International Journal of Multimedia and Ubiquitous Engineering
Vol.9, No.3 (2014), pp.331-340
http://dx.doi.org/10.14257/ijmue.2014.9.3.31

Innovation Process of Open Source Software based on Knowledge Mapping

Dan Tian

School of Management Science and Engineering,
Dongbei University of Finance & Economics, Dalian 116025, China

tiandan@dufe.edu.cn

Abstract

In recent years, open source software which is a typical form of open innovation has become an important business mode in software industry. However, the innovation process of open innovation in software products is quite different from traditional manufacturing industries. Taking Linux operating system as the research sample, the innovation process of open source software is studied based on knowledge mapping which is made by Citespace. The study shows that technical standards are technology platforms of open source software innovation. User participation is an effective way to knowledge transfer of open source software innovation. Innovative society is an important intermediary of open source software innovation.

Keywords: *open innovation; innovation process; software product; knowledge mapping*

1. Introduction

Open source software is computer software with its source code made available and licensed with a license in which the copyright holder provides the rights to study change and distribute the software to anyone and for any purpose [1]. The development of open source software is typical open innovation for open source can effectively match the widely distributed technical resources with market demand, making it become an important way to development software technology [2]. In recent years, open source software is increasingly been promoted and diffused in the worldwide. Linux operating system breaks the monopoly of PC operating system of Microsoft. Apache accounted for half of the market of web server software. MySQL is becoming the most widely used database system in the world.

Success in the marketplace demonstrates that open source software is not only the change of development thinking and technological development in the software development history, but also a successful business model. China's software enterprises are lagging behind foreign companies in technology base and innovation resource. Open source software provides valuable technical opportunities for the rise of China's software industry, especially the system software. However, the innovation process of open source software products are significant differences from the traditional software products which is in urgent need to carry out in-depth study.

The paper is organized as follows. Based on the state of the art in innovative measure indicators by screening the relevant literature, scientific literature is chosen to be the indicator to measure the open innovation in software industry. Then taking Linux operating system as the research sample, knowledge mapping about the innovation in Linux operating system is made by Citespace and analyzed. Finally, the results are summarized and discussed.

ISSN: 1975-0080 IJMUE

2. Literature Review

Open source software is open-source software and rose in the 1980s. Open source software became an emerging industry quickly for its new forms and competitive [3]. Osterloh and Rota suggests that open source software is an information product, user-driven innovation and highly modular design, but it does not achieve sustainable development just through integrated innovation [4]. Research on open source software main includes the motivation of open source software innovator, innovative organization and processes of open source software projects, competitiveness of open source software [5]. The source codes of open source software are open and any organization or individual can innovate. Therefore, open innovation theory is an important theory for open source software innovation.

Since the 1990s, the increasingly fierce global competition encourages enterprises to integrate internal and external technical resources to improve product development speed and reduce development risk [6]. Chesbrough and Vanhaverbeke pointed out that open innovation accelerate technological innovation and industrialization through making use of inