < 0.1 \tag{81.6}$$

5. Make judgments. If *C.R.* <0.1, accepts the judgement matrix as a consistent one. If *C.R.* > 0.1, then rebuild or adjust the judgement matrix.

Following the above procedures, we can generate the relative weight vectors of the alternatives:

$\overline{W_{B1}} = (0.241, 0.571, 0.065, 0.123)$
$\overline{W_{B2}} = (0.258, 0.105, 0.637)$
$\overline{W_{B3}} = (0.083, 0.068, 0.626, 0.223)$
$\overline{W_{B4}} = (0.315, 0.074, 0.373, 0.163, 0.074)$
$\overline{W_{B5}} = (0.75, 0.25)$

81.3.3 Fuzzy Comprehensive Evaluation

1. Establish the factor sets (Zhang Tienan and Li Jing-lei 2002; Wang Yun et al. 2006)

According to Table 81.1, the criteria set is defined as $U = \{u_1, u_2, u_3, u_4, u_5\}$, where u_1 refers to Corporate Strength, u_2 refers to Service Personnel, u_3 refers to Timeliness, u_4 refers to reliability, and u_5 refers to economics. The alternatives set is defined as $U_1 = \{u_{11}, u_{12}, u_{13}, u_{14}\}$, where u_{11} refers to Corporate recognition,

u_{12} refers to Corporate reputation, u_{13} refers to Possession of physical facilities, u_{14} refers to Coverage of delivery network. Similarly, the other alternatives set can also be established in this way.

2. Establish the evaluation sets

There are five grades of customer satisfaction in this research, namely $V = \{v_1, v_2, v_3, v_4, v_5\}$, where V is the evaluation set, v_1 refers to very satisfied, v_2 refers to satisfied, v_3 refers to basically satisfied, v_4 refers to dissatisfied, and v_5 refers to very dissatisfied.

3. First-Order Fuzzy Evaluation

First of all, we carry out fuzzy evaluation for each single factor $u_i(i = 1, 2 \ldots, m)$ of the factor set U and generate the membership degree of u_i to the evaluation set $v_j(j = 1, 2 \ldots, n)$, then yield the fuzzy set $r_i = (r_{i1}, r_{i2}, \ldots, r_{im})$ for each single factor(u_i). Since each criterion is determined by all of its subordinate alternatives, the single factor evaluation of each criterion is the comprehensive evaluation for all of its subordinate alternatives. The single factor evaluation matrix of the alternatives can be constructed as:

$$R = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ r_{21} & r_{22} & \cdots & r_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ r_{m1} & r_{m2} & \cdots & r_{mn} \end{bmatrix} \tag{81.7}$$

Where the number of rows is determined by the number of the alternatives, the number of columns is determined by the grades of the comment. Given the weight set A and the evaluation matrix R, we can transform the weight set A to the fuzzy subset B of the evaluation set V by applying fuzzy linear transformation. B is a fuzzy comprehensive evaluation set of the comment set V, and $b_j(j = 1, 2, \ldots, n)$ represents the membership degree of the evaluation set v_j to fuzzy comprehensive evaluation set B.

$$B_i = R_i \times \overline{W_{Bi}} \tag{81.8}$$

The fuzzy comprehensive evaluation vectors of the alternatives are given by formula (81.8):

$B_1 = (0.192, 0.265, 0.299, 0.163, 0.080)$
$B_2 = (0.156, 0.235, 0.322, 0.201, 0.082)$
$B_3 = (0.162, 0.220, 0.296, 0.205, 0.126)$
$B_4 = (0.110, 0.189, 0.291, 0.315, 0.097)$
$B_5 = (0.110, 0.178, 0.285, 0.265, 0.163)$

4. Second-Order Fuzzy Evaluation

In order to generate the fuzzy comprehensive evaluation of the criteria, we need to carry on second-order fuzzy evaluation. Given the weight vector

$\overline{W_A}$ and the single factor comprehensive evaluation matrix of the criteria $B = \{B_1, B_2, B_3, B_4, B_5\}$, we can yield the fuzzy comprehensive evaluation set of the criteria:

$$P = \overline{W_A} * B = \overline{W_A} * \begin{bmatrix} B_1 \\ B_2 \\ \vdots \\ B_5 \end{bmatrix} \tag{81.9}$$

$$P = (0.028,0.087,0.400,0.562,0.073)$$

After normalization, the score of selected express company is 2.31. As the customer satisfaction has been classified into five grades, namely very dissatisfied, dissatisfied, basically satisfied, satisfied, and very satisfied, we find out that the customer satisfaction of the selected express company belongs to "basically satisfied", which indicates that customers are basically satisfied with the service of the selected express company.

81.4 Conclusions

In order to generate greater customer satisfaction, express companies should evaluate their service quality objectively in the first place. The evaluation of express companies' service quality is one of the important issues in the study of service quality management. This article proposes the framework of express service customer satisfaction evaluation index system and evaluates the service of express enterprises based on AHP and FCE, which reflects the actual service level of the express companies and provides an objective basis for service improvement. In future works, we can increase or decrease the evaluation indexes of the proposed model as well as change the weights of the indexes according to specific requirements derived from the age of the customers, the types of the cargo and the region. In this way, we can evaluate the service quality of the express companies more accurately and provide objective base for service improvement, which will enables the express companies to satisfy the demand of their customers in a better way.

References

Bienstock C, Mentzer JT, Bird MM (1997) Measuring physical distribution service quality. J Acad Mark Sci 25(1):31–44

Chiu HC (2002) A study on the cognitive and affective components of service quality. Total Qual Manag 13(2):265–274

Huiskonen J, Pirttila T (1998) Sharpening logistics customer service strategy planning by applying kano, service quality element classification. Int J Prod Econ 56–57(2):253

Lalonde BJ, Zinszer PH (1976) Customer service: meaning and measurement. National Council of Physical Distribution Management, Chicago, pp 156–159

Mentzer JT, Gomes R, Krapfel RE Jr (1989) Physical distribution service: a fundamental marketing concept. J Acad Mark Sci 17(1):53–62

Mentzer JT, Flint DJ, Kent JL (1999) Developing a logistics service quality scale. J Bus Logist 20(1):9–32

Mentzer JT, Flint DJ, Hult GTM (2001) Logistics service quality as a segment-customized process. J Mark 65(4):82–101

Parasuraman A, Zeithamlva VA, Berry LL (1985) A conceptual model of service quality and its implications for future research. J Mark 49(3):41–50

Perrault WD, Russ F (1974) Physical distribution service: a neglected aspect of marketing management. MSU Bus Top 22:37–45

Sun Hua-li, Xie Jian ying, Xue Yao-feng (2006) A customer satisfaction degree evaluation model based on SVM in logistics. J Shanghai Jiaotong Univ 40(4):684–688 (in Chinese)

Wang Jing, Li Jin-fei (2006) Evaluation on customer satisfaction degree in third party logistics. Ind Eng J 9(2):91–94 (in Chinese)

Wang Yun, Mu Yong, Liu Qinghong (2006) Judging model of fuzzy cluster optimal dividing based on grey relational grade analysis. J Shandong Univ Technol 36(2):86–89

Xu Jincan, Ma Mouchao, Chen Yiwen (2002) Review on the research of service quality. J Dev Psychol 10(2):233–239 (in Chinese)

Zhang Tienan, Li Jing-lei (2002) Application of multi-step fuzzy comprehensive evaluation. J Harbin Eng Univ 23(3):132–135 (in Chinese)

Zhang Zhen, Yu Tianbiao, Liang Baozhu (2006) AHP-based vendor evaluation in combination with FCE. J Northeast Univ (Nat Sci) 27(10):1142–1145 (in Chinese)

Chapter 82
Service Enhancement Mode Study on Urban Rail Transportation Equipment Manufacturing Enterprises Cluster

Hui-qun Gong, Qing-jing Xu, and Chao Huang

Abstract In recent years services have been used to enhance the competitiveness of enterprise products and obtain new value in domestic and foreign manufacturing. In this paper, two enterprises service enhancement modes based on urban rail transportation equipment manufacturing cluster in China are proposed. And the mechanisms of these two modes are analyzed clearly. The results show that the two service enhancement modes can contribute to the sharing of enterprises knowledge, promote the learning process among enterprises and optimize the interaction between enterprises and their customers within the cluster.

Keywords Enterprises cluster • Rail transportation equipment manufacturing • Service enhancement • Service innovation

82.1 Introduction

With the continuous development of our economy and society in recent years, urban rail transportation construction has also developed rapidly. The construction plan has been drawn up in dozens of cities in China. And it is of great benefit to urban rail transportation equipment manufacturing enterprises. It brings tens of billions dollars annually. In addition, there are 118 Hyper-megalopolis with total population number over one million and 85 large cities with population of 50–100 million up to 2009. With the acceleration of urbanization, fast and convenient rail transportation will become an important means to solve these urban traffic problems. In early December 2009, the State Council of China approved 22 city subway construction plans and invested nearly 900 billion yuan in urban rail transportation manufacturing industry

H. Gong • Q. Xu (✉) • C. Huang
School of Economics and Management, Southeast University, Nanjing, China
e-mail: xuqingjing@126.com

E. Qi et al. (eds.), *The 19th International Conference on Industrial Engineering and Engineering Management*, DOI 10.1007/978-3-642-37270-4_82,

(Xi Liang 2009). The demand of Hyper-megalopolis and large cities for urban rail transit brings sustained and stable huge market space that urban rail transportation equipment manufacturing develops rapidly.

At the same time, urban rail transportation equipment manufacturing in China is also facing a huge challenge. It is a high-end manufacturing industry. Its products are technical complexity and higher level integration. In Changchun, Qingdao, Nanjing and other cities, the backbone of the local companies have formed an enterprises cluster with relatively complete industrial chain and a certain scale. On the whole, the scale of enterprise is small and the whole technical strength of industry is still weak. Although the degree of localization of the EMU has reached more than 75 % (Guohua Xi 2009), most of the key equipments are still mainly relying on imports. There is a big gap on competitiveness between our industry clusters and multinational companies such as Siemens and Bombardier.

In recent years, enterprises use services to enhance the competitiveness of their products and obtain new value in domestic and foreign manufacturing. The phenomenon known as "Service Enhancement" (Berger and Lester 1997; Gann and Salter 2002; Berger 2005; Camarinha-Matos et al. 2011). Studies have shown that the different levels of service enhancement phenomena are commonly observed in the manufacturing sector with different content of knowledge and different types of products. The phenomenon has shown that manufacturing companies began service competition widely. And these companies attempted to enhance the competitiveness of products by good service. It became a new value source of the manufacturing companies. Current researches on the service enhancement are mostly concentrated on the mode and mechanism of a single manufacturing enterprise service enhancement. Based on the characteristics of urban rail transportation project construction, operation and maintenance, we propose two service enhancement modes. They are the general contracting of urban rail transportation project construction and outsourcing of urban rail transportation operation and maintenance. The further study in this paper is about the mechanism of improving overall competitiveness of urban rail transportation equipment manufacturing enterprises cluster based on the two service enhancement modes.

82.2 Service Enhancement

82.2.1 Analysis on Service Enhancement Factors

Since Porter (1990, 1998) introduced his concept of clusters, the concept and other forms of collaborative networks have been related to competitiveness. The competitiveness of enterprise reflects its quick and appropriate response to external challenges through the development of new capabilities and innovations (Freitas and von Tunzelmann 2008; Ramsey et al. 2008).In 2003, van Dijk and Sverrission (2003) analyses the dynamics of clustered enterprise development in developing countries.

In recent years, urban rail transportation equipment manufacturing developed rapidly in our country. However, these companies all have generally small firm size and weak technical strength in each cluster. Therefore external economies of scale have not been well reflected, and strong overall competitiveness has not been formed in cluster. In order to improve the overall competitiveness of the enterprises cluster, on one hand, the enterprises should carry out technical innovation and breakthroughs in core technologies. On the other hand, the enterprises should give full play to the advantages of cluster and enhance the overall competitiveness of cluster by service. We believe that the key factors of service enhancement in the enterprises cluster include the following two aspects.

First of all, rail transportation equipment manufacturing is a high-tech industry in our country. The installation, use and maintenance of its products are all high technical complexity. According to the composition of technological systems, urban rail transportation is a complex system consisting of a number of different functional and technical subsystems. It includes eight subsystems. They are the driving, braking, steering, power supply, fan and air conditioning, interior, screen doors and roadbed (Yanliang Du and Xueqin Liu 2012). All of its products have high technological content and complicated installation. The safety of products is extremely demanded. Customers must maintain and inspect the products regularly to ensure their suitability and safety. In view of the technical complexity of products, customers need to get the perfect related support of services from the manufacturing enterprises in the process of urban rail transportation project construction, operation and maintenance.

Second, the process of urban rail transportation project construction, operation and maintenance is complicated. At present, the headquarters of subway construction or similar institutions setting up by the Government are in charge of construction of urban rail transportation project. And the Mass Transit Railway Corporation is responsible for operations and maintenance. The specific aspects including planning project, preparing project proposal and feasibility study report, designing project, purchasing equipment, constructing line, installing and commissioning equipment, line opened to traffic trial run, line delivery, line operation and maintenance (Ziqiang Zhu 2010). The entire process involves a number of relevant government departments and enterprises with high technical and management complexity. So customers also need a comprehensive service support.

82.2.2 The Mode of Service Enhancement

With the trend of service enhancement in manufacturing, the status of service in manufacturing companies has gone through three stages of evolution. In the initial stage, the service is regarded as append ant of product; it is only a necessary condition of the products. In the second stage, the service and products are closely integrated. The service became an important part of the enterprise product strategy. Then service became "embedded" of brand, and its value of the cost be embodied.

In the advanced stage, the service became an independent business and brand, and it turn into the new source of profits of manufacturing companies. At this time, service gradually separated from physical products. Enterprises relying on accumulated service will be able to create value independent.

The basic idea of service innovation on urban rail transportation equipment manufacturing enterprises cluster is to give full play to the advantage of the network of the cluster, integrate professional service capabilities based on the entity of the member enterprises cluster. The enterprise network in cluster will provide comprehensive and complete service instead of providing related products and services for one aspect of the urban rail transportation project by various companies. Thus service compensates the shortcoming of lack of core technical capabilities of a single enterprise and enhances the overall competitiveness of the cluster. Consequently, we propose to set up urban rail transportation construction project general contracting companies and urban transportation operation and maintenance outsourcing company which respectively use the two service enhancement modes.

Setting up a general contracting company of urban rail transportation construction is that regarding the backbone enterprises within the cluster as leader, together with the other related companies, and organizing the companies within the cluster in the form of corporate enterprises. This will help integrating professional technology and ability which originally dispersed in various enterprises to the comprehensive ability. And then it can provide overall service for the urban rail transportation project including consulting services, designing project, purchasing equipment, constructing and managing line, installing equipment, opening the line to traffic. General contracting company effectively integrate the related businesses within cluster to form the overall service capacity. Thus it makes up the defects of technical capacity of the subway construction headquarters and other relevant government departments and reduces complexity of its work. It will help these sectors coordinating with others in the process of project construction, thus contributing to reduce costs and shorten the construction cycle of construction project.

Setting up an outsourcing company of urban rail transportation operation and maintenance is that regarding the backbone enterprises within the cluster as leader and organizing the backbone enterprises in the rail transit equipment manufacturing network in the form of corporate enterprises. This will help integrating the business equipment operation and maintenance capacity to comprehensive capabilities of the outsourcing company. And then it can provide overall service for the urban rail transportation project including not only the operation and maintenance of the hardware (such as vehicle maintenance, equipment maintenance, line maintenance, etc.) but also the maintaining and upgrading of the software such as ticketing system, controlling system. As the carrier of operation and maintenance service, outsourcing company integrates the operation and maintenance capacities of related businesses which can improve response efficiency, reduce maintenance costs and improve maintenance quality. The operational side of the urban rail line entrusts maintenance work to the outsourcing company. It can focus on management and other service and play the social good utility of urban rail transportation.

In summary, the first mode provides general contracting services for the construction of urban rail transportation project side, the second one provides operation and maintenance outsourcing services for urban rail transportation operator. Its essence is to enhance the overall competitiveness of the industry cluster using the service.

82.3 Mechanism and Role of Service Enhancement Mode

Figure 82.1 shows the action mechanism of the rail transportation equipment manufacturing services innovation.

In Fig. 82.1, the core enterprises represent urban rail transportation construction project general contracting company or urban rail transportation operation and maintenance outsourcing company in the cluster. $P_1 \ldots P_n$ represent the enterprises which provide related products and service for rail transportation projects in the cluster. It belongs to the supplier dimension in terms of the users of the final products.

Each supplier may have a variety of ability or technology, so $C_1 \ldots C_q$ represent a variety of ability and $T_1 \ldots T_q$ represent a variety of techniques of each party. C stands for the construction side of the rail transportation project belonging to the customer dimension. $O_1 \ldots O_q$ represent the final products and service of the rail transportation project. Figure 82.1 contains three types of interaction:

1. The integration process of demand. The label 1 shows the information integration process. It is interaction of understanding demand of the customer (that is the construction side) and formulating goals of the construction side. The core

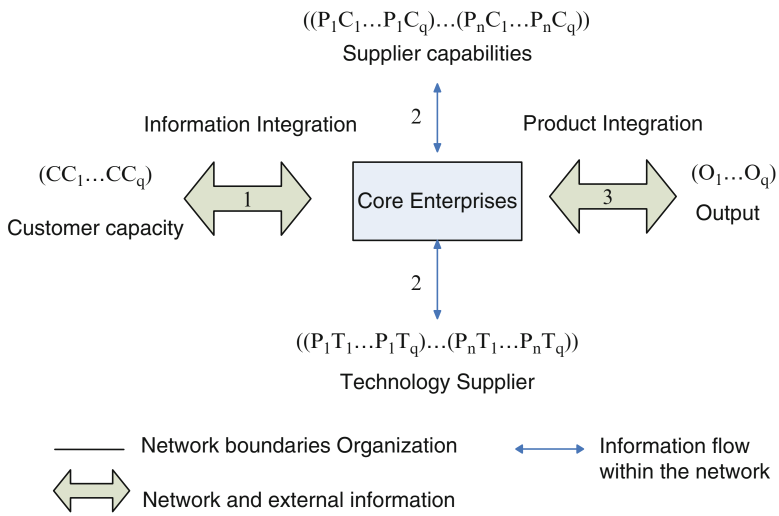


Fig. 82.1 Service enhancement mechanism in enterprises cluster

enterprises will obtain the demand effectively in this process, so that they can feedback the information to the customer. The information is about the products and service it provides and technical capability it has.

2. The integration process of capacity and technology. The label 2 shows information integration process between the core enterprises and other enterprises in the cluster. And in this process, the core enterprises deliver the demand of customers to other enterprises within the cluster. Customer demand will be decomposed and refined, and then they will be sent to specific enterprises. At the same time, the core enterprises will integrate the capacity and technology of related businesses within the cluster.
3. The integration process of the products. Lable 3 shows the products integration process. The core enterprises provide the final products and service for customers on the basis of the integration of capacity and technology in the related enterprises in cluster. From a customer perspective, these products and service as a whole are regarded as the general contracting service for construction projects and operational maintenance outsourcing service.

The above three interaction process occurred from 1 to 3, but it is not absolutely occurred one by one, each interaction process has a certain degree of overlap in time and space. In addition, each interaction process is not done overnight, it's a process which the information exchanged and adjusted continuously between the parties.

Further analysis on the action mechanism of service innovation in rail transportation equipment manufacturing shows that the two modes are useful for service enhancement. That is setting up urban rail transportation construction project general contracting company and urban rail transportation operation and maintenance outsourcing company can contribute to their service enhancement. The major role of the rail transportation industry cluster includes the following aspects:

1. It is conducive to enterprise knowledge sharing within the cluster. OECD proposed that the knowledge can be divided into four categories. They are factual knowledge (know-what), knowledge of the natural principles and laws (know-why), knowledge of skills and competencies on certain things (Know-how) and knowledge about who knows and who know how to do (know-who) (OECD 1998). In these four categories of knowledge, the first two are explicit knowledge; the latter two categories of knowledge are tacit knowledge. The tacit knowledge will mainly be obtained from practice. The capability dimensions and technology of suppliers in Fig. 82.1 are the third category of knowledge. It is tacit knowledge that explored and accumulated in the long-term practice. Due to the complexity of the knowledge that rail transportation required, the knowledge is distributed within the different organizations of the network. It is relatively limited to innovation of a single enterprise.

   The two service enhancement modes mentioned above not only conducive to the exchange of explicit knowledge within the cluster but also conducive to knowledge sharing and exchange of tacit knowledge within different organizations. The two modes are complementary, so that they promote greater and greater range innovation in the cluster.

2. It promotes the learning within enterprise cluster. Dale Neef generalized organizational learning in two ways: one is the "single-loop learning" in the stimulus-response mechanism, and the other is "Double Loop Learning" established in the open network (Neef 1998). The single-loop learning enterprise used the methods that proven to be successful and avoid that proved to be erroneous. This approach will help enterprises adapting to changes of the environment. But it will not be able to solve the problems the enterprises facing when the violent changes occurred in the external environment. And it even became barriers to enterprises innovation. It is a dynamic process on double-loop learning. The implicit assumptions in corporate behavior will be questioned when companies received new information. The enterprise will change this assumption as far as possible and make new assumption on the external environment. It is a fundamental transformation on strategic and organizational patterns instead of learning from past experience to passively adapt to the changing environment.

   It can be seen from Fig. 82.1, the enterprises within the cluster is in the open network environment which helps member companies to carry out the double-loop learning. The core enterprises act as a channel of communication between enterprises and external customers in the cluster. On one hand, they will do help to understand the customer demand and deliver the information to other enterprises within the cluster. They will guide the development and production of these enterprises. On the other hand, the core enterprises integrate the resources within the cluster. The originally loose partnership between enterprises is more stable in the integration process. And the mutual learning process is organized. The learning has turned form the individual, sporadic and linear to the collective, continuous and network.

3. It optimizes the interaction between businesses and their customers within the cluster. It can be seen from the two service enhancement modes on Fig. 82.1, the enterprises within the cluster are no longer communicating with the customer alone. The core enterprises will collect and deal with customers' demand, then they will interact with other enterprises in cluster based on these information.

   Core enterprises gather relevant information through communication and exchange with customers. It will help core enterprises to accurately grasp the customer's actual demand and analyze or sort demand in-depth. After that they will propose complete and systematic solutions. And then they will feedback actual demands of subsystem to the related businesses in the cluster. This will help related businesses in cluster accurately understanding the actual demand of the customers and provide targeted and applicable products. On one hand the two modes conducive to meet customers demand, on the other hand they are also conducive to the sustained development of the cluster itself.

82.4 Conclusion

In this paper, we proposed two enterprises service enhancement modes based on established urban rail transportation equipment manufacturing cluster in China. They are general contractor of construction of urban rail transportation project and outsourcing of urban rail transportation operation and maintenance. We analyzed the mechanism of urban rail transportation equipment manufacturing enterprise cluster service enhancement. It is shown that the two service enhancement modes can contribute to the sharing of enterprise knowledge, promote the learning process among enterprises and optimize the interaction between enterprises and their customers within the cluster.

Acknowledgment The research is supported by Project of National Soft Science Research Plan (Project No. 2010GXS1D033 and 2011GXS5K095), MOE (Ministry of Education in China) Project of Humanities and Social Sciences Funds for Young Scholar (Project No. 10YJCZH046) and Humanities Social Sciences Foundation Support Funds of College Basic Scientific Research Business Expense of Southeast University (Project No. SKCX20120023).

References

Berger S (2005) Building competitive strengths in a new international economy. East Gate Book, Taiwan

Berger S, Lester R (1997) Made by Hong Kong. 2nd edn, Oxford University Press, USA, (June 26, 1997)

Camarinha-Matos LM, Afsarmanesh H, Koelmel B (2011) Collaborative networks in support of service-enhanced products. Int Fed Inf Process 362:95–104

Freitas IMB, von Tunzelmann N (2008) Mapping public support for innovation: a comparison of policy alignment in the UK and France. Res Policy 37(9):1446–1464

Gann M, Salter J (2002) Innovation in project-based, service-enhanced firms: the construction of complex products and system. Res Policy 29:955–972

Guohua Xi (2009) Practice and exploration of the independent innovation of China's rail transportation equipment manufacturing. Study Times, 18 July 2009 (in Chinese)

Neef D (1998) The knowledge economy. Butterworth-Heinemann, Boston

OECD (1998) The knowledge-base economy. OECD Press, Paris

Porter M (1990) The competitive advantage of nation. The Free Press, New York

Porter M (1998) Clusters and the new economy of competition. Harv Bus Rev 76(6):77–91

Ramsey E, McCole P, Ibbotson P (2008) Factors that impact technology innovation adoption among Irish professional service sector SMEs. Int J Innov Manag 12(4):629–654

van Dijk MP, Sverrission A (2003) Enterprise clusters in developing countries: mechanisms of transition and stagnation. Entrep Reg Dev 9(15):183–206

Xi Liang (2009) Great leap of rail transportation. E-works, December 30 (in Chinese)

Yanliang Du, Xueqin Liu (2012) Technology and equipment of modern rail transport. Science Press, Beijing (in Chinese)

Ziqiang Zhu (2010) Urban Rail Transportation Construction project management guide. China Building Industry Press, Beijing, China

Chapter 83
Research on Comprehensive Evaluation of a Certain Type of Equipment's Service Life

Qing Li, Xiao-nan Ye, Rong Wang, and Li Wang

Abstract As an important indicator to measure the technical and tactical performance of certain equipment in its use and various security activities, service life must be taken strict control to prevent it from retirement or cannot work normally caused by the end of service life. Therefore, a control model of service life is set up to maximize its use efficiency on the basis of full study on the existing security model of the equipment. In this paper, a comprehensive evaluation method is used to study the equipment's service life, and a new protection method is provided to improve its use efficiency.

Keywords Comprehensive evaluation • Equipment • Life control • Service life

83.1 Introduction

Comprehensive evaluation is an important research content of management decisions and economic statistics areas, and it has a wide range of applications in the social sciences and natural sciences, the study has achieved fruitful results (Satty 1980; Hwang and Yoon 1981). Service life control is one of the important contents of quality control of a certain type of equipment. The equipment's service life is composed by P1, P2 and P3. How to use the equipment's service life and control it in the process of implementation of the tasks is one of the serious challenges currently the equipment support force has to face. In this paper, aiming atthe equipment's

Q. Li (✉) • X. Ye • R. Wang
The 3rd Department of Xuzhou Air Force Logistics College, Xuzhou, Jiangsu, China
e-mail: lqingx@163.com

L. Wang
93383 Unit, Mudanjiang, China

E. Qi et al. (eds.), *The 19th International Conference on Industrial Engineering and Engineering Management*, DOI 10.1007/978-3-642-37270-4_83,

service life, we construct the comprehensive evaluation index system, determine weight coefficients through the “improved” scatter-degree method, and use linear weighted method for comprehensive evaluation on the equipment’s service life.

83.2 Construct the Evaluation Index System for the Equipment’s Service Life

Access itself is the problem of evaluation of uncertainty systems (Huayuan Zhu et al. 2007). In a number of (similar) systems, that how to determine the operation (or development) is good or bad is so-called the comprehensive judge problem, namely the multiple attributes (or multiple indexes) comprehensive evaluation problem (Yajun Guo 2002). The process of comprehensive evaluation consists of construction of the index system, index pretreatment, determination of weight coefficients and information synthesis.

Choosing scientific and rational evaluation index system is the accurate, reliable foundation and guarantee of the system comprehensive evaluation, and also the important means to guide system development correctly. In actual evaluation, the selected indexes are not more is better, nor less is better. The key is the role size of the selected evaluation indexes in the comprehensive assessment. In general, choose the indexes which show main features and development status of evaluation object for practical evaluation. In this paper, we will use three components of the equipment’s service life P_1, P_2, P_3 as the research object to construct evaluation index system. Assume they are x_1, x_2, x_3 respectively, apply objective and subjective weighting method to determine the weight coefficients, establish a scientific and reasonable evaluation function and calculate comprehensive evaluation value. Finally, according to the different combat training tasks prefer to select the equipment whose comprehensive evaluation value is largest to perform tasks in order to control the equipment’s service life.

83.3 Index Pretreatment of the Equipment’s Service Life

In general, for the difference of index types, units, magnitude in evaluation system constructed in the comprehensive assessment, we needs to do some pretreatment for evaluation indexes, such as the consistency of evaluation index type, linear dimensionless etc. The evaluation indexes established in this paper are all extremely-large indexes; it does not need to carry on the evaluation types consistent processing. But the units of the index are different, we need to do linear dimensionless. Following we introduce the linear dimensionless processing for evaluation indexes of the equipment’s service life.

83.3.1 Linear Dimensionless Methods

Dimensionless is also called index data standardization that is a method which eliminates the influence of original index unit through a mathematical transformation (Yanyan Ren and Mingshun Li 2010). There are a number of linear dimensionless methods, of which such as standardization method, extremum method, normalization method, linear scaling method, vector normalization method and efficacy coefficient method are commonly used.

1. Standardization method

$$x_{ij}^{*} = \frac{x_{ij} - \overline{x_j}}{s_j} \tag{83.1}$$

$\overline{x_j}$ and s_j are respectively the average value and the standard deviation of the j-th index observations (samples), and x_{ij}^{*} is called the standard observation value.

Characteristic: the sample mean is 0, the variance is 1; interval is uncertain, the maximum value and the minimum value of each index being treated is not the same; inappropriate for the situation that index value is constant (namely $s_j = 0$); inappropriate for evaluation methods which require that index values $x_{ij}^{*} > 0$ (such as entropy method, weighted geometric mean method etc.).

2. Extremum method

$$x_{ij}^{*} = \frac{x_{ij} - m_j}{M_j - m_j} \tag{83.2}$$

In which, $M_j = \max\limits_{j}\{x_{ij}\}$, $m_j = \min\limits_{j}\{x_{ij}\}$ (applicable for the following formula).

Characteristic: $x_{ij}^{*} \in [0, 1]$, the maximum value is 1, the minimum value is 0; inappropriate for the situation that index value is constant (the denominator is 0).

3. Linear scaling method

$$x_{ij}^{*} = \frac{x_{ij}}{x'_j} \tag{83.3}$$

x'_j is taken as a special point which is a fixed value, and it generally can be taken such as m_j, M_j or $\overline{x_j}$.

Characteristic: $x'_j > 0$. When $x'_j = m_j > 0$, $x_{ij}^{*} \in [1, +\infty)$, there is a minimum 1 without the fixed maximum; when $x'_j = M_j > 0$, $x_{ij}^{*} \in (0, 1]$, there is a maximum 1 without the fixed minimum; when $x'_j = \overline{x_j} > 0$, $x_{ij}^{*} \in (-\infty, +\infty)$, the range is not fixed: $\sum\limits_{i} x_{ij}^{*} = n$.

4. Normalization method

$$x_{ij}^{*} = \frac{x_{ij}}{\sum\limits_{i=1}^{n} x_{ij}} \tag{83.4}$$

Characteristic: as a special case of linear scaling method, which require that $\sum_{i=1}^{n} x_{ij} > 0$. When $x_{ij} > 0, x_{ij}^{*} \in (0, 1)$, there is neither a fixed maximum value nor minimum value, and $\sum_{i} x_{ij}^{*} = 1$.

5. Vector normalization method

$$x_{ij}^{*} = \frac{x_{ij}}{\sqrt{\sum_{i=1}^{n} x_{ij}^{2}}} \tag{83.5}$$

Characteristic: when $x_{ij} > 0, x_{ij}^{*} \in (0, 1)$, there is no a fixed maximum value or minimum value, $\sum_{i=1}^{n} x_{ij}^{2} = 1$.

6. Efficacy coefficient method

$$x_{ij}^{*} = c + \frac{x_{ij} - m_j'}{M_j' - m_j'} \times d \tag{83.6}$$

M'_j and m'_j are respectively the satisfaction value and not allowed value of index x_j, c and d are known to be normal number. The role of c is to "translate" the transformed values; the role of d is to "amplify" or "narrow" the transformed values.

Characteristic: it can be regarded as the general extreme value method, of which the range is determined, and the maximum value is $c + d$, the minimum value is c.

83.3.2 Dimensionless Processing for the Equipment Service Life Indexes

Dimensionless processing is a crucial step in comprehensive evaluation, its results have great influence on the follow-up evaluation conclusions, and the process being reasonable or not directly relates to the final conclusion rationality. The equipment's service life indexes are x_1 (unit: years), x_2 (unit: HR), x_3 (unit: time). For the three units being different, it's not easy to build a comprehensive evaluation function $y = f(w, x)$ (w is a weight vector, x is the index vector), so it needs to carry on the linear dimensionless processing.

The ideal linear dimensionless methods generally satisfy six properties: monotone, difference ratio invariability, translation independence, scaling independence, interval stability and volume constancy, and standardization processing method, extremum method and efficacy coefficient method meet most properties. Literature (Yajun Guo et al. 2011) analyzed the influence on scatter-degree method from common dimensionless methods, and pointed out that scatter-degree method is sensitive to the non-dimensional methods. At the same time between linear proportion method and the extremum method, the one who makes the $\sum_{i}^{n} y = s^{*^2}$ maximum can expand the differences between the evaluated objects mostly. According to the characteristics of linear dimensionless methods, as well as in view of the later text adopting the "improved" scatter-degree method to determine the weight coefficients, this paper will use the extremum method to handle the equipment service life indexes.

83.4 Determination of Weight Coefficient of the Equipment Service Life Indexes

Compared with a certain evaluation objective, the relative importance of evaluation indexes is different. The size of relative importance of evaluation indexes can be portrayed by weight coefficients w_j ($w_j > 0$, $\sum_{j=1}^{m} w_j = 1$). That the determination of weight coefficients is reasonable or not relates to the credibility of the comprehensive evaluation results. Therefore, the weight distribution plays a vital role on the assessment results (Li Ma et al. 2004). In this paper, we use the "improved" scatter-degree method to determine weight coefficients of evaluation indexes of a certain equipment service life. Firstly apply the G1 method (Zhaoji Yu et al. 2010) to determine the weight coefficients which can reflect the relative importance of evaluation indexes, and get "linear weighting" for the evaluation indexes based on this evaluation, and finally determine the weight coefficients for the evaluation indexes by the "improved" scatter-degree method (Rongxing Li 2006) according to the "weighted" indexes. Consider the objective and subjective factors when determining the index weight, so it is more reasonable and realistic (Li Zhang et al. 2010).

83.4.1 Determine the Weight Coefficients Based by G1 Method

G1 method is a subjective weighting method. Compared with the analytic hierarchy process, G1 method has a comparative advantage: no need to build the judgment matrix; without the need for a consistency check; also no limit to the number of programs in the application; the given order relation fully expresses the wishes of the experts; the results are also completely trustworthy. Literature (Yajun Guo 2002) discussed the principles of G1 method and the weight calculation in detail, here we will not elaborate it and just explain the calculation steps.

1. Determine the order relation
   If the evaluation index x_i is relatively greater than (or less than) x_j compared to some certain evaluation criteria (or target), it's denoted by $x_i \succ x_j$.

Table 83.1 Assignment reference table of r_k

r_k	Illustration
1.0	x_{k-1} and x_k are equally important
1.2	x_{k-1} is slightly important than x_k
1.4	x_{k-1} is obviously important than x_k
1.6	x_{k-1} is strongly important than x_k
1.8	x_{k-1} is extremely important than x_k

2. Determine the relative importance between adjacent indexes
   The importance ratio of adjacent indexes is explained by $r_k = w_{k-1}/w_k$ from the experts, where $k = 2, 3, \cdots, n-1, n$. The assignment of r_k is shown an in Table 83.1.
3. Calculate the weight coefficient w_k
   If the rational assignment of r_k given by experts or policy makers satisfies the following formula.

$$r_{k-1} > 1/r_k, \quad k = 2, 3, \cdots, n-1, n$$

Then

$$w_n = \left(1 + \sum_{k=2}^{n} \prod_{i=k}^{n} r_i\right)^{-1} \tag{83.7}$$

$$w_{k-1} = r_k w_k, \quad k = 2, 3, \cdots, n-1, n \tag{83.8}$$

83.4.2 Determine the Weight Coefficients by Scatter-Degree Method

The basic idea of scatter-degree method is: the weight coefficient should be the measure of variation degree and the impact on other indexes of each index in the overall, the original information of empowerment should come directly from objective environment, and we can determine the weight coefficients of the corresponding indexes according to the size of information amount provided by each index. From a geometric point of view, n evaluated objects can be viewed as n points (or vectors) in the m-dimensional evaluation space constituted of m evaluation indexes. Seeking the evaluating values of n evaluation objects (scalar) is equivalent to projecting this point to its a-dimensional space. Select the evaluation index weight coefficients, and try to widen the difference between the evaluation objects. That is to construct a best one-dimensional space based on the m-dimensional evaluation space, and make the projection points in this one-dimensional space most decentralized, namely the dispersion is the largest.

Assuming that there are n evaluated objects $S = \{S_1, S_2, \cdots, S_n\}$, take a linear function of very large evaluation index $x_1, x_2, \cdots, x_m$ as a comprehensive evaluation function.

$$y = w_1 x_1 + w_2 x_2 + \cdots + w_m x_m = w^T x \tag{83.9}$$

$w = (w_1, w_2, \cdots, w_m)^T$ is the weight coefficient vector, $x = (x_1, x_2, \cdots, x_m)^T$ is the state vector of the evaluated system. Substitute m standard observations of the ith system S_i into the above equation,

$$y = w_1 x_{i1} + w_2 x_{i2} + \cdots + w_m x_{im}, \quad i = 1, 2, \cdots, n \tag{83.10}$$

If we denote $y = (y_1, y_2, \cdots, y_n)^T$

$$A = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1m} \\ x_{21} & x_{22} & \cdots & x_{2m} \\ \vdots & \vdots & & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nm} \end{bmatrix}$$

Then the above equation can be written as $y = Aw$.

The criteria to determine the weight coefficient vector w is to embody the difference between different systems. It is to obtain the linear function $w^T x$ of the index vector x by the language of mathematics, and make the dispersion or variance of the function values of these n systems as large as possible.

Therefore, the sample variance constituted by n evaluated objects values is

$$s^2 = \frac{1}{n} \sum_{j=1}^{n} (y_j - \overline{y})^2 = \frac{y^T y}{n} - \overline{y}^2 \tag{83.11}$$

Substitute $y = Aw$ into Eq. (83.11), and note that the standardization processing of original data, we know that $\overline{y} = 0$, so

$$ns^2 = w^T A^T A w = w^T H w \tag{83.12}$$

$H = A^T A$ is a real symmetric matrix.

Obviously, the above equation is preferable to arbitrarily large values without any restrictions on w. Here we limit that $w^T w = 1$, request the maximum value of Eq. (83.12). The resulting weight model of scatter-degree method as follows:

$$\begin{aligned} &\max \ w^T H w \\ &s.t. \ w^T w = 1 \\ &w > 0 \end{aligned} \tag{83.13}$$

83.5 Case Studies of Comprehensive Evaluation of Certain Equipment Service Life

Now select ten equipments as study objects from the equipment database, and use comprehensive evaluation method to study and analyze its service life situation. To save space, standardized data processed by extremum method are given directly as follows in Table 83.2, in which $x_1{}^*, x_2{}^*, x_3{}^*$ are standardized evaluation indexes.

1. Determine weight coefficients by G1 method
Firstly determine the order relation of evaluation indexes $x_1{}^*, x_2{}^*, x_3{}^*$. According to experts' experience and existing research data, we can determine that the order relation is $x_1{}^* \succ x_2{}^* \succ x_3{}^*$.Suppose the weight coefficients of $x_1{}^*, x_2{}^*, x_3{}^*$ are respectively w_1, w_2, w_3.
Simultaneously experts' opinion about the importance ratio of adjacent indexes are $r_2 = w_1/w_2 = 1.2$, $r_3 = w_2/w_3 = 1.2$, substitute them into Eqs. (83.7) and (83.8), we know that the weighting coefficients of $x_1{}^*, x_2{}^*, x_3{}^*$ are respectively as follows.

$$w_3 = (1 + 1.2 \times 1.2 + 1.2)^{-1} = 0.275$$
$$w_2 = r_3 w_3 = 1.2 \times 0.275 = 0.33$$
$$w_1 = r_2 w_2 = 1.2 \times 0.33 = 0.395$$

Therefore, the corresponding weight vector of the equipment service life indexes is

$$w = (0.395, 0.33, 0.275)$$

2. Determine index weight coefficients by scatter-degree method
Get "linear weighting" for the data in Table 83.2 to get the weighted data as follows in Table 83.3, in which $x_1{}^{*'}, x_2{}^{*'}, x_3{}^{*'}$ are respectively the data being weighted.

Table 83.2 Standardized data of evaluation indexes

Object	x1*	x2*	x3*
O1	1	1	1
O2	0.6	0.667	0.733
O3	0.875	0.778	0.933
O4	0.625	0.556	0.667
O5	0.75	0.778	0.8
O6	0.375	0.667	0.333
O7	0.125	0.222	0.2
O8	0.25	0.333	0.333
O9	0.375	0.444	0.2
O10	0	0	0

Table 83.3 The "weighted" data

Object	$x1^{*\prime}$	$x2^{*\prime}$	$x3^{*\prime}$
O1	0.395	0.33	0.275
O2	0.237	0.22	0.202
O3	0.346	0.257	0.257
O4	0.247	0.183	0.183
O5	0.296	0.257	0.22
O6	0.148	0.22	0.092
O7	0.049	0.073	0.055
O8	0.099	0.11	0.092
O9	0.148	0.147	0.055
O10	0	0	0

According to scatter-degree method, obviously we can know that

$$A = \begin{bmatrix} 0.395 & 0.33 & 0.275 \\ 0.237 & 0.22 & 0.202 \\ 0.346 & 0.257 & 0.257 \\ 0.247 & 0.183 & 0.183 \\ 0.296 & 0.257 & 0.22 \\ 0.148 & 0.22 & 0.092 \\ 0.049 & 0.073 & 0.055 \\ 0.099 & 0.11 & 0.092 \\ 0.148 & 0.147 & 0.055 \\ 0 & 0 & 0 \end{bmatrix} \quad H = A^T A = \begin{bmatrix} 0.5365 & 0.4615 & 0.4626 \\ 0.4615 & 0.4103 & 0.4065 \\ 0.4626 & 0.4065 & 0.5868 \end{bmatrix}$$

Using matlab to calculate the largest eigenvalue $\lambda_{\max}$ corresponding with H and its corresponding eigenvectors b, then we can get that

$$\lambda_{\max} = 1.4035$$

$$b = (0.6010, 0.5257, 0.6020)$$

Normalization processing for eigenvectors b will get weight vector w^* of evaluation indexes.

$$w^* = (0.348, 0.304, 0.348)$$

3. Calculate comprehensive evaluation value of certain equipment's service life
   Select linear weighting method model as a comprehensive evaluation function of the equipment and calculate its evaluation value, then we get

$$y = \sum_{j=1}^{m} w_j x_j \tag{83.14}$$

In which x_j and w_j are respectively the evaluation value and the corresponding weight coefficient.

Table 83.4 Comprehensive evaluation value of the equipment's service life

Object	Comprehensive evaluation value y	Sequence
O1	1	1
O2	0.667	4
O3	0.866	2
O4	0.619	5
O5	0.776	3
O6	0.449	6
O7	0.181	9
O8	0.304	8
O9	0.335	7
O10	0	10

Substitute w^* and standardized data in Table 83.2 into Eq. (83.14), then we can obtain comprehensive evaluation value of the equipment's service life in the following Table 83.4.

In summary, the result of sorting the comprehensive evaluation values of the equipment's service life is $y_1 > y_3 > y_5 > y_2 > y_4 > y_6 > y_9 > y_8 > y_7 > y_{10}$.Therefore, to use the equipment's service life rationally, and maximize the efficiency in the use of the equipment, and give priority to using the equipment whose comprehensive evaluation value is the largest to perform tasks depending on the different combat training tasks, the order of the equipment chosen in the duties is $O_1, O_3, O_5, O_2, O_4, O_6, O_9, O_8, O_7, O_{10}$.

83.6 Conclusion

Based on deep and systematical research of the comprehensive evaluation, this article uses the "improved" scatter-degree method to empower the evaluation indexes, avoiding the dual drawbacks of the objective weighting method which can not reflect the expertise's experience or the subjective weighting method which does not reflect changes in the objective conditions. And the case studies of the equipment's service life show that the method is simple, accurate and effective. At the same time, the method can artificially control the relative importance of life indexes according to different tasks in order to formulate different equipment use program so as to achieve the control of the equipment's service life.

References

Huayuan Zhu, Qinghua Chen, Qiang Han (2007) Study on fuzzy synthetic evaluation of supportability of a certain new aircraft equipment [J]. J Qingdao Univ (Eng Tech) 22(2):69–72

Hwang CL, Yoon K (1981) Multiple attribute decision making-methods and applications: a state-of-the-art survey [M]. Springer, New York

Li Ma, Donglin Chen, Yanfei Dong (2004) Assessment for one type missile system technical support with fuzzy theory [J]. J Missile Guid 4:290–293

Li Zhang, Xuanbing Leng, Ping Wang (2010) Evidence theory model of warship equipment technical supporting ability evaluation based on combined weight [J]. Firepower Command Control 35(8):162–165

Rongxing Li (2006) Research on the system of theories and methods of construction project evaluation [D]. Dalian University of Technology, Dalian, pp 27–30

Satty TL (1980) The analytic hierarchy process [M]. McGraw-Hill, New York

Yajun Guo (2002) Theory and methods of comprehensive evaluation [M], vol 8. Science Press, Beijing, p 1

Yajun Guo, Fengmei Ma, Qingxin Dong (2011) Analysis of influence of dimensionless methods on deviation maximization method [J]. J Manag Sci 14(5):19–28

Yanyan Ren, Mingshun Li (2010) Dimensionless research on the evaluation index of social in road construction project [J]. Enterp Technol Dev 29(7):63–65

Zhaoji Yu, Wenjuan Wang, Changhai Su (2010) A new way for comprehensive evaluation of a grassroots government's execution of public policy [J]. J Northeast Univ (Nat Sci) 31(10):1508–1511

Chapter 84
An Analysis and Empirical Study of the Impact of Customer Relationship Quality on Repurchase Intention in Online Shopping Context

Mi-yuan Shan, Yue-hong Luo, and Ren-long Zhang

Abstract Customer relationship quality is an important factor to affect customers' future continuous interaction with enterprises. The study analyzes the mechanism of the impact of customer relationship quality on repurchase intention from a new perspective, and develops a theory framework by taking post-purchase perceived risk as a mediating variable. Meanwhile, it collects data from the perspective of online customers, and has an empirical test of underlying assumptions. The empirical results indicate, customer satisfaction and trust both have an indirect positive effect on repurchase intention by taking post-purchase perceived risk, mainly economic, functional, service and psychological risk, as a medium. Findings of the study show the evaluation and mechanism of post-purchase perceived risk, and enhance our understanding of the inner drive mechanism between customer relationship quality and repurchase intention in online shopping context. The study also provides a guideline for online sellers to enhance customers' repurchase intention.

Keywords Customer relationship quality • Customer satisfaction • Trust • Post-purchase perceived risk • Repurchase intention

84.1 Introduction

With the popularity of the network in recent years, online shopping in e-commerce has developed very quickly, the transaction size of online shopping market has amounted to ¥ 766.6 billion in 2011, an annual growth rate of 66.3 %. Along with the switching costs and search costs are greatly reduced in network environment, the

M. Shan • Y. Luo (✉) • R. Zhang
College of Business Administration, Hunan University, Changsha, China
e-mail: luoyuehong0730@163.com; dingjing_103@126.com

E. Qi et al. (eds.), *The 19th International Conference on Industrial Engineering and Engineering Management*, DOI 10.1007/978-3-642-37270-4_84,

cost of retaining existing customers is much lower than obtaining new customers, so customer retention is a way for sellers to obtain a competitive advantage, and it is also the key of sellers' survival and development.

To maintain the long-term relationships with customers, relationship marketing has included how to establish, develop, maintain customer relationships and other aspects into its research scope, and carried out theoretical and empirical study to the correlation between customer relationship quality and customer loyalty in different industry background. It finds out that customer relationship quality will greatly affect the customers' post-purchase behavior, such as customer retention, continuous trading. Meanwhile, the level of uncertainty customer perceived in product/service purchase process is relatively high in network environment, and it is also difficult to make an objective evaluation even after some product/service has been used. Some literature studies show that good customer relationship quality can effectively reduce customers' uncertainty, thereby affecting consumer behavioral intentions in the future. Therefore, it has a significant management theoretical and practical significance to explore the transmission mechanism about how customer relationship quality affect repurchase intention through post-purchase perceived risk.

This article proceeds as follows: Section 84.2 is literature review and hypothesis. Section 84.3 is research methods and hypothesis testing. Section 84.4 concludes.

84.2 Literature Review and Hypothesis

Oliver delineated four stages for the development of customer loyalty including cognitive, affective, conative, and action loyalty, and considers that conative loyalty is behavioral intentions or attitudinal loyalty. Behavioral intentions have been defined as intention to spread positive word-of-mouth referrals, repurchase intention, and willingness to pay a premium price (Oliver 1999). Relative to other dimensions of customer loyalty, repeat purchase behavior will have direct impact on sellers' profit, and repurchase intention is a reliable psychological predictor of repeat purchase behavior (Jones and Sasser 1995). Therefore, large numbers of early studies suggest that repurchase intention is the core of customer loyalty (Day 1969; Jarvis and Wilcox 1977).

Customer relationship quality reflects the depth and atmosphere of a relationship, embodies the perception of customers' expectations and goal satisfaction degree for a relationship. Under the background of the retail industry, Crosby et al. first proposed trust and satisfaction is important dimensions of relationship quality structure (Crosby et al. 1990). Then there are scholars proposed more dimensions from the perspective of relationship interaction, relationship profitability, and overall quality perception. Satisfaction and trust is considered to be the basic core dimensions of the consumer market relationship quality.

In the field of service marketing and relationship marketing, relationship quality practice research focus on the effect of relationship quality on customer behavior intention. Consumer satisfaction is a key factor for establishing long-term relationships with customers and acquiring their repurchase intention (Lee et al. 2008). Khalifa and Liu explored the effect of trust and satisfactions on online customers' repurchase intention from the perspective of psychological factors (Khalifa and Liu 2003). The model tested by Ha, Janda, and Muthaly in which satisfaction and trust are found to have a direct effect upon repurchase intention (Hair et al. 2011). Palmatier et al. proposed that there is a positive correlation between the customer relationship quality and repeat purchase (Palmatier et al. 2006). Thus it can be seen that good relationship quality is a key factors to improve customer loyalty and enhance customers' repurchase intention.

H1: The better of relationship quality between customers and sellers, the higher of the customers' repurchase intention.

Cunningham thinks that perceived risk consists of two parts: uncertainty and consequence (Cunningham 1967). In consumer behavior research, studies often treat perceived risk as a multidimensional construct (Crespo et al. 2009). In a more relevant study, Featherman and Pavlou put forward six risk dimensions of online consumption, including economic, functional, psychological, social, privacy and time risk (Featherman and Pavlou 2003). The domestic scholars Jing Miao et al. proposed eight dimensions of perceived risk in online shopping through empirical research, including economic, functional, physical, time, privacy, service, social and psychological risk (Jing et al. 2006).

Product post-purchase perceived risk has been studied in product marketing literature, and it is considered to be strongly associated with consumers' behavior intention in future. The latest research on perceived risk by Grewal et al. shows that the evaluation and mechanism of perceived risk is different between pre-purchase and post-purchase, and the post-purchase perceived risk is a very important aspect of affecting customers' behavior intention in future (Grewal et al. 2007). Therefore, the level of post-purchase perceived risk provides a basis for future purchase decision making, and product post-purchase perceived risk has a negative impact on future repurchase intention.

H2: The lower of customers' post-purchased perceived risk, the higher of the customers' repurchase intention.

Researchers mainly discussed the impact of relationship quality on customers' uncertainty before purchasing. Crosby et al. points out that good relationship quality can effectively reduce customers' uncertainty in his study (Crosby et al. 1990). Kim et al. considers that the factor of trust is closely related to customer's perceived risks in electronic commerce (Kim et al. 2008). A high level of trust can contribute to the decrease of perceived risks. Several researchers describe trust as an antecedent of perceived risk (Cheung and Lee 2000). Therefore, the better the relationship quality between customers and sellers, the lower the perceived risk when purchasing product, there is a negative correlation between them.

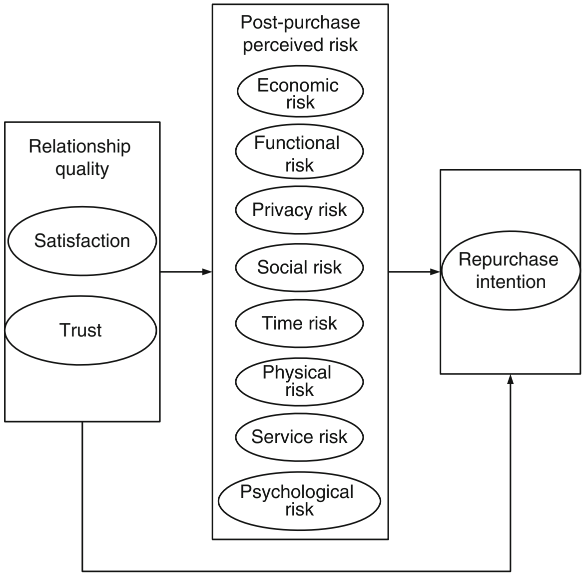


Fig. 84.1 Research model

H3: The better of relationship quality between customers and sellers, the lower of customers' post-purchased perceived risk.

Relevant researches in the past have examined the direct impact of customer relationship quality and its dimensions on repurchase intention, but in the process of the influence, some mediating variables that can enhance the explanatory power are not considered. As the level of uncertainty customer perceived in product/service purchase process is relatively high in online shopping environment, and it is even difficult to make an objective evaluation even after some product/service has been used, customers' post-purchase perceived risk is also relatively high. So, based on the above literature review, from the perspective of post-purchase perceived risk, we infer out the research model shown in Fig. 84.1.

84.3 Research Methods and Hypothesis Testing

84.3.1 Data Sources and Sample Selection

Through the collation of a large number of scales, we sum up the specific measure indicators of three variables in the model, and form initial questionnaire survey items. According to indicators extracted form the initial questionnaire design, based on the suggestions of experts and teachers of marketing, we have a small-scale test, then modify and determine the final questionnaire. The object of the questionnaire survey is customers that have purchased products/services in the B2C and C2C

Table 84.1 Scale reliability test

Latent variable	Factor	Measurement item	Cronbach α coefficient
Customer relationship quality	2	9	0.915
Post-purchase perceived risk	4	18	0.934
Repurchase intention	1	3	0.877

Table 84.2 Scale validity test

	KMO	0.907
Bartlett's test	chi-square value	5243.292
	df	406
	P value	0.000

online shop. We distribute 254 questionnaires, take 242 questionnaires back, including 232 valid questionnaires, and the questionnaires efficiency is 91.3 %.

84.3.2 Reliability and Validity Testing

To ensure the validity of the model fitting evaluation and hypothesis testing, it is necessary to pre-test reliability and validity of the variables measured. We have exploratory factor analysis for 254 valid samples by using SPSS software, conduct orthogonal rotation to the load matrix in analysis results, and have an understanding of the aggregate situation of various measurement items. The results show that the 35 measurement items were pooled into seven valid factors whose eigenvalues were greater than 1, and 5 of the items were excluded for its load capacity is less than the critical value of 0.5. Subsequently, we conduct the reliability and validity testing to the measurement model including the rest of the 30 measurement items.

The study test the reliability of the questionnaire by using Cronbach reliability coefficient, generally, it is believed that the reliability coefficient is above 0.6 can be considered the questionnaire has good reliability. The statistics results of SPSS17.0 are shown in Table 84.1, Cronbach reliability coefficient is greater than 0.8, indicate that the questionnaire has good reliability.

The study tests the validity of the questionnaire by using factor analysis. First, we test validity by using the KMO sample measure, shown in Table 84.2, the KMO test result is 0.907, indicating that it is suitable for factor analysis; Moreover, the chi-square value of spherical Bartlett's test for sample distribution is 5243.292, P value is .000, and it also indicates it is suitable for factor analysis.

At the same time, the situation of scale first-order confirmatory factor analysis fit index shows that the classification of the measurement items is consistent with theoretical expectation, the cross-load level of the measurement items is low, factor loading of each scale observed variables are greater than 0.5 which is the minimum acceptable level, and it indicates that the data achieves good validity.

Table 84.3 Overall goodness of fit index of SEM

Fitting index	Index value	Reference value	Test results
Absolute fit measures			
CMIN/DF	2.087	<3	Satisfied
GFI	0.902	>0.90	Satisfied
RMSEA	0.066	<0.08	Satisfied
Incremental fit measures			
NFI	0.911	>0.9	Satisfied
NNFI	0.937	>0.9	Satisfied
CFI	0.951	>0.9	Satisfied
Parsimonious fit measures			
PNFI	0.716	>0.50	Satisfied
PGFI	0.639	>0.50	Satisfied

84.3.3 The Goodness of Fit Evaluation of the Structural Model and Hypothesis Testing

The theoretical constructs presented by this study can not be observed or measured directly, it should be reflected by observed variables, and it is referred to as latent variables in statistical analysis. Traditional statistical analysis methods can not effectively deal with these latent variables, it needs to select appropriate confirmatory factor analysis method. We conduct confirmatory analysis of SEM by using Amos17.0 in this study. Under the premise of being in accordance with relevant theory and logic, we revise and improve the original model based on the significance of parameter estimation results and MI, and the final model classify four factors of Post-purchase perceived risk into four factors, namely economic, functional, service and psychological risk. Finally, we find that the survey data can fit the hypothetical structural equation model well after assessing the overall quality of the model.

The external quality of the structural equation model is been measured by the goodness of fit indices of the overall model, including absolute goodness of fit index, incremental goodness of fit index and parsimonious goodness of fit index, shown in Table 84.3. CMIN/DF value is 2.087, less than 3; incremental goodness of fit NFI, NNFI and CFI are greater than 0.9; parsimonious goodness of fit PNFI, PGFI are greater than 0.5. As can be seen from the comparison of the above indicators and reference values, the overall goodness of fit of the structural equation model is good.

Table 84.4 shows the construct reliability and average variance extracted of the model have reached the desired requirements, that is the construct reliability of each latent variable are greater than 0.6, and the average variance extracted of each latent variable are greater than 0.5, indicating that the model has a good internal quality.

In summary, it can be considered that the model has overall good quality, the empirical data can fit the hypothetical model well.

Table 84.4 Internal quality of the model

Latent variable	Construct reliability	Average variance extracted	Results
Customer relationship quality	0.9383	0.6299	Satisfied
Post-purchase perceived risk	0.942	0.6462	Satisfied
Repurchase intention	0.875	0.7001	Satisfied

84.4 Conclusion

Through the integration and combing of literatures, we construct a model of the impact of customer relationship quality on repurchase intention in e-commerce context. The focus of this study explores post-purchase perceived risk as an intermediary role in this influence process, and it enriches and supplements the relevant research in this area. By the empirical data, the study has verified the relationship between antecedent variable, mediating variable and outcome variable. According to the final validation results of the above structure equation model, we can find that all the assumptions of the study have got a strong validation. First, customer satisfaction and trust both have a direct positive effect upon repurchase intention, and it has got the same conclusion that has been confirmed by many researches. Second, customer trust has a significant negative effect on post-purchase perceived risk, mainly economic, functional, service and psychological risk. The effect of customer satisfaction on post-purchase perceived risk is not quite apparent. Finally, post-purchase perceived risk, mainly economic, functional, service and psychological risk, can affect repurchase intention negatively, and the effect of economic and functional risk is more significant.

Therefore, the study provides a new perspective for us to understand the influencing mechanism between customer relationship quality and repurchase intention. Based on these results, it can help online sellers have a full understanding of customers' concerns in sales interactive mode, provide a guideline for them to adopt targeted measures to reduce customers' post-purchase perceived risk, and increase the possibility of customers' continuous purchasing. Furthermore, it is significance to promote online sellers' survival and development in the fierce competitive environment.

References

Cheung C, Lee MKO (2000) Trust in internet shopping: a proposed model and measurement instrument. In: Proceedings of 6th Americas conference on information systems, Long Beach, California, pp 681–689

Crespo AH, del Bosque IR, de los Salmones Sanchez MMG (2009) The influence of perceived risk on Internet shopping behavior: a multidimensional perspective. J Risk Res 12(2):259–277

Crosby LA, Evans KR, Cowles D (1990) Relationship quality in services selling: an interpersonal influence perspective. J Mark 54(3):68–81

Cunningham SM (1967) The major dimensions of perceived risk, risk taking and information handling in consumer behavior. Harvard University Press, Boston

Day GS (1969) A two-dimensional concept of brand loyalty. J Advert Res 9(3):29–35

Featherman MS, Pavlou PA (2003) Predicting e-services adoption: a perceived risk facets perspective. Int J Hum-Comput Stud 59(4):451–474

Grewal D, Iyer GR, Gotlieb J, Levy M (2007) Developing a deeper understanding of post-purchase perceived risk and behavioral Intentions in a service setting. J Acad Mark Sci 35(2):250–258

Hair JF, Ringle CM, Sarstedt M (2011) PLS-SEM: indeed a silver bullet. J Mark Theory Pract 18(2):139–151

Jarvis LP, Wilcox JB (1977) True vendor loyalty or simply repeat purchase behavior? Ind Mark Manag 6:9–14

Jing M, Zhou Y, Lv W (2006) Customer perceived risk dimensions in online shopping environment. J Shanghai Jiao Tong Univ 40(4):805–808 (in Chinese)

Jones TO, Sasser WE (1995) Why satisfied customers defect? Harv Bus Rev 73(6):88–99

Khalifa M, Liu V (2003) Determinants of satisfaction at different adoption stages of internet-based services. J Assoc Inf Syst 4(5):206–232

Kim C, Zhao W, Yang KH (2008) An empirical study on the integrated framework of e-CRM in online shopping: evaluating the relationships among perceived value, satisfaction, and trust based on customers' perspectives. J Elect Commer Organ 6(3):1–19

Lee KL, Lee CK, Lee SK, Babin BJ (2008) Festivalscapes and patrons' emotions, satisfaction, and loyalty. J Bus Res 61(1):56–64

Oliver RL (1999) Whence consumer loyalty? J Mark 63:33–44

Palmatier RW, Dant RP, Grewal D, Evans KR (2006) Factors influencing the effectiveness of relationship marketing: a meta-analysis. J Mark 70(4):136–153

Chapter 85
An Empirical Research on Innovation Capability of Chinese Manufacturing Enterprises Based on Manufacturing Servitization

Xue-gang Shi

Abstract This paper establishes grey-fuzzy evaluation model, through the questionnaire survey method, and evaluates the innovation capability of China's manufacturing enterprises. The results indicate that the enterprises which take strategy of servitization would have higher performance than the enterprises don't do so, in the innovation opportunity recognition capability, innovation process execution capability, innovation results evaluation capability and comprehensive innovation capability. Therefore, we can come to the conclusion that manufacturing servitization can improve enterprise innovation capability.

Keywords Empirical study • Evaluation model • Innovation capability • Manufacturing servitization

85.1 Introduction

In today's high competitive market, innovation is the crucial success factor of manufacturing industry. In spite of the growing awareness that innovation is not confined to technical processes and products, literature research shows that innovative capability is still largely focused on technical processes and product innovation, and the effect of service activities to enhance the innovative capability has not been researched due to the intangibility of service activities and the difficulty to quantify the inputs and outputs of service activities (Gallouj and Weinstein 1997; Tether 2003; Koch and Strotmann 2008; Levitt 1986). Servitization has emerged

---

The paper is sponsored by the scientific research fund project of Civil Aviation University of China.

X. Shi (✉)
School of Economics and Management, Civil Aviation University of China, Tianjin, China
e-mail: shixgyl@163.com

E. Qi et al. (eds.), *The 19th International Conference on Industrial Engineering and Engineering Management*, DOI 10.1007/978-3-642-37270-4_85,

in recent years, and it has been adopted by a number of traditional manufacturers (Drucker 1990; Berger and Lester 1997; Pappas and Sheehan 1998; Quinn 1992), such as Rolls-Royce etc. Consequently, the link between industrial service and innovation has drawn the attention from both academia and industry.

Based on the manufacturing servitization, the paper constructs manufacturing enterprise innovation capability evaluation system, and makes an empirical study on Chinese manufacturing enterprises. I hope this paper can provide some reference for the relevant departments to develop innovative strategies and policies.

85.2 To Construct Enterprise Innovation Capability Evaluation System

85.2.1 The Selection of Evaluation Indicators

Innovative opportunity recognition capability is that in the changing market environment, the enterprises can discovery, identify and successfully capture the opportunity of innovation; innovation process execution capability is an enterprise can transfer innovative ideas into the products or services, avoid risk and bring profits for the enterprise; innovation results evaluation capability refers to the capability of scientific assessment and appraisal on innovation, it can accurately reflect the value of innovation capability. The above three dimensions from the view of innovation process can reflect the manufacturing enterprise innovation capability. Therefore, they are selected as the first indicators of evaluation index system.

Based on the related literature (Chundong Zheng and Jinsheng He 2000) and the trend of manufacturing servitization, combining principles of index system, such as theory, comprehensive, independence, representation, respectively determined the secondary indicators, as shown in Table 85.1.

85.2.2 The Gray Relation Analysis Method to Determine Index Weight (Yue 2007)

Using the improved gray correlation analysis method can determine the total influence degree of any index in each subsystem for other indexes in the same system, and thus determine the index weight. The basic steps are introduced briefly as follows:

For n enterprises under evaluation

$\mathrm{P}^{(1)}, \mathrm{P}^{(2)}, \cdots, \mathrm{P}^{(\mathrm{n})}$, $x_1, x_2, \cdots, x_m$ as m evaluation indexes, $P_{x_i}^{(j)}$ to show the value of Enterprise j is evaluated in a fixed point on index x_i.

Step one: dimensionless processing on the index- object sequence

In general, it need firstly to do non-dimensional treatment on sequences, different methods come out different effects. Considering the data is used to calculate the

Table 85.1 The enterprise innovation capability evaluation index system

Enterprise innovation capability	
Innovation opportunity recognition capability X	The innovation tendency of leaders and innovation decision-making sector X1
	The information collection ability of innovation department *X*2
	The information analysis and processing capacity of Innovation department X3
	The coordination situation between innovation decision department and market analysis department X4
	The knowledge of current market demand information X5
	The prediction and guidance on potential market demand information X6
	Pay attention to tacit knowledge X7
	Put internal/external knowledge or information into the R&D planning X8
Innovation process execution capability Y	Enterprise innovation platform structures Y1
	The enterprise inputs of R&D resource Y2
	The coordination status with external R&D agencies Y3
	Invite users to participate in the development of new products Y4
	Production organization and management capacity of Y5
	Bear innovation risk capability Y6
	The capability to obtain capital Y7
	The effectiveness of marketing management Y8
Innovation results evaluation capability Z	Patent, trademark and other intellectual property rights Z1
	Customer satisfaction measurement Z2
	Establish communication mechanism with customers to obtain rapid and accurate feedback information Z3
	Pay attention to the feedback information Z4
	The analysis capability on the feedback results Z5
	Invite collaborative network members to improve innovation results Z6

dimensionless correlation degree, its ultimate goal is to get the weight of each index, and among different indicators weights have certain differences is always hoped. Therefore, the standardization of dimensionless processing method is chose. To do standardization dimensionless processing on sequence $x_1^{(P)}, x_2^{(P)}, \cdots, x_m^{(P)}$ respectively, and gain $\phi_1^{(X)}, \phi_2^{(X)}, \cdots, \phi_m^{(X)}$.

Step two: calculate correlation matrix

To select respectively each sequence in $\phi_1^{(X)}, \phi_2^{(X)}, \cdots, \phi_m^{(X)}$ as parent sequence, the remaining sequences as subsidiary sequences, received a correlation with $(m-1)\times(m-1)$ degree of, can generate a m rank correlation matrix $\Psi_m^{(X)}$, and due to a target object sequence cannot do correlation calculation with its own, so the diagonal element of correlation matrix is empty.

$$\Psi_m^{(X)} = \begin{bmatrix} - & \xi_{12}^{(X)} & \cdots & \xi_{1m}^{(X)} \\ \xi_{21}^{(X)} & - & \cdots & \xi_{2m}^{(X)} \\ \vdots & \vdots & \ddots & \vdots \\ \xi_{m1}^{(X)} & \xi_{m2}^{(X)} & \cdots & - \end{bmatrix}$$

In which $\xi_{ij}^{(X)}$ expresses the correlation between subsequence $x_j^{(P)}$ against the parent sequence $x_i^{(P)}$, $(i = 1, 2, \cdots, m,\ j = 1, 2, \cdots, n,\ i \neq j)$, namely influence degree of the index x_i on the index x_j.

Step three: to summate and normalize the correlation degree
The matrix $\Psi_m^{(X)}$ can be obtained by adding at line i

$$\Delta_i^{(X)} = \sum_{j=1, j \neq i}^{m} \xi_{ij}^{(X)} \quad i = 1, 2, \cdots, m.$$

$\Delta_i^{(X)}$ shows total influence of index x_i on other indexes, or that $\Delta_i^{(X)}$ targets the degree of importance of x_i in the index system.

Then to normalize it,

$$a_i^{(X)} = \frac{\Delta_i^{(X)}}{\sum_{j=1}^{m} \Delta_j^{(X)}} \quad i = 1, 2, \cdots, n$$

So $a_i^{(X)}$ is the weight of index x_i.

In the same way, we can achieve the weight vector of indicators $x_1, \cdots, x_m$ as $A_m^{(X)} = \left(a_1^{(X)}, a_2^{(X)}, \cdots, a_m^{(X)}\right)$.

Use the same method can get weight vector the secondary indexed and first indexed respectively as:

$$A_X = \left(a_1^{(X)}, a_2^{(X)}, \cdots, a_8^{(X)}\right) \ A_Y = \left(a_1^{(Y)}, a_2^{(Y)}, \cdots, a_8^{(Y)}\right)$$

$$A_Z = \left(a_1^{(Z)}, a_2^{(Z)}, \cdots, a_6^{(Z)}\right)$$

$$A = (a_1, a_2, a_3)$$

85.2.3 The Calculate of Fuzzy Evaluation Matrix

Using the date in the index system of enterprise $P^{(j)}$ on the indexes of $x_1, x_2, \cdots, x_m$ data to generate an object – index sequence $P_X^{(j)} = \left(P_{x_1}^{(j)}, P_{x_2}^{(j)}, \cdots, P_{x_m}^{(j)}\right)$, wherein $P_{x_i}^{(j)}$ indicates data of enterprise $P^{(j)}$ on the index x_i, $i = 1, 2, \cdots, m$, $j = 1, 2, \cdots, n$. There can be n object-index sequences $P_X^{(1)}, P_X^{(2)}, \cdots, P_X^{(n)}$.

Definite a virtual optimal object-index sequence $P_X^* = \left(P_1^*, P_2^*, \cdots, P_m^*\right)$, the method are: if the value of index i is the maximum the better, then $P_i^* = \max\limits_{j=1,2,\cdots,n} \left\{P_{x_i}^{(j)}\right\}, \ i = 1, 2, \cdots, m$ is selected, if the value of index i is the minimum the better, then $P_i^* = \min\limits_{j=1,2,\cdots,n} \left\{P_{x_i}^{(j)}\right\}, \ i = 1, 2, \cdots, m$ is selected. If the various components of the sequence $P_X^*, P_X^{(1)}, P_X^{(2)}, \cdots, P_X^{(n)}$ have different dimensions, cannot calculate correlation coefficient, then they need to be dimensionless processing as $C_X^*, C_X^{(1)}, C_X^{(2)}, \cdots, C_X^{(n)}$.

Then select C_X^* as parent sequence, $C_X^{(1)}, C_X^{(2)}, \cdots, C_X^{(n)}$ as subsidiary sequences, calculating the correlation coefficient $\eta_{ji}^{(X)}$ between them $i = 1, 2, \cdots, m, \ j = 1, 2, \cdots, n$. Get the correlation coefficient matrix as follow:

$$\eta_X = \begin{bmatrix} \eta_{11}^{(X)} & \eta_{12}^{(X)} & \cdots & \eta_{1m}^{(X)} \\ \eta_{21}^{(X)} & \eta_{22}^{(X)} & \cdots & \eta_{2m}^{(X)} \\ \vdots & \vdots & \ddots & \vdots \\ \eta_{n1}^{(X)} & \eta_{n2}^{(X)} & \cdots & \eta_{nm}^{(X)} \end{bmatrix}$$

The elements in the matrix $\eta_{ji}^{(X)}$ $(i = 1, 2, \cdots, m, \ j = 1, 2, \cdots, n)$ is the correlation coefficient of index x_i for evaluation of enterprises $P^{(j)}$ against the virtual optimal object index sequence P_X^*. So the connotation $\eta_{ji}^{(X)}$ means the degree of evaluation indexes x_i for the enterprises $P^{(j)}$ belongs to the superior. Therefore, $\eta_{ji}^{(X)}$ can be the membership of enterprise $P^{(j)}$ for fuzzy sets $f\ (x_i)$, namely $\eta_{ji}^{(X)} = r_{ij}^{(X)}$.

So far, we have gained the membership $r_{ij}^{(X)}$ of evaluated enterprises, $i = 1, 2, \cdots, m, \ j = 1, 2, \cdots, n$. Accordingly, we obtained the fuzzy evaluation matrix from U_X to P.

$$R_X = \begin{bmatrix} r_{11}^{(X)} & r_{12}^{(X)} & \cdots & r_{1n}^{(X)} \\ r_{21}^{(X)} & r_{22}^{(X)} & \cdots & r_{2n}^{(X)} \\ \vdots & \vdots & \ddots & \vdots \\ r_{m1}^{(X)} & r_{m2}^{(X)} & \cdots & r_{mn}^{(X)} \end{bmatrix} = \begin{bmatrix} \eta_{11}^{(X)} & \eta_{12}^{(X)} & \cdots & \eta_{1m}^{(X)} \\ \eta_{21}^{(X)} & \eta_{22}^{(X)} & \cdots & \eta_{2m}^{(X)} \\ \vdots & \vdots & \ddots & \vdots \\ \eta_{n1}^{(X)} & \eta_{n2}^{(X)} & \cdots & \eta_{nm}^{(X)} \end{bmatrix}^T$$

Use the same method we can obtain the fuzzy evaluation matrix R_Y from U_Y to P, the fuzzy evaluation matrix R_Z from U_Z to P, and the fuzzy evaluation matrix R from U to P.

85.2.4 Two Level Grey-Fuzzy Comprehensive Evaluation Model

Using the fuzzy weight vector $A_X = \left(a_1^{(X)}, a_2^{(X)}, \cdots, a_m^{(X)}\right)$ and the fuzzy evaluation matrix R_X from U_X to P, it can be established the grey – fuzzy comprehensive evaluation model of evaluated enterprises as $B_X = \left(b_X^{(1)}, b_X^{(2)}, \cdots, b_X^{(n)}\right) = A_X \cdot R_X$. B_X is the evaluation of the vector of n enterprises, in which $b_X^{(j)}$ is index of the enterprise $P^{(j)}$ on first class index X (innovation opportunity recognition capability). In the same way, we can calculate index Y (innovation process execution capability), and index Z (innovation result evaluation capability), and comprehensive innovation capacity index respectively, as follow:

$$B_Y = \left(b_Y^{(1)}, b_Y^{(2)}, \cdots, b_Y^{(n)}\right) = A_Y \cdot R_Y$$

$$B_Z = \left(b_Z^{(1)}, b_Z^{(2)}, \cdots, b_Z^{(n)}\right) = A_Z \cdot R_Z$$

$$B = (b_1, b_2, \cdots, b_n) = A \cdot R$$

85.3 An Empirical Study on Chinese Manufacturing Enterprises Innovation Capability Evaluation

85.3.1 Date Resource

We select 60 enterprises from Beijing, Tianjin, Shanghai, Hebei, Jiangsu and other provinces and cities, they cover the state-owned enterprises, collective enterprises, private enterprises, joint ventures, foreign invested enterprises among which there are large, medium and small enterprises. The industries involved are quite extensive, covering the car, electronic equipment, household appliances, heavy machinery, electrical equipment, plastic products, instrumentation, metallurgy, cement, chemical, pharmaceutical, tobacco, food processing and so on, so the investigation statistic results have certain universality.

85.3.2 The Data Processing and Result Analysis

All the original data are from the investigation object according to the actual situation of the enterprise to give between 0 and 5 scores; therefore they do not need dimensionless processing. Using MATLB programming method to process the original data, obtain the index weight, fuzzy evaluation matrix and index value.

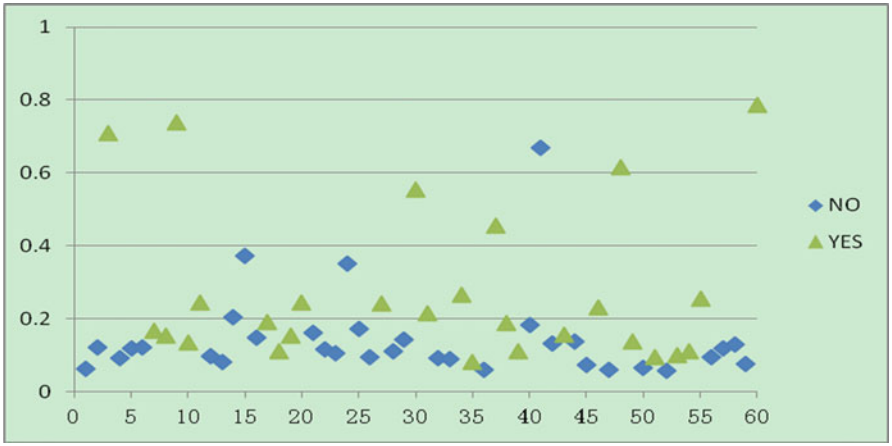


Fig. 85.1 The enterprise innovation opportunity recognition capability index scatter diagram

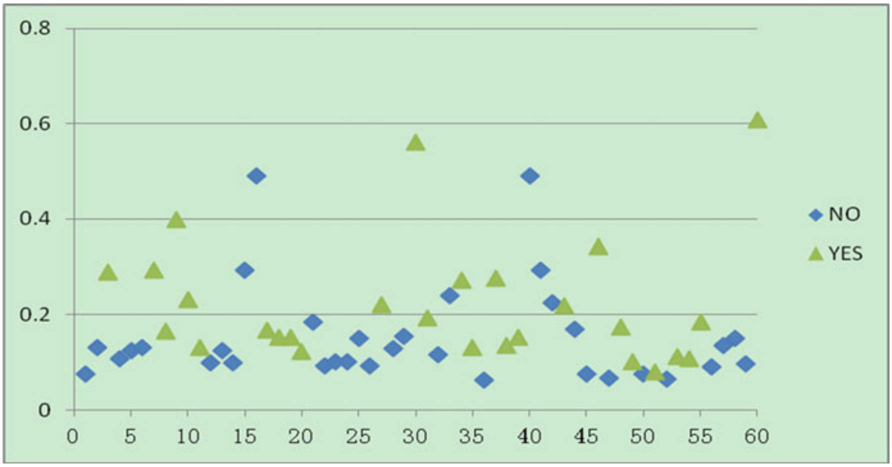


Fig. 85.2 The enterprise innovation process execution capability index scatter diagram

From Figs. 85.1, 85.2, and 85.3 we can clearly see the index results that the enterprises taking manufacturing servitization strategy have higher performance than those do not, in the innovation opportunity recognition capability, innovation process execution capability and innovation results evaluation capability.

Through the Fig. 85.4, we can clearly see that the innovation capability integrated index of the enterprises taking manufacturing servitization strategy are much higher than those do not; from the mean, the average index value of the enterprises taking the servitization strategy is 0.1258, the average index value of those did not is 0.0672. It can be concluded that manufacturing servitization can promote the enterprise innovation capability.

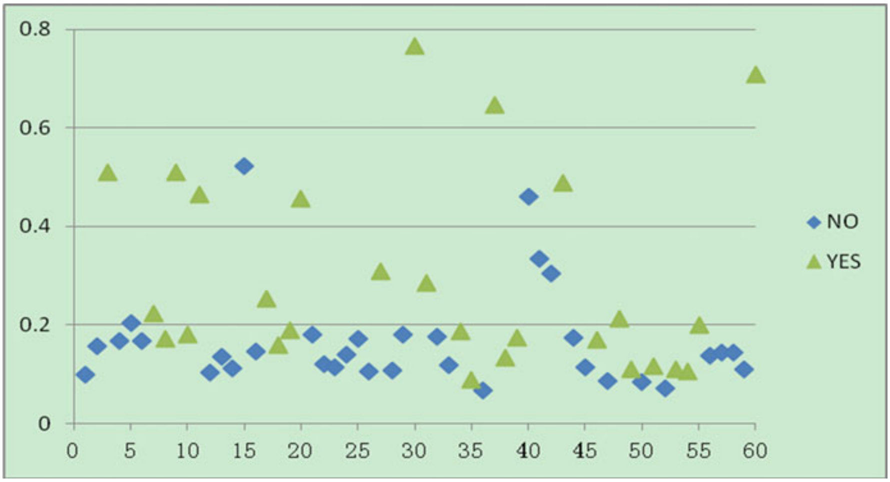


Fig. 85.3 The enterprise innovation results evaluation capability index scatter diagram

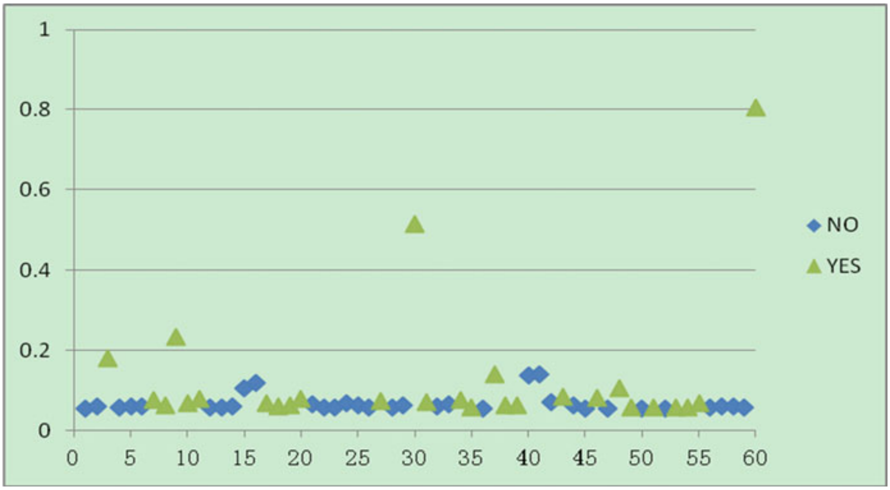


Fig. 85.4 The enterprise innovation integrated capability index scatter diagram

85.4 Conclusion

Based on the detailed analysis on calculated results, the following conclusions can be drawn:

Firstly, China's enterprise innovation capability overall level is not high, the majority of enterprises innovation capability integrated index are less than 0.2;

Secondly, the innovation opportunity recognition capability, innovation process execution capability and innovation results evaluation capability of the enterprises are not balanced, some enterprises in one area have a stronger capability, while in other areas are relatively weak.

Thirdly, manufacturing servitization can improve the enterprise innovation capability. The roles on enterprise innovation opportunity recognition capability, enterprise innovation results evaluation capability are more obvious; on the enterprise innovation process execution capability are the second, on the enterprise innovation comprehensive capability are more obvious too.

Manufacturing services have become the development trend of global manufacturing industry. In this paper, I build up the evaluation model of enterprise innovation capability; take the questionnaire survey on 60 enterprises and make empirical research on their innovation capability evaluation. The results show that manufacturing servitization can effectively enhance the enterprise innovation capability. Therefore, China's manufacturing enterprises should seize the opportunity, combine industry environment with product characteristics, customer demand, and other factors. They should adhere to the customers as the center, use network resources, extend the value chain of product, improve the innovative capacity of enterprises and then increase their competitive advantages with virtue of their professional advantages to expand their business.

References

Berger S, Lester R (1997) Made by Hong Kong [M]. Oxford University Press, Oxford
Chundong Zheng, Jinsheng He (2000) A new method of enterprise technological innovation capability evaluation [J]. Sci Technol Manag Res 03:41–44
Drucker PF (1990) The emerging theory of manufacturing [J]. Harv Bus Rev 68(3):94–102
Gallouj F, Weinstein O (1997) Innovation in services [J]. Res Policy 26(4–5):537–556
Koch A, Strotmann H (2008) Absorptive capacity and innovation in the knowledge intensive business service sector [J]. Econ Innov New Technol 17(6):511–531
Levitt T (1986) The marketing imagination [M]. The Free Press, New York
Pappas N, Sheehan P (1998) The new manufacturing: linkages between production and services activities, working for the future: technology and employment in the global knowledge economy [M]. Victoria University Press, Melbourne
Quinn JB (1992) Intelligent enterprise: a knowledge and service based paradigm for industry [M]. The Free Press, New York, pp 4–20
Tether BS (2003) The sources and aim of innovation in services: variety between and within sectors [J]. Econ Innov New Technol 12(6):481–505
Yue Liu (2007) Study on China regional science and technology development level comprehensive evaluation [D]. Master Thesis, Jiangsu University, Jiangsu, China

Chapter 86
Research on Policy Adjustment of Employment Relationships of Private Enterprises in Heilongjiang Province

Yong-hui Deng

Abstract Private economy in Heilongjiang Province has been booming since 1992. However, such problems like low assignment rate of labor contracts and differentiation treatment in social insurance remain amongst considerable private enterprises currently. The causes of the problems can be comprehended from the inevitability of economic and social development in Heilongjiang Province, the conclusiveness of development stage of productive forces, the regularity of employment relationships and its own development and the self-limitation of employers and labors. Thus, to promote the policy adjustment of employment relationships of private enterprises in Heilongjiang province, the construction of appropriate policy environment for the development of private enterprises, the maintenance function played by labor union as well as the gradual promotion of three party negotiation mechanism by government, the formation of a cooperative all-win cultural atmosphere and the implementation of effective individual incentive mechanism are needed.

Keywords Employment relationships • Labor union • Private enterprises • Productive forces

---

National soft science research project: 2005DGQ4D148; Heilongjiang Education Commission: periodic findings of research in labor relations in private enterprises of Heilongjiang province (12512077)

Y. Deng (✉)
Department of Labor and Social Security Teaching and Research, Harbin University of Commerce, Harbin, China
e-mail: 73dengyonghui@163.com

E. Qi et al. (eds.), *The 19th International Conference on Industrial Engineering and Engineering Management*, DOI 10.1007/978-3-642-37270-4_86,

86.1 Introduction

As is indicated in 1992 to 2010 in Heilongjiang Statistical Yearbook, the number of private enterprises and total employees in Heilongjiang Province is on the rise (Table 86.1), which increases from 1,600 and 22,800 in 1992 to 157,000 and 1,905,000 in 2010 respectively. What's more, the quantity of employers in urban and rural private enterprises rises from 19,000 and 3,800 in 1992 to 1,471,000 and 428,000 in 2010 respectively. The annual growth rate of the number of private enterprises in Heilongjiang Province is up to 511 %, and that of the number of the covering employees is 434 %, which exceeds the average employment growth.

Private enterprises are mainly engaged in the second and third industry in a small scale in Heilongjiang Province. For example, average capital for each enterprise (household) is merely 1,420,000 Yuan RMB, which is 14 % lower than the average level nationwide, and there are only 86 enterprises with over 100 million Yuan RMB registered capital (Gengll 2010). In the second industry, county private enterprises concentrate mainly in primary processing of resources and machinery manufacturing such as wood and food processing, while private enterprises are mainly engaged in commercial logistics, catering and comprehensive and community service in the third industry. Amongst the private enterprises, science and technology private enterprises occupy a smaller proportion. From the

Table 86.1 Reference from Heilongjiang statistic figure from 1992–2010

	Employees working for private enterprises by the end of year (10,000)	Employees working for private enterprises in urban areas by the end of year (10,000)	Employees working for private enterprises in rural areas by the end of year (10,000)	Number of private enterprises (10,000)
1992	2.28	1.9	0.38	0.16
1993	4.55	3.75	0.81	0.36
1994	8.5	7.2	1.4	0.7
1995	19.3	14.5	4.9	1.5
1996	23.6	17.4	6.2	2
1997	31.1	22.6	8.5	2.6
1998	40.6	30	10.6	3.3
1999	51.3	38	13.3	4.1
2000	53.7	40.5	13.2	4.2
2001	60.2	47.7	12.5	4.2
2002	65.7	52.6	13	4.4
2003	76.2	52.5	23.7	5.1
2004	84	68.2	15.8	5.8
2005	95.2	77.5	17.7	6.8
2006	102.8	82.8	20	7.8
2007	117	92.5	24.5	9.9
2008	141.5	108.5	33	12.6
2009	149.2	112.1	37.1	14
2010	190.5	147.1	42.8	15.7

prospect of employees' industry distribution of private enterprises, wholesale, retail, trade and catering absorb almost 50 % employees, to which are primary processing of resources and machinery manufacturing inferior, absorbing 30 % employees, and the rest employees scatter loosely, which indicate that most private enterprises in Heilongjiang Province lies in the downstream of the industrial value chain link and in a relatively low level of industry development.

86.2 Problems Existing in Employment Relationship of Private Enterprises in Heilongjiang Province

86.2.1 Low Assignment Rate of Labor Contracts

The Labor Law rules that the establishment of employment relationships should be based on the related labor contract. However, the assignment rate of labor contract of private enterprises in Heilongjiang Province is still relatively low (News Channel of Information Bay in Heilongjiang Province 2010). In reality, there is an awful amount of actual employment relationships existing. What's more, some enterprises are still using Employment Contracts or Service Contracts instead of Labor Contracts. Besides, some enterprises no longer sign labor contracts with these low-tech employees after the first fixed-term labor contracts fall due. Particularly, those enterprises engaged in architecture in a dangerous employment environment, accommodation and catering where the employees are poor at negotiation, and those enterprises with weak safety precautions, usually avoid the assignment of labor contracts so as to evade the obligation and duty ruled in *the Labor Law* (Lian-hai Xue 2007).

86.2.2 Differentiation Treatment in Social Insurance

Currently, employment relationship of private enterprises in Heilongjiang Province manifests the relationship between the single private enterprises owner and the individual labor, which is an individual atom of employment relationship. Besides few technological staff workers together with some medium and superior executives can share certain controlling reinforce over the enterprise because of their strong negotiation abilities, most staff members cannot participate in management and operation of the enterprise in some private enterprises. At the same time, enterprises will provide social insurance to these few technological staff workers together with some medium and superior executives, however, most staff members do not share the same treatment. Refusing to report the industrial injury of employees, trying to conceal the facts and never solving problems of employees legitimately do have a serious influence over the physical and emotional health of the employees.

86.2.3 The Shortage of Labor Union Construction

Labor union is an organization which maintains enterprise employees' legal rights and interests. Union Law gives labor union duty and power to maintain workers' legal rights and interests. Private Enterprises Labor Union in Heilongjiang Province Accordance's – Article 6 which effective 1st February, 2002 rules: If a private enterprise's labor union has more than 25 members, it should establish labor union committee; if it has less than 25 members, it can elect a chairman who can host it in the labor union; members from more than two units can also work together to establish labor union committee in basic level. However, many private enterprises in Heilongjiang province have not established labor unions yet, because "it is difficult to establish the labor union" and " it is difficult to charge labor union fees", and because of the large mobility of private enterprise's employees. For example, there are more than 10,000 private enterprises, which have a regular production and management and a certain scale, in Harbin now. They establish labor union organizations more than 1,800, accounting 16 % of the total.

86.2.4 The Infeasibility of Collective Bargaining System and Three Parties' Negotiation System in Reality

The shortage in construction in labor union results directly in collective bargaining system's pending, while the investors have few or no demands for the establishment of three parties negotiation system, and their present organization, the association of private enterprises, is imperfect and not typical in some area (Semoon Chang and Lynn Robinson 1990). Thus, chances are that employment and administration department handles the cases alone while quantities of arbitration committee are coping with the disputes. Consequently, it doesn't form the three parties' negotiation mechanism, and the mechanism doesn't work at all.

86.3 Reason Analysis of the Problems of Employment Relationship of Private Enterprises in Heilongjiang Province

86.3.1 The Inevitability of the Stage of Economic and Social Development in Heilongjiang Province

The progress of employment relationship is in accordance with the advancement of economy, society and technology (Hong-rui Han 2011). The urban-rural dualistic economic structure in Heilongjiang Province hasn't ever been fundamentally changed since the over 30 years' reformation and opening up in China, and the conventional social concept is relatively backward, besides, the standard of urbanization is still in the midair, what's more, the comprehensive strength of county economy is not that strong, in addition, the industrialization level in Heilongjiang Province is still low, and the alternative industries of resource-based city are still lagging behind (Yong-hui Deng 2011). The employment relationship in Heilongjiang Province is still in the atom stage depending on the conditions of economy and society, and thus, these problems existed is historically inevitable.

86.3.2 Conclusiveness of the Period of Productivity Development

Productive forces determine production relations, and then determine employment relationships. Most private enterprises are gradually advanced with a harsh entrepreneurial process in Heilongjiang Province, and they're at the bottom of the industrial value chain link with a low profit. Private economy in county economy is particularly underdeveloped, most of which are devoted to the field of commercial service with a small operation scale, few employees and are in a low level of development. There're less than 10 % of industrial enterprises in most county private economy. Therefore, the operation of private enterprises in Heilongjiang Province is still occupied in scale economy, branded products and efficient management instead of the simple pursuit of amount extension in product. The consciousness of development is superior to that of legislation in private enterprises, and there's not sufficient funds and energy to improve the staff members' treatment (An-bao Xiao 2010).

86.3.3 Regularity of the Self-Development of Employment Relationship

From the perspective of the hundreds of years western capitalist countries developing market economy, employment relationship is gradually evolved from direct deprived type of labor relationship in the primitive accumulation stage, conflict type of labor relationship to the direction of institutionalization and legalization (Kai Chang 2005). Therefore, there should be a recognition and interaction process for the main body of employment relationships of private enterprises in Heilongjiang Province in searching the method and countermeasures of employment relationship coordination.

86.3.4 Self-Limitation of Employers and Employees

On the one hand, few operators are in low quality and shortsighted, and lacking in the long-term development strategy. Because of the limitation of his own moral value, some individual operator of private enterprises will obtain his own interests on the cost of the employees'. On the other hand, with the increase of the labors' sense of self-protection, and the frequent occurring of labor drought, some employees emphasize only on their own interests. They unilateral slack and ask for the rise of salary and welfare and so on, when the economic benefit isn't better, which give rise to the tension of employment relationships.

86.4 Policy Adjustment Measures of the Employment Relationship of Private Enterprises in Heilongjiang Province

In practice, the private entrepreneurs in Heilongjiang province is also the investors, operators and managers, and they cope with the pioneering work and operation risks. So they are the builders of socialism with Chinese characteristics (Ze-min Jiang 2001). Overall, private entrepreneurs are a positive group and their work is a scarce work in Heilongjiang Province right now (Fan Feng 2007).

86.4.1 Constructing the Policy Environment Fit for the Development of Private Enterprise

According to the employment relationships system theory, created by Dunlop, an American scholar, the employment relationship system is a subsystem of social relations system, and the main body of employment relationships can be influenced by surroundings (Dunlop 1993). Most of the private enterprises are the traditional labor-intensive enterprises in Heilongjiang Province, and enterprises are under increasing management pressure with the international commodity and labor costs continuing to rise. Therefore, Heilongjiang province must to make a long-term development plan for private enterprises, improve the market as well as policy environment, decrease limitation in administration, perfect information service and make a sound socialization service system to enhance the competitive of private enterprises in Heilongjiang Province.

86.4.2 Labor Unions Bringing Maintenance Function into Play

At present, the main contradiction of the private enterprises in Heilongjiang Province is to develop the economy, and improve their competitiveness and the added value of products. Thus, the main task that private enterprises in Heilongjiang Province face is to regulate employment relationships, and realize the labor-management cooperation while maintaining employees' benefit. To realize the labor-management cooperation, we must change the organization system of labor union and promote the democratization and popularization process of the labor union (Whalen 2011).

Establish the federation of labor union of private enterprises in Heilongjiang Province led by the association of private enterprises based on the fact that most private enterprises in Heilongjiang Province are medium-sized and small enterprises. The staff members of the federation of labor union will be elected by employees and their employers, and form a specialized labor union team (Jun-yan Li 2009). The wage of the federation should be jointly paid by the union due and the enterprises. What's more, members of the federation should be based on the consistency of the interests between the employees and the employers and lead the internal cooperative game of enterprises (Ying Liu 2009).

86.4.3 Government Promoting the Three Parties' Negotiation Mechanism Gradually

Three parties' negotiation mechanism is a general concept of the principle and system that the government, employer organization and the labor union organized, negotiated and cooperated together on the wage-labor related policies of economy and society, employment legislation, the solution of the employment disputes and so on. The three parties are employment and administration department, the association of private enterprises and the federation of labor union of private enterprises in Heilongjiang Province.

Only led by the employment and administration department can advance the three parties' negotiation mechanism based on the imperfect conditions of private enterprises in Heilongjiang Province, which will help to extend the scope of adjustment of employment relationships from various grassroots economic organizations to the whole field of industry, solve and prevent some giant or destructive industrial conflicts and emergencies on time, which is useful for the employment and administration department to coordinate and mediate the daily disputes.

86.4.4 Constructing the Culture Atmosphere of Win-Win Cooperation

Labor is of dignity and emotions in the employment relationships of private enterprises in Heilongjiang Province, and private entrepreneurs are the socialist builders burdening the risks of pioneering and management. They both are cooperators of the same interests, thus, they should promote the sharing of information to bridge the gap between them in employment moral value and behavioral method, and eventually help to construct a new socialist employment relationship with win-win cooperation and sharing.

86.4.5 Implementing Effective Individual Incentive Mechanism

We should arouse employees' initiative from the perspective of wages and development. Local governments should found the training of performance bonus to overcome the short sight of private enterprises that emphasize only on the works employees do and regard training as a burden. And government should effect the training to the individuals with gizmos and creation in technology improving performance, make them upgrade the covering knowledge and technology, extend their access to promotion and stimulate their demands to challenge and realize themselves (Chun-hua Yang 2007).

86.5 Conclusion

Heilongjiang Province is in a transitional period from central planning economy to market economy, and the employment relationship of private enterprises in Heilongjiang Province is the wage labor relationship under the system of socialist market economy. Therefore, it not only should bring market mechanism in regulating employment relationships into play, but respect the truth and attach great attention to governments' policy adjustment measures' function in harmonious employment relationships.

References

An-bao Xiao (2010) Surplus labor and its allocation – contradicts between productivity and relations of production. J Guangxi Univ 12:58–61 (in Chinese)

Chun-hua Yang (2007) The design of incentive mechanism for skilled workers in enterprises-investigations on the empirical study of the current conditions in encouraging skilled worker in enterprises. Ind Technol Econ 11:23–25 (in Chinese)

Dunlop J (1993) Industrial relations systems (1958). Revised edn. Harvard Business School Press, Boston, US, pp 24–38

Fan Feng (2007) Investigation on economics of business organization. Sci Technol Manag Res 1:242–245 (in Chinese)

Gengll (2010) Heilongjiang province ranks 14 in the number of individual business and private enterprises. Northeast website. http://heilongjiang.dbw.cn/ (in Chinese)

Hong-rui Han (2011) The development progress of China's labor relationship from perspective of the labor contracts. Knowl Based Econ 20:110–111 (in Chinese)

Jun-yan Li (2009) The quality that the union leadership should possess in new period. Ind Sci Forum 12:241–243 (in Chinese)

Kai Chang (2005) Labor connections. China Labor Social Security Press, Beijing, pp 34–38 (in Chinese)

Lian-hai Xue (2007) Strategic thinking on the harmonious operation of employment relationship of private enterprises in Heilongjiang Province. Spec Zone Econ 2:104–105 (in Chinese)

News Channel of Information Bay in Heilongjiang Province (2010) The assignment rate of employment contracts is low of private enterprises in Heilongjiang province. http://news.hlj.net/100314/84224.html (in Chinese)

Semoon Chang, Lynn Robinson (1990) Perception of labor unions: by union members and their potential employers. Am J Bus 5:35–40

Whalen CJ (2011) Union engagement in Western New York economic development: an institutional analysis. Adv Ind Labor Relat 18:50–53

Ying Liu (2009) Discussion on the non-cooperative and cooperative games among enterprises. Econ Res Guid 14:280–282 (in Chinese)

Yong-hui Deng (2011) Research on problems of financial support of private economy by forest workers. For Econ 11:34–37 (in Chinese)

Ze-min Jiang (2001) Lecture at the meeting commemorating the 80th anniversary of the founding of the Communist Party of China. People Publishing Press, Shanghai, China, p 45 (in Chinese)

Chapter 87
Optimization of Foreground Support Facilities for Commercial Banks Based on Queuing Theory

Jianxin Shi and Jierong Wang

Abstract Many scholars have studied bank queuing problem through queuing model. Drawing on the service pack theory and queuing theory, this paper built a new cost-optimized model through the introduction of ATM variables. And the authors selected two bank outlets as the research objects to collect data, and then used MATLAB software for data analysis and processing. Finally, they optimized the support facilities of the selected two bank outlets through this model while trying to make the banks' total costs and customer waiting costs minimum. The paper also put forward some practical suggestions on the optimization of the bank's foreground support facilities.

Keywords ATM • Optimization • Queuing theory • Support facilities

87.1 Introduction

In recent years, with the rapid development of China's financial industry, bank queuing problem has emerged as a serious issue. The existence of this problem has exposed the inefficiency of commercial banking services, and has affected the image and competitiveness of china's commercial banks.

Some scholars believe that setting ATM in the bank outlets has shunt effect on the bank queuing problem. But in practical, many banks' service windows and ATM sets are unreasonable. From the perspective of service management, service packs consists of four aspects, including support facilities, aids, dominant and recessive service. Bank queuing problem has a great relationship with the set of bank support facilities and ancillary items. It can be said that the bank's service package set is not

J. Shi (✉) • J. Wang
School of Management, Harbin Institute of Technology, Harbin, China
e-mail: sjx@hit.edu.cn

E. Qi et al. (eds.), *The 19th International Conference on Industrial Engineering and Engineering Management*, DOI 10.1007/978-3-642-37270-4_87,

perfect. From the perspective of service package optimization, we believe that the service pack optimization of the commercial banks can improve service quality and customer satisfaction.

87.2 Theoretical Background

87.2.1 Service Pack Theory

James A. Fitzsimmons and Mona J. Fitzsimmons (2006) defined service pack as the combination of a range of products and services under certain circumstances in their book "Service management: operations, strategy and Information Technology". The combination consists of the following four elements: (1) Support facilities: the material resources must be in place before providing services. Once established, their structures are not easy to change. (2) Aids: material goods which customer purchase and consumption of, or customer-owned goods which services rely on. (3) The dominant service: refers to the interest which can be perceived with senses and it constitutes a basic characteristic of services. It is the most direct factor that impact customer satisfaction. (4) The recessive service: refers to the spiritual harvest that service brings which can be felt by customer or the nonessential features of services.

According to this definition, banks support facilities include bank office, service windows, computer networking systems and ATMs. Aids include deposit and withdrawal slips, applications, and even a pen or paper that needed, etc. The dominant service is the remittance or bank loans received from the counter. The recessive service includes staffs' warm services, comfortable environment, and high efficiency of business transactions which can give customers a good impression.

Among these support facilities, the settings of bank counter and ATM are most closely with queuing problem. Therefore, this paper mainly optimizes the settings of these two facilities.

87.2.2 Queuing Theory

87.2.2.1 Basic Theory of Queuing Systems

Queue is a line of customers waiting for service in front of one or more desks. In any service system, if the current demand for service overcomes the existing service capacity, queuing occurs. In the book "queuing theory" (Yinghui Tang and Xiaowo Tang 2006), Tang described the basic characteristics of the queuing system as follows: (1) Demand groups. Customers seeking services form the demand groups. It usually assumes that the demand group is infinite, or arrival is independent unless the demand group is very small. (2) Customer arrival process. It is the input process of queuing system. When customers' arrival time interval is independent and identically distributed, it's generally perceived as Poisson distribution. (3) Queuing structure: refers to the specific organizational structure of the queue. There are four main types: Single queue with a single window, single queue with multi-window, multi-queue with single-window and multi-queue with multi-window. (4) Queuing rules: refer to policies according to which to select next customer from the queue to accept service. The most commonly used rule is "First-come, first-served" (FCFS). This approach treats all customers equally. Thus it is the fairest rule for the customers waiting in line. LCFS-"Last-come, first served"; SIRO-service in random order. (5) Service process: refers to the service time used to service the customer. When the service is very quick and easy, the service time is usually exponential distribution.

87.2.2.2 Commercial Banks Queuing Model

Commercial banks queuing model have two forms: M/M/1 and M/M/s. M/M/1 model applies to the following situations: The total number of customers is infinite and the customers are independent and not subject to the impact of the queuing system; arrival time interval obeys a negative exponential distribution, or arrival rate obeys the Poisson distribution; there is only one waiting line, and no limit to its length; first arrived first served, only one desk and its service time obeys negative exponential distribution. The difference between M/M/s model and M/M/1 model is that the M/M/s model has more than one desk and the service time is exponentially distributed. Service rates of different services are independent and equal.

87.2.2.3 Queuing System Optimization

There are two ways to reduce customer queuing time. The first is to increase service counters and service staffs. The second is to improve service level. These two methods will increase the expenditures of service organizations more or less, but if do not take measures, waiting too long will lead to customers' dissatisfaction. Therefore, from an economic point of view, to optimize the service system should make the service costs and customer waiting cost minimum on the basis of a certain level of service quality (Wei Xiong 2005).

Commercial banks queuing model is M/M/1 or M/M/s. These two models have different optimization methods:

1. M/M/1 queuing system economic optimization model. M/M/1 model just has one desk, so it can only be optimized through improving the service rate u. Assume that C_S represents the unit time cost of the service organizations, C_W represents the cost of each customer residence time in the queuing system, then the total cost C is the expectation of the sum of services cost per unit and customers unit time expenses staying in the queuing system:

$$C = C_S * u + C_w * L_q \tag{87.1}$$

Substitute $L_q = \frac{\lambda}{u-\lambda}$ into Eq. (87.1):

$$C = C_S * \mu + C_w * \tfrac{\lambda}{\mu-\lambda}$$

The above equation indicates that the total system cost consists of two parts: service costs and customer waiting costs. Service costs are proportional to the service rate and customer waiting costs is related to the service rate and arrival rate. When the customer arrival rate is determined, it is inversely proportional to the service rate.

In order to get the optimal service rate to make service system at its best, cost formula differential method can be used:

$$\frac{dc}{d\mu} = C_s - C_w \frac{\lambda}{(\mu - \lambda)^2} = 0 \tag{87.2}$$

$$(\mu - \lambda)^2 = \tfrac{C_w \lambda}{C_s} \tag{87.3}$$

$$\because \mu > \lambda \quad \therefore \mu - \lambda > 0$$

$$\mu^* = \lambda + \sqrt{\frac{C_w \lambda}{C_s}} \tag{87.4}$$

For the M/M/1 queuing system with limited length queue, the optimal service rate can obtain by Eq. (87.5):

$$\rho^{N+1} \frac{N - (N+1)\rho + \rho^{N+1}}{(1 - \rho^{N+1})^2} = \frac{C_S}{G} \tag{87.5}$$

In Eq. (87.5): $\rho = \frac{\lambda}{\mu}$
G – Income received by the units of customer service.
N – The largest number of customers in the queuing system.

2. M/M/s queuing system economic optimization model. The objective of this model is to determine the optimal number of the service counter. Assume that $C(s)$ represents the total cost per unit time of queuing system, then:

$$C(s) = C_s * s + C_w * L \tag{87.6}$$

In Eq. (87.6): C_S represents the cost of each desk in unit time. Desk number s is a discrete variable. The best desk number s^* can be got through marginal analysis. The best desk number s^* make $C\ (s^*)$ take a minimum, namely

$$C(s^*) \le C(s^* + 1) \tag{87.7}$$

$$C(s^*) \le C(s^* - 1) \tag{87.8}$$

Substitute $C(s) = C_s * s + C_w * L$ into Eqs. (87.7) and (87.8):

$$C_s * s^* + C_w * L(s^*) \leq C_s * (s^* + 1) + C_w * L(s^* + 1) \tag{87.9}$$

$$C_s * s^* + C_w * L(s^*) \leq C_s * (s^* - 1) + C_w * L(s^* - 1) \tag{87.10}$$

After simplification, get

$$L(s^*) - L(s^* + 1) \leq \frac{Cs}{C_w} \leq L(s^* - 1) - L(s^*) \tag{87.11}$$

Based on the above equations, calculate the values when $s = 1, 2$ and so on, then calculate the difference of the adjacent L values to get a series of intervals. When $\frac{Cs}{C_w}$ falls into an interval, we can determine the best service counter number.

87.3 Model Development

87.3.1 Premises and Assumptions

1. First come first served. When customer reaches the bank, if there's a free help desk, customer immediately transacts business. If there's no free help desk, customer will wait in line;
2. The customer reaches the business site independently, and the arrival obeys the Poisson distribution with parameter λ;
3. Assume that the business site has c_1 desks, c_2 ATMs, and independently of each other service;
4. Service time of customers going to the service desk independently subject to the negative exponential distribution with parameter $u_1(u_1 > 0)$. And service time of customers going to the ATM independently obeys the negative exponential distribution with parameter $u_2(u_2 > 0)$;
5. Assume that the shunt rate of ATM is p, then customers to the service counter obey the Poisson distribution with parameter $\lambda_1 = \lambda(1-p)$, and customers to ATM obey the Poisson distribution with parameter $\lambda_2 = \lambda p$.

87.3.2 Model Parameters Definition

Symbols commonly used in this paper are as follows: among them, $i = 1, 2$, take 1 represents counter index and take 2 represents ATM index.

ρ_i – System services strength
$L_i(t)$ – the number of customers at time t

$\overline{L_i}$ – the average length of queue in the balance
$P_j^i(t)$ – the probability of queue length is j at time t
$W_i(t)$ – the customer's length of stay distribution function in the balance
$\overline{W_i}$ – Customer's average length of stay
$W_q^i(t)$ – the customer waiting time distribution function in the balance
$\overline{W_q^i}$ – the average customer waiting time

87.3.3 Model

The modeling idea is to minimize the total costs through the establishment of the ATM, help desk costs and customer waiting cost relationship, and get the optimal number of ATM and service windows.

Assume that each desk unit time cost is e_1u Yuan, the unit time cost of ATM is e_2u Yuan, and each customer stay costs per unit time at the service desk is e_3 Yuan, stay costs per unit time at ATM is e_4 Yuan. Then the average total cost per unit time is:

$$C(c_{1,}c_2) = c_1e_1u + e_3\overline{L_1} + c_2e_2u + e_4\overline{L_2} \tag{87.12}$$

Make $C_1(c_1) = c_1e_1u + e_3\overline{L_1}$, $C_2(c_2) = c_2e_2u + e_4\overline{L_2}$

Then: $\min\limits_{c_1,c_2} C(c_1, c_2) \Leftrightarrow \min\limits_{c_1} C_1(c_1), \quad \min\limits_{c_2} C_2(c_2)$, c_1 and c_2 are greater than zero and both of them are integers.

Using the marginal analysis method to obtain $c_{1,}$ c_2:

$$\begin{aligned} &C_1(c_1^*) \le C_1(c_1^* - 1), \quad C_1(c_1^*) \le C_1(c_1^* + 1) \\ &C_2(c_2^*) \le C_1(c_2^* - 1), \quad C_2(c_2^*) \le C_2(c_2^* + 1) \\ \Leftrightarrow &\begin{cases} \overline{L_1}(c_1^*) - \overline{L_1}(c_1^* + 1) \le \dfrac{e_1u}{e_3} \le \overline{L}_1(c_1^*) - \overline{L}_1(c_1^* - 1) \\ \overline{L_2}(c_2^*) - \overline{L_2}(c_2^* + 1) \le \dfrac{e_2u}{e_4} \le \overline{L}_1(c_2^*) - \overline{L}_2(c_2^* - 1) \end{cases} \end{aligned} \tag{87.13}$$

Make c_1, c_2 take the values 1, 2, 3, 4, 5 and so on respectively. Then calculate the values of $\overline{L_1}(c)$, $\quad \overline{L_2}(c)$ and calculate the difference between the respective adjacent. Then c_1, c_2 corresponding to the interval of which e_1u/e_3, e_2u/e_4 fall into is the optimal number of counter and ATM.

87.4 Data Collection

The model-related parameters of bank sites are as follows: the average customer arrival rate- λ; shunt rate of ATM- p; the average service rate of bank service window staff-μ_1; the average service rate of ATM-μ_2; the maximum waiting time that customer can accept-*Tmax*; the maximum daily workload of bank teller-*Qmax*. The data collection mainly used the method of field observation and data processing used the method of interval estimation, by taking the 95 % confidence level. Then determine the confidence interval and the distribution of the data through hypothesis testing.

This paper selected two different bank sites in Harbin city as the research objects, namely the sites A and B. And the following is the data collected from the two sites.

87.4.1 Data of Site A

87.4.1.1 Customer Arrival Rate λ and Test of Its Distribution

The paper used the field observation method to get the customer arrival rate λ. The authors recorded the number of customers reaching the bank once half hour and obtained the number of customers going to the service counter through observing the automatic calling machine. The authors used MATLAB (Mingsheng Yang et al. 2006) software for data processing and calculated the customer arrival rate $\lambda = 1.2301$ people/min.

The test of λ's distribution is mainly to test the fitting function based on queuing theory (Ozgecan and Tayfur 2009): $\bar{t} = \sum_{i=1}^{n} x_i f_i / n$ (x_i represents each time interval, f_i is the frequency), $\lambda = 1 / \bar{t}$, $\hat{p}_i = p\{x_i < \xi < x_{i+1}\} = e^{-\lambda x_i} - e^{-\lambda x_{i+1}}$, $\overline{f_i} = n\hat{p}_i$, $\chi^2 = (f_i - \overline{f_i})^2 / \overline{f_i}$.

Assume that customer arrival rate obeys negative exponential distribution. If $x^2 < x^2_{o.o5}$, accept the hypothesis. If $x^2 > x^2_{o.o5}$, reject the hypothesis.

The authors analyzed one day's observation data through MATLAB software for example and the results are as follows: Time period: 9:30am to 10:30am, $\overline{x} = 143.818$ s, $\lambda = 1/\ \overline{x} = 0.007$ people/s $= 0.417$ people/min, $n = 33$, $\chi^2 = 34.2064$. $r = 1$, χ^2 degree of freedom $n{-}r{-}1 = 33{-}1{-}1 = 31$. Take $\alpha = 0.005$, due to $\chi^2_{0.005}(n - r - 1) = \chi^2_{0.005}(31) = 56.328$, so $\chi^2_{0.005}(n - r - 1) > \chi^2 = 34.2064$. Therefore, accept the hypothesis under the dominant level of 0.005, namely customers arrival time intervals during 9:30–10:30 obey the parameter 0.007's negative exponential distribution.

It verifies that customers arrival time interval within an hour obeys negative exponential distribution. But it may reject the hypothesis at other times. Therefore, the customer arrival rate obeys Poisson distribution in the paper means customer arrival rate in different periods obeys Poisson distribution.

In addition, because the arrival of customers at different days and times are different, it is necessary to introduce a correction coefficient. Here, taking the relative weight of the negative exponential distribution parameter as the correction coefficient (Pengfei Guo 2009). The authors got λ values and correction coefficients of different time periods and different days. Then λ value of the negative exponential distribution can be calculated by the following formula: $\lambda(t) = \gamma_1(t) \times \gamma_2(t) \times \bar{\lambda}$, γ_1 is the correction coefficient of different time periods; γ_2 is the correction coefficient of different days; $\bar{\lambda}$ is the overall average and $\bar{\lambda} = 1.230$.

87.4.1.2 Average Shunt Rate of ATM: *p*

Through observation of the number of customers arrived every day and the number of people using ATM can get ATM's shunt rate of every day, then calculate the average shunt rate $p = 0.5010$. Then the rate of customers reaching the service counter $\lambda_1 = \lambda(1-p) = 0.6138$ people/min, the rate of customers reaching ATM $\lambda_2 = \lambda p = 0.6163$ people/min.

87.4.1.3 Average Service Rate μ

The average service rate includes two parts, the average service rate of bank teller and the average service rate of ATM. The data was obtained through the sample survey method. To the average service rate of bank teller, the authors make an random statistical of time the bank teller spent on handling 300 business, then calculate the average service rate of bank tellers: $\mu_1 = \frac{1}{t} = 0.325$ people/min. To the average service rate of ATM, the authors recorded the time ATM spend to handle 30 businesses, and then calculate the average service rate of ATM: $\mu_2 = \frac{1}{t} = 0.355$ people/min.

87.4.1.4 The Maximum Residence Time That Customer Can Accept

The maximum residence time that customers can accept (*Tmax*) refers to the total time that customer willing to pay to transact business from entering the business lobby to leaving. This time consists of the customers' waiting time and the service time. Statistics of different days' data show that: if the customer waiting time is less than 10 min, 100 % of customers could accept; if the customer waiting time is between 10 and 15 min, 52.72 % of customers are able to accept ; if the customer waiting time is between 15 and 20 min, only 30.95 % of customers could accept; if the waiting time is between 20 and 25 min, 14.85 % of customers may accept; if the customer waiting time is more than 25 min, there is only 4.6 % of customer acceptance. So the maximum waiting time that customers can accept is 10 min.

87.4.1.5 The Maximum Daily Workload of Bank Teller: *Qmax*

Data shows that the sample obeys non-normal distribution. Seeking *Qmax* means to seek the interval estimation of the overall mean with the parameter σ unknown

(Ivo et al. 2009). Through calculation: $\overline{X} = 110.667$, $S_{\overline{x}} = 3.064$, when $\alpha = 0.05$, the interval estimation is $\overline{X} \pm z_{0.05} \times s_{\overline{x}} = 110.667 \pm 6.005$, namely (105, 117). Therefore, when the confidence level is 0.05, the maximum daily workload of bank teller is between 105 and 117.

87.4.2 Data of Site B

Through analyzing and calculating, obtained data of site B: The customer arrival rate $\lambda = 1.35$ people/min; the average service rate = 0.25 people/min. When the value of c is between 1 and 5, its service intensity $\rho_c = \frac{\lambda}{c\mu} > 1$. In normal workdays, site B have 4 service windows in available. Therefore, this system is in extremely busy and crowded state.

The method used to get ATM shunt rate of site B is different from the front. Because the two ATMs in site B often break down, very few customers use ATM to transact business. This paper used the questionnaire method to determine the shunt rate of ATM. The authors distributed 480 questionnaires randomly to the customers reaching site B and 465 valid questionnaires were obtained. Among these customers, 257 customers have encountered ATM failure before, accounting for more than 50 %. And only 36 % of customers will choose to use ATM. Therefore, the shunt rate of ATM of site B in this paper is 36 %, namely $p = 36$ %.

87.5 Discussion

87.5.1 Optimization of Bank Foreground Support Facilities

87.5.1.1 Optimization of Support Facilities in Site A

As mentioned above, data of site A are as follows: the average customer arrival rate $\lambda = 1.2301$ people/min, the rate of customers reaching the service counter $\lambda_1 = 0.6138$ people/min, the rate of customers reaching ATM $\lambda_2 = 0.6163$ people/min. Then calculate the values of $\overline{L_1}(c)$, $\overline{L_2}(c)$ through MATLAB and the results are analyzed as follows:

When $c = 1$, the values of $L_1(c)$ and $L_2(c)$ are negative and has no practical meanings. Some scholars believe that the increased unit time cost e_1u of bank caused by increasing a service window is difficult to measure in money. And so is the unit time loss e_3 of each customer caused by waiting. Therefore assume that the costs of bank to open three new service windows are equal to the losses caused by increasing two queuing customers, namely $e_1u/e_3 = 0.6667$. As to the unit time cost e_2u of bank to increase an ATM and the unit time loss e_4 of each customer caused by waiting, assuming that $e_2u/e_4 = 0.6400$. The above analyses show that

e_1u/e_3 falls into the interval (0.5373 ∼ 14.9824) and e_2u/e_4 falls into the interval (0.3609 ∼ 4.8332). Therefore, opening three service windows and setting up three ATMs can make the total cost of bank minimum and its value is 2.3938. At the same time, calculate by MATLAB and get: the average residence time of customers in front of the service window is 7.8684; the average residence time of customers in front of ATM is 6.1589. Both of them are smaller than *Tmax*, the customers will not complain. $\lambda_1 = 0.6138$ people/min. Then the average number of customers is 294.62 and there are 98.21 customers go to each service window on average. It is less than the maximum daily workload of bank teller. Under this state, the bank teller can work efficiently and the system can run stably. Site A has three service windows and three ATMs. It is consistent with the number calculated through marginal analysis method in the paper. Therefore, the set of facilities in site A is reasonable.

87.5.1.2 Optimization of Support Facilities in Site B

The above has mentioned that: the average arrival rate of customers in site B is 1.35 people/min and customer arrival rate of the ATM $\lambda_2 = 0.36$ people/min. Therefore, customer arrival rate of the service window $\lambda_1 = 0.99$ people/min. There is no difference among ATMs. Therefore, take the value of ATM service rate in site B as same with site A, namely 0.355 people/min. Similarly, calculate by MATLAB and the results show that opening five service windows can make the cost $L_1(c)$ minimum and setting three ATMs can make the cost $L_2(c)$ minimum. Therefore, opening five service windows and setting up three ATMs in site B can make the total cost least. Site B has four service windows and two ATMs now, and its ATMs always break down so that ATMs cannot shunt customers normally. According to the above analysis, site B should open one more service window and an additional ATM while providing high-quality ATM service in order to solve the queuing problem effectively.

87.5.2 *The Influence of ATM on Queuing System*

Among the techniques used in banking industry, ATM has the biggest shunt effect when customers reaching the bank. In view of the shunt effect of ATM, the authors take site B for example and calculate the variables' values when the system is in optimal state through setting a certain shunt rate based on some estimations and assumptions about the ATM's cost. The results show that: As the shunt rate increases, the number of service counter and ATM change correspondingly. When the shunt rate rises to 20 % from 10 %, the configuration of service counters and ATMs have no change, but the total cost increased. After that, the total cost decreases with the shunt rate increases and the total cost is negatively correlated with the shunt rate.

When the shun rate is low, the bank must open more windows to meet the needs of customers. E.g. when the shunt rate is 30 %, four service windows and two ATMs are needed, but its cost is high; If the shunt rate rise to 40 %, only one more ATM is needed and the cost can also be reduced; When the shunt rate is 50 %, three ATMs can meet the needs of customers and it can also reduce a service window, what's more, the total cost is less. The purpose of commercial bank is profit, but it does not mean that the total cost as little as possible for the reason that customer satisfaction is also an important aspect to be considered. ATM has many functions, but is still not complete. Therefore banks can not open very few service windows in order to minimize the total cost.

87.5.3 Optimization Suggestions of Bank Support Facilities

In order to better ease the queuing problem, commercial banks should first pay attention to the potential shunt effect of ATM and try to improve the shunt rate of ATM through a series of measures. For example, by setting isolation area, remote monitoring center, or additional alarm devices to improve the security of ATM in order to reduce customers' security concerns; or banks can set up a dedicated ATM manager who is responsible for introducing and guiding customers to use ATM for service in order to reduce customers' waiting time. When the ATM breaks down, specialized person should be arranged to fix it timely in order to reduce customers' complaints and dissatisfaction; the bank can also introduce some value-added services which can be done through the ATM machine to attract more customers to use the ATM.

Second, commercial banks should improve the tangibility and reliability of services. Available measure to improve the tangibility of service is to improve the facilities and business environment of the bank which can reduce the customers' perceived waiting time effectively. E.g. set sufficient seats in the waiting area, set a water dispenser, and so on. To improve the reliability of the services mainly include two aspects. The first is to improve the service skills of employees and make the operation standardized in order to reduce errors. This can be achieved through appropriate training and incentives. The second aspect is to enhance the stability of the service facilities so as to ensure the security of system.

87.6 Conclusion

Nowadays, almost every bank outlet has set ATMs and other self-service equipments, but the settings are less than perfect and the queuing problem was not effectively solved. The model built in this study introduced ATM variable and added ATM cost to the total costs in order to find a reasonable set solution of bank service windows and ATMs. The model has strong applicability and can apply to all banks' support facilities optimization. The suggestions for optimization mentioned in the paper can help banks alleviating the queuing problem effectively. In addition, the results of the study on site A show that the shunt rate of ATM is not the bigger the better. Commercial banks should try to make the total service cost minimum based on the specific shunt rate of ATM. In conclusion, this study expanded the service pack optimization theory and further enriched the application research of queuing model. There is certain significance both theoretically and practically.

As with any research, there are limitations. First of all, the data related to ATM in this study is only a collection of banks' normal business hours, that is, the paper only considered ATM shunt rate during the normal business hours without analyzing ATMs' contribution rate in other time periods which will affect the accuracy of the empirical results of this study. Thus, future research can study the contribution rate of ATM during different time periods which can further show ATMs' shunt effect. Second, this study only chose two bank outlets in Harbin city as the subjects. However, the situations of banks in different regions may be different. Therefore, further research could consider expanding the research scope and doing more statistical analyses to verify the proposed model.

References

Fitzsimmons JA, Fitzsimmons MJ (2006) Service management: operations, strategy and information technology [M]. McGraw-Hill/Irwin, Boston, US

Ivo A, Antonis E, Stella K (2009) Synchronized reneging in queuing systems with vacations [J]. Queueing Syst 62:1–33

Mingsheng Yang, Xiwen Xiong, Jianhua Lin (2006) Matlab basis and mathematical software [M]. Dalian University of Technology Press, Dalian, pp 1–113

Ozgecan SU, Tayfur A (2009) Waiting time approximation in single-class queue in systems with multiple types of interruptions: modeling congestion at waterways entrances [J]. Ann Oper Res 172:291–313

Pengfei Guo ZP (2009) The effects of the availability of waiting-time information on a balking queue [J]. Eur J Oper Res 198:199–209

Wei Xiong (2005) Operations research [M]. China Machine Press, Beijing

Yinghui Tang, Xiaowo Tang (2006) Queuing theory [M]. Science Press, Beijing, pp 1–6

Chapter 88
Research of Sales Forecasting Model and Algorithm in Food and Beverage Enterprise

Li Gao, Ke-lin Xu, and Dao-xiao Mao

Abstract Dishes sales prediction is vital to Food & beverage Enterprise. Nevertheless determining the sales remains very difficult because there are many factors that may influence the Dishes sales. A novel time series forecasting model is established by evolving and clustering decision tree. This forecasting model integrates decision tree algorithms (DT) and genetic algorithms (CA) to construct a sales predictions system based on historical data and the most optimized decision tree. The example analysis and calculation result shows that the proposed GADT model is feasible and effective.

Keywords Dishes sales forecasting • Decision tree • Time series • Genetic algorithm

88.1 Introduction

It is necessary for food & beverage enterprises to solve the problem of raw material inventory, in order to win sales opportunities and avoid the loss of customers. And the best way to solve the problem is to construct a predictive model to predict enterprise's dishes sales accurately (Ying Zhong and Bingwen Wang 2002).

The dishes sales prediction belongs to time series sales forecasting for short shelf-term. At present, the traditional time series analysis method and the popular

L. Gao (✉)
College of Mechanical and Energy Engineering, Tongji University, Shanghai, China

Library of University of Shanghai for Science and Technology, Shanghai, China
e-mail: dolly2005@163.com

K. Xu • D. Mao
College of Mechanical and Energy Engineering, Tongji University, Shanghai, China
e-mail: tjklxu@163.com; maodaoxiao@126.com

E. Qi et al. (eds.), *The 19th International Conference on Industrial Engineering and Engineering Management*, DOI 10.1007/978-3-642-37270-4_88,

BP neural network prediction method are mainly used to deal with the sales forecasting for short shelf-term. Although, both of them can solve the sales prediction problem, there also exist some obvious shortcomings. For instance, Jenkins and Frantti et al. pointed out that the establishment of mathematical model is very difficult because the sales trend is nonlinear, so the use of time series analysis and modeling method to accurately predict the behavior of the system effects are far from satisfactory (Beenstock and Szpiro 2005). Chart-field gave warning to Hill and Green for they had blindly exaggerated the role of neural network. He thought, in the food & beverage industry, there is almost no systematic evidence to explain the neural network method is superior to the traditional time series prediction method (Beenstock and Szpiro 2005; Chris 1993). While doing data mining study, Michael J. A. Berry et al. found that although the neural network can predict accurately, it is with difficulties when using it because of the dishes varieties and large quantity of data. What's more, the neural network can not make a reasonable explanation for the results that are gained (Kuo et al. 2002).

Decision tree algorithm is the main technique used for classification and prediction. The famous decision tree algorithm C4.5 (Classification 4.5) proposed by Quinlan in 1993 has made remarkable achievements in sales volume prediction. However, C4.5 takes the strategy of local search which makes the decision tree not the best one (Rajkumar and Kumar 2006). Genetic algorithm (GA) is a universal global search algorithm, which simulates the natural evolution, and the advantage characteristics of global optimality and potential parallel give us a new prediction method (Li Liu and Xianping He 2008). Thus, this paper combines the genetic algorithm and the decision tree classification prediction method for constructing a GADT model to study this problem. Examples show that the method proposed has certain theoretical significance and application value.

88.2 Classification Prediction Model Based on Decision Tree

Dishes sales are often decided by many factors such as the average level, trend, seasonality, and error. Classification prediction model based on decision tree classifies the historical sales data according to the influence of many factors and decomposes the uncertain time series into many deterministic time series. Then, it will establish prediction model by using decision tree algorithm based on time series in deterministic time series (Zhihui Zhang 2008).

1. Data selection and classification. The first step is to classify the data according to the classification rules, and classify the predictive value of the tuple at the same time. The second step is to move corresponding values of tuples which are in the same class with the one required to be predicted. Moving average is what needs to be predicted (Zhihui Zhang 2008).

2. Relationship definition. Specific definition is as follows: for a given relationship $D(A, B, C\ldots)$, $D = \{T1, T2, \ldots, Tn\}$ and a set of class $C = \{C1, C2, \ldots, Cn\}$ (Ci for attribute field of D), the classification problem is to define a mapping of

$F{:}D \rightarrow C$, each tuple *Ti* is assigned to a class. A class of *Cj* contains all tuples that are mapped to this class, i.e., $Cj = \{Ti/F(Ti) = Cj,\ 1 <= i, j <= n, \text{ and } Ti \in D\}$. At the same time, assuming tuple *Tm* ($m > n$) appears in *D*, and the *Bm* value is unknown while the other values are known. Thus, according to mapping of $F{:}D \rightarrow C$, Tm belongs to a class of *Cj*, and *Bm* value can also be gained from the moving average of other tuples *Bi* that are belonged to class *Cj*.

3. Operational process. Now suppose that the attribute value of relationship *D* is ($A,B,C\ldots$), tuple is $T(T1, T2, \ldots, Tn)$, classification rule $C = \{C1, C2, \ldots, Cn\}$, prediction attribute values appear in *B*. $Tm(m > n)$ as predictors of tuples, and the *Bm* value is unknown while the other values are known.

The first step: decision tree classification. The basic algorithm of decision tree induction is a greedy algorithm, which use a top-down recursive and crush one by one way to structure a decision tree. Here, we adopted the famous C4.5 algorithm, and the basic procedures of the algorithm are shown as follows (Ming Zhou and Shudong Sun 2002):

Input: training sample data is (*T*); candidate tree set is (A, B, $C\ldots$). That is to say, the sample data set of candidate attributes of relationship *D* is (A, B, $C\ldots$).
Output: a decision tree, and the classification rules based on it.

4. The introduction of genetic algorithm. We know that group searching strategy and individual information exchange is the two big characteristics of genetic algorithm (Jun Qiu and Ruilin Zhang 2007). This paper proposed GADT model (Fig. 88.1), which uses genetic algorithm to obtain the optimal decision tree. Using C4.5 algorithm, SQL language to achieve the solution of GATD model genetic algorithm, an example of calculation shows that the algorithm and the model are effective.

88.2.1 The Definition of Genetic Operators and Fitness Function

In the scheme which was used by the genetic algorithm to explore the optimal decision tree discussed here, decision trees do not encode, all genetic operators are directly operated in decision tree. From the features of the structure of decision tree, we can definite appropriate genetic operators that can be easily operated.

There are three main kinds of genetic operators: reproduction operator, crossover operator and mutation operator (Yong Xiao and Yiyun Chen 1998). We may define the fitness function firstly in allusion to the structure and the characteristics of decision tree.

Introducing genetic algorithm to the evaluation of the tree, can take the depth of the tree, nodes and other factors into consideration.

Therefore, the fitness function is:

fitness_func=w_err*err_rate+c*(w_depth*ave_depth+w_num*node_num)

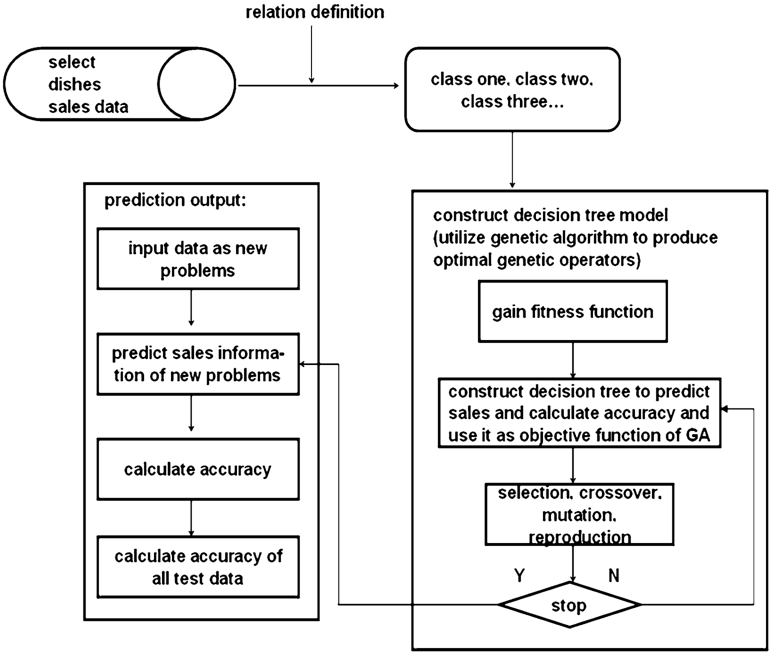


Fig. 88.1 GATD detail model

while,

w_err, w_depth, w_num are respectively the error rate, average depth, extreme value that the node number corresponds to. Because the error rate, the mean depth and node number have different units, a balance coefficient C has been added.

88.2.2 Operators Regrouping and Mutation

According to the structure characteristics of the decision tree, we have the following several kinds of reorganization:

(a) Use descendant node to replace ancestral node, similar to the pruning operation of tree. (b) Two nodes exchange within a tree (Fig. 88.2). Node A exchanged with node B, while A, B cannot be ancestor and descendant relationship. (c) Node exchange between two trees (Fig. 88.3), node A in tree1 exchanged with node B in tree2. Here, nodes exchange means exchanging the sub-trees which actually take the nodes as roots.

For the decision tree, we defined the mutation operators to change the structure of the tree or the information of the nodes in the tree (Yong Xiao and Yiyun Chen 1998). In allusion to internal nodes or leaf nodes mutation operations may have the following results (Kuo 2006; Doganis et al. 2006): (a) Change the attributes

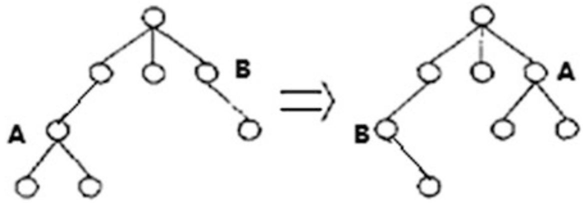


Fig. 88.2 Node exchange of a tree

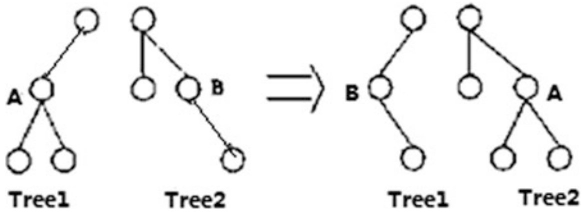


Fig. 88.3 Node exchange of two tree

judged on nodes; (b) Change the grouped cases of attribute values (if the attribute values of the nodes have been divided into groups); (c) Change the branch condition of the sub-trees following the node and the correspondence of attribute values and branching sub-tree; (d) change the branch number of the node (if the attribute values of the nodes have been divided into groups). For leaf nodes, mutation operation may change the conclusions represented by the leaf nodes.

The new decision tree required for structural integrity and consistency process after recombination and mutation operator. Adjustment the tree structure of variable nodes and the child nodes to make it a completed and correct decision tree; get rid of the duplicate attribute judgment from the root to the leaf node etc..

88.2.3 Model Operation Steps

The process of decision tree construction is to use genetic algorithm to produce the next generation of groups through genetic operator operation from the previous generation decision tree populations and evolve gradually until it meets genetic algorithm termination conditions.

1. The produce of the first generation of groups: setting the initial value

The performance of decision tree gained by using C4.5 algorithm is better; it shows a stronger adaptability in genetic process. In order to reduce the number of search, we use C4.5 method to generate the first generation of decision tree, and select different subsets of the set as a window, and construct decision tree respectively. These decision trees of better performance are regarded as the initial set which removes the blind exploration at the beginning.

2. Produce the next generation: determine the fitness function

Use the reproduction operator to choose the reproduction decision tree from the generation group as the reorganization of the candidate set. From the definition of fitness function we can know that the smaller the fitness function is, the bigger probability it will be chosen. Because the value of fitness function is in a certain range, so, we make the maximum and minimum values in a map.

The results of recombination and mutation may not move in the right direction, but we can add a recombination probability and mutation probability to control their operation.

3. Optimal decision tree: the determination of ultimate goal functions

Here optimal decision tree is actually a decision tree of minimal fitness function value (Frantti and Mahonen 2004; Pedrycz and Gomide 2001). In the process of evolution, the last generation of groups does not necessarily include the optimal decision tree, so we must save the best tree in the whole process. When producing the next generation, update and save the best tree constantly. The saved decision tree is the result that we get until the evolutionary process ends.

88.3 Calculation and Analysis

Next we will use the shrimp sales data of a dining Limited company in February and March 2009 (Table 88.1) to forecast sales in March 2010 and verify the accuracy of the model.

Table 88.1 Shrimp sales in Feb. and Mar. 2009

Date	Sales data of February	Weather	Day	Holiday	Sales data of March	Weather	Day	Holiday
1	25	Sunny	4		26	Sunny	4	
13	33	Sunny	2		34	Sunny	2	
14	40	Sunny	3	Valentine’s Day	30	Rainy	3	
15	26	Sunny	4		32	Rainy	4	
16	27	Sunny	5		40	Rainy	5	
17	46	Sunny	6	30th of lunner month	28	Rainy	6	
18	25	Sunny	7	Spring Festival	17	Sunny	7	
19	20	Sunny	1	2nd of lunner month	31	Sunny	1	
20	15	Rainy	2	3rd of lunner month	34	Sunny	2	
21	14	Rainy	3	4th of lunner month	33	Sunny	3	
22	10	Rainy	4	5th of lunner month	28	Sunny	4	
23	16	Rainy	5	6th of lunner month	29	Sunny	5	
24	18	Rainy	6	7th of lunner month	25	Sunny	6	
25	25	Rainy	7	8th of lunner month	14	Rainy	7	

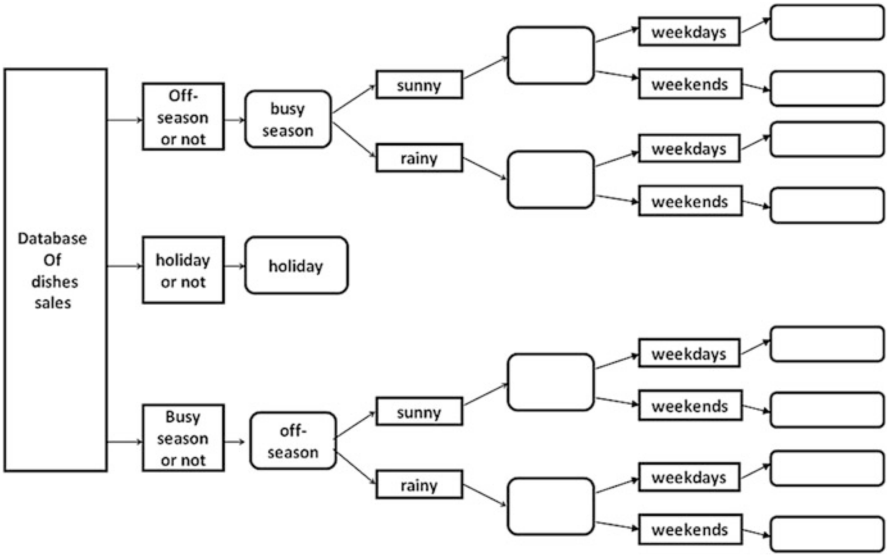


Fig. 88.4 Decision tree of sales data

88.3.1 Data Classification

According to the method of decision tree classification, we give the relationship: D (off-season, holiday, busy season . . .) and a set of class C = {weekend, weekdays} (*Ci* for attribute field of *D*), the classification problem is to define a mapping of F: D → C, a class of Ci contains all tuples that mapped to this class.

Next, we will use Microsoft Visual Studio (R) 2005 decision tree to classification mine model and build dishes sales data. First of all, create Analysis Services project in Visual Studio 2005, then create connection to Data Source objects of SQL Server example which includes sales classification data. And a data source view contains food grade properties and total daily sales of a single table must also be created. After adding the data source view, new mining structure and the mining model will be created for the classification exercises of historical sales data. The mining structure definition will be used to construct the structure of historical sales data classification model (Song and Chissom 2004). Finally, use the internal decision tree algorithm of SQL SERVER 2005 Analysis Service to gain the final overall decision tree. As shown in Fig. 88.4.

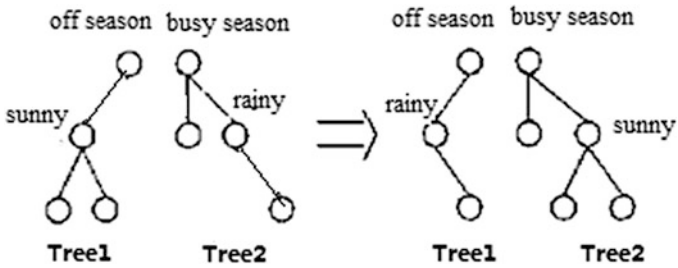


Fig. 88.5 Operators regrouping

Table 88.2 Forecasting results through the most optimized decision tree

Date	Actual sales quantity	Forecast quantity	Prediction error
March 2, 2010	26	25	1
March 3, 2010	20	19	1
March 4, 2010	31	30.5	0.5
March 5, 2010	30	30.5	–0.5
March 6, 2010	29	30	–1
March 7, 2010	31	30.5	0.5
March 8, 2010	32	32.5	–0.5
March 9, 2010	30	33	–3
March 10, 2010	27	29	0
March 11, 2010	31	30.5	0.5
March 12, 2010	35	36	–1

88.3.2 Classification Diagram (Fig. 88.4)

Exchange node A and node B, then we obtain an optimal tree (Fig. 88.5).

In order to highlight the gap of fitness function value and speed up the search process, we use the relative fitness function value, $-1 \leq \text{w\_err} \leq 1$, here we also make a mapping between maximum and minimum value. Finally, we get the prediction results through the optimal decision tree as shown in Table 88.2.

88.3.3 Comparison and Validation of Predicted Results

Next we choose the data from March 10 to 16 2010 for validation as shown in Table 88.3.

The mean absolute deviation based on the GADT model: $MAD = 10/7 = 1.43$

Accuracy rate of prediction based on the decision tree classification: $POA = 202/206 \times 100\,\% = 98\,\%$

Absolute deviation of GADT model: $MAD = 44.8/7 = 6.4$

Table 88.3 Results comparison of different algorithm

Date	Actual sale quantity	Prediction quantity of this model	Prediction quantity of decision tree classification
March 10, 2010	30	33	24.3
March 11, 2010	31	30.5	25.7
March 12, 2010	35	36	24.7
March 13, 2010	34	31	27.3
March 14, 2010	37	36	30.2
March 15, 2010	26	24	24.3
March 16, 2010	16	15.5	27.3
sum	206	202	183.8

Accuracy rate of prediction based on the decision tree classification: $POA = 183.8/206 \times 100\,\% = 89.2\,\%$

The mean absolute deviation based on the GADT model: $MAD = 10/7 = 1.43$

Accuracy rate of prediction based on the decision tree classification: $POA = 202/206 \times 100\,\% = 98\,\%$

Absolute deviation of GADT model: $MAD = 44.8/7 = 6.4$

Accuracy rate of prediction based on the decision tree classification: $POA = 183.8/206 \times 100\,\% = 89.2\,\%$

From the Table 88.3, we can see that the mean absolute deviation MAD of decision tree classification prediction method with genetic algorithm introduced in is far less than the one got by the decision tree method, and prediction accuracy POA is far higher than the one got by the decision tree method. Therefore, the decision tree classification prediction method based on genetic algorithm is more accurate than the simple decision tree prediction method.

88.4 Discussion

Because the tree is not easy to code, our genetic operator is directly operated based on the tree structure when using genetic algorithms to construct decision tree. Seeking a good coding method is what we should continue to study. In addition, a few things have been done when dealing with the impact caused by a small amount of error event in the definition of fitness function. So to introduce the concept of quotient in to adjust the fitness function expression is also the direction we are striving to.

References

Beenstock M, Szpiro G (2005) Speciation search in nonlinear time-series models using the genetic algorithm [J]. J Econ Dyn Control 18:811–835

Chartfield C (1993) Neural networks: breakthrough of prediction or fashion prediction [J]. Int J Forecast 4:70–71 (Translated by Ming Guo)

Doganis P, Alexandridis A, Patrinos P, Sarimveis H (2006) Time series sales forecasting for short shelf-life food products based on artificial neural networks and evolutionary computing [J]. J Food Eng 75:196–204

Frantti T, Mahonen P (2004) Fuzzy logic-based forecasting model [J]. English Appell Intel 14:189–201

Jun Qiu, Ruilin Zhang (2007) Recurrent neural networks in sales forecasting based on Genetic algorithm [J]. J Zhejiang Univ 24(3):266–270

Kuo RJ (2006) A sales forecasting system based on fuzzy neural network with initial weights generated by genetic algorithm [J]. Eur J Oper Res 129:496–517

Kuo RJ, Wu P, Wang CP (2002) An intelligent sales forecasting system through integration of artificial neural networks and fuzzy neural networks with fuzzy weight elimination [J]. Neural Netw 15:909–925

Li Liu, Xianping He (2008) Time series forecasting model based on genetic algorithm and fuzzy decision tree [J]. Comput Eng Des 10:5044–5056

Ming Zhou, Shudong Sun (2002) The principle and application of genetic algorithms [M]. National Defense Industry Press, Beijing, pp 33–64

Pedrycz W, Gomide F (2001) An introduction to Fuzzy Sets [J]. Anal Design 10:234–252, The MIT Press Cambridge, MA

Rajkumar V, Kumar V (2006) A genetic algorithms approach to growth phase forecasting of wireless subscribers [J]. Int J Forecast 18:625–645

Song Q, Chissom BS (2004) Forecasting enrolments with fuzzy time series-Part [J]. Fuzzy Set Syst 62:1–8

Ying Zhong, Bingwen Wang (2002) BP neural network prediction model of time series based on genetic algorithm [J]. Syst Eng Electron Technol 124(14):9–11

Yong Xiao, Yiyun Chen (1998) Using genetic algorithms to construct decision tree [J]. J Comput Res Dev 1:49–52

Zhihui Zhang (2008) The research and application of food & beverage sales forecasting system based on the data warehouse [D]. China Three Gorges University, Yichang

Chapter 89
Waste Collection Vehicle Routing Problem: Case Study in Alexandria, Egypt

A. Moustafa, A.A. Abdelhalim, A.B. Eltawil, and N. Fors

Abstract The problem of the solid waste collection is becoming one of the most important applications addressed as Vehicle Routing Problem with Time Windows (VRPTW). This paper presents a solution of a large scale vehicle routing problem for the waste collection in Alexandria city in Egypt. A GIS based model is developed and the VRPTW is solved using TransCAD® software. Different scenarios were developed and compared. The case was solved with three clustering scenarios and without clustering. The results showed superiority of clustering based scenarios over the solution without clustering. The developed model is used as a Spatial Decision Support System (SDSS) that allows investigating different scenarios for solid waste collection vehicle routing and improving the overall logistics performance of such systems.

Keywords Decision support system • Geographic information system • Vehicle routing problem

A. Moustafa (✉)
Nahdet Misr for Modern Environmental Services, Alexandria, Egypt
e-mail: amani.mostafa@nahdetmisr.com.eg

A.A. Abdelhalim • A.B. Eltawil
Industrial Engineering & Systems Management department, Egypt Japan University of Science & Technology, POBox 179, New Borg El-Arab City 21934 Alexandria, Egypt
e-mail: alyaa.adel@ejust.edu.eg; eltawil@ejust.edu.eg

N. Fors
Production Engineering Department, Faculty of Engineering, Alexandria University, Alexandria, Egypt
e-mail: fors@dataxprs.com.eg

E. Qi et al. (eds.), *The 19th International Conference on Industrial Engineering and Engineering Management*, DOI 10.1007/978-3-642-37270-4_89,

89.1 Introduction

Waste collection is a highly recognized municipal service that involves operational, tactical and strategic decisions that are a vital part of any city public service. Major portion of the expenditure in these systems is regarded to the distribution and logistics in the waste collection. Two main types of waste can be found, either it is a solid waste generation or medical or hazardous waste generation that after the appropriate treatment will become in itself a normal solid type of waste. This paper focuses on the solid waste generation type.

On the other hand, the advancement in Geographic Information Systems (GIS) changed the directions of research in many areas. In particular, in modeling the distribution networks with the vast amount of applications associated with it. Entities concerned with the routing of vehicles became even keener to invest in such tools to be able to have more accurate detailed information about the logistics networks. Hence, the optimization models are being more representing of the real world than the past assumptions were. Waste collection vehicle routing is one of the applications that took advantage of the advancement in GIS tools. Not only for the increased awareness towards the environmental concerns, but equally important is the concern on the public health as well. On the other hand, the rapid growth in population around the world, especially in the developing countries, with the natural consequence of having a rapid increase in the waste generations have taken major focus of research into waste collection management miscellaneous issues and problems.

Nonetheless, there are lots of tough challenges when it comes to the waste collection in Egypt. Egypt doesn't have a municipal solid waste management law. It only has scattered pieces of legislation as a legal framework of solid waste management (Tarek and Amine 2010). More importantly, one of the most difficult issues is being faced here in Egypt is not having a reliable digital map of the roads, services and facilities for each city in Egypt, which makes the GIS advanced technologies difficult to apply. Consequently, huge amount of geocoding efforts was done in this work to bring the map used in this work live and usable.

The objective is to use GIS and vehicle routing models to develop improved waste collection plans that will increase the efficiency of collection and yet reduce the cost in travel distance and time. An important side effect is the reduction of emissions and roads congestion.

89.2 Literature Review

In the recent decade, waste management and waste collection issues took the attention not only of the governments but the attention of researchers as well. This is mainly is due to the rapid increase in waste generation especially in the developing countries. A significant number of articles are published in the area considering the vehicle routing problem with time windows as one of the major issues in waste collection management. There is a dilemma of not visiting a collection point too early before a certain amount of waste is gathered as it is considered a waste of the vehicle' operational cost including the driver's time and visiting a collection point after the time specified that will generate too much waste that will impose hazardous effect on its surrounding area and this is also unacceptable. Hence, not only time windows consideration is a critical issue, but also updating the routes for the waste collection easily and frequently is of a critical issue. Thereupon, providing decision support system (DSS) for the waste collection organization is becoming a necessity.

Consequently, the advances in GIS have made a great impact in such applications. Several publications made use of these tools in developing waste collection. In an attempt to reduce the fuel consumption and atmospheric pollutant emissions, Tavares et al. (2009) used 3D GIS for the optimization of the municipal solid waste collection routes. The 3D-GIS had significant benefit in their case due to the vast variances of inclination in the roads of Praia city, the capital of Cape Verde, where they applied their model. The authors investigated two scenarios; the first one minimizes the fuel consumption and the second one gives the shortest path. The two solutions gave two completely different routes. And the question one shall ask is that which one of the two scenarios is better. The answer to this question will largely depend on the specific case and the objectives from such model with the priority given for each objective. In another case study in India in Asansol Municipality, Ghose et al. (2006) developed modules in Arc-GIS software to come up with the solution of routing for a heterogeneous fleet of vehicles which collect different sizes of bins. Alumur and Kara (2007) considered the hazardous waste location routing problem in three stages; treatment, disposal, and recycling. The authors built a model that determines where to open a treatment location and which technology to use, either chemical treatment or incinerator, and where to open a disposal center. Both decisions are being made while considering the routes of the hazardous waste to treatments locations and those to disposal centers. They presented a multi objective mixed integer programming formulation where the objectives were to minimize the total cost and transportation risk. Some demographic data were calculated using ArcView 3.1 software in the bandwidth of 800 m from one node to another. They solved a 92 nodes using CPLEX version 8.1 comparing their problem size to those found in literature of 10–15 candidate sites. However, the authors didn't consider time windows nor benefiting from the use of heuristics problem to solve larger problem size.

In Greece, the waste collection and transport (WC&T) accounts for 70–100 % of the total municipal solid waste management costs, as stated by Chalkias and Lasaridi (2009). Hence, they developed a model in ArcGIS Network Analyst to enhance the efficiency of the optimization of the vehicle routing. The authors studied two scenarios, the first was concerned with the collection vehicle routing optimization. The second was a scenario considering the reallocation of bins and routing optimization. Both scenarios provided better solutions in terms of collection time savings when compared to the current plan provided in Athens. However, the second scenario showed superior results over the first one. Another improvement in the total cost reduction by applying a GIS based vehicle routing model is obtained by the work of Apaydin and Gonullu (2007). The authors used Route View Pro™ as an optimization tool to solve their problem. Their optimization model proved success when compared to the current plan adopted by the city municipality and a 24.7 % benefit in total expenditure will be acquired. Some other work also focused on solving a solid waste collection vehicle routing and scheduling case for specific cities as found in Li et al. (2008) who solved a truck scheduling problem for the city of Porto in Brazil. Furthermore, Kyessi and (2009) considered the case of Dar es Salaam city in Tanzania. They gave an emphasis on the significant impact of using the GIS tools in modeling the waste collection problems efficiently. Likewise, Firinci et al. (2009) made a pilot study for optimizing the waste collection vehicle routes in Northern Cyprus with the sane emphasis on the better modeling of such problems as a GIS-based models.

In modeling one of the cases in Egypt, Badran and El-Haggar (2006) presented a Mixed Integer Linear Programming (MILP) formulation to solve the municipal solid waste management in Port Said, Egypt. The authors solved the problem using Modeling Programming Language (MPL) software V4.2. (Maximal Software, Inc.). Different scenarios were developed and evaluated with the best scenario, minimum total cost being the one with the mixed capacity with no limit in source waste flow. After the necessary calculation, the system total profit is 49,655.8 LE/day (US$ 8,418.23/day). However, there were some limitations on their work. They didn't consider the routing of the collection and transfer vehicles within the districts. Meanwhile, they experienced same problem we face is that the quality of data about the waste sources distribution was low and the detailed maps were not available though they didn't generate one with the streets and nodes locations as present in this work.

Although all the presented cases have done fairly good work in addressing such problem, however, most of them tackled a small to medium size problem. In one aspect this is not a flaw as long as the service is made and the objective of obtaining the plan is done with satisfaction from both the service provider and the client. However, considering large scale problems for a whole city or district may result in more cost savings and better service quality which was not extensively tackled in literature.

89.3 Logistics Systems in Waste Collection: Alexandria Case Description

89.3.1 Logistics in Waste Collection

The waste collection logistics systems impose significant hazardous risk on the public health and environment as well if managed poorly or improperly. Moreover, these systems contribute to the private and public sectors economy due to the large amount of operational expenses needed to run such systems. Briefly, the logistics system in waste collection consists of a set of facilities linked via transportation services. The facilities can be those representing sites where waste is either treated, recycled or dumped. The cycle starts from the waste generation nodes taken by the local transport to the transfer station and then being hauled to the disposal or landfill facility.

89.3.2 Alexandria Municipality

Alexandria is the second largest city in Egypt (2,118 km^2) with a population of about four million. Alexandria extends about 32 km along the coast of Mediterranean Sea. The total annual municipal solid waste in Egypt has increased more than 36 % since 2000 to the current level of 20.5 million tons (MT) per year where 1.35 MT/year is generated in Alexandria (Tarek and Amine 2010). A local company is currently taking the responsibility of the waste collection in Alexandria, Egypt. It is a promising national company that carries a vision for Egypt towards the waste treatment using the recent modern technological advancements. Waste collection involves high operational costs besides other large expenditure such as the labor spare part and fuel. As a result, the main problem is the routing of vehicles as it constitutes large part of the total operational cost. A problem description is given in the following section.

89.3.3 Problem Description

The problem is formulated as a vehicle routing problem with time windows and is concerned with collecting a number of geographically dispersed bins frequent enough to avoid the overflow of the waste and in the same time to avoid redundant visits. Out of 18 districts in Alexandria, this paper represents the solution of four districts. The number of nodes to be served in the selected area of study is 302 nodes. Compared to the presented cases in literature of 15, 20, and may be 90 nodes, this is a large scale network. Yet, with this large scale problem it is only part of Alexandria and not the whole city as shown in Fig. 89.1, which shows the geocoded area and gives a glimpse on the roads structure. As it can be seen from the figure, Alexandria doesn't have a grid like structure nor it has districts with similar shapes and not even a pattern to follow which makes it even harder to geocode it and crucial to be solved using a DSS rather than the manual random planning.

The total number of bins to be collected in those 302 nodes is about 1,089 and the bin size is 770 l. The vehicle capacity is 100 bins and there is only one single depot where all vehicles start and end at. Additionally, there are time windows allowed for each stop. In case of serving any customer after the latest allowable time, a penalty to the governorate of 20 EGP will be imposed for every uncollected bin and 200 LE for every street hasn't been cleaned, and in case of arriving to the stop before the earliest

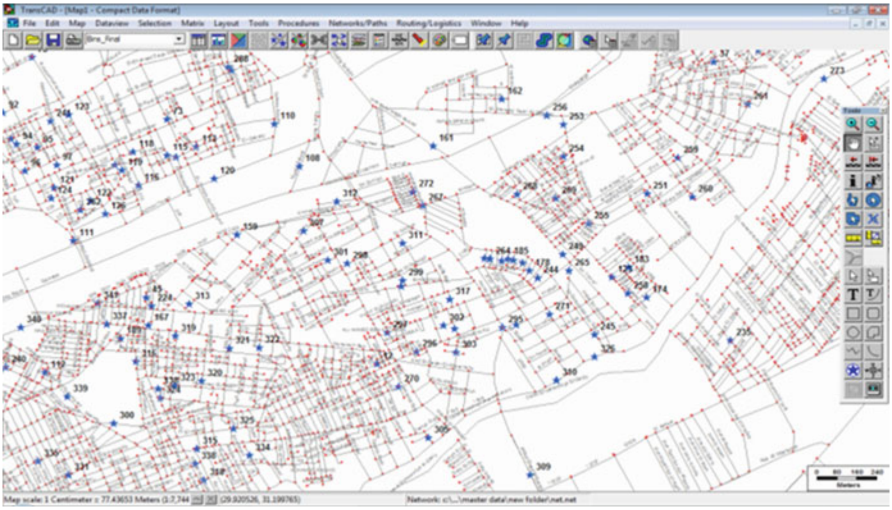


Fig. 89.1 The study area of Alexandria municipality after geocoding

time of its allowed time window this waiting time is assumed to have a penalty equal to the operating cost. The waste quantity of each stop is deterministically known and all waste must be collected.

The objective considered is to minimize the total time spent in collection, while determining concurrently the number of vehicles needed to accomplish the services, and preserving the service time within the allowable time windows of each stop. The problem is solved as a vehicle routing problem with time windows using TransCAD®. Five scenarios were considered and compared, first, the vehicle routing without clustering the nodes. Second, cluster first then dispatch the fleet of vehicles to the assigned clusters and this part contains three different clustering scenarios and the last scenario was considering the current plan of the company itself. TransCAD® clustering procedure allows modeling the cluster size to a maximum value which limits the number of features assigned to each cluster in order to achieve a clustering balance level. Accordingly, the model assigns feature to a cluster until the cluster capacity is reached (Caliper 2007). Due to the lack of digital maps, the plan was made to implement the following activities to achieve the goal of solving the collection problem:

- Preparing data collection templates
- Data collection
- Data verification and validation
- Data processing
- Problem detection
- Problem solving
- Solution implementation
- Service level monitoring.

89.4 Results and Discussions

The first scenario was solved, and the resulting routes were obtained with total amount of travelled distance equal to 113.9 km; Fig. 89.2. Meanwhile, with the clustering first scenario, three clustering iterations were investigated;

– First iteration, (SC1), with 20 clusters.
– Second iteration, (SC2), with 15 clusters.
– Third iteration, (SC3), with 11 clusters.

There are two main outputs from the clustering procedure: An assignment table where it indicates the clustering number where each stop is assigned to, and a new selection set of seeds. Seeds are those sets that contain those locations chosen as the best fit representing the clustered nodes in terms of proximity. In order to enhance the decision making a set of others parameters were also determined to compare between the different scenarios; the number of uncollected nodes, which is called the “orphans”, is calculated and more importantly the vehicle utilization in tons per hour is also determined in each scenario; as shown in Fig. 89.3.

Table 89.1 summarizes the resulting routes in each scenario with the values of the performance measures that were mentioned earlier and other parameters also calculated. The longest route scenario associated with the longest distance travelled as well was the result from adopting the VRP without clustering. Markedly, the number of trips in SC1 and SC2 is equal although the cluster size is different. The best solution that gives the minimum time spent in collection is (SC3); Fig. 89.4. There was a significant savings when the best clustering iteration SC3 compared with the results of VRPTW without clustering travelling time decreased by 4.7 %, travelling distance decreased by 51.45 %, and vehicle utilization improved by 5.26 %. All those results were compared to each other without mentioning what

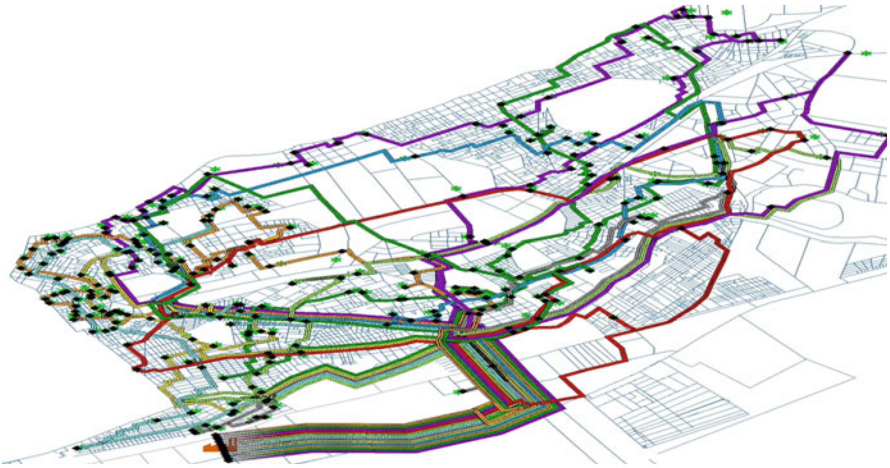


Fig. 89.2 Vehicle routing solution without clustering

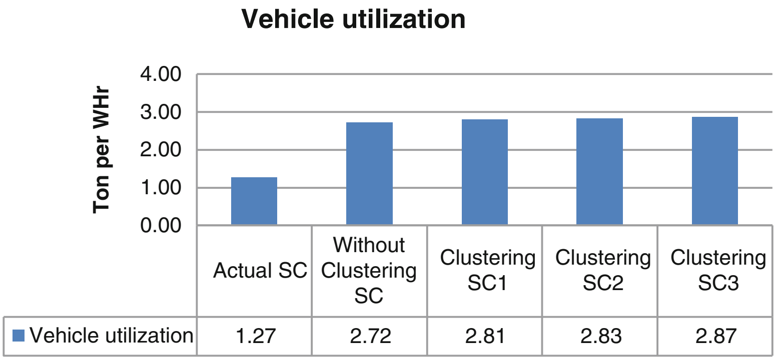


Fig. 89.3 Vehicle utilization comparison for the five scenarios

Table 89.1 Summary table of the four investigated scenarios in part Alexandria municipality

Total	Without clustering	SC1	SC2	SC3
Time (h)	39:58:00	38:51:00	38:34:00	38:05:00
Distance (Km)	113.9	79.1	70.4	55.3
Pickup (bin)	1089	1089	1089	1089
No. of trips	12	13	13	11
Vehicle utilization (ton/h)	2.72	2.81	2.83	2.87
No. of orphans	0	0	0	0

Fig. 89.4 The best solution scenario; SC3

happens in the actual running case of the company. Hence, and in this regard, it worth mentioning that a tremendous effect of clustering in SC3 is recorded when compared with the real or the actual routes designed manually and based on the same data TransCAD® scenarios were fed upon. Where there is a 55.5 % reduction in travelling time, 94.1 % reduction in travelling distance, and vehicle utilization improved by 125 %. This in itself emphasizes on the importance of having a decision support system tool to help in the decision making process in any organization.

89.5 Conclusions

The lack of good digital maps of Egypt and its governorates is making some difficulties faced by the Egyptian government and different organizations, specifically those logistics providers that would support in accurately planning the distributions models representing their organization' issues. In addition, the structure of Egypt map network is not an easy one as it is not a grid like nor it has some kind of districts similarity and not even alike in shape. Consequently, to digitize such maps, tremendous efforts must be executed to accomplish such hard task. This paper presents a decision support system to solve the waste collection problem in part of Alexandria municipality, Egypt.

In modeling such critical problems, it is not only the minimization of the total distance traveled or time spent in accomplishing a certain service that is crucial rather than focusing on its impact on the extent to reduce the harmful emissions.

This paper illustrates the application of Spatial Decision Support Systems SDSS in improving the waste collection system in a pilot study in Alexandria city, Egypt. The problem is modeled as a vehicle routing problem with time windows with single depot and homogeneous fleet of vehicles. There are four districts out of 18 districts considered in this case with 302 collection nodes and 1,089 bins. Four scenarios were investigated started with a scenario adopting no clustering before generating the routes and the other three scenarios considered a cluster first route second approach.

Furthermore, the best solution resulted from setting the cluster size to 11 clusters. More importantly, there is significant savings obtained from the best solution comparable to the current plan of the company. To the best of our knowledge, this is the first implementation for such technologies in such a large scale implementation in Egypt and in the region. An important result is the reduction of emissions resulting from the waste collection activity, also, the positive impact on the environment due to the better performance of the waste collection system. In the future, two major projects will be implemented based on the successful results of this work. First, a system for medical waste collection using GIS and Radio Frequency Identification (RFID), in the second project, a full implementation of GIS and SDSS technologies will be implemented in a full scale waste collection system in the city. One shall include more districts of Alexandria, if not the whole city. Additionally, applying the Radio Frequency Identification (RFID) technology more provide degree of dynamicity to the model. The number of clusters in this developed model was an input to the model it may be considered in future model development as a result.

Acknowledgment This research project is sponsored by the Egyptian ministry of higher education grant and the Japanese International Cooperation Agency (JICA) in the scope of the Egypt Japan University of Science and Technology. Moreover, the authors would like to thank Nahdet Misr For Modern Environmental Services for providing the data for the case study.

References

Alumur S, Kara BY (2007) A new model for the hazardous waste location-routing problem. Comput Oper Res 34:1406–1423

Apaydin O, Gonullu MT (2007) Route optimization for solid waste collection: Trabzon (Turkey) case study. Glob NEST J 9:6–11

Badran MF, EL-Haggar SM (2006) Optimization of municipal solid waste management in Port Said – Egypt. Waste Manag 26:534–545

Caliper (2007) TransCAD Transportation Planning Software, routing and logistics with TransCAD®, User Guide, version 5.0, Caliper Corporation

Chalkias C, Lasaridi K (2009) A GIS based model for the optimization of municipal solid waste collection: the case study of Nikeas, Athens, Greece. WSEAS Trans Environ Dev 10:5

Firinci N, Celik A, Akün E, Atif Khan M (2009) A pilot study for the optimization of routes for waste collection vehicles for the Göçmenköy district of Lefkoşa. Proc World Acad Sci Eng Technol 37:2070–3740

Ghose MK, Dikshit AK, Sharma SK (2006) A GIS based transportation model for solid waste disposal – a case study on Asansol municipality. Waste Manag 26:1287–1293

Jing-Quan Li, Borenstein D, Mirchandani PB (2008) Truck scheduling for solid waste collection in city of Porto Alegre, Brazil. Omega Int J Manag Sci 36:1133–1149

Kyessi A, Mwakalinga V (2009) GIS application in coordinating solid waste collection: the case of Sinza Neighbourhood in Kinondoni municipality, Dar es Salaam city, Tanzania. SDI in Municipality and Natural Resources Management

Tavares G, Zsigraiova Z, Semiao V, Carvalho MG (2009) Optimisation of the MSW collection routes for minimum fuel consumption using 3D GIS modeling. Waste Manag 29:1176–1185

Zaki T, Khayal A (2010) Country report on the solid waste management in Egypt. SWEEP-NET

Chapter 90
Determining Lean Practices in Health Care Service Systems

P. Punnakitikashem

Abstract Health care industry is a rapidly growing sector in most countries in the world. Hospitals strive to find efficient approaches to improve their service operations. Lean management has been widely recognized from manufacturing and service organizations as managerial philosophy to reduce wastes and increase values to customers. Recently, lean management was adopted in many health care organizations. To promote the success of lean implementation, it is crucial to understand the extent to which lean practices are implemented and their impacts in organizations. Thus, the purpose of this paper is to review lean management, lean practices, and their impact with respect to health care service system context. After reviewing literature, lean practices in health care, important service operations, were determined. In addition, this paper identifies five dimensions of operational performance in health care service organization. These five dimensions are speed, cost, productivity, quality, and customer satisfactions. For the next step of this research project, an empirical study investigating the relationship between lean practices and their impact on operational performance will be carried out.

Keywords Healthcare • Hospitals • Lean thinking • Lean • Lean management

90.1 Introduction

Health care is an important sector for every country as it involves health and well-being of population in the nation. The challenges of health care industry are from tremendously increasing cost and increased quality required from the patients. Many hospitals attempt to adopt innovative management concept with the aim to

P. Punnakitikashem (✉)
College of Management, Mahidol University, Bangkok, Thailand
e-mail: prattana.pun@mahidol.ac.th; prattana_p@yahoo.com

E. Qi et al. (eds.), *The 19th International Conference on Industrial Engineering and Engineering Management*, DOI 10.1007/978-3-642-37270-4_90,

improve their service operations. Many quality initiatives concepts have diffused to the health care sector. By implementing such management concept efficiently and effectively would ultimately help in delivering the highest value to patients. This leads to questions related to 'what' and 'how' management practices should be implemented to improve the overall hospital performance.

Lean management, originated from automotive industry, is one of the latest organizational innovations. Recently, many health care service organizations, ranging from those in developed countries, e.g., United Kingdom, the United States, Canada to emerging economies countries, e.g. Thailand, have adopted lean management to eliminate waste, thereby creating value of service delivered to the end customers.

Thus, it is crucial that the organizations understand impact of practices of lean implementation to the hospital performance. Research which provides a greater understanding of the impact of lean implementation in health care service organizations is needed. Thus, we start with the following research questions:

1. *What practices of lean thinking have been put into health care organization?*
2. *What are the impacts of lean implementation to the health care organization?*

This paper aims to identify practices of lean implementation in service operations, particularly in health care organizations. The impact of these management practices on hospital performance is investigated. Future research work will examine the extent to which lean practices are implemented by hospitals in Thailand and their impact on hospital performance.

This paper is organized as follows. The next section reviews literature on lean thinking concept, practices in lean management, lean management in health care, and performance of successful lean implementation. The Sect. 90.3 proposes lean practices in health care. The Sect. 90.4 illustrates the impact of lean management practices on hospital operational performance. Finally, the last section provides conclusion and future research.

90.2 Theoretical Background

90.2.1 Lean Thinking Concept

Lean thinking evolved from production philosophy at Toyota Motor Company, called Toyota Production System (TPS) over several decades (Ohno 1988). The TPS lies on the two basic components, which later characterizes lean productions; these two components are just-in-time (JIT) and automation (jidoka) (Ohno 1988). JIT production is delivering or producing when customer needs at the amount they need. JIT is a pull system in which production occurs when signaled by the customer's need with the minimum resources. Jidoka is the use of automation making problems visible when problems occur.

Lean thinking or lean management is the concept to improve organizational performance through waste elimination and values creation to customers. Lean management is defined ranged from philosophy, concepts, principles, bundles of practices, to a collection of tools for waste elimination. Lean thinking involves five important principles as follows: specify values, identify value stream mapping, emphasize continuous flow, use pull systems, and strive for perfection (Womack and Jones 1996, 2003). Womack et al. (2005) mentioned in that

> Lean thinking is not a manufacturing tactic or a cost-reduction program, but a management strategy that is applicable to all organizations because it has to do with improving processes. (Womack et al. 2005)

The major purpose of lean management is to eliminate waste (muda) from every activity performed in organizations. Several studies reported the benefits of lean implementation including higher level of quality and productivities, better customer responsiveness (Krafcik 1998). A number of studies investigated lean implementation in organizations in manufacturing industry. These studies revealed lean practices and their impact on performance (Dal Pont et al. 2008). In service sector, the practices are different, literature of management practices on lean implementation and its impact in service sector is still limited.

90.2.2 Practices in Lean Implementation

Lean practices were referred as a set of activities undertaken in an organization to promote effective management of its lean projects. In the past decade; there have been a number of studies that focused on lean thinking and their practices in manufacturing context. In manufacturing, Shah and Ward (2003) described 22 lean practices in lean systems, which they summarized from 16 major references. In addition to presenting lean practices in manufacturing, the previous studies also recommended to combine each individual practice into bundles in order to reflex the nature of multi-dimension of lean manufacturing. Previous research studied lean practices in manufacturing individually, lean practices combined to factors, and lean practices combined into bundles (Shah et al. 2008). Lean bundle is referred as a set of lean practices that are interrelated and internally consistent (Dal Pont et al. 2008; Shah and Ward 2003; Shah et al. 2008).

Shah and Ward (2003) described 22 lean practices which were combined into four lean bundles including (1) Just-in-Time (JIT), (2) Total Preventive Maintenance (TPM), (3) Total Quality Management (TQM), and (4) Human Resource Management (HRM). The "JIT bundle" included practices related to production flow, continuous flow, time reduction, agile manufacturing and production process. The "TPM bundle" consisted of practices associated to optimizing machine and maintenance through preventive maintenance program. The new process equipment and technology also belong in this TPM bundle. The "TQM bundle" composed of practices related to quality management program, total quality management,

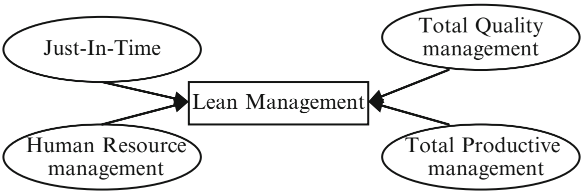


Fig. 90.1 The bundles of lean manufacturing (Source: Demeter and Matyusz 2011)

Table 90.1 Lean bundles and lean practices lean implementation

Bundles	Lean manufacturing practices
Just-in-Time (JIT) Shah and Ward	Lot size reduction
	JIT/Continuous flow production
	Pull system
	Cellular manufacturing
	Cycle time reductions
	Focused factory production systems
	Agile manufacturing strategies
	Quick changeover techniques
	Bottleneck/constraint removal
	Reengineering production process
TPM	Productive or preventive maintenance
	Maintenance optimization
	Safety improvement programs
	Planning and scheduling strategies
	New process equipment or technologies
TQM	Competitive benchmarking
	Quality management programs
	Total quality management
	Process capability measurement
HRM	Self-directed work team
	Flexible, cross-functional workforce

Shah and Ward (2003), Shah et al. (2008), Demeter and Matyusz (2011)

continuous improvement program, process capability measurement, and competitive benchmarking. Finally, practices related to self-directed work teams and cross-functional workforce was combined to form “HRM bundle” (Shah and Ward 2003). Figure 90.1 displays the bundles of lean manufacturing (Demeter and Matyusz 2011). Table 90.1 displays lean practices studied in previous literature. Shah et al. (2008) summarized lean 15 practices from previous research; these practices were related with quality management, pull production, preventive maintenance, and human resource management (Shah et al. 2008). In manufacturing, previous research examined the empirical development of lean construct, relationship between lean and organizational performance, and organizational factors that affect implementation of lean practices (Shah et al. 2008).

However, lean practices have not been identified outside the context of manufacturing. Research on lean practices in service organizations is still limited.

90.2.3 Lean Management in Health Care

In 2005, lean thinking was diffused to health care industry by James Womack and his team (Womack et al. 2005). They mentioned

> Lean thinking is not a manufacturing tactic or a cost-reduction program, but a management strategy that is applicable to all organizations because it has to do with improving processes. All organizations – including health care organizations – are composed of a series of process, or sets of actions intended to create value for those who use or depend on them (customer/patients). (Womack et al. 2005)

The lean management concept was successfully implemented at health care (Womack et al. 2005). Lean concept is incorporated within the quality improvement initiative for system transformation in public health (Riley et al. 2010). To date, lean management was widely implemented in several settings/departments in hospitals including emergency medicine, anesthesiology/intensive care, surgery, obstetrics-gynecology, family medicine, radiology, pathology, clinical chemistry, nursing, and pharmacy (King et al. 2006; Bush 2007; Mazzocato et al. 2010). However, the majority of research reported lean adaptation in one department. Not many studies reported implementing lean in crossed boundary of departments.

In 2010, DelliFraine et al. (2010) provided literature review of six sigma and lean projects in health care organizations to improve clinical outcome, processes of care and financial performance. They found that 34 out of 177 articles on six sigma and lean published in the past 10 years reported outcomes of the SS/S studied.

Lean interventions in health care organizations can be classified into four methods, which are

(a) Methods to understand processes in order to identify and analyze problems; these methods involve value stream mapping, process mapping, 5 whys, 5S,
(b) Methods to organize more effective and/or efficient process; these methods involve process orientation, specification of standard procedures with focus on waste elimination, one-piece continuous flow, kanban, 5S, process streaming, pull,
(c) Methods to improve error detection, relay information to problem solvers, and preventing errors from causing harm; these methods include visual management, 5S, patient safety alert system and stop the line, enhance adherence to standard procedures, and
(d) Methods to manage change and solve problem with a scientific approach; these method involves A3 reporting system, team approach to problem solving and rapid process improvement events (Mazzocato et al. 2010).

Recently, LaGanga (2011) conducted lean process improvement project to increase capacity. The study used lean to redesign service systems at outpatient service at a mental health center resulting in improved performance.

Lean principles, basic lean tools, and applications of these tools for health care setting is summarized in Zidel (2006). Details on lean thinking in hospital setting are provided by Mazzocato et al. (2010), Graban (2009).

Several cases of successful lean implementation in hospitals have been reported since 1990s. Lean projects conducted in hospitals improved patients care by reducing waiting time, medical errors and cost whereas improving employee satisfaction and interaction between departments (Graban 2009).

90.2.4 Performance Measures of Successful Lean Implementation

In general, success indicators are directly tied to the mission and objectives of the program which are often set by the primary stakeholders. Literature reveals that success performance measure/indicators of lean implementation usually include operational performance (Dal Pont et al. 2008; Shah and Ward 2003). The operational performance of a manufacturing plant are manufacturing cycle time, scrap and rework costs, labor productivity, unit manufacturing cost, first pass yield, and customer lead time (Shah and Ward 2003). Dal Pont et al. (2008) considered quality, dependability, speed, flexibility and cost. In summary, operational performance was assessed in term of (a) speed (Dal Pont et al. 2008; Shah and Ward 2003), (b) cost (Dal Pont et al. 2008; Shah and Ward 2003), (c) productivity (Shah and Ward 2003), and (d) quality (Dal Pont et al. 2008; Shah and Ward 2003).

90.3 Proposed Lean Practices in Health Care

90.3.1 Lean Practices in Health Care

In this paper, we identify the lean practices in health care service organizations. Lean practices in health care emphasized on patient flow, value-stream mapping and kaizen events (Bernstein 2008). Min et al. (2012) examined the extent of lean implementation in US hospitals in four dimensions involving patient focus, standardized work, seamless and coordinated work flow, and continuous improvement culture. It is important to note that these four dimensions are similar to four bundles stated in Table 90.1. Patient flow and workflow can be considered with JIT in manufacturing. Continuous improvement culture would represent TQM and HRM dimension in manufacturing.

Table 90.2 Proposed lean practices in health care

Lean practices in health care
Lot size reduction
JIT/Continuous flow/ Seamless and coordinated work flow
Pull system/Kanban
Cycle time reduction
Quick changeover techniques
Removing bottleneck
Preventive maintenance
Safety improvement program
Planning and Scheduling strategies
New process equipment/technologies
Competitive benchmarking
Quality management program
Continuous improvement programs, Kaizen event
Total quality management
Self-directed work teams
Flexible, cross-functional workforce
Value stream mapping
Patient focus
Patient flow
Standardized work
Continuous improvement culture

Based upon the classic paper, Shah and Ward (2003) and other studies (Min et al. 2012; Rahman et al. 2010), this paper identifies a number 21 operations in health care. Table 90.2 presented our summarized lean practices in health care.

90.3.2 Impact of Lean Practices on Performance of Hospitals

Health care organizations could gain benefits from lean implementation including cost saving, time saving and timeliness of service, productivity improvement, and quality enhancement (Mazzocato et al. 2010; Graban 2009).

The success of a lean implementation was measured by several aspects. Kollberg et al. (2006) developed the measurement framework called "the flow model" to measure lean initiatives or changes towards lean thinking in hospitals. The framework provided key performance indicators of each of five principles, i.e., specify value, value stream, flow, pull, and perfection. For instances, the first principle, specify value, were measured with related to accessibility, quality, and comfort. The second principle, value stream, was measured with respect to accessibility, capacity, preparation time, medical device down time, referral management, booking procedures. The third principle, flow, was assessed in term of JIT, scheduling, multi-skill teams. Pull principle were measures based on transparency of information accessibility.

Lastly, the perfection principle was measured on continuous improvement and process control (Kollberg et al. 2006).

Min et al. (2012) examined the impact of lean implementation to the hospital performance in term of quality and efficiency. The first measure was quality measures, which related to process quality measure, outcome quality measures, and perceived quality measure. Outcome quality measures were represented by risk-adjusted in hospital mortality and risk-adjusted readmission rate. Perceived quality measure was represented by patient satisfaction. The second measure is efficiency measure comprised of ratio of observed cost to adjusted expected cost.

DelliFraine et al. (2010) mentioned an impact of lean implementation on clinical outcome, processes of care and financial performance of health care service organizations. Graban (2009) reported the successful lean projects improve patients care by reducing errors, decreased cost, reduced waiting time, improved interdepartmental interaction and increase employee satisfaction. Womack et al. (2005) reported positive impact on productivity, cost, quality, and timely delivered service after having implementing lean management in the hospitals in the United States.

From a review of literature, we identify operational performance of health care organizations impacted from lean implementation. Operational performance can be summarized into five categories including speed, cost, overall productivity, quality, and customer satisfaction (Graban 2009). Brief summary of the impact from lean implementation is presented below.

“speed” – hospitals benefited from reducing waiting time (Womack et al. 2005; Graban 2009; Kollberg et al. 2006), improved delivery of outpatients service (LaGanga 2011).

“cost” – most hospitals reported reduced cost after implementing lean concept (Graban 2009). The daCare Inc. reported $10 million/year in cost saving (Womack et al. 2005).

“productivity” – literature reported improved patient access through flow improvement (Park Nicollet 2005), enhanced productivity (Mazzocato et al. 2010) after lean implementation.

“quality” – hospitals benefits from quality improvement through error reductions (Mazzocato et al. 2010; Graban 2009), mortality reduction (Mazzocato et al. 2010), defect reduction, clinical and service quality (Womack et al. 2005), better clinical outcome, improved processes of care (DelliFraine et al. 2010).

“customer satisfaction” – Not only lean helps to increase staff satisfaction, but also increased patient satisfaction (Mazzocato et al. 2010; Samson and Terziovski 1999)

Table 90.3 summarized operational performance of hospitals.

Table 90.3 Operational performance in health care

Operational performance in health care
Speed (Quick deliver compared with competitors)
Cost
Overall productivity
Quality
Overall customer satisfactions

90.4 Conclusion and Future Research

The challenges in high cost, low quality, and efficiency in health care systems called for novel management concepts. Recently, lean management is one of organizational innovation adopted by many health care organizations. This paper reviews lean thinking concepts, lean practices, and lean performance indicators from previous studies. In addition, lean management in health care is described. From existing literature, lean practices have not been identified outside the context of manufacturing.

The major contribution of this paper is to thoroughly review and summarize 21 lean practices in health care systems context. This paper also investigates the impact of lean implementation to health care operational performance. Five dimensions of health care operational performance affected by lean implementation are proposed; these are speed, cost, quality, productivity, and customer satisfactions.

The next phase of this research project is to empirically explore the relationship between lean practices and operational performance in health care service organization. By doing so, the results will report the extent to which lean practices conducting in health care service organizations and the relationship between lean practices and their impacts on operational performance.

Acknowledgement This research is supported by Thailand Research Fund (TRF) and Mahidol University.

References

Bernstein R (2008) Overbooking: the wrong solution for medical clinics. Available: http://leaninsider.productivitypress.com/2008/01/overbooking-wrong-solution-for-medical.html, 6 June 2008

Bush RW (2007) Reducing waste in US health care systems. J Am Med Assoc 297:871–874

Dal Pont G et al (2008) Interrelationships among lean bundles and their effects on operational performance. Oper Manag Res 1:150–158

DelliFraine JL et al (2010) Assessing the evidence of six sigma and lean in the health care industry. Qual Manag Health Care 19:211–225

Demeter K, Matyusz Z (2011) The impact of lean practices on inventory turnover. Int J Prod Econ 133:154–163

Graban M (2009) Lean hospitals. CRC Press/Taylor & Francis Group, Boca Raton

King DL et al (2006) Redesigning emergency department patient flows: application of lean thinking to health care. Emerg Med Australas 18:391–397

Kollberg B et al (2006) Measuring lean initiatives in health care services: issues and findings. Int J Prod Perform Manag 56:7–24

Krafcik JF (1998) Triumph of the lean production system. Sloan Manage Rev 30:41–52

LaGanga LR (2011) Lean service operations: reflections and new directions for capacity expansion in outpatient clinics. J Oper Manag 29:422–433

Mazzocato P et al (2010) Lean thinking in healthcare: a realist review of the literature. Qual Saf Health Care 19:376–382

Min YT et al (2012) The impact of lean implementation on quality and efficiency in US hospitals. Presented at the POMS 23rd annual conference, Chicago, USA

Ohno T (1988) Toyota production system – beyond large scale production. Productivity Press, Cambridge, MA

Park Nicollet (2005) Lean Production at Park Nicollet. Available: http://www.parknicollet.com/media/leanProduction.cfm

Rahman S et al (2010) Impact of lean strategy on operational performance: a study of Thai manufacturing companies. J Manuf Technol Manag 21:839–852

Riley WJ et al (2010) Realizing transformational change through quality improvement in public health. J Public Health Manag Pract 16:72–78

Samson D, Terziovski M (1999) The relationship between total quality management practices and operational performance. J Oper Manag 17:393–409

Shah R, Ward PT (2003) Lean manufacturing: context, practice bundles, and performance. J Oper Manag 21:129–149

Shah R et al (2008) In pursuit of implementation patterns: the context of lean and six sigma. Int J Prod Res 46:6679–6699

Womack JP, Jones DT (1996) Lean thinking: banish waste and create wealth in your corporation. Simon and Schuster, New York

Womack JP, Jones DT (2003) Lean thinking. Simon & Schuster, London

Womack JP et al (2005) Going lean in health care. Institute for Healthcare Improvement, Cambridge, MA

Zidel TG (2006) A lean toolbox—using lean principles and techniques in healthcare. J Healthc Qual 28:W1-7–W1-15

Chapter 91
A Proposed Heuristic for Solving the Container Pre-marshalling Problem

M.S. Gheith, A.B. El-Tawil, and N.A. Harraz

Abstract Containerized shipping is exhibiting increasing trends, which makes container terminal planning decisions a critical success factor to a container terminal. The efficient handling of containers has a significant and direct impact on the performance of the terminal exhibited in berthing time of vessels. In this paper we present a simple labeling and ordering heuristic to solve the container pre-marshalling problem. Preliminary results illustrate the good performance of the proposed heuristic.

Keywords Container pre-marshalling problem • Container retrieval problem • Container stacking problem • Container yard terminals • Heuristics

91.1 Introduction

A container terminal comprises three areas; the quay side, the land side, and the container yard. One of the most important performance indicators in a container terminal is the berthing time of vessels (Stahlbock and VoB 2008). It is desired to minimize the berthing time of a vessel which does not only depend on the performance of the Quay Cranes (QC), but also on the performance of Yard Cranes (YC). The work rate of the QC is usually double that of the YC (Ng and Mak 2005), and thus, YCs are a potential bottlenecks. This difference in rate exists mainly because the containers are not stacked in the yard in a configuration that facilitates the work of the QC, so that, the containers must be re-ordered in order to improve the retrieval rate of these containers when it is time to load them on the vessels, this re-ordering is called container pre-marshalling.

M.S. Gheith (✉) • A.B. El-Tawil • N.A. Harraz
Department of Industrial Engineering and Systems Management, Egypt-Japan University of Science and Technology (E-JUST), Alexandria, Egypt
e-mail: mohamed.gheith@ejust.edu.eg; eltawil@ejust.edu.eg; nermine.harraz@ejust.edu.eg

E. Qi et al. (eds.), *The 19th International Conference on Industrial Engineering and Engineering Management*, DOI 10.1007/978-3-642-37270-4_91,

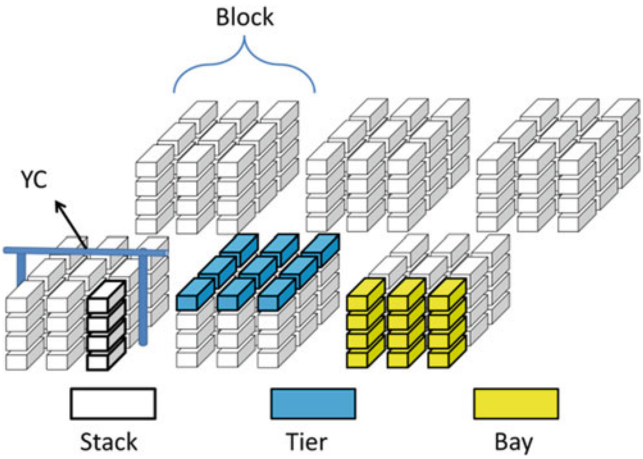


Fig. 91.1 Layout of storage area of a container terminal

As shown in Fig. 91.1, a container yard consists of blocks, each block consists of a set of bays, each bay consists of a set of stacks, and each stack consists of a set of tiers. When containers arrive to the yard, they are piled up above each other waiting for further operations. In order to get a container that is not on the top of the bay, additional movements have to be made to change the location of the containers that are above the desired container in order to get the desired one. These additional movements are called "reshuffles" which consume too much time leading to reduced work efficiency of YC and QC especially during vessel loading operations.

Generally, the container stacking problem is classified into three main types: the pre-marshalling problem, the remarshalling problem and the container retrieval problem (Steenken et al. 2004). The pre-marshaling problem, is the problem of converting an initial layout of a bay into a desired final layout within which containers are stacked above each other with the priority of stacking the containers that will be served first at the top of the stack. This will minimize or eliminate future additional reshuffles by the YC. In the case of the container retrieval problem, it is desired to remove a container from the bay with minimum number of reshuffles, and then remove another container and so on till the bay is empty. The three classes of the problem are of prime importance knowing that in large container terminals, the average number of movements made by yard cranes is 15,000 movements per day (Steenken et al. 2004), which means that the reduction of such moves will dramatically improve operations and efficiency. Containers arrive at the yard with the means of trucks and are staked in such a way that may not be compatible with the stowage plan of the vessel. Usually information about arriving of the trucks is not provided.

Following this introduction, the second section of the paper includes a review of the recent related literature. In the third section, a proposed methodology for solving the container pre-marshalling problem is presented. Finally, the conclusions and directions for future research are discussed in the fourth section.

91.2 Related Work

One major problem that we faced when reviewing the container pre-marshalling problem is that, the literature discussing this problem is very limited (Huang and Lin 2012). Also, it is difficult to classify the literature according to the methodologies used, because most of papers used heuristics and approximate algorithms as the problem is NP-hard (Caserta et al. 2012). In this paper, we track the previous work of the different authors in chronological order to elaborate the progress and improvement of their work.

Stahlbock and VoB (2008), provided an extensive review about container terminals operations, which they updated later (Steenken et al. 2004). In the section of the pre-marshalling problem, the review did not change. They described the environment of the pre-marshalling problem and the constraints that restrict it. Bortfeld (2004), described a tree search heuristic that is used to solve the pre-marshalling problem. Extending the work done by Bortfeld (2004), Bortfeldt and Foster (2012), provided a tree search heuristic for solving container pre-marshalling problem. The search was carried out by means of recursive procedure, and the procedures were coded using C programming language. The authors provide a comparison between the results obtained by their own heuristic and the results obtained by the algorithms of Lee and Hsu (2007) and Lee and Chao (2009) and in both cases, the proposed tree search heuristic was better.

Lee and Hsu (2007), proposed an integer programming model to solve pre-marshalling problem for small and medium size instances. In large size problems, the authors developed a simple heuristic that is capable of providing a solution while minimizing computation time. But, the authors did not give any comparison between their method and any other methods. Lee and Lee (2010), proposed a three phase heuristic based on heuristic rules to solve the container stacking problem. Computational experiments show that, the proposed heuristic is able to solve instances with more than 700 containers.

Based on the assumptions made by Kim et al. (2000) and Caserta and VoB (2009) developed three methods for solving the container relocation problem. The first one is the development of a binary integer programming model which is capable of solving small-size problem instances. Due to the large computation time, the model was simplified by some realistic assumptions to be able to solve medium-size problem instances. For large-size problem instances the authors developed heuristic rules based on the computation of a stack score, which helps to determine where a relocating container should be placed, the heuristic was coded in C++. Also, Caserta et al. (2011), proposed another heuristic approach for solving the problem of relocating containers, the proposed heuristic was based on the Corridor Method (CM). The basic idea of the CM is to use an exact method (e.g. Dynamic Programming (DP)) over a restricted portion of the solution space of a given problem in order to minimize the solution space of the problem. The authors used a modified DP algorithm as an exact method with CM. They solve their problem based on the assumptions of Kim and Hong (2006). The algorithm was coded in

C++ and computational experiments were made. Compared with the results of Kim and Hong, their proposed algorithm showed a decrease in the number of movements required in small to medium size instances, while in large instances, the proposed algorithm shows a decrease in the number of movements required, but, in a relatively larger time. In case of very large instances, the algorithm showed a tremendous decrease in the required number of movements, but, in very large amount of time compared with Kim and Hong.

Extending the work done by Molins et al. (2010, 2012), presented a new domain-dependent planning heuristic to solve the container stacking problem. The heuristic was coded using the Planning Domain Definition Language (PDDL) language. The advantage of this language is that, it is capable of representing the physical characteristics of the objects under study. The results of their heuristic were compared with the results of their old one (Molins et al. 2012), and the results show a significant decrease in computation time. Also, Molins et al. tried to integrate the container stacking problem with the Berth Allocation Problem (BAP) and the Quay Crane Assignment Problem (QCAP) (Salido et al. 2012). First they solve the container stacking problem with a method they developed earlier (Molins et al. 2010), after that they solve the integrated BAP + QCAP problem, then they proposed a planner that integrated the solution of the two problems, and the terminal operators are to decide which solution is the most appropriate in relation to a multi-objective function.

Exposito-Izoquierdo et al. (2012), present the Lowest Priority First Heuristic (LPFH) used to solve the pre-marshalling problem based on the assumption of Lee and Hsu (2007). Also they introduced a generator to generate instances for the pre-marshalling problem. The heuristic was coded using Java programming language, and the results were compared with the results obtained by Caserta and VoB (2009), and it was found to be relatively perfect. Also, the results were compared with the optimum solution gained by A^* search algorithm, and the results demonstrate the good performance of the proposed heuristic.

As a conclusion from the literature, there is no specific method that could be used in solving the problem under study. The mathematical models are capable of solving only small instances, while in large problem instances heuristics are appropriate. Also, there are limited comparisons between the results obtained from different solution methods. Based on these remarks, we developed a heuristic to solve the problem under study, and tried to avoid these remarks in the future work.

91.3 The Proposed Heuristic to Solve the Pre-marshalling Problem

The proposed methodology is based on developing a five stages heuristic in order to minimize the required movements with short computation time. The movement is counted if a container's location changed to another location. In case of small

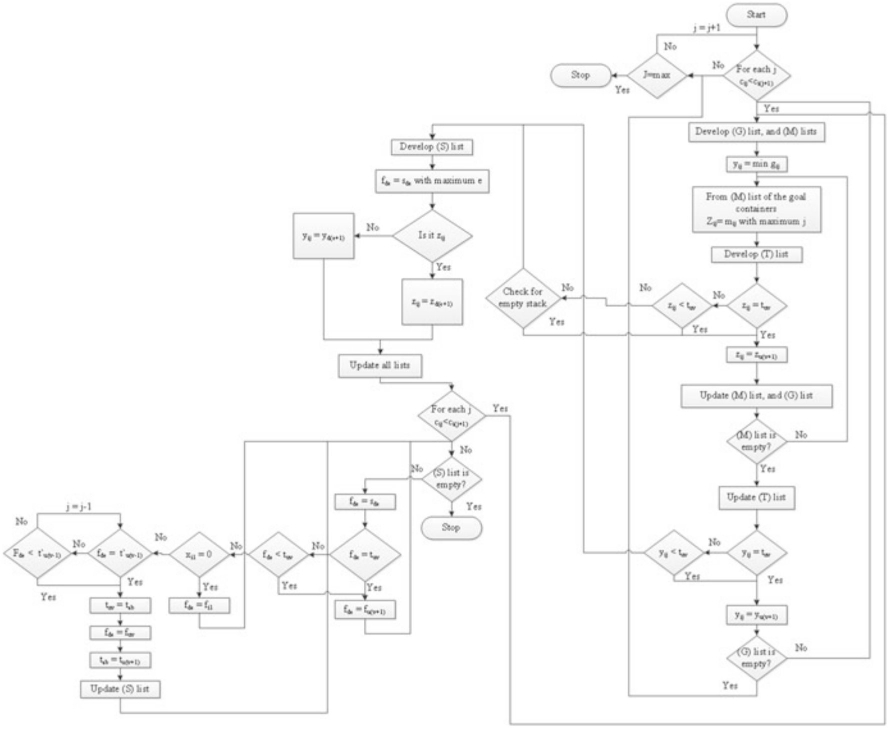


Fig. 91.2 Flow chart of the proposed heuristic

instances, integer programming model will be used and then, the results of the proposed heuristic will be compared with the results of the model. For validation, in case of large instances, the results obtained from the proposed heuristic will be compared with benchmark problems found in the literature, and the performance indicators will be calculated. The flow chart of the proposed heuristic is shown in Fig. 91.2. As described before, it is required to rearrange the initial layout of the containers so as to get a final desired layout that is confirmed with the stowage plan of the vessel.

For safety considerations, the yard crane usually serves one bay (Stahlbock and VoB 2008). Thus, our heuristic will focuses on solving the pre-marshalling problem in only one bay. The following mathematical notations will be used in the proposed heuristic.

$G =$ set of goal containers, $G = \{g_{11}, g_{12}, \ldots, g_{ij}\}$. $i =$ General index for stacks. $j =$ General index for tiers. $M_{gij} =$ set of moving containers of a goal container, $Mgij = \{m_{11}, m_{12}, \ldots, m_{ij}\}$. $T =$ set of top containers, $T = \{t_{12}, t_{24}, \ldots, t_{uv}\}$. $C =$ set of containers $\{c_{11}, c_{12}, \ldots, c_{ij}\}$. $y_{ij} =$ selected goal container from (G) set, $y_{ij} = \min \{g_{11}, g_{12}, \ldots, g_{ij}\}$. $z_{ij} =$ selected moving container, $z_{ij} = m_{ij}$ with maximum j. $t_{uv} =$ top container of stack u is at tier v. $u =$ special index for stacks for top containers. v = special index for tiers for top containers. S = set of containers in the extra stack, $S = \{s_{d1}, s_{d2}, \ldots, s_{de}\}$. f_{de} = selected container from (S) set. d = special index for stacks for the containers in the extra stack, $d = i + 1$. e = special index for tiers of the containers in the extra stack. $x_{i1} = 0$, indicates that a stack i has no containers. Note that, g, m, t, c, z, y, and s are numbers representing the priority of the container. The steps of the proposed heuristic are described below:

- Step 1: Check the priorities of the containers placed in the first tier and select a container of highest priority. If there is more than a container with the same high priority, then start with the container that the summation of the priorities of the containers above it is higher than of the other container. The selected container will be labeled as the goal container and the containers above it will be labeled as moving containers.
- Step 2: Move the moving containers to other stacks with top containers having a priority equal to or lower than the priority of the moving container. In this concern two cases may exist; the first one, if the height of the stacks is maximum, while the second case, if the moving container has a priority lower that the priorities of the top containers in all stacks. In both cases a new stack will be constructed. Only one extra stack could be constructed and the containers in the extra stack will be stacked in order of arrival.
- Step 3: Move the goal container to another stack that has a top container with a priority equal to or lower than the priority of the goal container.
- Step 4: Check the priorities of the containers of the first tier, if all the containers have lower priorities than of the containers above them, then, check the second tier and repeat steps 2, 3 and 4 till all tiers are considered.
- Step 5: In order to relocate the containers in the new constructed stack, the containers in the new stack are moved to other stacks. This will be done by checking the priorities of the containers that will be moved. First, the moving container will be moved to a stack has a top container has a priority equal to or lower than that of the moving one. If not, then check if there is any empty stack, If there is an empty stack, then construct a new stack with this moving container. If not, then check if there is any empty stack, If there is an empty stack, then construct a new stack with this moving container. If not, then check the priorities of the containers placed below the top containers of the stacks. If one of these containers has priorities equal to or lower than the priority of the moving container, then move this container to any other stack and place the moving container and then replace the container that moved earlier.

91.4 Numerical Example

Figures 91.3, 91.4, 91.5, 91.6, and 91.7 illustrate the application of the steps of the proposed heuristic with a numerical example. In this numerical example it is required to convert an initial layout of the containers (e.g. Fig. 91.3) into desired

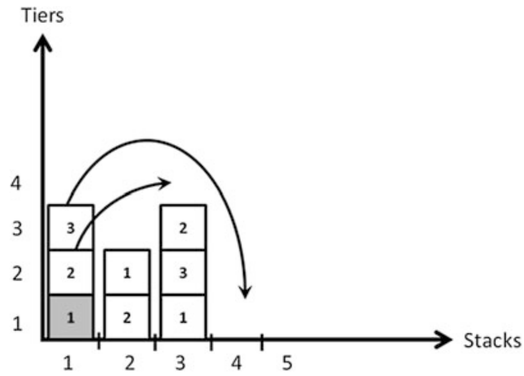


Fig. 91.3 Initial layout, and selecting the goal container

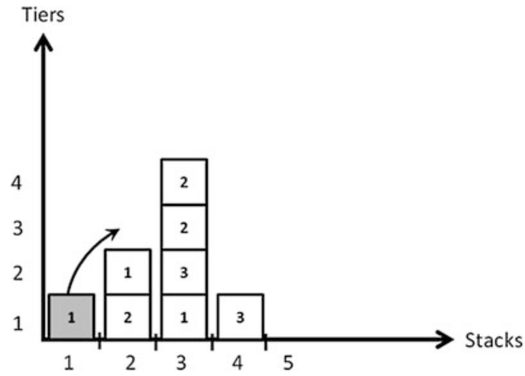


Fig. 91.4 Layout after moving the moving containers

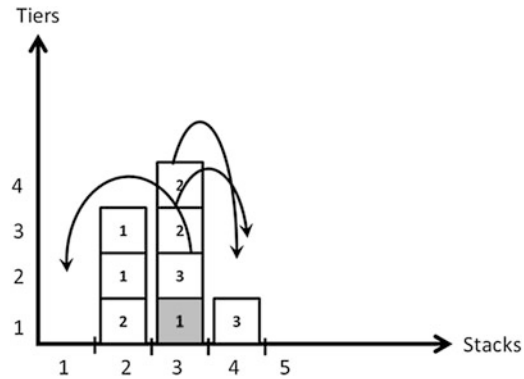


Fig. 91.5 Layout after moving the goal container and selecting a new goal container

final layout (e.g. Fig. 91.7), the maximum height of each stack is only four tiers, and the containers are allowed to be stored in only the first four stacks. The fifth stack is the extra stack that could be used temporarily during the pre-marshalling operation.

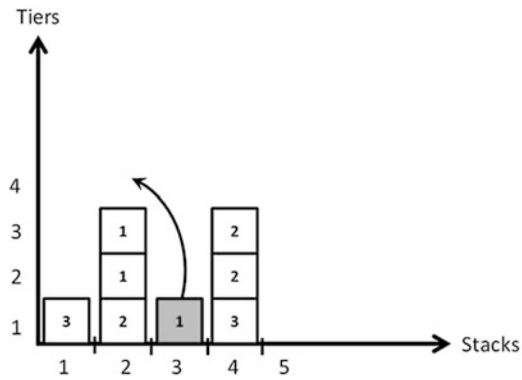


Fig. 91.6 Layout after moving the moving containers

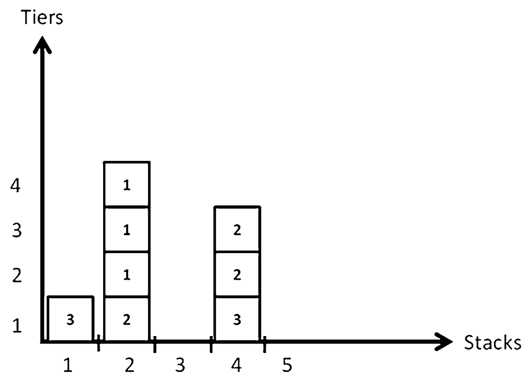


Fig. 91.7 Final layout

91.5 Analysis of the Performance of the Proposed Heuristic

At this stage, the only factor that could be computed is the number of movements made during the conversion from the initial layout into the desired final layout. The proposed heuristic is compared with the work done by Lee and Hsu (2007), and by Hung and Lin (2012). Lee and Hasu provide an optimization model to solve small problem instances and a heuristic to solve large problem instances. Two cases were selected to make the comparison, in the first case, the total number of containers are 14 containers grouped into three types, and the maximum height of each stack is four tiers, while in the second case the total number of containers are 45 grouped into six types, and the maximum height of each stack is five tiers. Hung and Lin propose a heuristic algorithm for solving the pre-marshalling problem and they compared their results with the results obtained by Lee and Hsu. Table 91.1 illustrates the comparison between the results of our proposed heuristic, the results got by Lee and Hsu, and the results by Hung and Lin.

As a conclusion from the comparison we can conclude that, comparing our heuristic to the optimum solution in case of small instance our solution quality

Table 91.1 Performance of the proposed heuristic

	Case 1	Case 2
Number of containers	14	45
The maximum height of a stack	4 tiers	5 tiers
Types of containers	3 types	6 types
Optimum solution	9 movements	–
Proposed heuristic	11 movements	42 movements
Lee and Hsu Heuristic	10 movements	47 movements
Hung and Lin	9 movements	35 movements

around 80 % of the optimum solution, and 90 % of Lee and Hsu solution using their heuristic. In case of large instance, our heuristic shows about 11 % improvement in the number of movements required to change the configuration of the containers.

91.6 Conclusion

This paper provided a review for the container pre-marshalling problem. The literature studying this type of problem is very limited. The paper also described a simple labeling and ordering heuristic procedure that is used to solve this problem. The solution will be validated and verified by comparing the results with benchmark problems. Concerning the future work, it is planned to implement the proposed methodology in a real world case study, where the heuristic is implemented in a dynamic environment to cope with the nature of work in container terminals. A possible extension is the integration of such a system with a Radio Frequency Identification RFID enabled system, that will help capture the current locations of target containers in real time. Also, comparing the results of the proposed heuristic with benchmarks will be considered. The comparison will be based on both number of movements and computation time required.

Acknowledgment This research project is sponsored by the Egyptian ministry of higher education grant and the Japanese International Cooperation Agency (JICA) in the scope of the Egypt Japan University of Science and Technology.

References

Bortfeldt A (2004) A heuristic for the container pre-marshalling problems. In: Proceedings of the 3rd international conference on computer and IT applications in maritime industries, 2004, Sigüenza, Spain, pp 419–429

Bortfeldt A, Foster F (2012) A tree search procedure for the container pre-marshalling problem. Eur J Oper Res 217:531–540

Caserta M, VoB S (2009) A corridor method-based algorithm for the pre-marshalling problem. Lect Notes Comput Sci 5484:788–797
Caserta M, VoB S, Sniedovich M (2011) Applying the corridor method to a blocks relocation problem. OR Spectr 33:915–929
Caserta M, Schwarze S, VoB S (2012) A mathematical formulation and complexity considerations for the blocks relocations problem. Eur J Oper Res 219:96–104
Exposito-Izoquierdo C, Melian-Batista N, Morena-Vaga M (2012) Pre-marshalling problem: heuristic solution method and instance generator. Expert Syst Appl 39:8337–8349
Huang S, Lin T (2012) Heuristic algorithm for container pre-marshalling problems. Comput Ind Eng 62:13–20
Kim KH, Hong GP (2006) A heuristic rule for relocating blocks. Comput Oper Res 33:940–954
Kim KH, Park YM, Ryu KR (2000) Deriving decision rules to locate export containers in container yards. Eur J Oper Res 124:89–101
Lee Y, Chao S (2009) A neighborhood search heuristic for pre-marshalling export containers. Eur J Oper Res 196:468–574
Lee Y, Hsu N (2007) An optimization model for the container pre-marshalling problem. Comput Oper Res 34:3295–3313
Lee Y, Lee Y (2010) A heuristic for retrieving containers from a yard. Comput Oper Res 37: 1139–1147
Molins M, Salido M, Barber F (2010) Domain-dependent planning heuristics for locating containers in maritime terminals. IEA/AIE 6096(1):742–751
Molins M, Salido M, Barber F (2012) Intelligent planning for allocating containers in maritime terminals. Expert Syst Appl 39:978–989
Ng W, Mak K (2005) Yard crane scheduling in port container terminals. Appl Math Model 29: 263–275
Salido M, Molins M, Barber F (2012) A decision support system for managing combinatorial problems in container terminals. Knowl Based Syst 29:63–74
Stahlbock R, VoB S (2008) Operations research at container terminals: a literature update. OR Spectr 30:1–52
Steenken D, VoB S, Stahlbock R (2004) Container terminal operations and operations research – a classification and literature review. OR Spect 26:3–49

Chapter 92
Research on Optimization of Explosive Products Supply Support in Certain Base

Jun Peng, Jian-jun Zhao, Xiao-yan Zhang, and Yang Sun

Abstract In order to reduce the total costs, a minimum cost model (EOQ) was proposed based on the arrangement of explosive products supply support. Analyzing the whole flow including purchase, storage and transportation, explosive products were transported from military plants to depots and from depots to combat troops. Concerning all the constraints, the minimum cost model was transformed into mathematical programming model, and then optimal solutions can be solved. The model provides a theoretical basis to take measures to raise the efficiency of equipment support and improves the order capability in certain base.

Keywords EOQ model • Model optimization • Supply support

92.1 Introduction

Supply support is the main component of integrated logistic support. Its major task is to control the requirement of spares and consumables (GJB 376-87 1987). Combined with the actual equipment operational condition, the user revise equipment spares promptly, adjust inventory and supply network, improve supply method by initial supply list and management demand (Ting-xue Xu et al. 2004).

J. Peng • Y. Sun
New Equipment Training Center, Naval Aeronautical and Astronautical University, Yanta, China

J. Zhao
Department of Ordnance Science Technology, Naval Aeronautical and Astronautical University, Yantai, China

X. Zhang (✉)
Department of Aerocraft Engineering, Naval Aeronautical and Astronautical University, Yantai, China
e-mail: summeryan@163.com

E. Qi et al. (eds.), *The 19th International Conference on Industrial Engineering and Engineering Management*, DOI 10.1007/978-3-642-37270-4_92,

There are many explosive products models in certain base. So its annual consumption is also very large, besides, the link from order to storage and distribution is much more. In order to minimize the total cost of supply support (Tai-liang Song 2008), it is very meaningful to research the optimization of support mode on explosive products.

92.2 Problem of Supply Support on Explosive Products

There are five explosive products depots (numbered B1 to B5) on certain base, which supply eight combat troops (numbered C1 to C8) with explosive products support. The support concept is mainly ten kinds of explosive products, and they are all ordered by three military plants (numbered A1 to A3).

The explosive products flow is as follows: (1) placing an order to the military plants for explosive products based on operational training requirement of the eight combat troops, (2) transporting the various ordered explosive products to the corresponding depots, and (3) delivering the explosive products from depots to the combat troops (Feng-chun Zhu et al. 2007). The combat troops cannot store the explosive products and just use them by reason of the combustibility and explosively.

The base charges a million yuan for each order time, and the available liquid fund top is ¥100 million. At the same time, it includes seven known conditions: (1) annual output of ten kinds of explosive products by three military plants, (2) ex-factory price of ten kinds of explosive products, (3) unit transportation cost from the three military plants to the five explosive products depots, (4) capacity of five depots, (5) unit storage charge and volume of ten kinds of explosive products, (6) unit freight from the five explosive products depots to the eight combat troops, and (7) requirement of ten kinds of explosive products by eight combat troops.

So the problem will focus on the optimal solution according to the arrangement of explosive products support, such as order times of various explosive products, order quantity, transportation program and the selection of purchasing party.

92.3 Mathematical Model of Supply Support

92.3.1 Establishment of Objective Function

1. *Daily purchase cost of the explosive products:* On the assumption that certain case order and lade the explosive products in a cycle T. There are three military plants and ten kinds of explosive products. The variable $\alpha_{\iota\varphi}$ is the total quantity of the number φ explosive products from the number ι military plant, and the

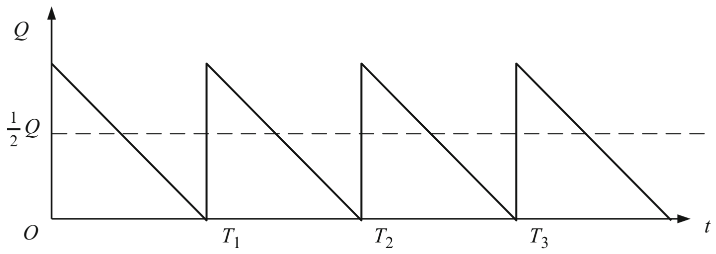


Fig. 92.1 Storage quantity curve of *EOΘ* model

variable $\beta_{\iota\varphi}$ is the corresponding unit price. In a cycle *T*, the daily purchase cost of explosive products is:

$$\sum_{\tau=1}^{3}\sum_{\varphi=1}^{10}\alpha_{\tau\varphi}\beta_{\tau\varphi}/T \tag{92.1}$$

2. *Daily freight from military plants to the depots:* Because all sorts of explosive products need to transport to the explosive products depots, relevant freights are inevitable. In certain base there are three military plants and five depots, so the freights differ from different military plant to different depot. The variable $\delta_{\iota\kappa}$ is the total quantity of all kinds of explosive products from the number ι plant to the number κ depot in a cycle *T*, and the variable $\phi_{\iota\kappa}$ is the unit freight also from the number ι plant to the number κ depot. then the daily freight from military plants to the explosive products depots is:

$$\sum_{\tau=1}^{3}\sum_{\kappa=1}^{5}\delta_{\tau\kappa}\phi_{\tau\kappa}/T \tag{92.2}$$

3. *Storage cost in the depots*: The various explosive products are stored in various depots, so the daily storage costs vary subsequently. The variable $\varepsilon_{\varphi\kappa}$ is the total quantity of number φ explosive product which stored in number κ depot in a cycle *T*, and the variable η_{φ} is the unit storage cost of number φ explosive product in a cycle *T*. Considering the requirement of explosive product is even, and the storage quantities in the depots reduce evenly too. The *EOΘ* model can be used to calculate the average daily storage cost of explosive products (Qi-yuan Jiang et al. 2003; Sheng Ju 1989).

   Suppose that the maximal storage quantity is Θ, the minimal storage quantity is zero, the change of requirement is continuous and even. Then the time history plot of storage quantity is according to "Fig. 92.1". In Fig. 92.1 the lateral axis represents time, and the longitudinal axis represents storage quantity (Croft 2002).

It is known from "Fig. 92.1" that the average storage quantity is $\Theta/2$, and the storage cost is $C_p\Theta/2$ (Chun-he Liu et al. 2001).

Based on the analysis above, the daily storage cost of explosive products in a cycle T is:

$$1/2\sum_{\varphi=1}^{10}\sum_{\kappa=1}^{5}\eta_{\varphi}\varepsilon_{\varphi\kappa}/365 \tag{92.3}$$

4. *Daily freight from explosive products depots to combat troops*: Explosive products are supplied for eight combat troops, and the freights differ from different depots to different troops. On the assumption that in a cycle T, the variable $\mu_{\kappa\lambda}$ represents the total quantity of explosive products transported from the number κ depot to the number λ troop, the variable $\nu_{\kappa\lambda}$ represents the corresponding freight of unit explosive product from the number κ depot to the number λ troop. Then the average daily freight from different depot to different troop is:

$$\sum_{\kappa=1}^{5}\sum_{\lambda=1}^{8}\mu_{\kappa\lambda}\nu_{\kappa\lambda}/T \tag{92.4}$$

5. *Fixed cost for ordering*: In the above-mentioned problems the base charges a million yuan for each order time. The average daily order cost in a cycle T is:

$$10,000/T \tag{92.5}$$

Synthesizing the relevant costs, the total costs Ω of each day averagely in a cycle T can be calculated. In case that the variable Ω fetch the minimal value, i.e. the total costs of each day averagely are:

$$\begin{aligned}\Omega &= \sum_{\tau=1}^{3}\sum_{\varphi=1}^{10}\alpha_{\tau\varphi}\beta_{\tau\varphi}/T + \sum_{\tau=1}^{3}\sum_{\kappa=1}^{5}\delta_{\tau\kappa}\phi_{\tau\kappa}/T + 10,000/T \\ \Omega &= +1/2\sum_{\varphi=1}^{10}\sum_{\kappa=1}^{5}\eta_{\varphi}\varepsilon_{\varphi\kappa}/365 + \sum_{\kappa=1}^{5}\sum_{\lambda=1}^{8}\mu_{\kappa\lambda}\nu_{\kappa\lambda}/T\end{aligned} \tag{92.6}$$

92.3.2 Constraints of Model

In the order problems above it requires minimizing the total costs, so the solutions to the annual order times, the order concept and the transport concept are inevitable

(Velocci 2002). Considering the objective function, the optimum value of the variable Ω can be sought out under all the constraints of the base (Yu-quan Wen et al. 2007). During solutions of the objective function all the variables can be solved, such as order cycle T, annual order times, order concept $\alpha_{\iota\varphi}$ of the number φ explosive products ordered from the number ι military plant, initial transport concept $\delta_{\iota\kappa}$ of all kinds of explosive products from the number ι plant to the number κ depot, and the total quantity $o_{\varphi\lambda}$ of the number φ explosive products transported to the number λ combat troop. The restraints consist of six classes.

1. *Constraint between production output and order quantity*: Suppose that the variable $\chi_{\iota\varphi}$ is the annual output of the number φ explosive product produced by the number ι military plant, and the production is uniform distributed at the same time (Joseph 2004). In certain case, annual order quantity cannot exceed the production outputs of three military plants in a cycle T.

$$\alpha_{\tau\varphi} \leq T\chi_{\tau\varphi}/365(\tau = 1, 2, 3; \varphi = 1, \cdots, 10) \tag{92.7}$$

2. *Constraint between order quantity and freight volume*: After the base placing an order for explosive products, the ordered goods will be all transported to the depots. The total quantity of the number φ explosive products transported to the five depots amount to the order quantity of three military plants. Its equation is:

$$\sum_{\tau=1}^{3} \alpha_{\tau\varphi} = \sum_{\kappa=1}^{5} \varepsilon_{\varphi\kappa} \tag{92.8}$$

The total quantity of the ten sorts of explosive products transported to the five depots is equal to the order quantity of the base. The equation is:

$$\sum_{\tau=1}^{3}\sum_{\varphi=1}^{10} \alpha_{\tau\varphi} = \sum_{\tau=1}^{3}\sum_{\kappa=1}^{5} \varepsilon_{\varphi\kappa} \tag{92.9}$$

3. *Constraint between depot capacity and explosive product volume*: On the assumption that the variable γ_{κ} is the capacity of the number κ depot, and the variable π_{φ} is the unit volume of the number φ explosive product. It's certain that the gross volume of all the explosive products which stored in any depot cannot exceed the capacity of the depot (Peng Wang and Zhi-ming Du 2005). The inequation is:

$$\sum_{\varphi=1}^{10} \pi_{\varphi}\varepsilon_{\varphi\kappa} \leq \gamma_{\kappa}(\kappa = 1, \ldots, 5) \tag{92.10}$$

4. *Constraint between memory capacity and freight volume*: Suppose that the variable $o_{\varphi\lambda}$ is the quantity of the number φ explosive product transported to the number λ military troop, and the total quantity of the number φ explosive product transported to various troops are just the gross memory capacity of the number φ explosive product from all depots. The equation is:

$$\sum_{\lambda=1}^{8} o_{\varphi\lambda} = \sum_{\kappa=1}^{5} \varepsilon_{\varphi\kappa} (\varphi = 1, \ldots, 10) \tag{92.11}$$

5. *Constraint between freight volume and requirement*: Suppose that the variable $\theta_{\varphi\lambda}$ is the requirement of the number φ explosive product in the number λ troop. In certain case the condition of out of stock is unallowed, and at the same time the explosive products cannot be stored in any troop. So the freight volume transported to the number λ troop is needed to satisfy the requirement of the number φ explosive product. The equation is:

$$o_{\varphi\lambda} = \theta_{\varphi\lambda} T/365 (\varphi = 1, \ldots, 10; \lambda = 1, \ldots, 8) \tag{92.12}$$

The total freight volume transported to the combat troops is equal to the general requirement for explosive products. The equation is:

$$\sum_{\kappa=1}^{5} \sum_{\lambda=1}^{8} \mu_{\kappa\lambda} = \sum_{\varphi=1}^{10} \sum_{\lambda=1}^{8} o_{\varphi\lambda} \tag{92.13}$$

6. *Constraint of order fund*: During the order process of the certain base, the upper limit of order fund is ¥100 million. That is:

$$\sum_{\tau=1}^{3} \sum_{\varphi=1}^{10} \alpha_{\tau\varphi} \beta_{\tau\varphi} \leq 100,000,000 \tag{92.14}$$

Sum up all the constraints above-mentioned the mathematical programming model will be transformed into another form. The transformed model is:

$$\min\ \Omega = \sum_{\tau=1}^{3} \sum_{\varphi=1}^{10} \alpha_{\tau\varphi} \beta_{\tau\varphi} / T + \sum_{\tau=1}^{3} \sum_{\kappa=1}^{5} \delta_{\tau\kappa} \phi_{\tau\kappa} / T + 10,000/T$$

$$\min\ \Omega = + 1/2 \sum_{\varphi=1}^{10} \sum_{\kappa=1}^{5} \eta_{\varphi} \varepsilon_{\varphi\kappa} / 365 + \sum_{\kappa=1}^{5} \sum_{\lambda=1}^{8} \mu_{\kappa\lambda} \nu_{\kappa\lambda} / T$$

$$s.t.\begin{cases} \alpha_{\tau\varphi} \leq T\chi_{\tau\varphi}/365\ (\tau = 1,2,3; \varphi = 1,\ldots,10) \\ \sum_{\tau=1}^{3} \alpha_{\tau\varphi} = \sum_{\kappa=1}^{5} \varepsilon_{\varphi\kappa} \\ \sum_{\tau=1}^{3}\sum_{\varphi=1}^{10} \alpha_{\tau\varphi} = \sum_{\tau=1}^{3}\sum_{\kappa=1}^{5} \varepsilon_{\varphi\kappa} \\ \sum_{\varphi=1}^{10} \pi_{\varphi}\varepsilon_{\varphi\kappa} \leq \gamma_{\kappa}(\kappa = 1,\ldots,5) \\ \sum_{\lambda=1}^{8} o_{\varphi\lambda} = \sum_{\kappa=1}^{5} \varepsilon_{\varphi\kappa}(\varphi = 1,\ldots,10) \\ o_{\varphi\lambda} = \theta_{\varphi\lambda}T/365\ (\varphi = 1,\ldots,10; \lambda = 1,\ldots 8) \\ \sum_{\kappa=1}^{5}\sum_{\lambda=1}^{8} \mu_{\kappa\lambda} = \sum_{\varphi=1}^{10}\sum_{\lambda=1}^{8} o_{\varphi\lambda} \\ \sum_{\tau=1}^{3}\sum_{\varphi=1}^{10} \alpha_{\tau\varphi}\beta_{\tau\varphi} \leq 100,000,000 \end{cases} \tag{92.15}$$

92.3.3 Abbreviations and Acronyms

Equation (92.15) can solve such problems as order cycle, concrete programming of which explosive product ordered from which military plant. But in the aspect of transport concept from plants to depots, the model only can solve two problems, one is the total quantity $\delta_{\iota\kappa}$ transported from the number ι plant to the number κ depot, the other is the total quantity $\varepsilon_{\varphi\kappa}$ of the number φ explosive product transported from the number κ depot. Except the two problems, other concrete problems such as which plant is transported from, which explosive product is transported and which depot is transported to, "(92.15)" can't give a clear answer. Considering the conditions the variables, $\varepsilon_{\varphi\kappa}$ and $\alpha_{\iota\varphi}$ are used to solve the problems farther.

On the assumption that the variable $\rho_{\iota\varphi\kappa}$ is the total quantity of the number φ explosive product transported from the number ι plant to the number κ depot. The equation is:

$$\sum_{\varphi=1}^{10} \rho_{\tau\varphi\kappa} = \delta_{\tau\kappa}, \sum_{\tau=1}^{3} \rho_{\tau\varphi\kappa} = \varepsilon_{\varphi\kappa}, \sum_{\kappa=1}^{5} \rho_{\tau\varphi\kappa} = \alpha_{\tau\varphi} \tag{92.16}$$

In the aspect of transport concept from depots to troops, also only can solve two problems, one is the total quantity $\mu_{\kappa\lambda}$ of the number φ explosive product transported from the number κ depot, the other is the total quantity $o_{\varphi\lambda}$ of the number φ explosive product transported to the number λ military troop. Except the two problems, other concrete problems such as which depot is transported from, which explosive product is transported and which troop is transported to, "(92.15)" can't give a clear answer. Considering the conditions the variables $\mu_{\kappa\lambda}$, $o_{\varphi\kappa}$ and $\varepsilon_{\varphi\kappa}$ are used to solve the problems farther.

On the assumption that the variable $\sigma_{\kappa\varphi\lambda}$ is the total quantity of the number φ explosive product transported from the number κ depot to the number λ troop. The equation is:

$$\sum_{\varphi=1}^{10} \sigma_{\kappa\varphi\lambda} = \mu_{\kappa\lambda}, \sum_{\kappa=1}^{5} \sigma_{\kappa\varphi\lambda} = o_{\varphi\lambda}, \sum_{\lambda=1}^{8} \sigma_{\kappa\varphi\lambda} = \varepsilon_{\varphi\kappa} \tag{92.17}$$

92.4 Solutions to Model

Concerning all the constraints of the model, the objective function can be optimized, and after solving the model by program LINGO, the values such as order quantity, freight volume are not integers with the great majority. But as to the actual condition, it's just the reverse, the values are all integers. So it's required to make a post data processing. Rounding-off method is used to process the relevant values (Clement et al. 2004). After processing, the values of the variables below are all integers. The variables include: (1) order quantity $\alpha_{\iota\varphi}$ (the total quantity of the number φ explosive products ordered from the number ι military plant), (2) freight volume $\delta_{\iota\kappa}$(the total quantity of all kinds of explosive products from the number ι plant to the number κ depot), (3) freight volume $o_{\varphi\lambda}$ (the total quantity of the number φ explosive products transported to the number λ combat troop), (4) $\rho_{\iota\varphi\kappa}$ (the total quantity of the number φ explosive product transported from the number ι plant to the number κ depot), (5) $\sigma_{\kappa\varphi\lambda}$ (the total quantity of the number φ explosive product transported from the number κ depot to the number λ troop).

Thus the model has been built and the solutions to the model including post data processing have been solved.

92.5 Conclusion

To minimize the total cost of the base, the programming model from ordering to storing is studied. And during the course of study a setting objective function Ω is used to change annual cost minimization into daily average cost minimization. With the establishment of the objective function Ω (Schmierer et al. 2008), the problems of ordering, transporting and storing are transformed to the problems of optimum solutions to the objective function. It is practically significant to study the optimization of supply support mathematical model, which raises the efficiency of equipment support and improves the capability of ordering in certain base.

In view of the actual state analysis the values of order must be integers. But in the course of data processing it's unrealizable, not all the values can be converted, so rounding-off method is taken up to solve the problem. In the meantime the rounding-off method can bring on a series of problems such as depots capacity overflow with regard to the issue of explosive products storage, freight volume overrun memory capacity concerning transport concept. All in all, in the aspect of each exceeding limitation it's required to taking proper measures to adjust decisions to fit in with the needs. It will be meaningful to improve the model through optimizing the process. And it also can yield returns on explosive products order, giving full scope to the process optimization analysis (Jun-ming Yuan and Yu-cun Liu 2005).

Acknowledgment The author is grateful to Peng Shaoxiong for whom he works as an equipment support consultant. The assistance provided by Liu Yong in preparing the manuscript of this paper is acknowledged.

References

Chun-he Liu, Zu-jian Lu, Yu-hua Yuan (2001) The storage reliability assessment for missile. Mathematics in Practice and Theory 31:416–420 (in Chinese)

Clement D, Diener J, Kovalev D (2004) In: Proceedings of 35th international annual conference of ICT, Karlsruhe, Germany, 29 June–2 July, 2004, pp 5/1–1/11

Croft J (2002) Baggage deadline to refocus standards. Aviat Week Space Technol 156(1):51–52

Feng-chun Zhu, Xiao-ming Wang, Da-wei Cui (2007) An evaluation method of reliable stockpile lifetime of air ammunitions. J Ballist 19(1):68–71 (in Chinese)

GJB 376-87 (1987) Assessment method of reliability of initiating devices. National Standard Press, Beijing, pp 45–55 (in Chinese)

Joseph VR (2004) Efficient robbins-monro procedure for binary data. Biometrika 91:461–470

Jun-ming Yuan, Yu-cun Liu (2005) Study on new never optimal sensitivity test. Initiat Pyrotech 2:24–27 (in Chinese)

Peng Wang, Zhi-ming Du (2005) Smmarize of margin research and design method of pyrotechnic devices. Initiat Pyrotech 2:34–38 (in Chinese)

Qi-yuan Jiang, Jing-xing Xie, Jun Ye (2003) Mathematical model. Higher Education Press, Beijing, pp 156–157 (in Chinese)

Schmierer B, Tournier AL, Bates PA, Hill CS (2008) Mathematical modeling identifies smad nucleocytoplasmic shuttling as a dynamic signal-interpreting system. Proc Natl Acad Sci USA 105(18):6608–6613

Sheng Ju (1989) Probability theory and mathematical statistics. Higher Education Press, Beijing, pp 39–41

Tai-liang Song (2008) Equipment supportability engineering. Haichao Publishing House, Beijing, pp 201–212 (in Chinese)

Ting-xue Xu, Hui-guang Deng, Qi-ming Li (2004) Theory and method on integrated logistic support of missile equipment. Haichao Publishing House, Beijing

Velocci A (2002) L-3 positions itself to dominate explosives detection business. Aviat Week Space Technol 156(3):38–39

Yu-quan Wen, Dong-pao Hong, Wei Wang (2007) Study on theory and method of reliability assessment of explosive initiator based on testing entropy. Explo Shock Waves 27:553–556 (in Chinese)

Chapter 93
An Exploration on the Organizational Structure Design of the Hong Kong-Zhuhai-Macao Bridge Authority

Jia-qing Yu, Jie Huo, Qiang Shi, Gang Sheng, and Hao Zhang

Abstract The in-construction Hong Kong-Zhuhai-Macao Bridge (HZMB) will be the longest cross-sea bridge and tunnel combined highway in the world. Firstly, the related projects and the project features of HZMB are investigated and summarized in this paper. Then the construction scheme, management system and business of HZMB are described in detail. Finally, based on the analysis mentioned above and the four principles of organization structure design, the total organizational structure of HZMB is presented.

Keywords HZMB • Management system • Organization structure design

93.1 Introduction

The in-construction Hong Kong-Zhuhai-Macao Bridge (HZMB) is a series of bridges and tunnels that will connect Hong Kong, Macau and Zhuhai, three major cities situated on the Pearl River Delta in southern China (Hussain et al. 2011). The proposed 50 km link, with an estimated investment of 72.94 billion yuan and a designed service life of 120 years, consists three parts: Offshore Bridge-cum-Tunnel, Boundary Crossing Facilities (BCF) at Hong Kong, Zhuhai and Macao, and Link Roads in these three regions (Yu Lie 2011). Construction formally began on 15 December 2009, and expected to be put into service in 2016 (Zhang and Zhu 2012). When completed, it will become the longest six-lane immersed tube tunnel, and the longest cross-sea bridge and tunnel combined highway in the world.

J. Yu (✉) • J. Huo • Q. Shi • G. Sheng • H. Zhang
Research Institute of Highway Ministry of Transport, Beijing 100088, China

RIOH Transport Consultants Ltd, Beijing 100191, China
e-mail: jq.yu@rioh.cn; jiaqing.yu@ttu.edu

E. Qi et al. (eds.), *The 19th International Conference on Industrial Engineering and Engineering Management*, DOI 10.1007/978-3-642-37270-4_93,

The purpose of an organization is to enable ordinary people to do extraordinary things (Levitt 1991). While a study by Harvard Business School and McKinsey & Company concluded that organizational structure is one of the key points of project management. Mintzberg has written extensively and significantly on the importance of organizational structure (Mintzberg 1989). Further, organizational structure and design are closely entwined (Mabey et al. 2001) with many aspects of human resource management.

In this paper, the organizational structure design of HZMB is studied to achieve the construction and management goals. Because of the unique feature that project engages three regional governments administered under the policy of "One Country, Two Systems", this paper could give a novel perspective for the organization design.

The rest of this paper is organized as follows: In Sect. 93.1, the related projects and the project features of HZMB are investigated and summarized. Requirement analysis is taken into account in Sect. 93.2. Section 93.3 proposes the organizational structure of HZMB. Section 93.4 summarizes the paper.

93.2 Related Projects and Project Features of Hzmb

93.2.1 Related Projects

1. Øresund Bridge

The 16.4 km long Øresund Bridge consists of a bridge, an artificial island and an underwater tunnel, which connect the Danish capital Copenhagen with Malmö in Sweden across the Øresund strait. The bridge includes the longest immersed concrete tunnel and double-deck road and rail bridge in the world and was designed to have minimal impact on the sensitive marine environment (Peeters et al. 2009).

The Danish-Swedish company, Øresundsbro Konsortiet, is the operator of the Øresund Bridge, whose organization is shown in the Fig. 93.1:

Property operates the infrastructure - highway, railway, tunnel, toll station and so on - and manages the company's environmental and safety efforts.

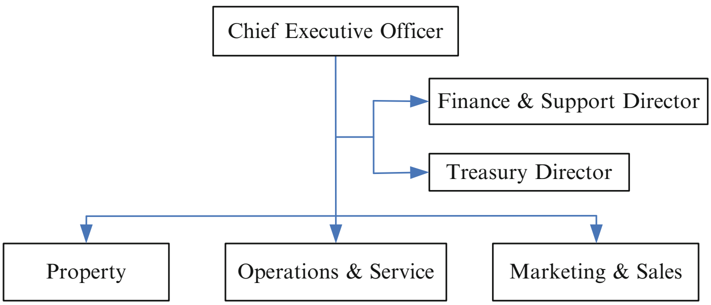


Fig. 93.1 Organizational structure of the Øresund bridge

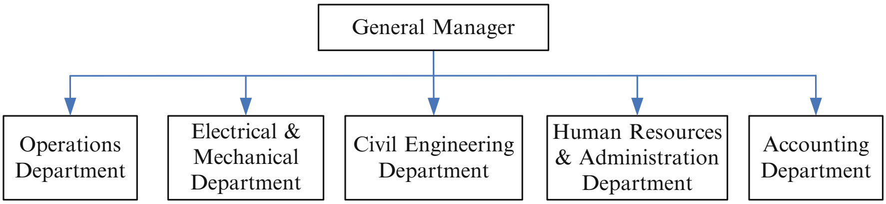


Fig. 93.2 Organizational structure of the EHT

Operations & Service is responsible for the customer experience of the trip across the bridge and the unit is manning the toll station, the traffic center and the customer center.

Marketing & Sales markets Øresundsbron and work with sales and service communications, branding and public relations, including contacts with the media.

Finance & Support includes support functions such as internal service, accounting, HR and IT.

Treasury administers the loans for the Øresund Bridge and several other companies (Molina-Morales and Mas-Verdu 2010).

2. Eastern Harbour Tunnel (EHT)

The EHT is a combined road and MTR rail link under Victoria Harbour between Quarry Bay on Hong Kong Island and Cha Kwo Ling on Kowloon. Constructed using the immersed-tube-technique, this 2.2 km long tunnel was opened to traffic on 21 September 1989. It is one of the largest tunnels of its type in the world and has five separate conduits running side by side within the tunnel: two for the road link, two for the rail and a fifth to house environmental control systems plus electrical and mechanical equipment (Lo 2004).

New Hong Kong Tunnel Company Limited is the operator of the EHT, whose organization is shown in the Fig. 93.2:

Operations Department is responsible for the traffic management, law enforcement, emergency response, incident handling, security, toll collection, etc.

Electrical & Mechanical Department is responsible for the maintenance of E&M systems, including mechanical, electrical, electronics, computer, vehicle fleet, etc.

Civil Engineering Department is responsible for the maintenance of civil structures and building, including main bridges, viaducts, slopes, road surfaces, building structure, general cleansing, building services, etc.

Human Resources & Administration Department includes the functions of human resources, general administration, IT.

Accounting Department is responsible for the accounting, purchasing & store.

3. Hong Kong-Shenzhen Western Corridor (HK-SWC)

HK-SWC (also known as the Shenzhen Bay Bridge) is a 5.5-km long dual 3-lane carriageway of which the Hong Kong section is 3.5-km. Spanning across Deep Bay, HK-SWC links up the northern end of Deep Bay Link at Ngau Hom Shek of Hong Kong with Shekou of Shenzhen (Lee et al. 2004).

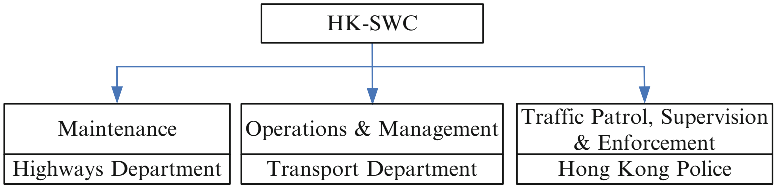


Fig. 93.3 Organizational structure of HK-SWC

This is a joint project between the governments of Hong Kong Special Administrative Region (HKSAR) and Shenzhen, each side responsible for the design and construction of the portion of HK-SWC within its own territory. Upon its opening, the Shenzhen Section of the bridge, along with the Shenzhen Bay Control Point, were leased to Hong Kong and made under Hong Kong's jurisdiction for an initial period until 30 June 2047. As such, Hong Kong laws, including traffic laws, instead of the laws of the PRC, apply on the whole bridge and within the control point (Report of the Bills Committee on Shenzhen Bay Port Hong Kong Port Area Bill). The organizational structure of HK-SWC can be shown below:

The Highways Department of the HKSAR Government is the maintenance authority of the bridge for both the Hong Kong and Shenzhen portions of HK-SWC (Fig. 93.3).

The Transport Department of the HKSAR Government is responsible for the routine traffic operation and management, giving support to the traffic control and surveillance system (TCSS).

And it's Hong Kong police's duties to make traffic patrol, supervision and enforcement. If necessary, Hong Kong police will help supervise the execution of TCSS traffic planning.

93.2.2 Project Features of HZMB

The HZMB is one of the most technically difficult projects in the history of the transportation industry of China, because of the high design standards and construction requirements adopted. The main structure of the bridge is on deep and soft ground with many constraints, including significant variation of stratigraphic distribution, high seismic design code, a stringent limit on the water resistance ratio, and a severe and complicated hydrodynamic environment. Moreover, HZMB crosses the conservation area of the Chinese White Dolphin as well as the busiest marine navigation channels of the Pearl River. The level of design and construction difficulty of the HZMB can be ranked the highest in the world in terms of the project scale, expertise and disciplines involved (Yu Lie 2011).

1. Engineering Complexity

The construction conditions can be affected by a number of factors with high uncertainty, such as typhoons, navigation channels, hydrology and hydrodynamic issues, geotechnical issues, aviation limit, floods, etc. since the project is composed of multiple types of structures such as the sea-crossing bridges, submerged tunnel and artificial islands, it requires compatible design criteria that meet common and basic requirements, but also allow for no important differences for the criteria of the three regions.

2. Difficulties in Coordination

Because the project engages three regional governments administered under the policy of "One Country, Two Systems", the coordination of efforts within respective departments (navigation, maritime, water resource agency, environment agency, BCF, etc.) can be enormous and difficult.

3. Environmental Sensitivity

Because the bridge crosses over two environmentally sensitive areas, the conservation area of Chinese White Dolphin (Indo-pacific Humpback Dolphin) and the conservation area of Parr/Juvenile Prawn at the mouth of the Pearl River, the standard of environmental protection is high. Close attention is required to the protection of the marine ecosystem and fishery resources.

93.3 Requirement Analysis

93.3.1 Construction Scheme

The alignment of the offshore bridge and tunnel starts from San Shek Wan of Latau Island in Hong Kong, connects to the BCF in Hong Kong, passes through Hong Kong waters, extends to the west along the north side of 23-DY anchorage, then goes across several navigational channels (Tonggu Channel in the pearl river, Lingding West Channel, Qingzhou Channel and Jiuzhou Channel) to reach the end at the artificial island of Zhuhai and Macao BCF. The total length is about 35.6 km, of which 6 km of the bridge is in Hong Kong territory and 29.6 km of the bridge is in the territory of Guangdong.

A scheme for a bridge and tunnel combination was adopted for the main part of the project; the tunnel section is approximately 6.7 km long across the Lingding West Channel and Tonggu Channel, and the bridge section is about 22.9 km. For the transitions of the bridge to tunnel sections and to accommodate the tunnel ventilation shafts, artificial islands are built at each end of the tunnel.

For the Management and Maintenance facilities of HZMB, there's a administration center (also the control, communication, toll, and joint rescue command centre) and a maintenance center near the Hongwan Interchange at the end of the link road to Zhuhai, a control room (also the rescue command centre on spot) on the west artificial island, a mainline toll station on the Zhuhai-Macao artificial island, four communication stations at the administration center, Zhuhai-Macao artificial island,

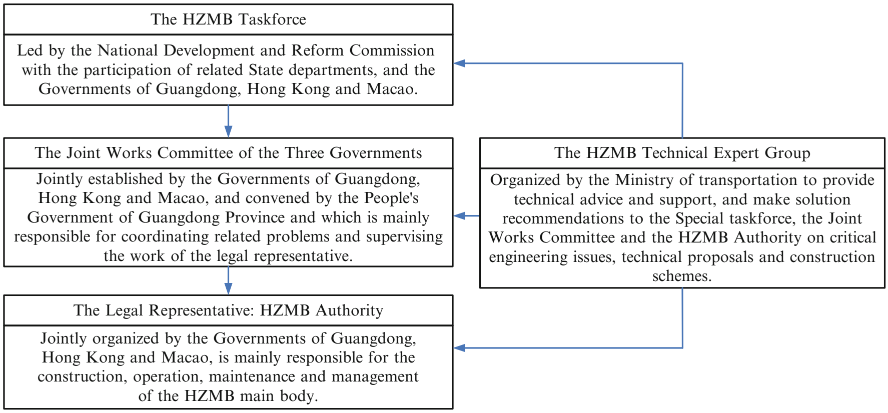


Fig. 93.4 Administrative organization chart of the HZMB

west artificial island and east artificial island separately, three rescue stations on the Zhuhai-Macao artificial island, west artificial island and east artificial island separately, two maintenance stations on the Zhuhai-Macao artificial island and the west artificial island separately.

93.3.2 Management System Analysis

The administrative organization of the HZMB is structured in three levels (Website of Hong Kong-Zhuhai-Macao Bridge): the HZMB Taskforce, the Joint Works Committee of the Three Governments (JWCoTG), and the legal representative of the project, which can be shown in Fig. 93.4.

The HZMB Authority will take the integrative mode in construction, operation and management, with the layer of hierarchical business management.

1. Routine Business

Two-level management system is taken for the routine business: administration center- control room (toll station or communication station).

2. Emergency Rescue Business

Three-level management system is taken for the emergency rescue business: joint rescue command centre- rescue command centre on spot-rescue station.

3. Maintenance Business

Two-level management system is taken for the maintenance business: maintenance center-maintenance station.

93.3.3 Business Analysis

The main business of the HZMB Authority is understood to fall into three categories: operation management, maintenance management and administrative management.

1. Operation Management

   The main contents of operation management include: toll, traffic management, highway law enforcement, emergency rescue, safety and environment protection, auxiliary business, etc.

2. Maintenance Management

   The main contents of maintenance management include: the main structure and infrastructure, Electro-mechanical system and equipments, roadside facilities, house-building engineering and facilities, landscape engineering, etc.

3. Administrative Management

   The main contents of administration management include: human resource, project planning, finance, logistics, party-masses, etc.

93.4 Organizational Structure Design

93.4.1 Principles of Organization Structure Design

1. Unity of command

   Every person in an organization should be responsible to one superior and receive orders from that person only. Fayol (1949) considered this to be the most important principle for efficient working and increased productivity in an organization.

2. The Scalar Principle

   Decision making authority and the chain of command in an organization should flow in a straight line from the highest level to the lowest. The principle evolves from the principle of unity of command. However, this may not always be possible, particularly in large organizations or in research institutions. Therefore, Fayol (1949) felt that members in such organizations could also communicate directly at the same level of hierarchy, with prior intimation to their superiors.

3. Division and Coordination

   According to the classical approach, work can be performed much better if it is divided into components and people are encouraged to specialize by components.

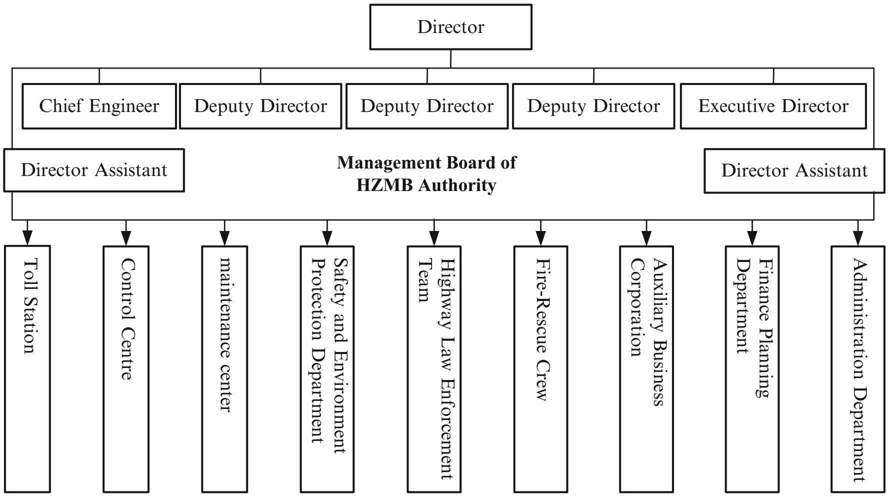


Fig. 93.5 Organization structure of the HZMB

Coordination refers to integrating the objectives and activities of specialized departments to realize broad strategic objectives of the organization. Hierarchy facilitates vertical coordination of various departments and their activities.

4. Localization and Preconscious

Based on the contents and requirements of management, local conditions should be considered in the organization structure design. Moreover, with the rapid development of contemporary science and technology, the design of organization structure must be preconscious and extensibility so as to adapt to the future requirements.

93.4.2 Setup of the Organization Structure

The HZMB Authority will take the line-functional organization structure. In the structure, the Director, Deputy Director, Chief Engineer, Executive Director, and Director Assistant compose the management board, followed by nine departments. The total organization structure of the HZMB is shown in Fig. 93.5.

The administrative organization of the HZMB is introduced in Sect. 93.2.2. This section mainly describes the nine departments of the HZMB Authority.

1. Toll Station

Toll Station is responsible for the business of toll collection and assisting the Finance Planning Department in the establishment and executive oversight of the annual toll collection plan.

The major responsibilities of the Toll Station include: integrated toll management, management of toll pass ticket (card), toll audit, toll information management, toll scene management.

2. Control Centre

Control Centre is in charge of the supervision of the traffic conditions, monitoring of the running status of the infrastructure and equipments, and maintenance of the Electro-mechanical system and equipments.

The major responsibilities of the Control Centre include: traffic supervision, operations monitoring, public service, dispatch and command, external coordination, information sharing, maintenance and outsourcing service of the Electro-mechanical system and equipments.

3. Maintenance Center

Maintenance Center is responsible for the maintenance of main structures, traffic safety facilities and building, including the establishment and implement of the maintenance schedule, detection and its data analysis of the main infrastructure, road surfaces, traffic safety facilities and building, use and maintenance of the structural health monitoring system, management of maintenance facilities and passageways, emergency maintenance, outsourcing service, technical documents management, etc.

4. Safety and Environment Protection Department

The major responsibilities of the Safety and Environment Protection Department include: safety management in production and environment, health, safety & environment (HSE) management, maintenance management of landscape engineering.

5. Highway Law Enforcement Team

In accordance with the "Administrative Management Regulations on Highways" issued by the Ministry of Transport, the major responsibilities of the Highway Law Enforcement Team include: highway property rights, out-of-gauge transportation, control areas on both sides of roads, patrol and obstacles removing, maintenance of traffic order within jurisdiction, etc.

6. Fire-Rescue Crew

The major responsibilities of the Fire-Rescue Crew include: conducting the fire fighting, routine exercises and regular fire drills, checking and maintaining of all fire services installations and equipments, training of security procedures and fire fighting skills, coordination with the Highway and traffic police, providing support to the adjacent road section, etc.

7. Auxiliary Business Corporation

The Auxiliary Business Corporation mainly operates the service facilities and advertising business of HZMB.

8. Finance Planning Department

Finance Planning Department is responsible for the financial planning and management, include: establishment and executive oversight of the annual finance plan, register of toll, purchase of toll pass ticket (card), routine accounting, cooperation with the finance audit, etc.

9. Administration Department

Administration Department is responsible for the management of the human resource, logistics, party-masses, asset, etc.

93.5 Conclusion

This paper tries to introduce the HZMB from the project features, construction scheme, management system and business. Based on the requirements analysis and the four principles of organization structure design, this paper proposes the total organizational structure of HZMB. Though the proposed organization structure satisfies the current operation and management requirements, it should be finally demonstrated and improved in the actual operation after open of the HZMB.

Because of the unique feature that project engages three regional governments administered under the policy of “One Country, Two Systems”, this paper could give a novel perspective for the organization design.

Acknowledgment This work is supported by No. 4 Special Subject Study of the design and construction drawings for traffic engineering of the HZMB: Preparation of the Operation and Maintenance Manual (Traffic Engineering).

References

Fayol H (1949) General and industrial management (trans: Storrs C). Pitman, London

Hussain N, Wong C, Carter M, Kwan S, Mak T-W (2011) Hong Kong Zhuhai macao link. Proc Eng 14:1485–1492

Lee J, Kong A, Chan E (2004) Hong Kong-Shenzhen Western Corridor and deep bay link: design development and environmental challenges. In: Proceedings of IABSE symposium Shanghai, metropolitan habitats and infrastructure, Shanghai, 2004, pp 25–30 (6)

Levitt T (1991) Thinking about management. The Free Press, New York

Lo C-w (2004) A case study of build-operate-transfer road project in Hong Kong: route 3 Country park section. Master dissertation, University of Hong Kong, Pokfulam Road, Hong Kong

Mabey C, Salaman G, Storey J (2001) Organizational structuring and restructuring. In: Salaman G (ed) Understanding business organisations. Routledge, London

Mintzberg H (1989) The structuring of organizations. In: Asch D, Bowman C (eds) Readings in strategic management. Macmillan Education, London

Molina-Morales FX, Mas-Verdu F (2010) The role of supporting organizations in regional innovation policies: lessons learnt from the cases of Oresund and Veneto. Int J Innov Reg Dev 2(3):221–239

Peeters B et al (2009) Continuous monitoring of the Øresund bridge: system and data analysis. Struct Infrastruct Eng Maint Manage Life-Cycle Des Perform 5(5):395–405

Report of the Bills Committee on Shenzhen Bay Port Hong Kong Port Area Bill (2007). Available at: http://www.legco.gov.hk/yr0607/english/hc/papers/hc0413cb2-1534-e.pdf

Website of Hong Kong-Zhuhai-Macao Bridge (HZMB) Organizational structure (in Chinese) (2010). Available at: http://www.hzmb.org/cn/bencandy.asp?id=5

Yu Lie (2011) Crossing the lingdingyang. Hong Kong-Zhuhai-Macao Bridge 1:10–17

Zhang J-w, Zhu Y-l (2012) Project management planning for construction of Hong Kong-Zhuhai-Macao Bridge (HZMB) main bridge project. Highway 3:143–147 (in Chinese)

Chapter 94
Co-innovation and Code of Value Appropriation

Ping Xu and Yong-cheng Xu

Abstract There would be a fair code to decide the value that had been created by assets such as technology. The code would encourage the owner of assets to innovate and to maximize the firm value. The game theory is introduced into the research of creating and distributing firm value and the game model based on the rule of distribution proportion between capital owner and the technology owner was set up. The total maximum value of the firm is counteracted by action of players' maximizing self-value. A new code was given that the innovator captures the whole value caused by innovation to maximize the total maximum value of the firm.

Keywords Co-innovation • Game • Technology • Value

94.1 Introduction

The competitive environment faced by the firm has been greatly changed with the rapid development of emerging technologies and the shortening of the product life cycle (Bettis and Hitt 1995). In this situation, technological innovation is becoming increasingly important to the survival and development of a firm (Fujita 1997). The firm can develop and introduce new products and reduce production costs through technical innovation to improve product competitiveness in the market (Guth 1990). Technical innovation is the way of getting the core competitiveness and the main factor of measuring business performance (Stopford 1994). But along with the development of the emerging technologies becoming more and more

P. Xu (✉)
Department of Management, Huaiyin Institute of Technology, Huaian, China
e-mail: xu1ping1@xinhuanet.com

Y. Xu
The School of Architecture, South China University of Technology, Guangzhou, China

E. Qi et al. (eds.), *The 19th International Conference on Industrial Engineering and Engineering Management*, DOI 10.1007/978-3-642-37270-4_94,

complex, the trend of cross and fusion of different subjects and technology is more and more obvious. Technology innovation becomes more and more difficult within an organization if the enterprise merely relies on its internal resources and ability. Cooperation between organizations is widely used in emerging technology development and commercialization (Parkhe 1993). One of the foundations for Organization cooperation is resource's complementary. If these complementary resources can be gathered up, it would be able to create an organizational synergy effect and to make innovation activities more effective. Innovation is to introduce "new combination" of production factor to the production system (Schurnpeter 1934). Science and technology achievements appraised as capital stock has become an important pattern of commercialization of scientific and technological achievements (Zhao Jie et al. 2011). The survey made by American scholar showed that 70 % of total respondents had exchanged the technical achievements for the capital stock in 2000 (Feldman 2002). Stanford University has acquired 145 company shares by the way of technology transactions until 2005 (Stanford University 2010).

Cooperation innovation will fail if lack of mutual trust and cooperation among its members (Peng and Shenkar 1997). Moral risk can be produced when there are the asymmetric information and uncertainty (David Pérez-Castrillo and Sandonís 1997). An enterprise's aim is not only to create value, but able to obtain the value which had been created (Saloner and Podolny 2001). Both sides of cooperation consider how to distribute the created value while they consider how to create value. Proper value distribution mechanism can effectively promote enterprise partners to create value. So during the process of cooperative innovation, the rules of profit distribution among owners who have different resources are crucial and prominent problems. The owners of different resources often made the cooperation become a failure because of improper distribution of profit. Therefore, a reasonable, fair interest distribution mechanism could become the key to create maximum enterprise value in cooperative innovation. Brandenburger and Stuart (1996) proposed the strategy based on value. The strategy put the creation and access of value as its basic contents. He (Brandenburger and Stuart 2007) used two-stage game of no cooperation and cooperative to describe the creation and distribution of value. Makowski and Ostroy (2001) used value distribution to analyze a perfectly competitive market and proposed an idea that innovators should fully gain value caused by innovation. This paper applies two-stage game model researching creation and distribution of enterprise's value when a technological achievement has been changed into enterprise share and establishes a game model about the rules of proportion of value distribution and find the way of realizing the maximization of enterprise creating value.

94.2 Basic Assumptions

A firm comprises of two shareholders and uses technical achievements A to produce the product during the process of cooperation. One shareholder provides technical achievements A and another shareholder provides capital K. The labor input and business management of the firm are decided by the side that provides capital. Further assume that the labor input of the firm is L. Let $V(A, K, L)$ denote the firm's production function that is closely related to technical achievements A, capital K and labor inputs L. The firm's production cost is denoted by $C(K, L)$. The technology provider and the capital provider share the value created by the enterprise during the co-operation period. The gain obtained in value is denoted by P_A and P_K respectively. The amount of value created by the enterprise is decided by capital, the enterprise's management and technology. Two investment sides jointly influence the creation of value of the enterprise during the process of cooperation. One side that provides technology relies on advanced and unique technology and the ability to solve the technical problems appeared to bargain with another side that provides capital and management in their cooperation. Both sides rely on what they have to game and they will game about how to distribute value. In this paper, let r to denote the proportion of the total value of the enterprise obtained by the side that provide technology, so the proportion of the total value of the enterprise obtained by other side who provide capital is $1-r$. And $r \in (0, 1)$. $R\ (=R\ (A, K))$ represents a distribution mechanism. Both sides will game to decide about r.

94.3 The Game About Proportion of Distribution

In the cooperation process, the game process between technology provider and capital provider can be divided into two steps: the first step, technology provider and capital provider decided to provide what technology and how much capital respectively according to their assets' value and the maximization of the income distribution. The second step: technology provider and capital provider decide together how to distribute value on the basis that enterprise has already created value. In fact, the two steps of the game process are closely related, the decision and investment of two sides in the first step is the basis of distribution in the second step, while the situation of distribution of the second step directly decide input of the first step. In the first step, technology provider and capital provider decide technology A, the amount of capital K and labor inputs L respectively and decide the total value that enterprise create together:

$$V = V(A, K, L) - C(K, L) \tag{94.1}$$

In the first step, technology provider and capital provider agree on proportion of value distribution on the basis of their inherent bargaining ability. It is mentioned above that technology provider and capital provider obtaining distribution proportion are respectively $1-r$ and r, and $r = R(A, K)$, then the benefit of these two sides is respectively following:

$$\begin{cases} P_A^0 = rV = R(A, K)[V(A, K, L) - C(K, L)] \\ P_K^0 = (1-r)V = [1 - R(A, K)][V(A, K, L) - C(K, L)] \end{cases} \tag{94.2}$$

The decision-making of investment of technology provider and capital provider is made according to their maximum profit. Thus:

$$\frac{\partial p_A^o}{\partial A} = 0, \frac{\partial p_K^o}{\partial K} = 0, \frac{\partial p_K^o}{\partial L} = 0 \tag{94.3}$$

Now we can solve above three equations for the three unknown.

$$A = A_0, K = K_0, L = L_0 \tag{94.4}$$

The total value that enterprise creates is following at this moment:

$$V_0 = V(A_0, K_0, L_0) - C(K_0, L_0) \tag{94.5}$$

Could the sum of the firm value created be the maximum based on non-cooperative game in which two players maximize their own profit? According to the formula (94.1) we can know the necessary conditions for a maximum

$$\frac{\partial p}{\partial A} = 0, \frac{\partial p}{\partial K} = 0, \frac{\partial p}{\partial L} = 0 \tag{94.6}$$

We can get the answer by solving Eq. (94.6):

$$A = A_{max}, K = K_{max}, L = L_{max} \tag{94.7}$$

Now following is the maximum value of the enterprise created.

$$V_{max} = V(A_{max}, K_{max}, L_{max}) - C(A_{max}, L_{max}).$$

So:

$$\begin{cases} \frac{d}{dA}[V(A, K_{max}, L_{max}) - C(K_{max}, L_{max})]|_{A=A_{max}} & = 0 \\ \frac{d}{dL}[V(A_{max}, K_{max}, L) - C(A_{max}, L)]|_{L=L_{max}} & = 0 \\ \frac{d}{dk}[V(A_{max}, K, L_{max}) - C(K, L_{max})]|_{A=A_{max}} & = 0 \end{cases} \tag{94.8}$$

According to the formula (94.3), we can get :

$$\begin{cases} \frac{d}{dA}R(A, K_o)[V(A, K_o, L_o) - C(K_o, L_o)]_{A=A_o} & = 0 \\ \frac{d}{dL}[1 - R(A_o, K_o)][V(A_o, K_o, L) - C(K_o, L)]_{L=L_0} & = 0 \\ \frac{d}{dk}[1 - R(A_o, K)][V(A_o, K, L_o,) - C(K, L_o)]_{K=K_0} & = 0 \end{cases} \tag{94.9}$$

Comparing Eqs. (94.9) and (94.8), we can know that:

$$\begin{cases} \dfrac{d}{dA}[V(A, K_o, L_o) - C(K_o, L_o)]|_{A=A_o} \neq 0 \\ \dfrac{d}{dL}[V(A_o, K_o, L) - C(K_o, L)]|_{L=L_0} \neq 0 \\ \dfrac{d}{dk}[V(A_o, K, L_o,) - C(K, L_o)]|_{K=K_0} \neq 0 \end{cases} \tag{94.10}$$

So:

$$A_0 \neq A_{max}, K_0 \neq K_{max}, L_0 \neq L_{max} \tag{94.11}$$

From the above analysis we can see that enterprise's sum maximum value could not be realized when the two players seek for their own maximum interest by means of non-cooperative game. In the following we will design a distribution mechanism to encourage both players looking for the innovation of input elements in cooperation and to realize finally enterprise's maximum value by means of innovation of both players.

94.4 Mechanism Design to Encourage Both Sides of Cooperation to Innovate

Players of gamer sacrifice the maximum value of the whole enterprise while participating in value creation of enterprise based on their own maximum profit obtained. When the game reached a balance, both sides can't make change. Once one side change the input, the other side will correspondingly move and get back to the equilibrium. How could the maximum sum value created by the firm realize? The only way to change the balance of original situation is to design a new value distribution rules under the foundation of the original game balance. But is there any motivation of maximizing the sum of the firm value created through a dynamic innovation process existed in the minds of two players according to the Makowski's view that innovators obtain the value brought by innovation?

94.4.1 Analyzing Innovation Motivation on Both Sides of the Cooperation

From the Eq. (94.11) we can know the following:

$$A_0 \neq A_{max}, K_0 \neq K_{max}\ L_0 \neq L_{max};$$

Table 94.1 Return matrix

		s	
		Y	N
b	Y	$2(1 - R_0)(\Delta - 1)B$, $2R_0(\Delta - 1)B$	$(1 - R_0)(\Delta - 1)B$, $R_0\Delta B$
	N	$(1 - R_0)\Delta B$, $R_0(\Delta - 1)B$	0, 0

Contrast Eqs. (94.8) with (94.9) we found that $A_0 = A_{max}, K_0 = K_{max}$ $L_0 = L_{max}$ as long as $r = R(A, K) = R_0$ (R_0 is constant). But according to the new distribution rules, i.e., when $r = R_0$, is there any constant desire for innovation in the mind of both players? Is the technology provider willing to improve technology to add the total creating value of enterprise when the capital provider's input remains unchanged? And is the capital provider willing to carry out business management innovation and to adjust continuously capital supply for adding the total creating value of enterprise on the basis that the technology provider's technology remains unchanged. That is in the condition that one side's input remains unchanged, whether the other side is willing to make the amount of enterprise creating values changed through change their own inputs. Suppose $r = R_0$. $0.5 > R_0 > 0$. In another word, shareholders are divided into big shareholders (denoted by b) and small shareholders (denoted by s). Further suppose the efficiency of innovation's output on both sides is same and the costs of innovation is B, output efficiency is Δ ($\Delta > 0$), so the innovation output is ΔB. There are several situations as the following:

1. The cost of any player's innovation will be borne by itself and the enterprise value brought by innovation will be distributed according to the proportion of R_0. Table 94.1 is the return matrix; Y and N represent the two choices of the both players respectively: Innovation and not innovation. From the Table 94.1 we can see that:

   When $R_0\Delta > 1$, the Nash equilibrium of the two players is that big shareholder innovates and small shareholder innovates.

   When $(1 - R_0)\ \Delta > 1 > R_0\Delta$, the Nash equilibrium of the two players is that big shareholder innovates and small shareholder does not innovate.

   When $1 > (1 - R_0)\ \Delta$, the Nash equilibrium of the two players is that big shareholder and small shareholder do not innovate.

   Because $0.5 > R_0 > 0$, So when $\Delta > 1$, there is $1 > R_0\Delta$. So there is a possibility that shareholders sacrifice the enterprise's interests (i.e., non-innovation) while seeking for maximum personal interests. Distributing value according to the fixed ratio ($r = R_0$) does not cause the enterprise to create maximum value necessarily.
2. If the innovation cost B is assumed by the innovator itself and the innovator get all value brought by innovation. The return matrix is Table 94.2. The meaning of Y& N is the same as above. We can know from Table 94.2 that:

Table 94.2 Return matrix

		s	
		Y	N
b	Y	$(\Delta - 1)\,B,\ (\Delta - 1)\,B$	$(\Delta - 1)\,B,\ 0$
	N	$0,\ (\Delta - 1)\,B$	$0,\ 0$

Table 94.3 Return matrix

		s	
		Y	N
b	Y	$2\,(1 - R_o)\,(\Delta - 1)\,B,$ $2R_o\,(\Delta - 1)\,B$	$(1 - R_0)\,(\Delta - 1)\,B,$ $R_0\,(\Delta - 1)\,B$
	N	$(1 - R_0)\,(\Delta - 1)\,B,$ $R_0\,(\Delta - 1)\,B$	$0,$ 0

When$1 > \varDelta$, the Nash equilibrium of the two players is that big shareholder and small shareholder do not innovate. In this case the innovation activity is of no significance.

When $\varDelta > 1$, the Nash equilibrium of the two players is that big shareholder innovates and small shareholder innovates. There is no possibility that shareholders sacrifice the enterprise's interests (not innovation) while seeking for maximum personal interests.

3. When the situation is that innovation cost is completely assumed by the enterprise, the adding enterprise value $\varDelta B$ which brought by innovation is distributed according to proportion $R_{0\,.}$ Its return matrix is Table 94.3. The meaning of Y& N is the same as above. In this case both sides have innovation motive, even if when $\varDelta{-}1 < 0$, the side who innovates bear little loss. There is moral risk. Furthermore, that the cost of innovation B is assumed by the enterprise while innovator obtains all the value brought by innovation will not be accepted.

94.4.2 Convergence

In condition that both sides have innovation motivation, just like situation (94.2), does the sum value of the firm eventually converge on the maximum value caused by both sides dynamic innovation activities?

Definition An innovation cycle is that the two players (Technology provider A& capital provider K) are engaged in innovation only once respectively and the innovation activity could cause the value of the firm to add $P_{A(i)}\left(P_{A(i)} \geq 0\right)$, $P_{K(i)}\left(P_{K(i)} \geq 0\right)$ respectively.

Table 94.4 The firm value change caused by innovation

Cycle	The enterprise value change caused by innovation of side K	The enterprise value change caused by innovation of side A	The enterprise value change in end of circle
0	V_0	V_0	V_0
1	$V_0 + P_{K(1)}$	$V_0 + P_{A(1)}$	$V_1 = V_0 + P_{K(1)} + P_{A(1)}$
2	$V_1 + P_{K(2)}$	$V_1 + P_{A(2)}$	$V_2 = V_1 + P_{K(2)} + P_{A(2)}$
...	...	...	...
i	$V_{i-1} + P_{K(i)}$	$V_{i-1} + P_{K(i)}$	$V_i = V_{i-1} + P_{K(i)}+P_{A(i)}$
...	...	...	...
∞	V_{max}	V_{max}	V_{max}

In any innovation cycle, suppose that any player's innovation activity is based on the state of the last cycle and the innovation activity of a player is independent with another player. After the *i*-th innovation cycle, the sum value is V_{i}:

$$V_i = V_{i-1} + P_{A(i)} + P_{K(i)} \tag{94.12}$$

According to the Eq. (94.12), we can know the firm value after any innovation cycle. We can see concrete numerical value in Table 94.4.

We use value to construct a sequence of numbers I:

$$a_0 = V_0$$

$$a_1 = min\left\{V_0 + P_{A(1)}, V_0 + P_{K(1)}\right\}$$

$$a_2 = max\left\{V_1 + P_{A(1)}, V_0 + P_{K(1)}\right\}$$

$$a_3 = V_1$$

$$a_4 = min\left\{V_1 + P_{A(2)} V_1 + P_{K(2)}\right\}$$

$$a_5 = max\left\{V_1 + P_{A(2)} V_1 + P_{K(2)}\right\}$$

$$a_6 = V_2$$

$$\ldots$$

$$a_{3i-2} = min\left\{V_{i-1} = P_{A(i)}, V_{i-1} + P_{K(i-1)}\right\}$$

$$a_{3i-2} = max\left\{V_{i-1} + P_{A(i)}, V_{i-1} + P_{Ki-1}\right\}$$

$$a_{3i} = V_i$$

$$\ldots$$

$$a_{3-j}$$

V_{max}. Among them i $= 1, 2, \ldots;$ j $= 0, 1, 2$

Proposition 94.1 *The sequence of numbers (denoted by I) is convergent, and its limit value is the extreme value of the value creation function.*

Prove Because the sequence of numbers (denoted by I) have a limit, the sequence of numbers I is bounded according to the nature of the sequence of numbers. And because the sequence of numbers (denoted by I) is monotonous increasing, sequence of numbers (denoted by I) has a supremum. That is $V_{max} = \text{Sup } I$. We can know $\forall : \varepsilon > 0$, from the definition of the supremum, $V_{max} - \varepsilon$ is no longer the upper bound of a sequence of numbers (denoted by I). $\exists : a_N > V_{max} - \varepsilon$.

The sequence of numbers (denoted by I) is monotonous increasing when $3i - j > N$. So: a $a_{3i-j} \geq a_N$. Thus: $|a_{3_{i-j}} - V_{max}| = V_{max} - a_{3i-j} \leq V_{max} - a_N < \varepsilon$

So the sequence of numbers converges at V_{max}. And proposition 94.1 holds water.

Proposition 94.2 *With any innovation cycle, no matter which one of the two players who can first innovate, the maximum of enterprise value created can be realized.*

Prove It is not difficult to know that in any innovation cycle $P_{A(i)} > 0$ $K_{A(i)} > 0$ from observing the sequence of numbers I. No matter which side of the two players participating in the innovation first innovates, the number of innovation path is m and its sequence of numbers value creation series $\mathrm{I_m}$ is the subsequence of the sequence (denoted by I). The sequence of numbers (denoted by I) converges at V_{max}. According to the relationship of sequence limit and subsequence limit, all the sequence of numbers $\mathrm{I_m}$ converges at V_{max}. Proposition 94.2 holds water.

Particularly, when $P_{A(i)} \equiv 0$,

$$\sum_{i=1}^{n} P_{K(i)} = V(A_0, K_{\max}, L_{\max}) - \mathrm{C}(K_{\max}, L_{\max}) - [V(A_0, K_0, L_0) - \mathrm{C}(K_0, L_0)]$$

This is enterprise value created by the capital provider after cooperation (or the value of the innovation obtained by capital provider).

When $P_{K(i)} \equiv 0$,

$$\sum_{i=1}^{n} P_{A(i)} = V(A_{\max}, K_0, L_0) - C(K_0, L_0) - [V(A_0, K_0, L_0) - C(K_0, L_0)]$$

This is enterprise value created by technology provider after cooperation (or the value of the innovation obtained by technology provider).

94.5 Conclusion

Realizing industrialization of the technology achievement needs every participator jointly to work together, and there must be a proper value distribution mechanism to encourage participants to innovate constantly. This paper establishes the game model about the value distribution of both players and finds that participator may sacrifice maximum value of the whole enterprise for the maximum of their own interest while participating in the creation of value. Therefore, this paper builds new distribution rules for firm to create maximum value of the firm by participators' innovation of many cycles. This paper demonstrates that when participants seek for their own interests, the maximum sum value of a firm can be realized if there is a reasonable distribution mechanism.

References

Bettis R, Hitt M (1995) The new competitive landscape. Strateg Manage J 16(16):7–19
Brandenburger A, Stuart HW (1996) Value-based business strategy. J Econ Manage Strategy 5(1):5–24
Brandenburger A, Stuart HW (2007) Biform games. Manage Sci 53(4):537–549
David Pérez-Castrillo J, Sandonís J (1997) Disclosure of knowhow in research joint ventures. Inter J Ind Org 15(1):5l–75l
Feldman M (2002) Equity and the technology transfer strategies of American research universities. Manage Sci 48(1):105–121
Fujita M (1997) Small and medium-sized enterprises in foreign direct investment. In: Buckley PJ, Campos J, Mirza H (eds) International technology transfer by small and medium-sized enterprises. St Martin's Press, New York, pp 9–70
Guth WD (1990) Guest editor's introduction corporate entrepreneurship. Strateg Manage J 11:5–15
Makowski L, Ostroy JM (2001) Perfect competition and the creativity of the market. J Econ Lit 39(2):479–535
Parkhe A (1993) Strategic alliance structuring: a game theoretic and transaction cost examination of inter-firm cooperation. Acad Manage J 36(4):794–829
Peng MW, Shenkar O (1997) The meltdown of trust: a process model of strategic alliance dissolution. In: Proceedings of the academy of management annual meeting, Boston, 1997, pp 143–167
Saloner GA, Podolny J (2001) Strategic management. Wiley, New York, p 47
Schurnpeter JA (1934) The theory of economic development, (1912) (trans: Opie R). Cambridge, MA, Harvard University Press, pp 12–16
Stanford University (2010) Taking stock of equity. http://www.stanford.edu/group/OTL/about/brainstorm/1203_equity.html. Accessed 12 Oct 2010
Stopford JM (1994) Creating corporate entrepreneurship. Strateg Manage J 24(1):213–241
Zhao Jie, Zhang Jie-Jun, Tang Shi-Guo, Di Xiao-Yan (2011) The study of technology equity in the transformation of scientific and technological achievements. Stud Sci Sci 29(10):1485–1490 (in Chinese)

Chapter 95
Study on Product Information Model for Collaborative Product Design Based on Ontology

You-yuan Wang, Peng-biao Nie, Fa-lin Wang, Chao-yang Guo, and Xian-ming Dai

Abstract For the semantics of existing product information models is too simple, the lack of adequate semantic description of the product information model during collaborative product design process, and the information sharing and interoperability in the semantic level between the collaborative designers are difficulties and other issues, this paper presents a product information model for collaborative product design based on ontology. To achieve the description of semantic information of products, ontology is introduced in this paper. Then which can enable the collaborative product designers to carry out the semantic level of information sharing and interoperability. Finally, an instance of the presented product information model is proposed.

Keywords Collaborative design • Ontology • OWL • Product information model

95.1 Introduction

Collaborative product design (*CPD*) is a typical knowledge-intensive activity that includes conceptual design, detailed design, engineering analysis, assembly design, process design, and performance evaluation; each design task involves various

Y. Wang (✉)
Institute of Industrial Engineering, Nanchang Hangkong University, Nanchang, China
e-mail: yywnc@sina.com

P. Nie • F. Wang
School of Aeronautical Manufacturing Engineering, Nanchang Hangkong University, Nanchang, China

C. Guo
Hongdu Aviation Industry Group LTD, Nanchang, China

X. Dai
China Aviation Industry Archives, Beijing, China

E. Qi et al. (eds.), *The 19th International Conference on Industrial Engineering and Engineering Management*, DOI 10.1007/978-3-642-37270-4_95,

disciplines of design knowledge and experience, and is considered as one of the most promising business strategies for enterprises to response to global competition (Chen et al. 2008). Collaborative product design is a regional, cross-enterprise, and across time and space product design patterns, which emphasis on enterprises should have the collaboration capabilities of fast, flexible and dynamic response. Product development based on collaborative product design model not only improves the capacity of product innovation of enterprises, but also greatly shortens the product development cycle.

Collaborative product design is a process that all the designers are come from different disciplines, and spread across multiple enterprises in multiple sectors, which through consultations among their to complete the assigned task, during the design process requires for a large number of product data exchange and the communication of reliable information within the enterprise and between enterprises (Chen et al. 2005). However, the application systems and operating platform in enterprises and within the various departments of the enterprise are existed heterogeneity, the sources of product data also have a diversity, which are generate constraints on the achievement of the overall optimization of product design, conflict coordination and decision-making cooperation (Li et al. 2000). Therefore, to solve the problem of interoperability among heterogeneous systems in collaborative product design environment, and realize the efficient exchange of data and product design information sharing between the enterprise sectors and between collaboration enterprises, it is necessary to establish a complete, semantically rich product information model.

Based on the analysis of the collaborative product design work process and the theory of ontology modeling, this paper presents a product information model for collaborative product design based on ontology. To achieve the description of semantic information of products, ontology is introduced in this paper. Then which can enable the collaborative product designers to carry out the semantic level of information sharing and the interoperability of information model. Finally, an ontology-based application instance of the presented product information model is proposed.

95.2 Related Works

There have some scholars study on product information modeling for collaborative product design. Based on the analysis of collaborative product design data exchange demand among heterogeneous data sources and product design characteristics model, Chen et al. (2005) proposed a documentation model of product design information based on XML Schema and methods of XML documents interactive information processing. Rao et al. (2007) proposed use the generic standard SP95 (the 95th Standard Project) of management and control system integration to product information modeling in the collaborative design, which make up the deficiencies of the enterprise management system and control system integration capabilities. Wu et al. (2005) proposed a product data model which is suitable for distributed collaborative design and product data integration solutions in different places, and they also constructed a XML Web services-based PDM system, so the exchange and sharing of collaborative design data has been simplified, the problem of product data integration in heterogeneous collaborative enterprise also has been solved. Guo et al. (2004) proposed a product information modeling point of view based on Ontology and Semantic Web, by introducing the middleware layer in the model to achieve consistency of product information semantics, and enhance interoperability of the model. Wang (2009) proposed and implemented an ontology-based conceptual design of product information model, using the *F-B-S* (Function-Behavior-Structure) model to express non-geometry information, using the EXPRESS modeling language to describe the geometry information, and realize mapping from EXPRESS language to *OWL* language.

95.3 The Work Process of Collaborative Product Design

In the development process of complex product, the task decomposition and collaboration of product development is an important prerequisite for the successful completion of product development. Collaborative product design process is involved in many complicated interactions among multidisciplinary design teams in a distributed, heterogeneous and dynamic environment, including communication, cooperation, and coordination (Shen et al. 2000). The collaborative product design work process shown in Fig. 95.1.

In Fig. 95.1, the collaborative product design can be decomposed into a general problem-solving task, through the following 6 steps to complete the task of product design (Ertas and Jones 1993; Pahl and Beitz 1984):

Step 1: Acquiring requirements of product. Including market research, quality of understanding, and what the customer for such products, etc.
Step 2: Determining the overall goal and task. Determine the overall goal and task of the product development on the basis of master product demands.
Step 3: Decomposing the overall task. Decomposing the overall task into hierarchical subtasks (subtask 1, subtask 2, … , subtask *n*).
Step 4: Distributing subtasks to the all collaborative enterprises (collaborative enterprise 1, collaborative enterprise 2,…, collaborative enterprise *n*), which established on the basis of common consensus, trust and cooperation through the Internet/Intranet.
Step 5: Each collaborative enterprise solutions the assigned subtask.
Step 6: Synthesizing the sub-solutions, and ultimately produce an overall artifact solution.

Carry out the semantic level of information sharing and interoperability by the collaborative product designers is an important factor to successful completes the above steps. Establish a complete, rich semantics, and ontology-based product

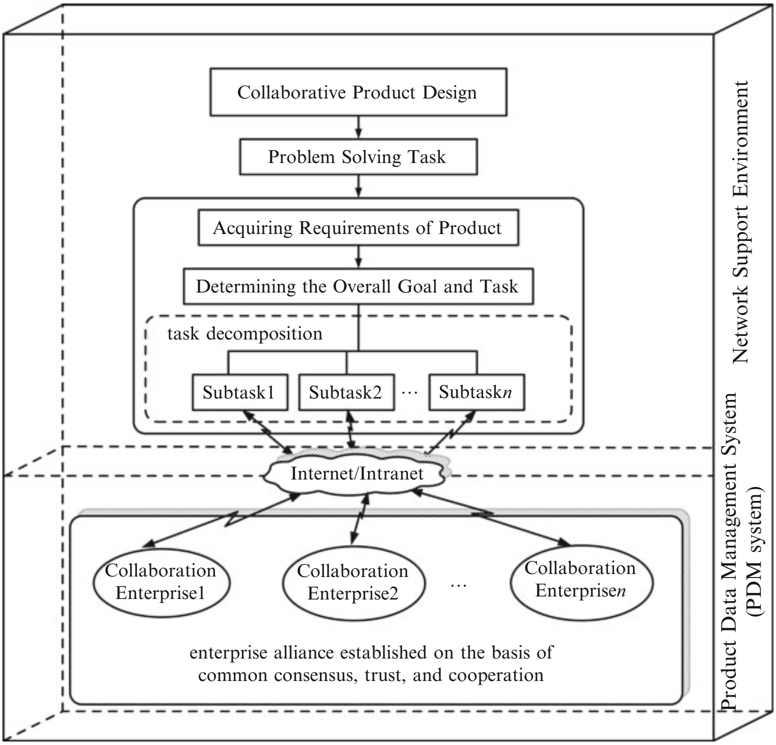


Fig. 95.1 The work process of collaborative product design

information model is intended to solve the problem of interoperability among heterogeneous systems in collaborative product design environment, and realize the efficient exchange of data and product design information sharing between the enterprise sectors and between collaboration enterprises.

95.4 Ontology Modeling

95.4.1 STEP Standard and EXPRESS Modeling Language

The standard for the exchange of product model data (*STEP*) has been developed by the International Organization for Standardization (*ISO*); it is to define a uniform representation of product information and to provide the neutral mechanisms that enable the exchange of product data between different computer systems over the entire product life cycle (Spidy 1994; Pratt 2000; ISO 10303-21 1994; ISO 10303-28 1998; ISO 10303-25 2002). *STEP* standard provides the foundation for the information integration of product model and general description mechanisms.

EXPRESS information modeling language (*ISO* 10303-11) is an object-oriented description language for information model which is developed by *STEP* standard used to describe the integration of resources and application protocols. Use the *EXPRESS* language to define the *STEP* model structure can complete and constrained to describe the product data information. However, the information model established with the *EXPRESS* cannot fully meet the needs of product knowledge representation and sharing for collaborative enterprises, and the product model established is not easy to understand too.

95.4.2 Ontology Modeling Primitives

Ontology is an explicit, formal specification of a shared conceptual model (Studer et al. 1998). Perez and Benjamins (1999) considered that the ontology can be organized by classification, and summarized ontology contains five basic modeling primitives. These primitives are: classes or concepts, relations, functions, axioms and instances.

1. Classes or concepts: Semantically speaking, the classes are a collection of instances of the same nature; in knowledge engineering, we take the concepts interpreted as objects or a model for the knowledge of events.
2. Relations: An association between the concepts, formally defined as a subset of the n dimensions of Cartesian product: $R : C_1 \times C_2 \times \cdots \times C_n$. There are 4 basic relationships between the concepts and shown as follows:

   - *part-of*: express the relationship of the whole and part between the concepts.
   - *kind-of*: express the inheritance relationship between the concepts, similar to the relationships between super class and subclass in object-oriented.
   - *attribute-of*: express a concept is another concept's property.
   - *instance-of*: express the relationships between the concept's instances and concepts, similar to the relationships between objects and classes in object-oriented.

3. Functions: A special relationship, which the first n element relative to the previous n − 1 elements is unique. Formally defined as: $F : C_1 \times C_2 \times \ldots \times C_{n-1} \to C_n$.
4. Axioms: Axioms are a recognized fact or inference rule within ontology, used for knowledge reasoning.
5. Instances: Are the basic elements of a concept class, from the semantic analysis, instance that is the object.

There are some advantages through the introduction of ontology in the product information modeling: ontology providing a common understanding of knowledge in specific domains; determine the common recognition of the vocabulary; and given clearly defined for these terms (terminology) and the relationship between lexical from different level of formal models. So the information model can support the semantic information description of product and product-related process and planning and management, which is particularly important for collaborative product design.

95.4.3 Semantic Web

For the situation of how to add contents on Web resources which the computer can understand, and then the computer can better handle, Berners-Lee proposed the Semantic Web (2000) concept. The Semantic Web's goal is to make the information on the Web possess semantics which computers can understand.

The *OWL* Web Ontology Language was developed by the W3C (World Wide Web Consortium), used to describe the semantics of Ontology, and is the standards of ontology description language in Semantic Web by W3C recommended. *OWL* based on *RDF* and *RDFS*, using XML-based *RDF* syntax, and described by the definition of RDF instance. *OWL* language has well-defined syntax and semantics, efficient reasoning support, adequate expression and expression of convenience (Petrinja et al. 2007). The semantic information and relationships between the terms of design knowledge can be articulated by the *OWL*.

95.5 Product Information Model Based on Ontology in CPD Environment

Ontology study has laid a good foundation for the establishment of product information model which support semantic interoperability. Figure 95.2 shown in this paper is the established ontology-based product information model. The logical structure of this model consists of four levels, namely, the resource layer, ontology layer, the information access layer and application layer.

1. Resource layer: Storage various data to support the collaborative design of products, including drawing files, office documents, database, structured data, unstructured data, and domain experts, etc.
2. Ontology layer: The ontology layer is the core layer of this product information model. There has a mapping relationship between local ontology and global shared ontology. Different collaborative enterprises can share the common understanding of knowledge by ontology translation between them. Collaborative enterprises can communicate with each other by communication platform on the base of global shared ontology (Su et al. 2009).
3. Information access layer: The access layer is the portal for each collaborative enterprise to access the ontology layer, it providing ontology-based product model to support access to information, and can be implemented with wide

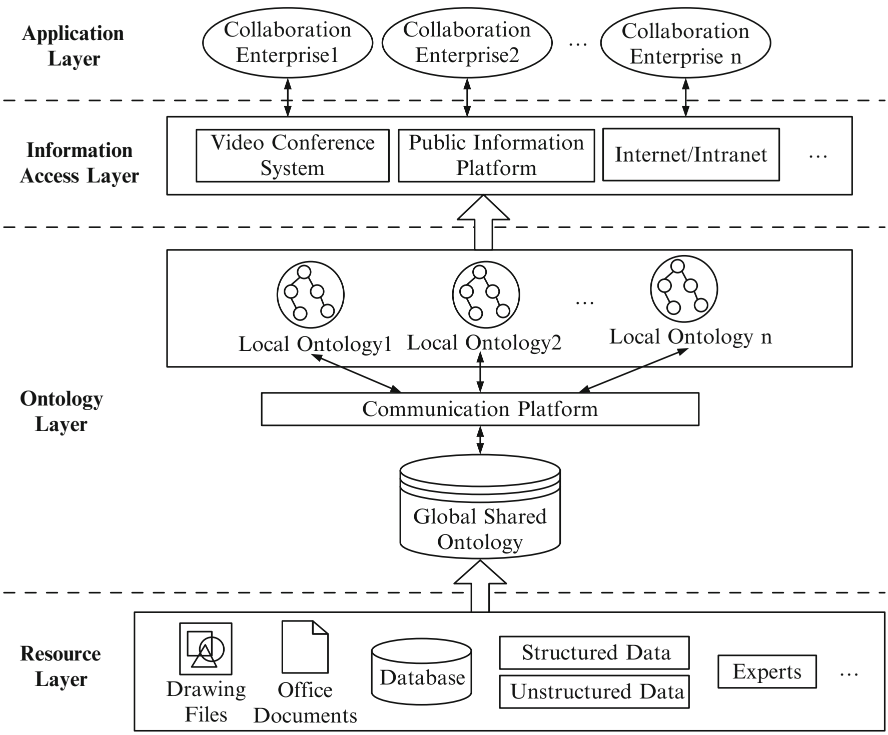


Fig. 95.2 Product information model for collaborative product design based on ontology

range of information access mode based on CORBA, Agent technology, etc. The collaborating enterprises can be accessed through a variety of terminal equipments, such as Internet/Intranet, video conferencing systems and public information platform, etc.

4. Application layer: The application layer provides collaborative design services for each collaborative enterprise, such as CAD, PDM, ERP and other enterprise distributed application systems.

Current product model can be summarized as the following three types, namely, geometry-oriented product information model, feature-oriented product information model and integrated product information model. There exist some deficiencies above models: (1) semantic model is too simple; (2) difficult to achieve for a description of knowledge and intelligent reasoning on the basis of that; (3) unable to express non-standardized data, etc. The introduction of ontology layer in the product information model based on ontology in CPD environment can make up the above deficiencies. By constructing the same semantic context in the ontology layer, the collaborative product designers who use ontology can carry out the semantic level of information sharing and interoperability.

95.6 Application Instance of Ontology-Based Product Information Model

This paper takes the collaborative design of rolling bearing as an example to establish an ontology-based rolling bearing information model. Rolling bearing generally include the following aspects: structure, characteristics, classification, code, main dimensions and tolerances, materials, clearance, bearing selection, fit, lubrication and other parts. Now the following three characteristics are taken to classify the information content of rolling bearing: performance characteristics, structure characteristics and auxiliary characteristics. Part of the information model of rolling bearing shown in Fig. 95.3.

The ontology development tool Protégé (Knublauch et al. 2004) developed by Stanford University is currently one of the most widely used ontology editor. Input the structured or semi-structured entity concept and attribute information by the users, and then output ontological, semantic knowledge model and based on that model for semantic reasoning; support multiple inheritance, can check on the consistency and expanded for the new data; by use a different plug-ins can be converted internal representation into various forms of text formats, including XML, RDF (S), OIL, DAML, DAML + OIL, OWL, and other languages. In this paper, we use the Protégé 3.1.1 version to create a rolling bearing ontology; Fig. 95.4 shows the part of OWL ontology model diagram of rolling bearing.

Using the OWL to describe the ontology means taking the defined meta-ontology to describe the concepts and relationships in formal. In Fig. 95.4, the ellipse represents a class, the connection indicates the property, "is-a" property represents the inheritance relationship between the classes.

Figure 95.5 shows the hierarchical structure of the rolling bearing ontology.

The hierarchical structure of ontology reflects the relationships between the super class and subclass of the rolling bearing concepts. In Protégé 3.1.1of the Asserted Hierarchy, the classes of ten have only one super class, then the inference engine (e.g. Racer) can help calculate and maintain multiple inheritances, making the class's modules of ontology and their relationships become clear.

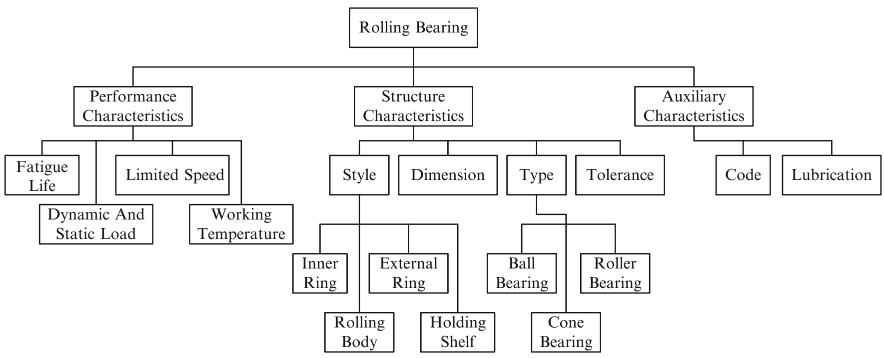


Fig. 95.3 Part of the information model of rolling bearing

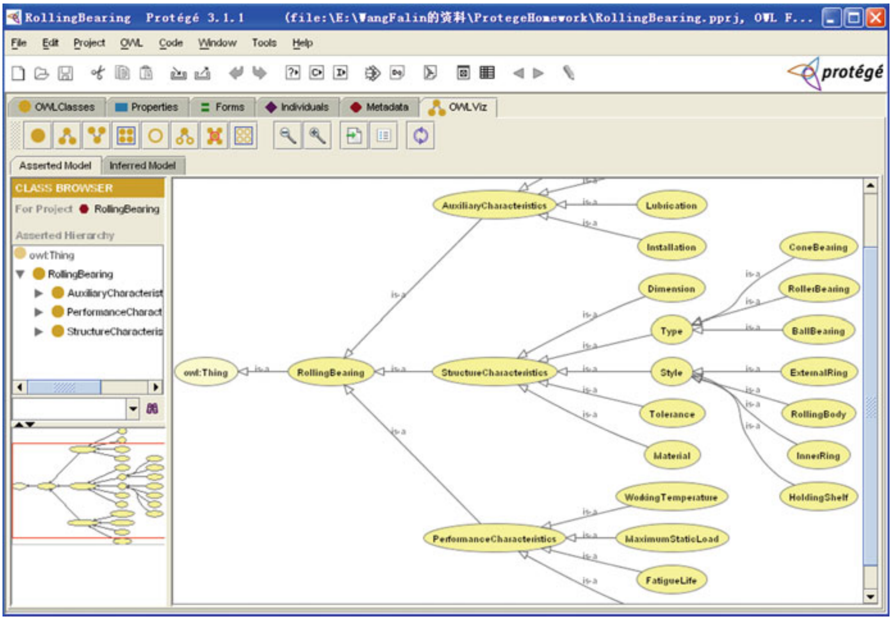
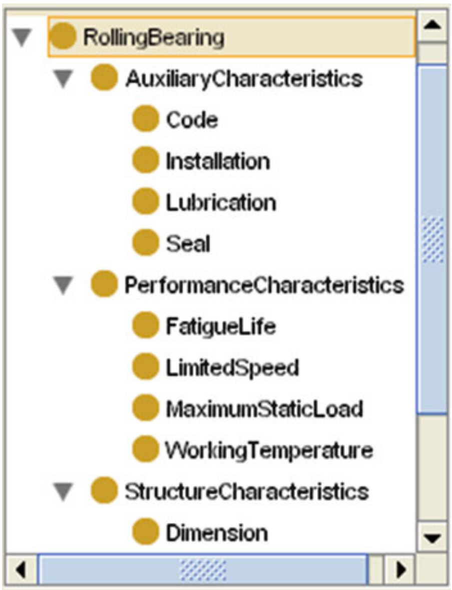


Fig. 95.4 The part of OWL ontology model diagram of rolling bearing

Fig. 95.5 The hierarchical structure of the rolling bearing ontology



The ontology fragment of the rolling bearing design information described by OWL shown in following.

```
……
<owl:Classrdf:ID=”PerformanceCharacteristics”>
    <rdfs:subClassOf>
    <owl:Classrdf:ID=”RollingBearing”/>
    </rdfs:subClassOf>
</owl:Class>
<owl:Classrdf:ID=”StructureCharacteristics”>
    <rdfs:subClassOf>
<owl:Classrdf:ID=”RollingBearing”/>
    </rdfs:subClassOf>
</owl:Class>
<owl:Classrdf:ID=”AuxiliaryCharacteristics”>
    <rdfs:subClassOf
            rdf:resource=”#RollingBearing”/>
</owl:Class>
<owl:Classrdf:ID=”FatigueLife”>
    <rdfs:subClassOf
    rdf:resource=”#PerformanceCharacteristics”/>
</owl:Class>
<owl:Classrdf:ID=”LimitedSpeed”>
    <rdfs:subClassOf
    rdf:resource=”#PerformanceCharacteristics”/>
</owl:Class>
……
```

95.7 Conclusion

Ontology study has laid a good foundation for the establishment of product information model which support semantic interoperability. This paper first analyzes the work process of collaborative product design and the theory of ontology modeling, and then has proposed a product information model for collaborative product design based on ontology. To achieve the description of semantic information of products, ontology is introduced in this paper. Finally, this paper takes the collaborative design of rolling bearing as an example, establishes an ontology-based rolling bearing information model. Based on this information model, implement the semantic level of information sharing and interoperability for the collaborative designers. And also effectively solve the problems of heterogeneous data exchange and interoperability and other issues during the collaborative product design process of heterogeneous enterprises.

Acknowledgment This paper is supported by Natural Science Foundation of Jiangxi Province of China (20114BA201045), Scientific and Technological Support Projects of Jiangxi Province of China (No. 2009BGB03100), Aeronautical Science Foundation of China (No. 2009ZG3900, No. 2010ZG56027) and Research Project of Jiangxi Provincial Department Education of China (No. GJJ11511).

References

Berners-Lee and the Semantic Web Vision (2000). http://www.xml.com/pub/a/2000/12/xml200/timlb.html. Accessed 6 Dec 2000

Chen J-X, Chen H, Zhang P-W (2005) Research on XML-based document model for product cooperation design. J Inform 24:49–52

Chen Y-J, Chen Y-M, Chu H-C (2008) Enabling collaborative product design through distributed engineering knowledge management. Comput Ind 59:395–409

Ertas A, Jones J-C (1993) The engineering design process. Wiley, New York

Guo M, Li S-P, Dong J-X, Fu X-J (2004) Research on product information model based on Ontology and Semantic Web. J Zhejiang Univ (Eng Sci) 38:22–28

ISO 10303-21 (1994) Industrial automation systems and integration Product data representation and exchange. Part 21. Clear text encoding of the exchange structure, China Standards Press

ISO 10303-28 (1998) Industrial automation systems and integration. Product data representation and exchange. Part 28. Implementation methods: XML representations of EXPRESS schemas and data, China Standards Press

ISO 10303-25 (2002) Industrial automation systems and integration Product data representation and exchange. Part 25. Implementation methods: EXPRESS to XML Binding, China Standards Press

Knublauch H, Musen MA, Rector AL (2004) Editing description logics ontologies with the Protégé owl plugin. In: International workshop on description logics, Whistler, BC, Canada, 2004

Li J, Zhang H-L, Wang J-B, Xiong G-L (2000) Collaborative design method on network. J Tsinghua Univ (Sci Technol) 40:93–96

Pahl G, Beitz W (1984) Engineering design: a systematic approach. Springer, Berlin

Perez A-G, Benjamins V-R (1999) Overview of knowledge sharing and reuse components: ontologies and problem-solving methods. In: Stockholm VR, Benjamins B, Chandrasekaran A (eds) Proceedings of the IJCAI-99 workshop on ontologies and problem-solving methods (KRR5), Stockholm, 1999, pp 1–15

Petrinja E, Stankovski V, Turk Ž (2007) A provenance data management system for improving the product modelling process. Automat Constr 16:485–497

Pratt M (2000) Introduction to ISO 10303-the STEP standard for product data exchange. ASME J Comput Inform Sci Eng 1:102–103

Rao Y, Chen D-K, Zhou Z-C, Wan X-L (2007) Product information model based on SP95 standard in collaborative design. Comput Integr Manuf Syst 13:1528–1533

Shen W, Norrie D-H, Barthes J-P (2000) Multi-agent systems for concurrent intelligent design and manufacturing. Taylor & Francis, London

Spidy P (1994) STEP Part 11: the EXPRESS language reference manual. ISO TC184/SC4 N65

Studer R, Benjamins V-R, Fensel D (1998) Knowledge engineering: principles and methods. Data Knowl Eng 25:161–197

Su T-M, Qiu X-P, Yu Y-L (2009) An ontology-based collaborative design system, vol 5738. Springer, Heidelberg, pp 69–76

Wang Y-Y (2009) Ontology-based product information modeling of conceptual design. In: Proceedings-2009 international conference on electronic commerce and business intelligence, ECBI 2009, Beijing, pp 53–56. doi:10.1109/ECBI.2009.118

Wu J-W, Qiu Q-Y, Feng P-E, Gao F (2005) Management strategy of product data in distributed collaborative design environment. J Zhejiang Univ (Eng Sci) 39:1465–1469

Chapter 96
An Analysis of Evaluating Enterprises' Ecological Management Information Systems

Jin-yan Sang, Zhen-fa Qi, and Ying-hua Zhang

Abstract Enterprises' management information systems performance appraisal traditionally focused on overall efficiency evaluation. This paper brings ecological performance appraisal for enterprises' management information systems with ANN and GA, which can provide the sequence of the subsystems and to find out the feeblest subsystem, and take the corresponding measure to improve it according to the expert's appraisal. The evaluation of ecological management information system should be implemented step by step: First of all, carry on overall analysis and appraisal to each subsystem that is operated; then, draw the value of evaluative index of each subsystem and arrange them in an order based on that. Finally, draw the lowest value in each subsystem according to the result of arranging in an order, judging by that we can improve the subsystem purposefully. The improved mechanism will harmonize the enterprises' interior information system and make enterprises ecologically adapt to the changing environment.

Keywords Artificial neural networks • Enterprises • Evaluation • Genetic algorithm • Management information system

96.1 Introduction

At present, both seller and user have paid more attention to the most outstanding subsystem, but neglected to optimize the feeblest subsystem when they evaluate on management information system. As we know, it is deadly sometimes. Therefore

J. Sang (✉) • Z. Qi
School of Business, Shandong University of Technology, Zibo, People's Republic of China
e-mail: sangjy@126.com

Y. Zhang
School of Business, Tianjin University of Finance and Economics, Tianjin, People's Republic of China

E. Qi et al. (eds.), *The 19th International Conference on Industrial Engineering and Engineering Management*, DOI 10.1007/978-3-642-37270-4_96,

this paper intends to adopt the model that combines Artificial Neural Networks (ANN) with Genetic Algorithm (GA) to carry on comprehensive appraisal to each subsystem of MIS, so that we can find out the subsystem that is most feeble and need to be improved among them to enhance the capability of whole ecological Management Information System (Gupta 2010; Oreski et al. 2012). The ecological MIS behold enterprises and series of them of the perspective of ecology. Like the interdependence of ecosystem, the management information subsystems are also interdependent. The adaptability and promotability of ecological information subsystems are helpful to promote the evolution of management information system (Summers et al. 2012). And the evolutionary mechanism will enhance the adaptability and competition of enterprises.

The evaluation of ecological management information system should be implemented step by step: First of all, carry on overall analysis and appraisal to each subsystem that is operated; then, draw the value of evaluation index of each subsystem and arrange them in an order based on that. Finally, draw the lowest value in each subsystem according to the result of arranging in an order, judging by that we can improve the subsystem purposefully. In this paper, we will combine GA with ANN to utilize the ability of GA to optimize whole system to combine GA with ANN to set up the genetic neural network (Rasmussen 2011). To analyze and evaluate enterprises' ecological MIS through studying cases, we can find out the feasibility and higher scientific theoretical foundation of this research (Li et al. 2012).

96.2 Confirming the Evaluation Index Systems of Enterprise's Ecological Management Information Systems

According analyzing characteristic of ecological management information system, it is workable to set up the decision evaluation index to each subsystem. There are 12 indexes in all, and it must be pointed out that they are not totally separate. Some have certain dependence, so they are must be considered synthetically (Montiel et al. 2012). The concrete indexes are as follows:

Dependability of the system: Dependability is decided by the dependability of the hardware system and the software system, etc. It is usually primary problem that a user cares about, especially in finance, traffic, security system, etc.

Expansibility of the system: The expansibility refers to the extendible degree of systematic treatment ability and system function; it can be divided into expandability of the structure of the system, hardware equipment, software function, etc.

Maintainability of the system: The systematic maintainability refers to the expense of confirming and correcting the mistake in the system. It is determined by some factors such as degree of modularize, simplicity, consistency, etc.

Systemic sharing: The sharing refers to degree the system shares information with outside.

Practicability of the system: Practicability means using the MIS to support to the decision and management in the daily affairs.

Systematic amity: Amity means that users use the MIS very convenient, and the man–machine interface is good.

Security of the system: The reason to endanger the security of system is as follows: the unreliability of systemic software and hardware, user's unintentional wrong operation, natural calamity and many means taken by the opponent to filch secret and destroy systemic normal running. The effective countermeasure and safety measure must be taken.

Advance of the system: Advance means building the MIS is advanced on the whole, so the system can produce greater benefit, and have longer lifecycle.

Systemic economical efficiency: Economical efficiency refers to the expenses that users use the MIS.

Degree of user's satisfaction to System: Degree of user satisfaction means the degree of user's satisfaction of all respects in subsystem such as function, performance, user interface, etc. (Gokmen and Hamşioglu 2011). In addition, the strong fault-tolerant, system apt using, reasonable design, helpful function, operating easily should also be considered. Value of subsystem is reflected through user using. Subsystem can come into operation through approval of user.

Support of Systematic leader: The support of the person in charge of the MIS is an extremely important factor to maintain in construction.

Systematic benefit: Benefit refers to social benefit and economic benefit that the MIS produces to the outside society. Usually social benefit of MIS is much greater than the direct economic benefits (Ghosh et al. 2012). It is embodied in promoting economy to develop in harmony, improving the level of science and technology, realizing the scientific decision, improving the level of production, promotes the social welfare and rationally and protecting ecological environment, etc.

96.3 Building Genetic Neural Network Model to Evaluate the Enterprise's Ecological Management Information System

96.3.1 Figures and Tables

The artificial neural network is a dynamics system formed by extensive interconnection of a lot of non-linear components. The hereditary algorithm is one kind optimizes algorithms overall situation on the basis of natural selection and naturally inheritance, and have essential parallel computation characteristics, adopt several kinds of operators that comes out from natural selection mechanism to operate the parameter code character bunch. The operation to go on to the colony formed by a lot of feasible answers, so it can afford to walk abreast in replacing in generation, and it search for towards more likely finding optimum direction of the overall situation and not unlikely making the search be put into min. Adopting the GA to train the forward network can study the connection right value and structure of the network at the same time and get better results of learning, the networks got have good adaptability (Ozcelik and Sarac 2012).

96.3.2 Structure and Algorithm of Genetic Neural Network

Adopting Genetic Algorithms to study weight value in network:

Beside input layer suppose the total input of the node in any layer as:

$$u_i = \sum_j W_{ij} \cdot V_j \tag{96.1}$$

Where V_j stands for the output of the node j in former layer, W_{ij} stands for the weight value in network. The output of node i is expressed as:

$$U_i = g(U_i) \tag{96.2}$$

Where g(·) is encourage function, it is expressed as:

$$V_j = \frac{1}{1 + \exp\left(-\left(U_j - \theta_j\right)\right)} \tag{96.3}$$

In the formula (96.7), θ_j stands for the node's threshold.

Here, we define error function of training network's weigh value as:

$$E_k = \sum_j \left(t_{kj} - v_{kj}\right)^2 \tag{96.4}$$

In T training sets, the total errors are expressed as:

$$E = \frac{1}{k}\sum_k E_k = \frac{1}{k}\sum_k \sum_j \left(t_{kj} - v_{kj}\right)^2 \tag{96.5}$$

Where, K = T refers to the number of sample. If there are L layers in multiplayer forward network, including an input layer and an output layer (Vara and Marcos 2012). Using P_i to express the number of nodes in i layer and every node includes a threshold called θ_j, then every network has parameter as:

$$P = \sum_{i=1}^{l-1} P_{i+1}\left(P_i + 1\right) \tag{96.6}$$

For example, if there a network including two layer, L = 2, every layer include two nodes, ($P_1 = P_2 = 2$), then the parameter $P = 2 \times (2 + 1) = 6$.

Generally, every network expressed by a chromosome in binary. In the network, every connection is expressed as ten genes in binary. There are six parameters in the above example, coding at random:

1100010101	0100011010	0111110001
W11	W11	W11

One of important problem when GA is used to optimize the weight value of neural network is how to define appropriate goal function. Error function is often used as a solution.

$$F(E) = E_{\max} - E \tag{96.7}$$

In the formula (96.7), E_{max} stands for the maximum of error function.

96.3.3 Genetic Neural Network Model to Evaluate the Ecological Management Information System of Enterprises

Genetic neural network model that will evaluate the MIS for enterprises includes two steps. Appraise the subsystem at first, then execute comprehensive appraise on enterprises' MIS. The structures and functions of two systems are different.

Enterprises' management information subsystem appraises in genetic neural network model is confirmed according to the amount of the subsystem, but the neural network structure of each subsystem is the same (Peifeng 2011). The network is divided into three layers, and input layer has 12 units. They are dependability, expansibility, maintainability, and security of the system, advance, and economic efficiency of the system, sharing of the system, practicability of the system, amity of the system, systematic user's satisfaction, system leader support and systematic benefit. We can obtain structure weight among units in network. In order to obtain actual weight of input and output unit, we define a group of formulae here:

We have the coefficient of correlation significance in formula (96.8). $x = w_{jk}$ in the formula.

$$r_{ij} = \sum_{k=1}^{k=p} \frac{w_{ki}\,(1 - e^{-x})}{(1 + e^{-x})} \tag{96.8}$$

Relative effect coefficient:

$$F_{ij} = \frac{rij}{\sum\limits_{i=1}^{i=m} \sum\limits_{j=1}^{j=n} r_{ij}} \tag{96.9}$$

Relevance exponent is in formula (96.10). $y = r_{ij}$ in the formula.

$$R_{ij} = \frac{(1 - e^{-y})}{(1 + e^{-y})} \tag{96.10}$$

Absolute effect coefficient:

$$s_{ij} = \frac{R_{ij}}{\sum_{i=1}^{i=m} R_{ij}} \tag{96.11}$$

Where, i is input unit of the neural network, i is from 1 to m; j is output unit of the neural network, j is from 1 to n; k is hidden unit of the neural network, k is from 1 to p; w_{ki} is weight coefficient between input neuron i and hidden neuron.; w_{jk} is weight coefficient between output neuron j and hidden neuron.

The absolute influence coefficient Sij defined above shows the relative weight comes from the input factor contrasting to output. Utilize a group of formulae to sum up trained network weight coefficient and get every factor weight when appraised of enterprises' ecological information subsystems (Kwak et al. 2012). The details of indexes and weights on experts are dependability (0.15), expandability (0.10), maintain ability (0.09), sharing (0.09), practicability (0.06), amity (0.06), security (0.06), advance (0.10), economical efficiency (0.07), satisfaction of user (0.06), support of leader (0.05), systematic benefit (0.11).

Genetic neural network model on comprehensive evaluation of enterprises' MIS is decided by the number of subsystems. Input unit is every subsystem's utility value that can be got from the appraisal above. Output unit is comprehensive appraisal value of Enterprises' MIS.

96.3.4 Standardizing the Network Input

In the system of evaluation index, the most indexes are qualitative. It is difficult to quantify. Though some indexes can show by qualitative analysis, such as the economic index, it is very difficult to calculate precisely the cost in the real work. It's the same to benefit index, especially the social benefit. It is may be a bit more convenient to adopt qualitative appraisal to kind of index. So in the actual appraisal work, the expert should adopt qualitative appraisal to all indexes at first, then try quantification.

Principle for the expert evaluating is to adopt the fuzzy comprehensive decision method to analyze primitive information, and sort them to five kinds: best, better, good, generally, bad, worse, worst (Allahviranloo and Ghanbari 2012). For example: In some subsystem, information of checking is examined index. Fuzzy comprehensive decision method produces the following analysis message: (best, better, good, generally, bad, worse, worst) corresponding to be (1/7, 1/7, 1/7, 1/7, 1/7, 1/7, 1/7). Then "generally" and the degree of the above is about $1/7 + 1/7 + 1/7 + 1/7 = 4/7$. Therefore the frequency of "generally" appearing is greater than 50 %, so we think the grade of the examination as "generally". After each expert examine each index of each subsystem and determining the qualitative grade, we go on to do corresponding quantification to the comment. Quantitative standards of experts to evaluate are best (100 points), better (90 points), good (80 points), generally (70 points), bad (60 points), worse (50 points) and worst (40 points).

In order to standard the index value, we suppose the following formula.

$$X_{ij} = \frac{a_{ij}}{\max\{a_{ij}\}} \tag{96.12}$$

In the formula (96.12), $1 \le i \le m$, and x_{ij} is the input value network need, and a_{ij} is the standard score of some index after some experts appraise.

96.4 Case Study

Take enterprise's ecological management information system as an example now, then to carry on comprehensive appraisal to its each subsystem and explain the application of the model (Liaskos and Khan 2012). Further, This enterprise' ecological information system includes 9 subsystems, because of the security of enterprise's information, nine subsystems are expressed respectively with A, B, C, D, E, F, G, H, I. And 12 evaluation indexes, ten experts will be exam respectively each index. In order to state the question, the evaluate subsystem D is done first, do the same to other subsystems.

According fuzzy evaluation, we can obtain result of appraisal to each index in each subsystem (Lu et al. 2012). According the formula (96.12), we can obtain the evaluation result of each subsystem. The details of scores (A, B, C, D, E, F, G, H, I) of each subsystem are dependability (1.0, 0.9, 0.9, 1.0, 0.6, 0.7, 0.9, 0.4, 0.7), expandability (1.0, 0.89, 0.67, 0.67, 0.44, 0.89, 0.78, 0.89, 0.78), maintain ability (1.0, 0.78, 0.78, 1.0, 0.67, 1.0, 0.67, 0.78, 0.89), sharing (0.75, 0.875, 0.75, 1.0, 0.5, 1.0, 0.625, 0.625, 0.5), practicability (1.0, 0.67, 0.78, 1.0, 0.67, 1.0, 1.0, 0.67, 1.0), amity (1.0, 0.625, 0.625, 0.625, 0.625, 1.0, 0.875, 1.0, 1.0), security (1.0, 0.67, 0.78, 0.44, 0.89, 1.0, 0.78, 0.67, 1.0), advance (1.0, 0.75, 0.875, 0.75, 0.5, 1.0, 0.75, 0.875, 0.75), economical efficiency (1.0, 0.57, 0.86, 0.86, 0.71, 1.0, 0.86, 1.0, 0.86), satisfaction of user (1.0, 0.625, 1.0, 0.75, 0.5, 1.0, 0.75, 1.0, 0.875), support of leader (1.0, 0.78, 0.78, 1.0, 0.56, 1.0, 0.44, 0.78, 1.0), systematic benefit (1.0, 0.5, 0.75, 0.875, 0.5, 0.75, 0.875, 1.0, 0.875). Input data to neural network and analysis, we can see that subsystem B and subsystem E are feeble links in implementing ecological MIS of enterprise. They should be strengthened and improved in daily management.

96.5 Conclusion

It is a very important job for the improvement of the system to comprehensively appraise on every subsystem of MIS; feeble link of system can be found out and improved. In this paper, to use ANN and GA can provide the sequence of the subsystems and to find out the feeblest subsystem, and take the corresponding measure to improve it according to the expert's appraisal. Beyond all questions the improved information system will ecologically adapt to the changing environment.

Acknowledgment The research was sponsored by the Shandong Provincial Natural Science Foundation of China (Project No. ZR2010GM002) and the Shandong Provincial Soft Science Foundation of China (Project No. 2010RKGA 3018).

References

Allahviranloo T, Ghanbari M (2012) On the algebraic solution of fuzzy linear systems based on interval theory. Appl Math Model 36(11):5360–5379

Du Peifeng (2011) Analysis on the development trend and cause of business MIS strategy. Manage World 167(8):144 (in Chinese)

Ghosh S, Moonan PK, Cowan L, Grant J, Kammerer S (2012) Tuberculosis genotyping information management system: enhancing tuberculosis surveillance in the United States. Genet Evol 12(4):782–788

Gokmen A, Hamşioglu AB (2011) The effect of knowledge management, technological capability and innovation on the enterprise performance: a comprehensive empirical study of the Turkish Textile Sector. J Inf Knowl Manage 10(1):1–10

Gupta AK (2010) Predictive modeling of turning operations using response surface methodology, artificial neural networks and support vector regression. Int J Prod Res 48(3):763–778

Kwak B, Ro BT, Suk I (2012) The composition of top management with general counsel and voluntary information disclosure. J Acc Econ 54(1):19–41

Li C, Peters GF, Richardson VJ, Marcia WW (2012) The consequences of information technology control weaknesses on management information systems: the case of sarbanes-oxley internal control reports. MIS Q 36(1):179–204

Liaskos S, Khan SM (2012) Behavioral adaptation of information systems through goal models. Inf Syst 37(8):767–783

Lu CY, Gu W, Dai AH (2012) Assessing habitat suitability based on geographic information system and fuzzy: a case study of Schisandra sphenanthera Rehd. Ecol Model 242(3):105–115

Montiel I, Husted BW, Christmann P (2012) Using private management standard certification to reduce information asymmetries in corrupt environments. Strateg Manage J 33(9):1103–1113

Oreski S, Oreski D, Oreski G (2012) Hybrid system with genetic algorithm and artificial neural networks and its application to retail credit risk assessment. Expert Syst Appl 39(16):12605–12617

Ozcelik F, Sarac T (2012) A genetic algorithm extended modified sub-gradient algorithm for cell formation problem with alternative routings. Int J Prod Res 50(15):4025–4037

Rasmussen W (2011) Use of the local comprehensive plan in the appraisal process. Apprais J 79(3):223–228

Summers JK, Humphrey SE, Ferris GR (2012) Team member change, flux in coordination, and performance: effects of strategic core roles, information transfer, and cognitive ability. Acad Manage J 55(2):314–338

Vara JM, Marcos E (2012) A framework for model-driven development of information systems: technical decisions and lessons learned. J Syst Softw 85(10):2368–2384

Chapter 97
Knowledge Workers' Competence Evaluation Model Based on Extenics

Liang Tang

Abstract In relation to knowledge workers' competence evaluation, knowledge workers' competency evaluation model based on the Extenics was established. First of all, on the basis of literature review, according to certain principles, knowledge workers' competence evaluation index system was built, and then based on Extenics, knowledge workers' competence evaluation method was given, and finally by an applied example, the feasibility and validity of this method was verified.

Keywords Competence • Countermeasure • Extenics • Knowledge workers

97.1 Introduction

At present, the knowledge-based economy is emerging as a new economic form. Knowledge and technology–intensive companies become the most dynamic form of economic organization in economic activity; the knowledge-based economy represents the future direction of economic development. As the main body of companies whose key resource is knowledge, it is particularly important whether the knowledge workers are competent in their jobs, and the key point is to build the competence model in line with the nature of their works and their characteristics. On the one hand, companies evaluate, select and develop knowledge workers and design

---

Supported by Social Science Foundation of Education Department of Heilongjiang Province of China (No. 12522345).

L. Tang (✉)
School of Economics and Management, Qiqihar University, Qiqihar, China
e-mail: neutang@163.com

E. Qi et al. (eds.), *The 19th International Conference on Industrial Engineering and Engineering Management*, DOI 10.1007/978-3-642-37270-4_97,

their career plans according to competence model; On the other hand, knowledge workers identify their own inadequacies according to the competence evaluation results, self-learn and promote themselves.

According to the relationship between competence and performance, Huang Zhepeng and Yang Zongyan (2005) build knowledge workers competence evaluation index system, which includes qualifications, work experiences, project experiences, professional qualifications, technical titles and comprehensive interview, and use the gray relational analysis method to evaluate the knowledge workers competence.

Lin Lijie (2006) obtain university knowledge workers competence through interviews, which includes five dimensions of the meta-capability dimension, knowledge and skills dimension etc., totally 36 elements.

From two components of competence perspective, combining with characteristics of knowledge workers, Zhang Lanxia (2006) proposes knowledge workers competence evaluation index system. The threshold competence includes educational background, knowledge structure, work experiences and job performance. The differentiating competence includes personal qualities, innovation, relationships ability, conceptual ability, and knowledge management capability. The paper uses the method of triangular fuzzy numbers to evaluate the competence of knowledge workers.

Lin Lijie (2007) through literature review and previous studies, put forward the university knowledge workers competence, including families of personality factor family, necessary knowledge family and skills family, a total of 36 elements, and give the evaluation method based on complementary and synergies effects.

Yang Jieqiong (2007) through behavioral event interviews and questionnaires, constructs the competence model of human resources supervisor and knowledge workers in quality improvement sector of the household electrical appliance enterprises.

On the basis of the literature review, Jia Jianfeng (2009) present the overall framework of competency model of knowledge workers, the three-dimensional framework of knowledge workers of competency model, consisting of three main dimensions, namely knowledge and technical ability, job behavior ability and basic behavior ability.

Yu Huaren and Li Ma (2010) argue that knowledge workers competence evaluation index includes knowledge quality, ability quality, personality and motivation, and use analytic hierarchy process and fuzzy comprehensive evaluation method to evaluate knowledge workers competence.

Colomo-Palacios (2010) establish the competence evaluation index system of IT knowledge workers including 20 indicators, such as analysis and synthesis ability, problem solving and decision-making ability.

Extenics with formal tools studies laws and methods of resolving contradictory problems from two perspectives of qualitative and quantitative, and takes into account things, characteristics, and values together to describe appropriately the process of which objective things change. This paper establishes the competence evaluation model of knowledge workers based on Extenics.

97.2 Knowledge Workers Competence Evaluation Index System

In the late 1990s, human resources management theory and practice fields adopt the concept of competence, and promote competence evaluation model to be extensively used. However, many enterprises do not achieve the expected effects. Because the competence evaluation model is not every enterprise common evaluation model, which is closely related to each enterprise's strategic objectives, corporate culture, job nature. Therefore, enterprises need to build and use the competence evaluation model from the point of view of adaptability.

To design knowledge workers competence index system should comply with the following principles:

1. *Pertinent*. Different jobs have different properties, characteristics, and professional technology. The design of the knowledge workers competence takes into account the object and purpose of the evaluation indicators, in addition to the basic quality and management capabilities, as well as decision-making capabilities, job performance, and other special factors.
2. *Scientific*. Evaluation index system should be set up with regard to the scientific principles of psychology, management, and leadership science and based on scientific analysis methods.
3. *Concise*. In theory, more comprehensive the index system is, more precise the evaluation results are. But in fact it is hard to do that and it costs too much. Therefore, the competence evaluation should extract the most representative indexes.
4. *Indexes avoid being overlapped*. If indexes are overlapped mutually, it makes the competence evaluation results are not accurate, and generally lead to the distortion of information.
5. *Indexes should be hierarchical*. Knowledge workers competence evaluation is complex, the single-level index are often not enough. Thus, we usually build a multi-level indexes (two, three, etc.) for the evaluated object.

In accordance with the above principles, we integrate the past literature (Cai Wen 1994; Chen Wei 2006; Lewis 2004; Straub and Karahanna 2008; Kleinman and Vallas 2001; Davenport et al. 2002; Thompson and Heron 2005) to build a knowledge worker competence evaluation index system, which is as Table 97.1.

97.3 Competence Evaluation Method of Knowledge Workers Based on Extenics

The method based on Extenics is as follows:

1. *Basic model*. R_{ij} denotes matter-element model for knowledge workers competence, knowledge workers competence includes three dimensions, represented

Table 97.1 Knowledge worker competence evaluation index

Goal	Dimension	Subdimension
Knowledge worker competence evaluation index	Knowledge and technical ability	Job knowledge technology
		Job process skill
		Industry knowledge
	Basic behavior ability	Motivation
		Personality
		Internal driving force
	Job behavior ability	Learning capacity
		Innovative capacity
		Team cooperation

with C_{ik}. Each dimension has three subdimensions, represented with C_{ik}. Suppose each competence dimension has J levels: $N_{i1}, N_{i2}, \ldots, N_{i\mathrm{J}}$, thus the matter-element

$$R_{ij} = \begin{bmatrix} N_{ij}, & C_{i1}, & V_{ij1} \\ & C_{i2}, & V_{ij2} \\ & \vdots & \vdots \\ & C_{im}, & V_{ijm} \end{bmatrix} = \begin{bmatrix} N_{ij}, & C_{i1}, & < a_{ij1}, b_{ij1} > \\ & C_{i2}, & < a_{ij1}, b_{ij1} > \\ & \vdots & \vdots \\ & C_{im}, & < a_{ijm}, b_{ijm} > \end{bmatrix},$$

$$i = 1, 2, 3, j = 1, 2, 3, \ldots, J,$$

R_{1j} represents the j-level matter-element model of knowledge and skills dimension, N_{1j} represents the j th level knowledge and skills dimension, C_{1k} represents the kth aspect of knowledge and skills dimension. $V_{1jk} =< a_{1jk}, b_{1jk} >$ represents the value range of C_{1k} for the n-level knowledge and skills. V_{ijk} represents the classical field of subdimension C_{ik} for dimension N_{ij}.

Suppose $R_{ip} \supset R_{ij}$, R_{ip} represents all possible samples;

$$R_{ip} = \begin{bmatrix} N_{ij}, & C_{i1}, & V_{ip1} \\ & C_{i2}, & V_{ip2} \\ & \vdots & \vdots \\ & C_{im}, & V_{ipm} \end{bmatrix} = \begin{bmatrix} N_{ij}, & C_{i1}, & < a_{ip1}, b_{ip1} > \\ & C_{i2}, & < a_{ip1}, b_{ip1} > \\ & \vdots & \vdots \\ & C_{im}, & < a_{ipm}, b_{ipm} > \end{bmatrix},$$

$V_{ipk} =< a_{ipk}, b_{ipk} >$, $k = 1, 2, \ldots, m$, represents the segment field of subdimension C_{ik}.

2. *Distance and correlation function*. After the establishment of matter-element model of knowledge worker competency, it is necessary to calculate the proximity between matter-element sample and matter-element model. The proximity $\rho(v_k, V_{ipk})$ is represented by the distance between v_k and the interval V_{ipk}. $\rho(v_k, V_{ipk}) \geq 0$ denotes v_k is not in V_{ipk}, $\rho(v_k, V_{ipk}) < 0$ denotes v_k is in V_{ipk}, and different negative values represent a different position of v_k in V_{ipk}.

$$\rho(v_k, V_{ijk}) = \left| v_k - \frac{a_{ijk} + b_{ijk}}{2} \right| - \frac{1}{2}(b_{ijk} - a_{ijk}), \tag{97.1}$$

$$\rho(v_k, V_{ipk}) = \left| v_k - \frac{a_{ipk} + b_{ipk}}{2} \right| - \frac{1}{2}(b_{ipk} - a_{ipk}), \tag{97.2}$$

The C_{ik}'s correlation function about the j th-level dimension of the matter-element is as follows:

$$K_{ij}(v_k) = \begin{cases} -\dfrac{\rho(v_k, V_{ijk})}{\left|V_{ijk}\right|}, & v_k \in V_{ijk} \\ \dfrac{\rho(v_k, V_{ijk})}{\rho(v_k, V_{ipk}) - \rho(v_k, V_{ijk})}, & v_k \notin V_{ijk}, \rho(v_k, V_{ipk}) \neq \rho(v_k, V_{ijk}), \\ -\rho(v_k, V_{ijk}) - 1\,, & v_k \notin V_{ijk}, \rho(v_k, V_{ipk}) = \rho(v_k, V_{ijk}) \end{cases} \tag{97.3}$$

$K_{ij}(v_k)$ denotes the C_{ik}'s correlation function of matter-element sample about the j th-level dimension. $K_{ij}(v_k) \geq 0$ denotes v_k is belong to V_{ijk},the bigger $K_{ij}(v_k)$ is, the more v_k has V_{ijk}'s properties, and vice versa .

3. *Determining weights*. Each subdimension's contribution to the dimensions is different, thus, it is necessary to determine the subdimension weight with regard to the dimension. It can be determined by experts' survey or analytical hierarchy process.
4. *Determining the level*. The correlation degree about the j th-level of knowledge and skill dimension is as follows:

$$\mathrm{K_{1j}(R)} = \sum_{\mathrm{k=1}}^{\mathrm{m}} \alpha_{\mathrm{1k}} K_{1j}(v_k),\ j = 1, 2, 3, \ldots, J. \tag{97.4}$$

Suppose $K_{1j_1} = \max\limits_{1 \leq j \leq 3} K_{1j}, \quad K_{2j_1} = \max\limits_{1 \leq j \leq 3} K_{2j}, \quad K_{3j_1} = \max\limits_{1 \leq j \leq 3} K_{3j},$

$$\bar{K}_{ij}(R) = \frac{K_{ij}(R) - \min\limits_{j} K_{ij}(R)}{\max\limits_{j} K_{ij}(R) - \min\limits_{j} K_{ij}(R)} \tag{97.5}$$

$$j_i{}^* = \frac{\sum\limits_{j=1}^{J} j \times \bar{K}_{ij}(R)}{\bar{K}_{ij}(R)}. \tag{97.6}$$

$j_i{}^*$ denotes the characteristic value of level variable for the ith dimension, and reflects the extent that how each dimension is close to the j th level.

5. *Levels of comprehensive evaluation values*. Sort comprehensive evaluation values by the level if they are not in the same level, while in the same level, sort comprehensive evaluation values according to the characteristic values.

97.4 Numerical Example

Now, evaluate the competence of three candidates of knowledge worker in a software company, and choose the best knowledge worker as a supervisor, the steps are as follows:

1. *Building models*. Each dimension in competence evaluation has three levels N_{i1} (strong), N_{i2} (general) and N_{i3} (weak), then the matter element model is:

$$R_{ij} = \begin{bmatrix} N_{ij}, & C_{i1}, & V_{ij1} \\ & C_{i2}, & V_{ij2} \\ & \vdots & \vdots \\ & C_{im}, & V_{ijm} \end{bmatrix} = \begin{bmatrix} N_{ij}, & C_{i1}, & < a_{ij1}, b_{ij1} > \\ & C_{i2}, & < a_{ij1}, b_{ij1} > \\ & \vdots & \vdots \\ & C_{im}, & < a_{ijm}, b_{ijm} > \end{bmatrix},$$
$$i = 1, 2, 3, \; j = 1, 2, 3.$$

i represents the i th dimension of competence, j indicates that the dimension is the j th level. When to determine classical domain and sectional domain of the index in the competence evaluation, if the sample fully meets the requirements of an index, the index value is 10 and completely not satisfies with the value 0. In theory, 10 is the ideal value, as long as the value is close to 10, it can be considered "strong" level, 8 and 6 is "general" and "weak" level, respectively. Based on the actual situation of the enterprise, combined with the literature and the views of experts group, we get the matter element model's the classical field and sectional domain for knowledge workers competence:

$$R_{11} = \begin{bmatrix} N_{11}, & C_{11}, & < 8.5, 10 > \\ & C_{12}, & < 9.0, 10 > \\ & C_{13}, & < 8.0, 10 > \end{bmatrix}, \quad R_{12} = \begin{bmatrix} N_{11}, & C_{11}, & < 6.5, 8.5 > \\ & C_{12}, & < 7.0, 9.0 > \\ & C_{13}, & < 6.0, 8.0 > \end{bmatrix},$$

$$R_{13} = \begin{bmatrix} N_{11}, & C_{11}, & < 0, 6.5 > \\ & C_{12}, & < 0, 7.0 > \\ & C_{13}, & < 0, 6.0 > \end{bmatrix}, \quad R_{21} = \begin{bmatrix} N_{11}, & C_{11}, & < 9.0, 10 > \\ & C_{12}, & < 8.5, 10 > \\ & C_{13}, & < 9.0, 10 > \end{bmatrix},$$

$$R_{22} = \begin{bmatrix} N_{11}, & C_{11}, & < 7.0, 9.0 > \\ & C_{12}, & < 6.5, 8.5 > \\ & C_{13}, & < 7.0, 9.0 > \end{bmatrix}, \quad R_{23} = \begin{bmatrix} N_{11}, & C_{11}, & < 0, 7.0 > \\ & C_{12}, & < 0, 6.5 > \\ & C_{13}, & < 0, 7.0 > \end{bmatrix},$$

$$R_{31} = \begin{bmatrix} N_{11}, & C_{11}, & < 8.5, 10 > \\ & C_{12}, & < 8.5, 10 > \\ & C_{13}, & < 9.0, 10 > \end{bmatrix}, \quad R_{32} = \begin{bmatrix} N_{11}, & C_{11}, & < 6.5, 8.5 > \\ & C_{12}, & < 6.5, 8.5 > \\ & C_{13}, & < 7.0, 9.0 > \end{bmatrix},$$

$$R_{33} = \begin{bmatrix} N_{11}, & C_{11}, & < 0, 6.5 > \\ & C_{12}, & < 0, 6.5 > \\ & C_{13}, & < 0, 7.0 > \end{bmatrix}, \quad R_{1p} = \begin{bmatrix} N_{1p}, & C_{11}, & < 0, 10 > \\ & C_{12}, & < 0, 10 > \\ & C_{13}, & < 0, 10 > \end{bmatrix},$$

$$R_{2p} = \begin{bmatrix} N_{2p}, & C_{11}, & <0,10> \\ & C_{12}, & <0,10> \\ & C_{13}, & <0,10> \end{bmatrix}, \quad R_{3p} = \begin{bmatrix} N_{3p}, & C_{11}, & <0,10> \\ & C_{12}, & <0,10> \\ & C_{13}, & <0,10> \end{bmatrix}.$$

2. *Determining the weights*. Due to the different influence of every subdimension on the dimensions, it is necessary to determine the subdimension weights. By analysis hierarchy process, each subdimension weights are obtained (the calculation process is omitted).

$$\alpha_1 = [0.3, 0.3, 0.4], \alpha_2 = [0.4, 0.3, 0.3], \alpha_3 = [0.3, 0.4, 0.3].$$

3. *Dimensions' levels*. Experts group give a mark to the three candidates, the three candidates of matter-element model for knowledge workers are as follows:

$$R_1 = \begin{bmatrix} N_1, & C_{11}, & 9.52 \\ & C_{12}, & 8.60 \\ & C_{13}, & 7.23 \end{bmatrix}, \quad R_2 = \begin{bmatrix} N_2, & C_{11}, & 8.82 \\ & C_{12}, & 9.10 \\ & C_{13}, & 8.40 \end{bmatrix},$$

$$R_3 = \begin{bmatrix} N_3, & C_{11}, & 8.20 \\ & C_{12}, & 5.73 \\ & C_{13}, & 7.15 \end{bmatrix}, \quad R'_1 = \begin{bmatrix} N'_1, & C_{11}, & 8.58 \\ & C_{12}, & 7.40 \\ & C_{13}, & 6.45 \end{bmatrix},$$

$$R'_2 = \begin{bmatrix} N'_2, & C_{11}, & 7.02 \\ & C_{12}, & 8.56 \\ & C_{13}, & 7.88 \end{bmatrix}, \quad R'_3 = \begin{bmatrix} N'_3, & C_{11}, & 7.10 \\ & C_{12}, & 8.22 \\ & C_{13}, & 8.38 \end{bmatrix},$$

$$R''_1 = \begin{bmatrix} N''_1, & C_{11}, & 7.54 \\ & C_{12}, & 7.60 \\ & C_{13}, & 8.47 \end{bmatrix}, \quad R''_2 = \begin{bmatrix} N''_2, & C_{11}, & 8.24 \\ & C_{12}, & 8.30 \\ & C_{13}, & 8.66 \end{bmatrix},$$

$$R''_3 = \begin{bmatrix} N''_3, & C_{11}, & 8.64 \\ & C_{12}, & 7.46 \\ & C_{13}, & 9.37 \end{bmatrix}.$$

According to the formula (97.1), (97.2) and (97.3), calculate the candidate A's correlation degree of the first subdimension of the first dimension with "strong" level:

$$\because 9.52 \in <8.5, 10>,$$

$$\therefore \rho(9.52, V_{111}) = \left| 9.52 - \frac{8.5 + 10}{2} \right| - \frac{10 - 8.5}{2} = -0.48.$$

Then $K_{11}(v_1) = -\frac{\rho(9.52, V_{111})}{|V_{111}|} = \frac{0.48}{1.5} = 0.32$, the rest may be deduced by analogy, Candidate A, B and C's correlation matrix are as follows:

$$K_1 = \begin{bmatrix} 0.32 & -0.68 & -0.86 \\ -0.22 & 0.20 & -0.53 \\ -0.21 & 0.39 & -0.31 \end{bmatrix}, \quad K_2 = \begin{bmatrix} -0.13 & 0.09 & -0.61 \\ 0.40 & -0.40 & -0.74 \\ -0.27 & 0.30 & -0.47 \end{bmatrix},$$

$$K_3 = \begin{bmatrix} -0.14 & 0.15 & -0.49 \\ -0.39 & -0.15 & 0.12 \\ -0.39 & 0.08 & -0.05 \end{bmatrix}, \quad K'_1 = \begin{bmatrix} 0.05 & -0.05 & -0.59 \\ -0.38 & 0.20 & -0.13 \\ -0.30 & 0.23 & -0.11 \end{bmatrix},$$

$$K'_2 = \begin{bmatrix} -0.04 & 0.01 & -0.01 \\ 0.04 & -0.04 & -0.59 \\ -0.35 & 0.44 & -0.29 \end{bmatrix}, \quad K'_3 = \begin{bmatrix} -0.33 & 0.30 & -0.17 \\ -0.14 & 0.14 & -0.49 \\ -0.28 & 0.31 & -0.46 \end{bmatrix},$$

$$K''_1 = \begin{bmatrix} -0.28 & 0.48 & -0.30 \\ -0.37 & 0.30 & -0.20 \\ 0.24 & -0.24 & -0.62 \end{bmatrix}, \quad K''_2 = \begin{bmatrix} -0.30 & 0.38 & -0.41 \\ -0.11 & 0.10 & -0.51 \\ -0.20 & 0.17 & -0.55 \end{bmatrix},$$

$$K''_3 = \begin{bmatrix} 0.09 & -0.09 & -0.61 \\ -0.29 & 0.48 & -0.27 \\ 0.37 & -0.37 & -0.79 \end{bmatrix}.$$

According to the formula (97.4), calculate the candidate A's correlation degree of knowledge and technical ability dimension with "strong" level:

$$K_1(R) = \sum_{k=1}^{3} a_{1k} K_1(v_k) = -0.055,$$

Thus, Candidate A, B and C's correlation matrix are as follows:

$$K = \begin{bmatrix} -0.054 & 0.012 & -0.541 \\ -0.013 & 0.006 & -0.607 \\ -0.315 & 0.009 & -0.114 \end{bmatrix}, \quad K' = \begin{bmatrix} -0.219 & 0.137 & -0.260 \\ -0.109 & 0.124 & -0.268 \\ -0.239 & 0.239 & -0.385 \end{bmatrix},$$

$$K'' = \begin{bmatrix} -0.099 & 0.138 & -0.398 \\ -0.213 & 0.233 & -0.482 \\ 0.022 & 0.054 & -0.528 \end{bmatrix}.$$

So, $\max\limits_{1 \le j \le 3} K_{1j} = 0.012$, $\max\limits_{1 \le j \le 3} K_{2j} = 0.006$, $\max\limits_{1 \le j \le 3} K_{3j} = 0.009$, $\max\limits_{1 \le j \le 3} K'_{1j} = 0.137$, $\max\limits_{1 \le j \le 3} K'_{2j} = 0.124$, $\max\limits_{1 \le j \le 3} K'_{3j} = 0.239$ $\max\limits_{1 \le j \le 3} K''_{1j} = 0.138$, $\max\limits_{1 \le j \le 3} K''_{2j} = 0.233$, $\max\limits_{1 \le j \le 3} K''_{3j} = 0.054$.

According to the formula (97.5) and (97.6), calculate the level characteristics variables values:

$$\bar{K}_{11}(R) = \frac{-0.054 + 0.541}{0.0120 + 0.541} = 0.88,$$
$$\bar{K}_{12}(R) = 1, \bar{K}_{13}(R) = 0, \bar{K}_{21}(R) = 0.82,$$
$$\bar{K}_{22}(R) = 1, \bar{K}_{23}(R) = 0, \bar{K}_{31}(R) = 0, \bar{K}_{32}(R) = 1, \bar{K}_{33}(R) = 0.62;$$

$$j_1{}^* = \frac{\sum_{j=1}^{3} j \times \bar{K}_{1j}(R)}{\sum_{j=1}^{3} \bar{K}_{1j}(R)} = \frac{1 \times 0.88 + 2 \times 1}{1 + 0.88} = 1.53, j_2{}^* = 1.55, j_3{}^* = 2.38.$$

Candidate B and C's level characteristics variables values are as follows:

$$j_1'^* = 1.93, j_2'^* = 1.71, j_3'^* = 1.81,$$
$$j_1''^* = 1.64, j_2''^* = 1.72, j_3''^* = 1.51,$$

In regard to knowledge and skill dimension, B $\succ$ C $\succ$ A; basic job capability, C $\succ$ B $\succ$ A and job behavior capability, A $\succ$ B $\succ$ C, respectively. If you are using this issue as a group decision making, the ranking is B $\succ$ C $\succ$ A through Borda rule.

97.5 Conclusion

On the basis of the prior study, the study builds the competence evaluation index system of knowledge workers, applies extension theory to evaluate competence of knowledge workers, and the sample shows the feasibility and validity of the method. How to build a more rational and comprehensive knowledge worker competency evaluation index system and evaluation method need the further study.

Acknowledgment I would like to thank my supervisor for his good advice and thank my classmates too.

References

Cai Wen (1994) Matter element model and its application. Science and Technology Literature Press, Beijing (in Chinese)

Chen Wei (2006) Factors affecting assessment based on dependent function. J Harbin Inst Technol 38(7):1143–1145 (in Chinese)

Colomo-Palacios R (2010) Generic competences for the IT knowledge workers: a study from the field. Commun Comput Inf Sci 11(111):1–7

Davenport TH, Thomas RJ, Cantrell S (2002) The mysterious art and science of knowledge worker performance. MIT Sloan Manage Rev 44(1):23–30

Huang Zhepeng, Yang Zongyan (2005) Grey relational analysis of the knowledge worker competencies. Sci Technol Ind 5(2):25–28 (in Chinese)

Jia Jianfeng (2009) A framework for research on competency model of knowledge workers. Sci Sci Manage S&T 30(8):166–171 (in Chinese)

Kleinman DL, Vallas SP (2001) Science, capitalism, and the rise of the "knowledge worker", the changing structure of knowledge production in the United States. Theory Soc 30(4):451–492

Lewis K (2004) Knowledge and performance in knowledge worker teams: a longitudinal study of transactive memory systems. Manage Sci 50(11):1519–1533

Lin Lijie (2006) Study on competence element of college knowledge workers. Mod Univ Educ 23(2):5–11 (in Chinese)

Lin Lijie (2007) Empirical study on the relationship between knowledge worker competence elements and his/her individual performances in colleges and universities. Chin J Manage 4(2):230–234 (in Chinese)

Straub D, Karahanna E (2008) Knowledge worker communications and recipient availability: toward a task closure explanation of media choice. Org Sci 9(2):160–175

Thompson M, Heron P (2005) The difference a manager can make: organizational justice and knowledge worker commitment. Int J Hum Resour Manage 16(3):383–404

Yang Jieqiong (2007) Research on competency of knowledge workers in manufacturing enterprises of household appliances. Hefei University of Technology, Hefei (in Chinese)

Yu Huaren, Li Ma (2010) Research on knowledge worker competency evaluation based on AHP. Economist 1(12):225–226 (in Chinese)

Zhang Lanxia (2006) Evaluation indexes and methodology for competence of knowledge worker. J Northeast Univ (Nat Sci) 1(8):941–944 (in Chinese)

Chapter 98
Knowledge Know-How Management Model of High-Tech Enterprises – Dual Perspectives of Knowledge Management and Strategic Human Resource Management

Jing-yu Chen and Zu-guang Hu

Abstract The achievement of competitive advantage in high-tech enterprise depended on the development and management of technological know-how and market pioneering know-how heavily. Based on the perspectives of knowledge management and strategic human resource management, a model on how to manage the two kinds of key know-how effectively was set up to develop and utilize competitive advantage in high-tech enterprise.

Keywords High-tech enterprises • Rechnological know-how • Market pioneering know-how • Knowledge management • Strategic human resource management

98.1 Introduction

In high-tech market where there is plenty of confusion, enterprises' ability to grasp new opportunities or prevent from lagging behind with other competitors depends on their ability to relocate resources prior to opponents, and the latter relies on their ability on research and development as well as marketing (Daneels 2002). These kinds of knowledge themselves itself or these utilization are crucial to the survival as well as development of high-tech enterprises. And they are also the origin of competitive advantage of enterprises. Besides, the effective development and management of these two kinds of knowledge depends on enterprises' effective of knowledge management activities and strategic human resources activities, but the existing literatures often ignore this aspect (Zhou Fu-zhan and Chen Shu-wen 2011). Based on dual perspectives of knowledge management as well as strategic human resource management, our work constructed a frame model to research and

J. Chen (✉) • Z. Hu
School of Business Administration, Zhe Jiang Gong Shang University, Hang Zhou, China
e-mail: cjywho@126.com

E. Qi et al. (eds.), *The 19th International Conference on Industrial Engineering and Engineering Management*, DOI 10.1007/978-3-642-37270-4_98,

explain how knowledge effectively manages technological know-how and market pioneering know-how with the support of strategic human resource and expounds the emphasis points of knowledge management and strategic human resource management in high-tech enterprises, as well as providing new ideas for theoretical research in the future.

98.2 Theoretical Commentary

98.2.1 Source of Competitive Advantage in High-Tech Enterprises – Synergy of Two Kinds of Knowledge Know-How

Knowledge-based view of the firm considers that knowledge is the most crucial competitive asset owned by business organization. Compared with market, the advantage of business organization lies in its stronger capacity in creating as well as maintaining knowledge (Grant 1996). Knowledge is the most important strategic resource and the basic principle of its operation as well as competition is generation and combination-recombination and development of knowledge (Conner and Prahalad 1996). Faced with more complex and dynamic marketing environment, high-tech enterprises, in order to keep competitive advantage, need not only to have technical innovations but also their potential of technical innovation must be released and utilized in market only by rapidly understanding as well as satisfying new requirements of customers. Therefore, this paper has a specific consideration in two kinds of know-how in high-tech enterprise knowledge, that is, technological knowledge know-how and market pioneering knowledge know-how and technological knowledge know-how produces knowledge and capacity of new scientific discovery and new breakthrough. Moreover, market pioneering know-how indicates whether enterprises can commercialize theirs technological innovation prior to opponents. Knowledge and capacity of marketing pioneering are especially important in high-tech product market with short life cycle. These two kinds of knowledge ability form the foundation of high-tech enterprises innovation and marketing-based differentiation. And two kinds of knowledge are mutually complementary, in which enterprises' research productivity can have returns and really build dynamic competitive advantage of enterprises only by technological innovation translates into killer application. For example, the success of Apple Inc. is realized synergically by excellent product design as well as excellent marketing ability.

98.2.2 Development and Utilization of Knowledge Management Strategy and Competitive Advantage

In reality, although some enterprises which own strong ability of knowledge creation and accumulation can understand various complicated technological problems, however, they cannot translate their various knowledge into enterprises' competitive advantage (Baker et al. 2003). Which manifests that enterprises need to distinguish to the creation as well as accumulation ability of knowledge and application ability of knowledge. It is a very crucial challenge for enterprises to reconcile dichotomy between knowledge creation and knowledge application activity. Besides, it is very difficult (Daneels 2002). If enterprises cannot solve the problem of complicated interdependence between the two kinds of knowledge of technological know-how and market pioneering know-how, rich knowledge may lead to opposite results. If enterprises emphasize one knowledge only, then their knowledges cannot be utilized fully, which not only means the value of enterprises' research and development as well as business process is not fully occupied, but signifies the increase of competition (Rajshree et al. 2004). High-tech enterprises must adopt appropriate knowledge management strategies if they want to reconcile the dichotomy and effectively manage the two kinds of knowledge, in order to develop and utilize their competitive advantage indeed.

Knowledge management strategy is a directional selection on basic framework, process and emphasis point of knowledge management activity which is made by enterprises (Zack 2002). Faced with industrial structure, scientific and technological progress and the rapid changes in consumer behaviour, high-tech enterprises must do exploration, aiming at keep their own position in competition. Enterprises must do effective exploitation in order to consolidate and keep advantage when their knowledge resource and ability have an obvious advantage over opponents (Zack 1999). And for different situations, high-tech enterprises can implement changes and obtain as well as keep competitive advantage by utilizing three basic knowledge management strategies. These knowledge management strategies are knowledge innovation, knowledge transfer and knowledge protection (Bloodgood and Salisbury 2001). Enterprises adopting knowledge innovation strategy focus on innovating, applying and devoting themselves to creating a kind of consensus so as to develop new knowledge of new products and service within innovation group; organizations adopting knowledge transfer strategy focus on rapidly transmit knowledge within organization, in order to take full use of these knowledge as far as possible; those adopting knowledge protection strategy focus on protecting the original as well as positive status of knowledge and preventing its transfer to other organization without authority. When knowledge gap occurs, enterprises can organize staff to create and transmit new knowledge within organization by internal learning. Besides, enterprises can obtain as well as simulate from the outside by external learning, so as to transmit knowledge into interior of organization (Bierly and Chakrabarti 1996). In order to coordinate the development and utilization of technological and market pioneering know-how, high-tech enterprises can choose effective knowledge management strategies combination in terms of different conditions.

98.2.3 Strategic Human Resource Management and Knowledge Know-How Management

The complicated and key parts of technological and market knowledge know-how are often owned by staff themselves in forms of tacit knowledge and competent capital, so the knowledge capital owned by enterprises are more shown as human capital. Human capital is strategic resource owned by enterprises and the ability originated from human capital enabling enterprises to make full use of strategic opportunities to create values (Lee and Miller 1999). The fixed position of Strategic Human Resource Management (SHRM) is to provide support for the function and role of human resource of enterprise strategies. Essentially, SHRM is to utilize human resource values of enterprises through sufficient development of human resource management application, thus to obtain competitive advantage (Laursen and Mahnke 2001). In essence, behaviors like knowledge creation, management and transfer are closely related to human resource management application (Wright et al. 2001). If strategies of high-tech enterprises with abundant knowledge only emphasize one aspect of technological know-how or market know-how, the result will be that core workers find that enterprises ignore knowledge values at the aspects of creating or utilizing, making these staff have frustrations and resulting in a larger and larger difference between staff ambition and enterprise prospects. This target difference will reduce the occupation satisfaction degree and enlarge liquidity. Furthermore, the uncommercialized advanced technology and unutilized major market opportunities will increase workers' confidence in initiating enterprises and increase the probabilities for staff to do pioneering work (Rajshree et al. 2004). On this condition, not only will enterprises strategic human resource drain be caused, but unnecessary market competition will increase. In view of these conditions, enterprises can organically combine human resource management applications like recruitment, training, selecting, organizational design as well as working design, staff involvement, evaluation, as well as remuneration and stock option incentive, corporate culture, internal venture platform construction, etc. through effective strategic human resource management and conform with enterprises' strategies to emphasize the synergic development of the two kinds of knowledge know-how, thus to create a social climate appropriate for creation and transfer of key knowledge in enterprises (Collins and Snith 2006) and this climate can support knowledge management strategy of organization effectively.

98.3 Development and Utilization Model of Knowledge Know-How

98.3.1 Construction of Model

In order to keep continuous competitiveness, high-tech enterprises must effectively develop and synergically utilize their knowledge know-how at the two aspects of technology and market pioneering, which requires enterprises to adopt appropriate knowledge management strategy and matching human resource strategies. In terms of the specific requirement to high-tech enterprises, this paper builded a theoretical model of dual perspectives based on knowledge management and strategic human resource management, as is shown in Fig. 98.1. Faced with different stock status of knowledge know-how, high-tech enterprises needs to choose different strategies to obtain and manage knowledge. Moreover, knowledge management strategy and process must be effectively combined with human resource strategy and process, through organically integration of each practice module of human resource, giving effective Consilience and support to knowledge management strategy, thus to form effective development, utilization, combination and protection of technological know-how and market pioneering know-how. Thus enterprises can achieve the competitive advantages.

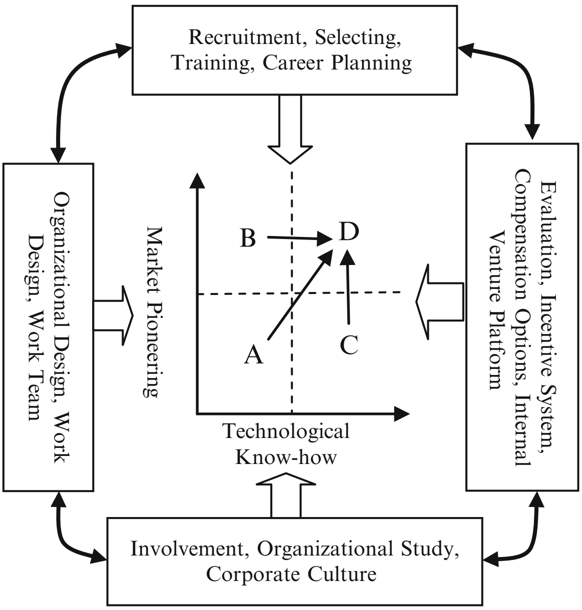


Fig. 98.1 Frame model of knowledge know-how in high-tech enterprises

98.3.2 Knowledge Management Strategy with Different Knowledge Inventory

As is shown in Fig. 98.1, high-tech enterprises face four different kinds of knowledge inventory and competitive status which are A, knowledge deficiency, competitive deficiency; B, technological knowledge gap due to failure to utilize market advantage; C: market knowledge gap due to failure to utilize technological advantage; D, abundant knowledge and strong competitiveness.

In status A, enterprises are both deficient in technology and marketing knowledge, so they should adopt knowledge acquisition strategy mainly by external learning and acquire knowledge from clients, consumers, suppliers and opponents by various ways; enterprises should confirm knowledge management strategy mainly by knowledge creation and widely encourage to develop knowledge about new products and services which are used to satisfy markets. In status B, enterprises have good marketing insight and marketing ability, but they lack enough technology breakthrough to realize them, so knowledge acquisition strategy combining internal learning with external learning should be adopted in order to improve enterprises technical capacity and technical knowledge reserves; knowledge management strategy which combines knowledge utilization and knowledge protection and knowledge alliance are used by enterprises to fully develop relevant technology to realize the effective utilization of market capacity. At the same time, enterprises should try to avoid the spillover of market knowledge. In status C, enterprises master technological breakthrough, but there are not enough market knowledge to make it commercialized. Enterprises should adopt knowledge acquisition strategy which combines external learning with internal learning to collect market information from external and cultivate tacit knowledge of market within enterprises; knowledge utilizing strategy of market orientation should be built and communication between Proscenium and Backstage should be strengthened, making knowledge transfer among each department. In status D, enterprises are in a status with abundant knowledge and high competitiveness, in which condition enterprises should mainly take internal learning and keep knowledge vitality and mainly take knowledge protection strategy to prevent knowledge from transferring to other organizations without authority.

98.3.3 The Supportive Function of Strategic Human Resource Management

According to the view of human resource advantage proposed by Boxall (1996), high-tech enterprises adopt strategic human resource management to do development and utilization to its knowledge know-how. And the supportive functions of Human Resource Management are represented as the following two aspects. First, outstanding human resource are used to increase enterprise knowledge inventory;

second, a series of environment and process for learning, cooperation and innovation are provided to enhance the efficiency of enterprise knowledge transfer, utilization and protection. To be specific, what enterprises need to do are as follows: (1) Enterprises must find knowledge gap according to knowledge stock status, carry on the strategic human resource planning, and select as well as train by recruitment to have rapid learning from external and effectively absorb the knowledge. (2) Enterprises can build high-level and specialized staff and have different employment relationship and management means according to different uniqueness and strategic value of human resources (Lepak and Snell 1999). (3) Enterprises can effectively realize the development and utilization of staff Specificity knowledge and ability by effective career planning and career development path and they effectively combine staff Specificity knowledge and ability with enterprises' strategies through proper evaluation and remuneration system. (4) Enterprises need to adjust organizational design based on customer value, adjust job design based on knowledge management, strengthen communication and transfer of tacit knowledge based on job team for improving knowledge spread and utilization efficiency. (5) Enterprises can widely adopt staff involvement in decision-making and management to improve knowledge process and promote transmission and sharing of knowledge, especially marketing knowledge. (6) Enterprises should change the incentive structures of internal knowledge market by a series of human resource practice like effective selection, vision incentive, organizational design, a check-up system and remuneration as well as option incentive, thus to build cooperate culture which is in pursuit of excellence and encouraging knowledge sharing, strengthen organizational learning and improve knowledge transfer and utilization efficiency. (7) Enterprises need to distinguish the creation, employee stock ownership plan and i Internal Venture Platform to retain staff, which not only avoid core staff and key knowledge drain, but coordinate the difference between the personal goal of core staff and organizational goal, thus to strengthen the development and utilization of key knowledge know-how.

In summary, human resource management's supportive function on knowledge know-how management is by creating a kind of environment and a condition to stimulate the continuous emerging of relevant technological as well as market core knowledge know-how. Besides, it also encourages and promotes communication, spread and transfer of these knowledge. Strategic human resource management not only contributes to creating and obtaining these key knowledge know-how, but translates the knowledge know-how into organizational routines, thus to form competitive advantage of high-tech enterprises indeed.

98.4 Conclusion

Competitive power of high-tech enterprises is not only from the key technological know-how as well as market pioneering know-how, but from the utilization of these two kinds of know-how. These two kinds of knowledge must play a role concurrently and competitive advantage cannot really be obtained by emphasizing one kind of knowledge only. In order to effectively develop and utilize these two kinds of knowledge, this paper tried to develop a frame model of high-tech enterprise know-how management which is based on dual perspectives of knowledge management and strategic human resource management. This model considers that enterprises need to adopt relevant knowledge management strategy to coordinate dichotomy between knowledge creation and knowledge utilization. In addition, under different knowledge status, it has an emphasized focus on three different knowledge management strategies, which are knowledge creation, knowledge utilization and knowledge protection. And each human resource management practice module needs to be effectively integrated according to knowledge management strategy. Besides, this model summarizes the strategic human resource management's supportive function on key knowledge management of enterprises. Furthermore, it provides a new angle and view to key knowledge know-how of high-tech enterprise management and a new way for future theoretical research is also provided.

References

Agarwal R, Echambadi R, Franco A, Sarkar MB (2004) Knowledge transfer through inheritance: spin-out generation, development, and survival. J Acad Manage 47:501–522

Baker T, Miner AS, Eesley DA (2003) Improvising firms: bricolage, account giving and improvisational competencies in the founding process. Res Policy 32(2):255–276

Bierly P, Chakrabarti A (1996) Generic knowledge strategies in the U. S. pharmaceutical industry. Strateg Manage J 17(Winter Special Issue):123–135

Bloodgood JM, Salisbury WD (2001) Understanding the influence of organizational change strategies on information technology and knowledge management strategies. J Decis Support Syst 31:55–69

Boxall PF (1996) The strategic HRM debate and the resource-based view of the firm. Hum Res Manage J 6:59–76

Collins CJ, Snith KG (2006) Knowledge exchange and combination: the role of human resource practices in the performance of high-technology firms. Acad Manage J 49(3):544–560

Conner R, Prahalad RK (1996) A resource-based theory of the firm: knowledge versus opportunism. J Org Sci 7:477–501

Daneels E (2002) The dynamics of product innovation and firm competences. Strateg Manage 23:1095–1121

Grant RM (1996) Toward a knowledge-based theory of the firm. Strateg Manage J 17(Winter Special Issue):109–122

Laursen K, Mahnke V (2001) Knowledge strategies, firm types, and complementarity in human-resource practices. J Manage Gov 5(1):1–27

Lee J, Miller D (1999) People matter: commitment to employees, strategy and performance in Korean firms. Strateg Manage J 20(6):579–593

Lepak DP, Snell SA (1999) The human resource architecture: toward a theory of human capital allocation and development. Acad Manage Rev 24:31–48

Wright PM, Dunford BB, Snell SA (2001) Human resources and the resource based view of the firm. J Manage 27(6):701–721

Zack MH (1999) Developing a knowledge strategy. Calif Manage Rev 41(3):125–145

Zack M (2002) Epilogue: developing a knowledge strategy. In: Choo CW, Bontis N (eds) The strategic management of intellectual capital and organizational knowledge. Oxford University Press, New York, pp 268–276

Zhou Fu-zhan, Chen Shu-wen (2011) Knowledge-based study of human resource development model in high-tech enterprises. J Bus Econ 1(12):38–44, Chinese

Chapter 99
Ontology Technology-Supported English Teaching Scheme for TESOL

Xin-xing Gong and Ming-yang Zhong

Abstract In research of TESOL, the most important issue is how to determine the course content scientifically and how to solve students' problems accurately, and the existing researches have not provided effective solutions yet, therefore the Ontology Technology-supported English Teaching Scheme has been proposed, in which ontology-based problem detection and resource recommendation method has been used to fulfill this goal. Finally, the results of the application experiment have verified the feasibility and superiority of the proposed scheme.

Keywords English teaching scheme • Ontology • Semantics • TESOL

99.1 Introduction

TESOL which means Teaching English to Speakers of Other language emphasizes English teaching and learning rather than English itself, and its purpose is to let every learner to enjoy excellent learning experience. With the internationalization of English, TESOL has gained more and more attention. Many scholars have devoted their energy to the research of fundamental teaching patterns, while others have concentrated on the research of evaluation system or the formation of teaching (May 2011; Canagarajah 2006; Richard 2005). As the research continues, we have found that the essence of TESOL – the course content which helps teachers to convey their knowledge and correct students' existing problems, has always been neglected.

X. Gong (✉)
Application of Foreign Language Institute, Sichuan International Studies University, Chongqing, China
e-mail: 274807397@qq.com

M. Zhong
College of Software, Chongqing University, Chongqing, China
e-mail: my.zhong@hotmail.com

E. Qi et al. (eds.), *The 19th International Conference on Industrial Engineering and Engineering Management*, DOI 10.1007/978-3-642-37270-4_99,

Therefore, we have decided to use the advanced ontology technology (Batet et al. 2011; Jin et al. 2009; Van Rijsbergen 1998; Qiu and Frei 1993) in order to make the organization of course content more scientific and effective, then the ontology-based problem detection and resource recommendation method has been proposed, and in order to guarantee the teaching and learning quality, a full-scale scheme is required, we have enhanced and fused many traditional methods and, finally, formed the Ontology Technology-supported English Teaching Scheme (OTETS).

99.2 Ontology Technology-Supported English Teaching Scheme

OTETS has considered the characteristic of TESOL objects – speakers of other languages, and includes five levels of English teaching: (1) Teaching Content, (2) Teaching Pattern, (3) Course Content, (4) Homework and Quiz, (5) Extracurricular Atmosphere.

99.2.1 Teaching Content (Ortega 2009; Pennycook 1999)

TESOL has classified language teaching into several lesson types as Fig. 99.1:

Considering receptive skills, for example, we have found that skills, such as listening and reading, have similar processes with which include in pre-listening/reading, whilst listening/reading and post-listening/reading, hence corresponding training methods can be carried out.

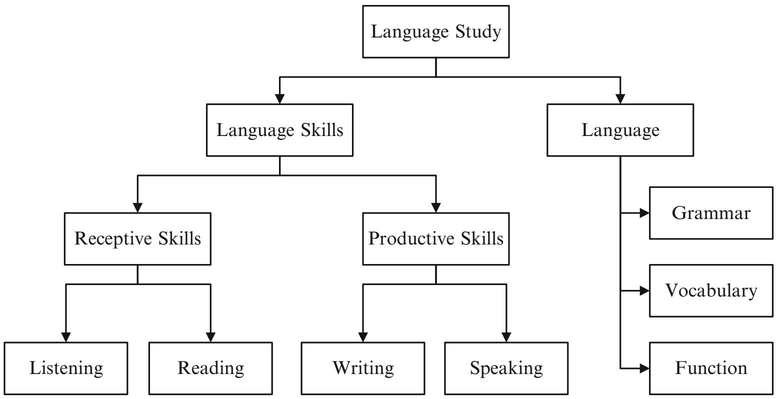


Fig. 99.1 Classification of lesson types

99.2.2 Teaching Pattern

Because of the current situation of Chinese education, most of the students have not formed self-consciousness and initiative learning habits yet, moderate supervisions are required in order to regulate and oversee students' learning process, then we have integrated interaction, supervision and initiative into 'Interaction and Supervised-Initiative Teaching' (ISIT).

1. Interaction includes the blowing three types (Berg 2011; May 2007; Hyland 2006):

   - Teacher-to-Student
     'Teachers teach, students learn' is the rudimentary rule of education, but in traditional Chinese concept, as teachers have been regarded as seniors, students have to listen and do what they have been told to, as a consequence, students start to learn things passively. In order to change such an issue, teachers are informed to be more amiable for the sake of being regarded as supporters and cooperators.
   - Teacher-to-Teacher
     With the development of education industry, much attention has been paid to the importance of the interaction between teachers and students. This kind of interaction can not only improve the fusion of distinctive teaching methodologies, but also enhance different teachers' understanding to students' performances
   - Student-to-Student
     This kind of interaction has always been neglected in traditional Chinese education, which severely hinders students' language communication abilities in different environments, and even impedes the development of other meaningful skills, like teamwork.

2. Supervised-Initiative
   Supervised-Initiative means to improve students' initiative under teachers' or teaching assistants' supervision. As most of the students may not know their strengths and weaknesses clearly, and the schedules they made may not be effectively, appropriate guidance and supervision are needed to help students to learn efficiently and initiatively.

99.2.3 Course Content

Course content, basing on ontology technology, is the most significant part of OTETS, it divides the whole teaching and learning process into many stages which can be determined by teachers, and in the end of every stage, we have proposed the ontology-based problem detection and resource recommendation method which

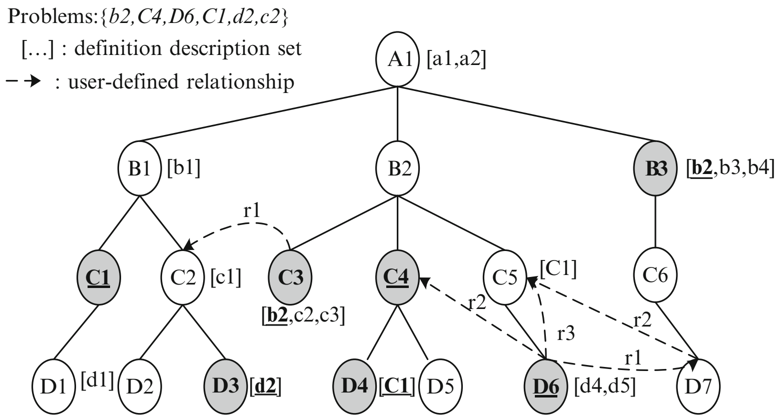


Fig. 99.2 Problem mapping

contains four steps to detect students' unique problems which may not only due to one separate concept, but a serial of related ones, and then recommend the most proper teaching resources to teachers.

***Step 1*. Problem extraction**

According to the results which come from the quiz (see next section), the problems will be extracted from the wrong answers, then synonymous expansion will be implemented to get the expanded problem set – $Q = \{q_1, q_2 \ldots q_i\}$. Similarly, synonymous expansion will be carried out on the concepts in domain ontology for sake of the highest mapping probability, the expanded concepts has been described as $K = \{k_{1c1}, k_{2c1} \ldots k_{ic1}, k_{1c2}, k_{2c2} \ldots k_{jc2}, \ldots, k_{1cx} \ldots k_{mcx}\} = \{K_{c1}, K_{c2} \ldots K_{cx}\}$.

Definition 99.1 Basic concept: $\exists\, c_i \in C$ and $K_{ci} \cap Q \notin \varphi$, then c_i is a basic concept.

If one concept and its synonymous description exists in Q, or result calculated by the Hownet similarity calculation is bigger than the pre-set threshold value, then the concept is called basic concept. Therefore, Q can be divided to two subsets: Q_{match} and $Q_{mismatch}$.

***Step2*. Problem mapping**

Definition 99.2 Concept occupancy $O(c)$: if concept c_i and the number of K is a, in K_{ci} and Q_{match}, the number of matched concepts is β, then:

$$O(c_i) = \frac{\beta}{\alpha} = \frac{|K_{c_i} \cap Q_{match}|}{|K_{c_i}|} \tag{99.1}$$

Though the mapping process from problem set to ontology concepts, this formula can reflect the students' existing problems. If $Q = \{b2, C4, D6, C1, d2, c2\}$, after step2, the problems will be mapped to the gray-marked concepts in Fig. 99.2.

***Step3*. Problem tree construction**

As some problems may map different concepts and show different context, such as C1 and D4 in Fig. 99.2, and as for students' problems which always incline to some concepts, construction of minimized concepts set which reflect students' problems appropriately is required.

Definition 99.3 Depth and width of concept node: based on tree, set the depth of c_i as $Dep(c_i) = Dep(Parent(c_i)) + 1$, the depth of root is 1, the width of c_i is $Wid(c_i)$ which means the number of branches.

Definition 99.4 Weight of concept $\omega(c_i)$: as the depth and width of concept will influence the weight of concept, the formula is as below:

$$\omega\left(c_i\right) = \frac{\omega(Parent(c_i))}{Dep(c_i) + Wid(c_i)} \tag{99.2}$$

Especially, set $\omega(root) = 1$. Semantic distance $D(c_i,c_j)$ can be calculated, which means the sum of the weights connecting the two concepts in the ontology tree, the formula is as below:

$$D\left(c_i,\ c_j\right) = \sum_{m=0}^{n} \omega_m \tag{99.3}$$

Definition 99.5 Degree of concept $Deg(c_i)$: the number of the child nodes. If c_i is the leaf node, $Deg(c_i) = 0$.

As the hierarchical relationship is not the only factor which influence the semantic similarity, objects' attributes should also be taken into consideration. Therefore, we have introduced the formula of relationship similarity, set the self-defined objects' attributes set as $R = (r_1, r_2\ \ldots\ r_n) \subseteq P$, then the value of the relationship r between c_i and c_j concepts – if r exists between c_i and c_j, $RV_r(c_i,\ c_j) = 1$, else $RV_r(c_i,\ c_j) = 0$.

According to the definition of relationship, the following formula can be used to calculate the relationship similarity between c_i and c_j:

$$Sim_r(c_i, c_j) = \frac{\sum_{k=1}^{n} RV_{r_k}(c_i, c_j)}{m} \tag{99.4}$$

In formula 99.4, n means the number of the set R, m means the number of the distinct self-defined relationship between c_i and other concepts. So, we can conclude the similarity of concepts c_i and c_j:

$$\mathrm{Sim}(c_i, c_j) = \lambda\left(\frac{\alpha}{\alpha + D(c_i, c_j)}\right) + (1-\lambda)Sim_r(c_i, c_j) \tag{99.5}$$

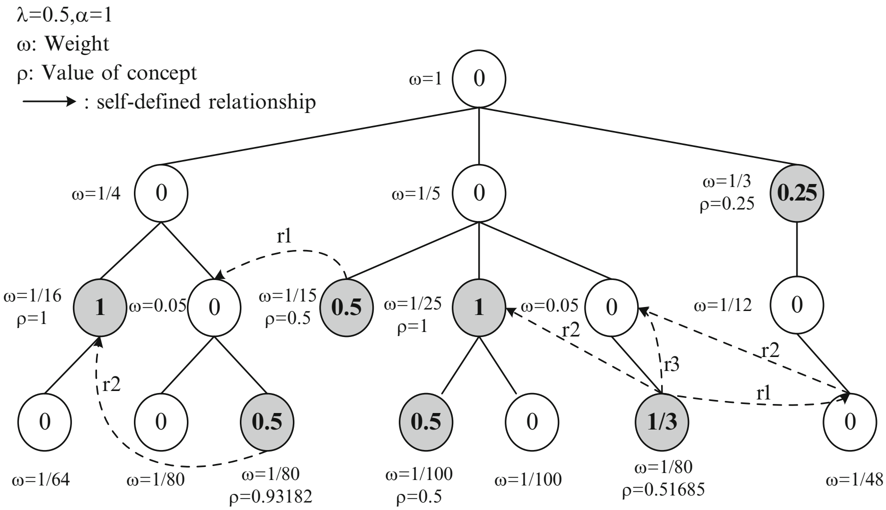


Fig. 99.3 Calculation of weight and value of concept

In formula 99.5, α and λ are the regulatory factors which are used to adjust the similarity in order to meet the requirement of different applications.

Then value of basic concept c_i can be quantified as ρ:

$$\begin{aligned}\rho(c_i) &= O(c_i) + \sum_{j=1, j\neq i} Sim(c_i, c_j)\times O(c_i) \times O(c_j)\\ &= O(c_i) \times \left[1 + \sum_{j=1, j\neq i} Sim(c_i, c_j)\times O(c_j)\right]\end{aligned} \tag{99.6}$$

c_i and $c_j \in$ basic concepts, and r exists between c_i and c_j, $r \in R$.

For example, in Fig. 99.3, set $\lambda = 0.5$ and $\alpha = 1$, the relationships $|R|$ which relate concept D6 to other concepts is 3, and $\omega(D6) = 1/80$ and $O(D6) = 1/3$, r2 is the relationship between D6 and C4, then $\rho(D6) = O(D6)*(1 + Sim(D6, C4)*O(C4)) \approx 0.51685$.

Based on formula 99.6, the influential value $I(c_p)$ of child node c_i to its parent node c_p can be calculated:

$$I\left(c_p\right) = \left\{\sum \rho(c_i) \times \frac{Deg(c_i) + 1}{Deg(c_p) + 1} \middle| \quad \begin{array}{l} c_i \in \text{Concept} \\ \text{c}_p \notin \text{Concept and Parent}(c_i) = \text{c}_p \end{array}\right\} \tag{99.7}$$

The '1's added both in numerator and denominator are used to avoid miscalculation in the situation of 0 degree. In Fig. 99.4, the $O(c)$ of A1, B1, B2, C2 and C5 are all 0, their direct children contain basic concept, so they are considered as candidate nodes, then every $I(c_p)$ will be calculated respectively according to formula 99.7.

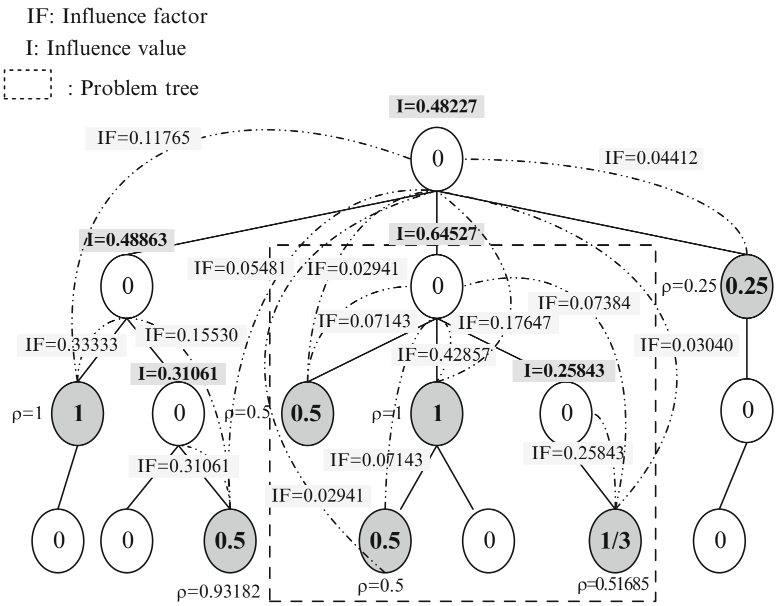


Fig. 99.4 Comparison of results

For instance, $I(B2) = IF_{B2}(C3) + IF_{B2}(C4) + IF_{B2}(D4) + IF_{B2}(D6) = \rho(C3)^* \dfrac{0+1}{16+1} + \rho(C4)^* \dfrac{2+1}{16+1} + \rho(D4)^* \dfrac{0+1}{16+1} + \rho(D6)^* \dfrac{0+1}{16+1} \approx 0.64527$.

After getting all the influential values of candidate nodes, the maximum one will be chosen as the finial root by going through all these values in order to construct the problem tree.

***Step4*. Resource recommendation**

The last step is to recommend the resources related to the concepts on the constructed problem tree to the teachers for sake of guiding their course content scientifically and effectively.

99.2.4 Homework and Quiz

Homework and Quiz are indispensable parts of education, so we still need to use the traditional ways continually. But, more importantly, we added 10 min computer quiz in the end of every teaching and learning stage determined by teachers, and the questions of the computer quiz are related to the concepts in English teaching ontology database, and the results will be used by *Step1* problem extraction.

99.2.5 Extracurricular Atmosphere

As English is the second language for non-native learners, and especially for Chinese, people have become accustomed to Chinese thinking pattern which has a totally different grammar and order of sentence components, when they learn, translation processes exist in the receptive and productive skills, which will definitely take time. Therefore, the best way to shorten this time is to exposure students to English atmosphere as long as possible and use English as much as possible. On one hand, students' cognition to English will enhanced, on the other hand, students' skills will not only be polished, and their enthusiasm will also be cultivated.

99.3 Application Analysis

The application experiment of OTETS has been implemented in an English training institution – Baff Global English (Chongqing), consuming 2 month time period and involving 6 teachers', 4 teaching assistants' and 50 students' participation. The students have been divided into two groups (Group T: using traditional teaching methods; Group O: using OTETS) with equal number, and each of the two groups has been divided into four sub-group (domestic junior tests – 10 students; domestic senior tests – 5; TOEFL – 5; IELTS – 5). During the 2 months, we have monitored the results of every sub-group. As the scoring standards of each test are different, we have converted the results unified hundred-mark system. Figure 99.5 clearly shows that, the improvements contributed by the sub-groups using OTETS are the most significant ones amongst the four categories.

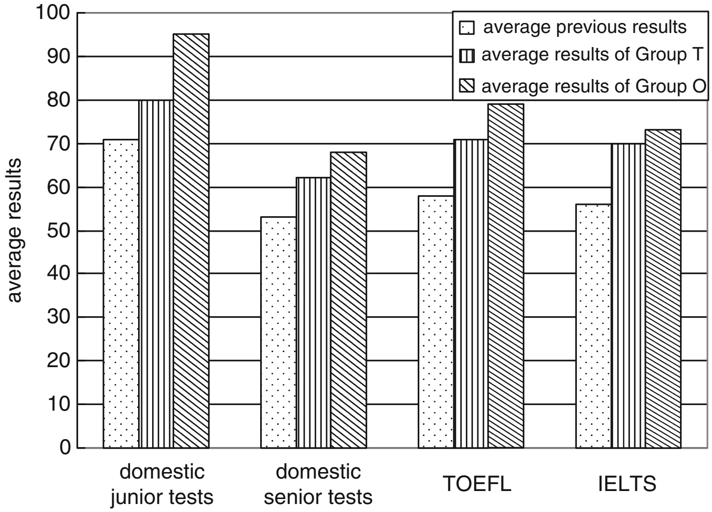


Fig. 99.5 Construction of problem tree

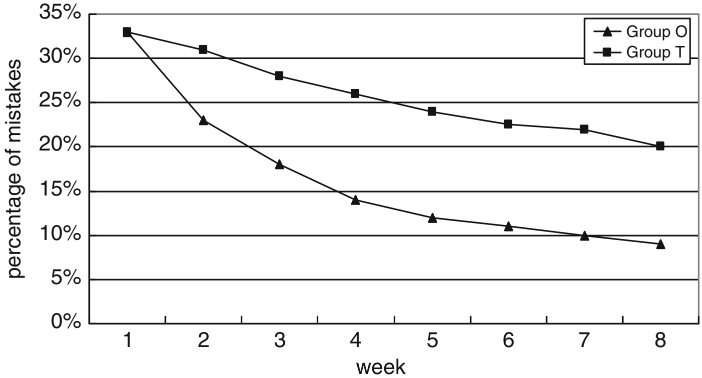


Fig. 99.6 Comparison of mistake-making

The 10 min computer quiz has been given to every sub-group including non-OTETS ones in order to know students' mistake-making situation, but only group O has received the resource recommendation. Figure 99.6 shows the performance – the mistakes made by group O declined sharply during the 2 month time, compared with group T, the number of which also decreased, but constantly and slightly.

99.4 Conclusion

To tackle the inaccuracy as to detection of students' distinct problems in English Study and to recommend corresponding resources scientifically, the Ontology Technology-supported English Teaching Scheme has been proposed. In OTETS, there are five levels scheme in order to guarantee the quality of language learning, particularly in *Course content* and *Homework/Quiz*, ontology-based problem detection and resource recommendation method has been introduced to fulfill our goal. Finally, the application experiment has been implemented in an English training institution, and the results have shown the feasibility and superiority of OTETS. Therefore, we can conclude that OTETS is effective and superior in TESOL.

Acknowledgment We'd like to thank Baff Global English (Chongqing) regarding the support to the application experiment, and the staffs working there as well as the voluntary students studying there.

References

Batet M, Sanchez D et al (2011) An ontology-based measure to compute semantic similarity in biomedicine. J Biomed Inf 44(1):118–125

Berg BC (2011) The effects of trained peer response on ESL students' revision types and writing quality. J Sec Lang Writ 22:233–247

Canagarajah AS (2006) TESOL at forty: what are the issues? TESOL Q 40(1):9–34

Hyland K (2006) Feedback on second language students' writing. Lang Teach 2:83–101

Jin H, Ning XM, Jia W, Yuan P (2009) Practical and effective IR-style keyword search over semantic web. Inf Process Manage 45(2):263–271

May S (2007) Sustaining effective literacy practices over time in secondary schools: school organizational and change issues. Lang Educ 21(5):387–405

May S (2011) The disciplinary constraints SLA and TESOL: additive bilingualism and second language acquisition, teaching and learning. Linguist Educ 22:233–247

Ortega L (2009) Understanding second language acquisition. Hodder Arnold, London

Pennycook A (1999) Introduction: critical approaches to TESOL. TESOL Q 33(3):329–348

Qiu YG, Frei HP (1993) Concept based query expansion. In: Proceedings of the 16th annual international ACM SIGIR conference on research and development in information retrieval, New York, NY, USA, pp 160–169

Richard JC (2005) 30 years of TEEL/TESL: a personal reflection. http://www.professorjackrichards.com/pdfs/30-years-of-TEEL.pdf

Van Rijsbergen CJ (1998) A new theoretical framework for information retrieval. In: Proceedings of the AMC conference on research and development in information retrieval, Pisa, Italy, 1986, pp 194–200

Chapter 100
Strategy Planning for State-Owned New Energy Vehicle Industry of China

Li-ran Chen, Guo Yan, and Ying Zhang

Abstract Due to the tasks of the national energy saving policy and other reasons, the sales quantity of 2011 for traditional car shows a downward trend, after the rapid growth phase of 2009–2010. Comparatively, the new energy vehicles are more suitable for national energy reduction policies. The new energy vehicle industry should be the trend of the future development for the automobile industry. Therefore, a reasonable plan for state-owned new energy vehicle industry is quite important. This article developed an adaptive balanced scorecard knowledge-based system for industry strategic planning. This knowledge-based system also uses the partial least squares regression analysis to determine the specific strategy weights. So this adaptive BSC knowledge-based system can help providing recommendations for the industry development plan.

Keywords Adaptive BSC system • New energy vehicle • Strategy plan • State-owned industry

100.1 Introduction

New energy vehicle industry in China generally started late, and there is a big gap between China and the world's advanced level. China has successfully developed a series of products for pure electric vehicles and dimethyl ether motor vehicles

L. Chen (✉)
School of Management, Tianjin University, Tianjin, China

School of Environmental Science and Safety Engineering, Tianjin University of Technology, Tianjin, China
e-mail: lisa.liranchen@gmail.com

G. Yan • Y. Zhang
School of Management, Tianjin University, Tianjin, China

E. Qi et al. (eds.), *The 19th International Conference on Industrial Engineering and Engineering Management*, DOI 10.1007/978-3-642-37270-4_100,

and has conducted demonstration industrial runs to entering the ranks of the international leading. For industrial applications, China has started to promote ethanol fuel vehicles, bio-diesel cars and natural gas vehicles. Especially ethanol gasoline and bio-diesel car have gradually expanded the scope of the experiment. 2008 Beijing Olympic Games, the 2010 Shanghai World Expo and the action that "10 City 1000 cars" energy saving and new energy electric vehicles mass demonstration by the Ministry of Science have further promoted China's lithium-ion battery development process.

100.2 Characteristics of State Owned Assets in China Vehicle Industry

State owned assets mentioned here refers to the generalized state-owned assets definition, which includes central government control, local government control, national equity shares control, local government equity shares control and so on.

After this 20 years' opening up, the market mechanism has played a huge role, and most large state-owned automotive groups have some subsidiary branches involved in joint venture. (Table 100.1 shows some examples of joint venture companies.) Because the starting point of Chinese reform is highly single national structure of property rights, the state-owned property is still occupying a dominant position in China automobile industry although.

Table 100.1 Some automotive joint venture proportion at the end of 2008

Joint venture enterprise	Foreign part	Share scale (%)
TFTM (under FAW group)	Toyota	40
Guang Zhou HONDA (under Guangzhou Automobile Industry Group)	Toyota	50
Honda Export Base (under Guangzhou Automobile Industry Group)	Toyota	60
Engine factory of FAW–Volkswagen (under FAW group)	Volkswagen	60
Engine factory of DongFeng Honda (under DFM group)	Honda	50
DongFeng Motor Company Limited (under DFM group)	Nissan	50
Dongfeng Cummins Engine Co., Ltd (under DFM group)	Cummins	50
Beijing Foton Cummins Engine Co., Ltd. (under Beijing Automotive Group)	Foton	50
Dongfeng Peugeot Citroen Automobile Company Ltd. (under DFM group)	PSA Peugeot Citroen	50
Shanghai Volkswagen (under SAIC group)	Volkswagen	50
SAIC Motor Co., Ltd. (under SAIC group)	Volkswagen	60
BMW Brilliance Automotive (under Brilliance China Automotive Holdings Ltd.)	BMW	50

Source: China Automotive Industry Yearbook, 2009

Poor performance is a long plagued prominent problem for state-owned property company. Many studies have attributed the low performance problem to state-owned property. But Lin Yifu et al. (1997) think the crux of this problem for state-owned enterprises is not property, but rather competition. The statistical result does not prove high proportion of state-owned enterprises will inevitably lead to poor performance. Andrew Delios and Xu (2008) observed ownership infections for the performance of Chinese auto enterprises. The major difference of Chinese auto industry from the common enterprise is its property feature. State-owned enterprises are more impacted by the government and their performances are not relying solely on the market to determine.

In researches of Chinese automotive industry, Luo Yunhui and Lin Jie (2001) did empirical analysis for Chinese Automotive Industry market performance and analyzed the relationship between profit margins and centralization degree on the auto industry basing on SCP theory. Liu Guoxin (2001) made a regression analysis from an empirical point of view by data of several major Chinese automotive group, then explored the interaction between industrial organization and technological innovation mechanism. Gao Yan and Wang Xiuting (2004) selected ten evaluated indexes, made a comprehensive evaluation of the auto industry performance level for 18 provinces, municipalities and autonomous regions using SPSS statistical software and principal component analysis method. Wang Ying (2007) explore the performance of China's auto industry problems basing on geographical distribution pattern of China, and several major Chinese automotive industry cluster were the target. Guo Haitao (2005, 2008) explored the relationship between the automotive industry market performance and property rights, and inspected effects of property rights and market structure on industry profits by using multiple regressive analysis method. The conclusion was made that the property right reformation on large state-owned enterprises should actively attract Chinese private capital besides overseas investment, and only this can improve business efficiency. Deng Zhituan (2010) studied Chinese automobile industry performance in regional comparison. Quotient of location method, GE value and relative value of efficiency were used in that article to measure differences level of performance in each sample region. The proposal was put forward that how to further promote the performance of regional industries.

100.3 Proposed Model

In this research, we first base on the four perspectives of the BSC (Kaplan and Norton 1992) to prepare a list of strategy performance evaluation indicators under the guidance of experts in Chinese automotive industry. The evaluation indicators are distributed to performance statistics data of the new energy vehicle factories in the automobile industry, and the collected data are analyzed through a constructed PLS (Chin and Newsted 1998) program to obtain the relative importance of the four perspectives and the relative importance of the key performance indicators under

Table 100.2 Strategic performance evaluation frame for state owned assets of new energy vehicle industry

Goal	Perspectives	Performance indicators
Strategic performance of state-owned new energy vehicle industry (Goal)	Financial perspective (FP)	Return on main operating profit
		ratio of industrial added value
	Social responsibility (SR)	Market power
		Stock control
		Technology control
	Innovation perspective (IP)	R & D devoted intensity
		R&D staff rate
	Internal business (IB)	Revealed comparative advantages
		Degree of industry concentration

each perspective. All data are taken from the statistics data of China Automobile Industry Yearbook and annual statistics from National Bureau of statistics (Center CATR 2010; China Ministry of Environmental Protection 2010). Then the relationship of this evaluation system is tested by PLS method using SmartPLS 2.0 program. The results can provide some suggestions to new energy vehicle industry in developing future strategies, development objectives and performance evaluation.

100.3.1 Framework System Established

Based on the concept of the BSC, review of new energy vehicle industry strategic development literature and interview with Chinese automobile industry experts, a strategic performance evaluation frame for state owned assets of new energy vehicle industry is constructed as in Table 100.2.

100.3.1.1 Goal in BSC System

The ultimate goal of BSC system for new energy automotive industry is to measure the strategic performance of the industry, and to enhance the international competitiveness of whole industry. Finally this could promote the overall industrial development and growth. Since PLS regression analysis need concrete data to reflect the industry's strategic performance, and the ultimate development target of state-owned new energy vehicle industry, Industry Trade Competitiveness Index is chosen as the measurement variable for latent variable "Strategic Performance". Industry Trade Competitiveness Index is calculated as the ratio of an industry's net export value to total import and export value of this industry. If the ratio approaches 1, this will indicate the stronger international competitiveness of this industry.

100.3.1.2 Financial Perspective

Financial performance measures indicate the extent to that the strategy, implementation, and execution of new energy vehicle industry are contributing to bottom-line improvement. The achievement of financial targets reflects the outcome of dimensions of performance captured by the BSC's other three perspectives.

Return on Main Operating Profit (RMOP) and Ratio of Industrial Added Value are the two performance indicators for financial perspective. RMOP is operating profit with net income ratio of organization's main business in a certain period. It shows how many operating profit that the main business income per unit of the organization could bring to. This is the main indicator to evaluate business benefits and reflect the profitability of the main business. Ratio of Industrial Added Value means the ratio of industrial added value to GDP. This ratio reflects the essentiality of the target industry in national economy.

100.3.1.3 Social Responsibility

In the operation of state-owned assets, the organization must serve the higher social goals, such as supporting social economic stability, improving the core competitiveness of state-owned economy and meeting the needs of the public. These overriding goals should be placed in BSC's top position for state-owned new energy automotive industry, and in order to guide the industry complete the social mission. Social responsibility is the starting point for management of state-owned assets, so the industry decision-making department should review what is the whole industrial mission, what is the reached degree of social responsibility. The achievement of social responsibility could maximize the assets value of state-owned industries.

Market Power, Stock Control and Technology Control of state-owned assets are the three indicators of social responsibility perspective for new energy automotive industry. Market Power reflects domestic market control degree of state-owned enterprises. It can be measured with the ratio of domestic market share of state-owned enterprises to the whole market share. Stock Control of state-owned assets could reflect control complexion on domestic industries from the angle of internal equity capital controls. Technology Control can reflect the domestic controllable situation from a technical view. The control rate of foreign technology and localization rate sum to 1.

100.3.1.4 Innovation Perspective

Innovation Perspective concern whether this industry can continue to improve and create value. So it can be said that innovation is the driver for the other three perspectives (Financial Perspective, Social Responsibility and Internal Business) achieving good result. Considering ease of statistical data collection and industry characteristics, R&D Devoted Intensity and R&D Staff Rate are selected to reflect the effect of industrial innovation for whole industrial strategy performance.

R & D Devoted Intensity is assessed by two sub-indicators. One is accounted ratio of industry research and development expenditures in sales, and the other is increased percentage point over the previous year which is accounted ratio of industry research and development expenditures in sales. R&D Staff Rate is the ratio of R & D employees to the total number of employees.

100.3.1.5 Internal Business

Internal business process perspectives of the BSC concern the use of business processes to achieve shareholder and customer satisfaction. Considering new energy automotive industry characteristics, this process fully reflects how the industry exceeds new energy automotive industry in other country. Internal perspective aspect of the BSC is closely aligned with an organization's innovation (Kaplan and Norton 1996). Revealed Comparative Advantages and Degree of Industry Concentration are selected to reflect internal business perspectives for new energy automotive industrial strategy performance.

RCA means the ratio of an industry exports share in its country total exports value to world this industry exports share in total world exports value. This RCA ratio reflects the industry's competitiveness from export data view. Degree of Industry Concentration reflects the international competitiveness of the industry from internal organization situation.

100.4 Analysis and Results

100.4.1 Data Collection

New energy automotive industry in China is still at the starting stage, and the statistics of operation is quite inadequate. Since characteristics and general state-owned new energy automotive industry are significant similarities with features of state-owned automotive industry in China. The overall objective of these two state-owned assets is to achieve a strong automotive industry in China. Also these two parts regard industry comprehensive development, which include financial perspective, social responsibility, technological innovation and enhancing industry competitiveness. Therefore the statistical data of state-owned automobile industry are selected to explain relationship and weight of the whole system and various indicators within the framework. But enormous changes have taken place and statistics caliber for automotive industry have changed after the attraction of foreign investment since reform and opening up in China. Basing on the same statistic caliber, 2001–2007 statistics data of state-owned automotive industry are used to analysis the relationship between the indicators by partial least squares method.

Table 100.3 Suggest latent variable correlations

	IB	FP	IP	SR	Goal
IB	1.000				
FP	0.828	1.000			
IP	0.948	0.929	1.000		
SR	−0.922	−0.815	−0.958	1.000	
Goal	0.970	0.861	0.969	−0.970	1.000

100.4.2 PLS Analysis and Results

PLS is used to generate the relationship of the four perspectives of the BSC and the relationship of the performance indicators. In particular, RCA indicator is composites formed by the corresponding observed variables, while the four performance dimensions constitute unobserved variables reflected in a set of manifest variables. We estimated our model using SmartPLS 2.0 Build (Ringle et al. 2005). The level of statistical significance of the coefficients of both the measurement and the construct models was determined through a bootstrap re-sampling procedure (200 sub-samples were randomly generated). The standard deviation and generate an approximate t-statistic can be calculated. This overcomes non-parametric methods' disadvantage of having no formal significance tests for the estimated parameters. Construct model evaluation is conducted by examining the size and significance of the path coefficients and the R^2 values of the dependent variables. To describe conveniently, all the process variables and latent variables are standardized data, that average value is 0 and standard deviation is 1.

100.4.2.1 Latent Variable Correlations

The findings shown in Table 100.3 suggest latent variable correlations. All diagonal elements are greater than the off-diagonal elements in the corresponding rows and columns. In this model program, the mean square error extracted of latent variable rate meet the discriminate validity of the test conditions, the square root of the value are greater than all of the correlation coefficient between latent variables. On the other hand, the high correlation coefficients of variables, shows a relatively significant correlation.

100.4.2.2 Communality

The equation of communality in PLS model is shown below. Generally believed when the communality greater than 0.5, the variance explained by the latent variable is greater than the variance of measurement error caused in outer model. The greater the communality value, the effect of system model is better. Table 100.4 shows the communality result.

Table 100.4 Communality

	Communality	Average
IB	0.502330	0.780857
FP	0.512053	
IP	0.918073	
SR	0.971829	
Goal	1.000000	

Table 100.5 Construct model total effects

	Original sample (O)	Sample mean (M)	Standard deviation (STDEV)	Standard error (STERR)	T statistics (|O/STERR|)
IB- > Goal	0.582	0.529	0.092	0.092	6.326
FP- > Goal	0.337	0.282	0.112	0.112	3.005
IP- > Goal	−0.591	−0.435	0.286	0.286	2.070
SR- > Goal	−0.725	−0.672	0.144	0.144	5.036

The communality of all variables performs well. From the overall effect of matrix model, the average communality factor is 0.780857, meeting the requirements of the general 0.5 standard. Above index value indicates that latent variables have better reflective effect respecting to their measurement variables.

$$\text{communality}_{jk} = \frac{1}{p_j}\sum_{k=1}^{p_j} \text{cor}^2\left(x_{jk}, y_j\right)$$

100.4.2.3 Construct Model Total Effects

Bootstrap re-sampling method is commonly used in PLS path modeling. In this paper, Bootstrap (re-sampling frequency of 200 times) is tested by SmartPLS 2.0 software for Loading, Weight and Path Coefficient of strategic performance evaluation model in new energy automotive industry. Through value list, t-values is 1.97 at 95 % statistical probability and df = 200. From Bootstrap test results (Table 100.5), Loading, Weight and Path Coefficient of strategic performance evaluation model are all pass this statistical test. Bootstrapped standard deviations and t-values confirm the significance of hypotheses.

100.4.2.4 R^2 and Redundancy

R^2 is an important indicator for evaluation the explanative effect of internal relations in PLS path model. Unexplained variance of latent variable by construct model is smaller with greater R^2 values. The minimum standards required R^2 is 0.65 for evaluating the fitting results of internal model. The minimum required standard of Redundancy is 0.325, which means a product result of communality and R^2.

Table 100.6 R^2 and redundancy of the model

	R^2	Redundancy
Goal	0.985918	0.790261

Table 100.7 Outer loadings and t-statistic

	Performance indicators	Loadings	T-statistic
IB	Revealed comparative advantages	0.998	114.374
	Degree of industry concentration	0.138	0.722
FP	Return on main operating profit	−0.251	1.510
	Ratio of industrial added value	0.849	15.823
IP	R & D devoted intensity	0.963	349.942
	R&D staff rate	0.959	298.732
SR	Market power	0.943	291.768
	Stock control	0.975	311.202
	Technology control	−0.769	24.284
Goal	Industry trade competitiveness index	1.000	–

From the result of Table 100.6, R^2 value is greater than 0.65. Internal relationship can be considered as an ideal fitting result. The comprehensive evaluation of latent variable for each latent variable has a good representation.

100.4.2.5 Outer Model (Loadings and T-Statistic)

In PLS path model outer loadings between each measurement variable and the respective structure variables can be used for reliability measurement. A standard that accepted and adopted by many scholars is the outer loadings factor should greater than 0.7, which means a latent variable could explain more than 50 % of the variance of its measured variable group.

In Table 100.7 most outer loadings of measure variables between their latent variables are more than a general standard except “Degree of Industry Concentration”, “Return On Main Operating Profit” and “Technology Control” three measure variables. Most performance indicators except “Degree of Industry Concentration” and “Return On Main Operating Profit” have pass the t-values examination. So these two indicators should leave outside of the integrated variables. While the “Technology Control” factor is lower, but the average load factor of overall dimensions comprehensive evaluation is 0.71 after elimination of the two indicators. This is still higher than the general standard of 0.7. And the t-value of “Technology Control” factor is 24.284. Therefore the “Technology Control” variable can be retained in the integrated variables.

100.5 Conclusion

This paper proposes an approach based on the PLS and BSC for evaluating the strategic performance of new energy automotive industry in China. The PLS is structured by the four major perspectives of the BSC including internal business, financial perspective, innovation perspective, and social responsibility, followed by performance indicators. Because the number of sample sizes are quite small and the high correlation coefficients of variables, the PLS path model is adopted to solve these problem.

The PLS results show that the comprehensive evaluation of variables has a good representation. And the internal relationship within BSC system can be considered as an ideal fitting result. The variable relationships in construct model confirm hypotheses. Social Responsibility (SR) and Innovation Perspective (IP) perspectives have higher absolute weighting values and are found to be negative with the strategic performance variable. This means these two perspectives need substantial investment to improve performance. Performance increasing of these two perspectives is quite important for upgrading the whole strategic performance goal.

References

Andrew Delios NZ, Xu WW (2008) Ownership structure and the diversification and performance of publicly-listed companies in China. Bus Horiz 51(6):473–483

Center CATR (2010) China automotive industry yearbook. CATR, Tianjin, pp 35–46

Chin WW, Newsted PR (1998) Structural equation modeling analysis with small samples using partial least squares. In: Hoyle IR (ed) Statistical strategies for small samples research. Sage, Thousand Oaks

China Ministry of Environmental Protection (2010) China Vehicle Emission Control annual report. Ministry of Environmental Protection of the People's Republic of China, 2010

Deng Zhituan (2010) China's auto manufacturing industry, the performance of regional comparative studies. Stat Decis 5:100–103

Gao Yan, Wang Xiuting (2004) Hubei automotive industry performance evaluation. Automob Res Dev 9:20–23

Guo Haitao (2005) Research on state property, market structure and market performance in China's auto industry. J Beijing Technol Bus Univ Soc Sci 20(6):8–13

Guo Haitao (2008) Market competition, industry concentration and industry performance. Reform Strateg 24(1):107–110

Kaplan RS, Norton DP (1992) The balanced scorecard: measures that drive performance. Harv Bus Rev 70(1):71–79

Kaplan RS, Norton DP (1996) Using the balanced scorecard as a strategic management systems. Harv Bus Rev 74:75–85

Lin Yifu et al (1997) Full information and state-owned enterprise reform. Shanghai Joint Publishing, Shang Hai

Liu Guoxin (2001) China's auto industry market analysis of market conduct and market performance. China Mech Eng 12(7):821–824

Luo Yunhui, Lin Jie (2001) China's auto industry market structure and the performance research. Shanghai Auto 8:8–10

Ringle CMW, Will S, Alexander S (2005) SmartPLS 2.0 software. Hamburg

Wang Ying (2007) An analysis of the performance level of auto industry in China based on the industrial cluster theory. J Anhui Univ (Philos Soc Sci) 31(1):124–128

Chapter 101
Research on Organization Tacit Knowledge Sharing: A Game Theoretic Analysis

Li Dan

Abstract Tacit knowledge sharing is one of the important ways to improve the tacit knowledge stock of organization and realize knowledge innovation. The tacit knowledge sharing of organization is a dynamic game process, which is working on the social network of knowledge management among kinds of social units. Based on the theory of games, this paper constructs a model of organization tacit knowledge sharing. The possibility of tacit knowledge sharing from the viewpoint of static games is analyzed. Strategy and methods of how to realize tacit knowledge sharing among organizations are presented.

Keywords Tacit knowledge sharing • Game theoretic • Knowledge innovation

101.1 Introduction

There is broad recognition that effective knowledge management is essential to the success of modern organizations. Nonaka distinguishes knowledge to tacit and explicit knowledge. Explicit knowledge is transmittable in formal, systematic language and takes the form of documents, reports, "white papers," catalogues, presentations, patents, formulas, etc. (Nonaka and Konno 1998, 1995; Holste and Fields 2005). In contrast, tacit knowledge has a personal quality, which makes it hard to formalize and communicate. Tacit knowledge is rooted in an individual's experience and the values he or she holds. It involves abilities, experience, undocumented processes and so on (Brockmann and Anthony 1998; Qian et al. 2008). As we know, tacit knowledge may play an important role in the strategic planning performance of knowledge management and innovation.

L. Dan (✉)
Business School, East China University of Political Science and Law, Shanghai, China
e-mail: huazhengld@126.com

E. Qi et al. (eds.), *The 19th International Conference on Industrial Engineering and Engineering Management*, DOI 10.1007/978-3-642-37270-4_101,

The shares and transfers of tacit knowledge are complicated process, resulting from job transfers, interdepartmental task forces, and informal networks and employee interactions (Marquardt 1996). Key to both types of knowledge transfer is the willingness and capacity of individuals or organizations to share what they know and to use what they transfer and learn. The willingness of members to share and use tacit knowledge may depend on the extent that partners are trusted. Organizations participate in tacit knowledge sharing activities to realize knowledge innovation and to achieve their own economic interests. One of the most important characters of tacit knowledge sharing operation is collaborative competition. This is also one important form of the tacit knowledge sharing behavior. Knowledge sharing behavior among partners is affected by potential competition mechanism. The members of implementing knowledge sharing cooperation are extremely sensitive to the cooperation behavior and the income apportionment involved in the whole knowledge sharing process. With the expansion of the knowledge sharing cooperation scope, partners can make full use of geographical proximity and industrial association to achieve the goal of mutually beneficial and win-win eventually by sharing resources, advantage complementary and joint investment. Now the cooperation philosophy and mechanism will play the positive role.

Based on the sharing mechanism and competitive mechanism, tacit knowledge sharing operation is usual in a stable equilibrium. In this paper, the possibility of knowledge sharing from the viewpoint of static games is analyzed. And then the methods of how to realize tacit knowledge sharing among organizations are presented.

101.2 The Game Analysis of Tacit Knowledge Sharing Behavior

In this paper, the complicated tacit knowledge sharing behavior is simplified to the infinitely repeated game problem of the coordination behavior between two organizations. The simple cooperative competition game model is established. The knowledge sharing main bodies are the mutual cooperative partners, respectively accept responsibility of the knowledge resource providers and the knowledge resource recipients (organization A and organization B).

101.2.1 Game Model of Tacit Knowledge Sharing Behavior

It is the requirement for constructing the model and the simplified calculation, following the related research results, our first study hypothesis in this paper are as follow (Yu and Shi 2011; Cheng and Ning 2005; Li 2009; Liu 2010):

H1 *knowledge sharing main bodies need to provide certain knowledge resource, It is assumed that the knowledge level of organizations for sharing strategy is K_i, the knowledge level invested by organization A is K_1, the knowledge level invested by organization B is K_2.*

H2 *knowledge sharing main bodies can get two kinds of benefit during knowledge sharing process, including motivational revenue and additional revenue.*

(a) *Motivational Revenue*

Knowledge sharing main bodies can get a certain proportion of revenue as reward according to individual efforts on cooperation. It is assumed that the incentive coefficient is π_i, $0 < \pi_i < 1$. The reward of organization A provided for implementing knowledge sharing behavior is $\pi_1 K_1$, The reward of organization B provided for implementing knowledge sharing behavior is $\pi_2 K_2$.

(b) *Additional Revenue*

Additional revenue is created by applying the new tacit knowledge received during the cooperation process for each partner.

The revenue of tacit knowledge sharing cooperation for each organization is U, it can be expressed as a function g.

$$U = g\left(K_i, \gamma_i, \omega_i\right) \tag{101.1}$$

γ_i represents the absorption coefficient of tacit knowledge for the knowledge cooperation for organization A and organization B. $\gamma_i \cdot K_i$ represents the tacit knowledge accumulation achieved during the knowledge sharing cooperation process. ω_i represents the efficiency for shifting knowledge accumulation into the economic returns of each partner. In this paper, the revenue of knowledge sharing cooperation can be expressed as

$$U = \gamma_i \cdot \omega_i \cdot K_i \tag{101.2}$$

H3 *knowledge sharing main bodies need to pay for cooperation cost, such as the opportunity cost for providing the shared tacit knowledge, the learning cost for acquiring new tacit knowledge, the control cost for the whole cooperation process and so on.*

It is assumed that the cost coefficient is C_i, $0 < C_i < 1$. The cost paid by organization A is $C_1 K_1$, cost paid by organization B is $C_2 K_2$.

Now we analyze the tacit knowledge sharing behavior and the revenue. There are four cases.

The first case, if cooperative strategy is adopted for both organization A and organization B, The revenue of organization A is the sum of motivational revenue and additional revenue that can be express as

$$G_{11} = \gamma_1 \cdot \omega_1 \cdot K_2 + \pi_1 K_1 - C_1 K_1 \tag{101.3}$$

Table 101.1 Win matrix for tacit knowledge sharing cooperation

		Organization B	
		Sharing	Defense
Organization A	Sharing	G_{11}, G_{12}	G_{13}, G_{14}
	Defense	G_{21}, G_{22}	0 0

The revenue of organization B can be express as

$$G_{12} = \gamma_2 \cdot \omega_2 \cdot K_1 + \pi_2 K_2 - C_2 K_2 \tag{101.4}$$

In the second case, if cooperative strategy is adopted for organization A, and organization B takes the defense strategy, the revenue of organization A is

$$G_{13} = \pi_1 K_1 - C_1 K_1 \tag{101.5}$$

The revenue for organization B is

$$G_{14} = 0 \tag{101.6}$$

In the third case, if cooperative strategy is adopted for organization B, and organization A takes the defense strategy, the revenue for organization A is

$$G_{21} = 0 \tag{101.7}$$

The revenue of organization B is

$$G_{22} = \pi_2 K_2 - C_2 K_2 \tag{101.8}$$

In fourth case, if defense strategy is adopted for organization A and organization B, the revenue of both two organizations is zero, that is

$$G_{14} = G_{24} = 0 \tag{101.9}$$

The win matrix of knowledge sharing cooperation for organization A and organization B can be shown as Table 101.1.

101.2.2 The Condition for Tacit Knowledge Sharing Behavior

Whether organization A adopts cooperation strategy or not, it depends on balancing the revenue of cooperation strategy against defense strategy.

When organization B chooses the sharing strategy, the condition for organization A adopting cooperation strategy is $G_{11} > G_{21}$, that is

$$\gamma_1 \cdot \omega_1 \cdot K_2 + \pi_1 K_1 - C_1 K_1 > 0 \tag{101.10}$$

When organization B chooses the defense strategy, the condition for organization A adopting cooperation strategy is $G_{13} > G_{23}$, that is

$$\pi_1 K_1 - C_1 K_1 > 0 \tag{101.11}$$

When organization A chooses the sharing strategy, the condition for organization B adopting cooperation strategy is $G_{12} > G_{14}$, that is

$$\gamma_2 \cdot \omega_2 \cdot K_1 + \pi_2 K_2 - C_2 K_2 > 0 \tag{101.12}$$

When organization A chooses the defense strategy, the condition for organization B adopting cooperation strategy is $G_{22} > G_{24}$, that is

$$\pi_2 K_2 - C_2 K_2 > 0 \tag{101.13}$$

Now we analyze the cooperation condition of organization A.

- If $\pi_1 - C_1 > 0$, then $\gamma_1 \cdot \omega_1 \cdot K_2 + \pi_1 K_1 - C_1 K_1 > 0$, organization A will sure choose the sharing strategy regardless of whether organization B make the same choice. Now the reward achieved from cooperation can make up for the sharing cost for organization A.
- If $\pi_1 - C_1 < 0$, but $\gamma_1 \cdot \omega_1 \cdot K_2 + \pi_1 K_1 - C_1 K_1 > 0$, organization A will sure choose the sharing strategy when organization B makes the same choice. If organization B chooses defense strategy, and then organization A will refuse to cooperation. Now the reward achieved from cooperation cannot make up for the sharing cost for organization A, so it necessary for organization A get revenue from acquiring knowledge resource provided by organization B.
- If $\gamma_1 \cdot \omega_1 \cdot K_2 + \pi_1 K_1 - C_1 K_1 < 0$, organization A will sure choose the defense strategy regardless of whether organization B make the same choice.
- Now we analyze the cooperation condition of organization B.
- If $\pi_2 - C_2 > 0$, then $\gamma_2 \cdot \omega_2 \cdot K_1 + \pi_2 K_2 - C_2 K_2 > 0$, organization B will sure choose the sharing strategy regardless of whether organization A make the same choice. Now the reward achieved from cooperation can make up for the sharing cost for organization A.
- If $\pi_2 - C_2 < 0$, but $\gamma_2 \cdot \omega_2 \cdot K_1 + \pi_2 K_2 - C_2 K_2 > 0$, organization B will sure choose the sharing strategy when organization A makes the same choice. If organization B chooses defense strategy, and then organization B will refuse to cooperation. Now the reward achieved from cooperation cannot make up for the sharing cost for organization B, so it necessary for organization B get revenue from acquiring knowledge resource provided by organization A.
- If $\gamma_2 \cdot \omega_2 \cdot K_1 + \pi_2 K_2 - C_2 K_2 < 0$, organization B will sure choose the defense strategy regardless of whether organization A make the same choice.

The strategies matrix of knowledge sharing cooperation for organization A and organization B can be shown as Table 101.2.

Table 101.2 Strategies matrix for tacit knowledge sharing cooperation

		Organization B		
			$\pi_2 < C_2$	
		$\pi_2 > C_2$	$\gamma_2 \cdot \omega_2 \cdot K_1 + \pi_2 K_2 - C_2 K_2 > 0$	$\gamma_2 \cdot \omega_2 \cdot K_1 + \pi_2 K_2 - C_2 K_2 < 0$
Organization A	$\pi_1 > C_1$	(sharing, sharing)		(sharing, defense)
	$\pi_1 < C_1$			
	$\gamma_1 \cdot \omega_1 \cdot K_2 + \pi_1 K_1 - C_1 K_1 > 0$	(sharing, sharing)	(sharing, sharing) (defense, defense)	(defense, defense)
	$\gamma_1 \cdot \omega_1 \cdot K_2 + \pi_1 K_1 - C_1 K_1 < 0$	(defense, sharing)	(defense, defense)	

101.3 The Strategy for Exciting Tacit Knowledge Sharing

Several facets need to be synthetically considered to designs for tacit knowledge sharing process.

101.3.1 Improving the Absorption Coefficient of Tacit Knowledge

Improving the absorption coefficient of tacit knowledge can stimulate the cooperative desire. Organizations should improve intensity and correlation degree of synergistic tacit knowledge resource which be provided by organizations.

Actively guide and inspire organizations that have the predominance knowledge resource to engage in tacit knowledge sharing cooperative operation, and then select the suitable cooperative mode. It helps organizations to provide the knowledge resources closely related to and maximize meet the needs of knowledge sharing cooperation. Strengthen the exchange and communication among organizations and enhance mutual trust degree.

101.3.2 Encouraging Cooperation for Many Times

There are more revenue stability and balance for repeated knowledge sharing cooperation than just one time of cooperation. Because of cooperation between organizations more times, organizations more incline to establish a stable, long-term cooperation relationship. It benefit for exchanging and interaction, and reducing the uncertainty and risk.

101.3.3 Improve the Incentive Intensity (Tu 2002; Jiang 2003)

Organizations that participate in tacit knowledge sharing cooperation is mostly operation in the cooperative environment, organizations' behavior get further conditioned by groups of formal and informal institutional constraints, they cannot betray the punishment. In this case, basing on the various preferential policies and interest drive, the larger incentive intensity is, the more it attracted to organizations involved in cooperative operation.

It can be embodied in groups that the knowledge collaborative opportunities, cooperative degree and synergistic contribution of each member organization. Secondly, it can be embodied in the range of decision-making and operation rights, and the harmonious degree of collaborative operation environment. Thirdly, it can be embodied in emotion factors, including collaborative subject cooperative status, prestige and so on. Then it also can be embodied in cooperation culture, the relationship among members, and the ability to resolve the conflict and so on.

101.3.4 Improve the Profitability of Knowledge Innovation

It is necessary for organizations to improve the ability of knowledge transferring and innovation. It needs to conform to exterior market and to improve the capacity for the commercialization of knowledge and research findings. It can be guided by establishing incentive mechanism and internal innovation culture.

101.4 Conclusions

It is important for organizations to share and use of tacit knowledge. In this paper, the game model of knowledge sharing cooperation is analyzed. And then the methods of how to realize tacit knowledge sharing among organizations are presented. It is the effective approach for organizations to implement the knowledge cooperation and improve the knowledge sharing performance. Further research need to be made to make the conclusion generalization.

References

Brockmann E, Anthony W (1998) The influence of tacit knowledge and collective mind on strategic planning. J Manage Issues 10(2):204–222

Cheng GF, Ning XX (2005) The game analysis of knowledge cooperation behavior of members of virtual enterprise. Sci Technol Prog Policy 5:10–12

Holste J, Fields D (2005) The relationship of affect-and cognition-based trust with sharing and use of tacit knowledge. In: Academy of Management best conference paper, Honolulu (On the Island of Oahu), Hawaii, USA, pp B1–B9

Jiang RH (2003) An analysis of incentive on team cooperation and innovation. Syst Engg 21(5):38–42

Li D (2009) Research on the knowledge collaboration behavior and management mechanism based on industrial cluster. Law Press, China

Liu JS (2010) Study on incentive strategies of organization knowledge sharing based on static game analysis. Econ Res Guide 16:126–128

Marquardt M (1996) Building the learning organization: a system approach to quantum improvement and global success. McGraw-Hill, New York

Nonaka I, Konno N (1995) The knowledge creating company: how Japanese companies create the dynamics of innovation. Oxford University Press, New York

Nonaka I, Konno N (1998) The concept of “ba”: building a foundation for knowledge creation. Calif Manage Rev 40(3):1–15

Qian H, Davison R, Jibao G (2008) Impact of personal and cultural factors on knowledge sharing in China. Asia Pac J Manage 25(3):451–471

Tu HQ (2002) The incentive mechanism of knowledge employees. Product Res 2:157–159

Yu Y, Shi Q (2011) The game theoretical study on the tacit knowledge sharing from the bounded rationality perspective. J Intell 30(11):110–114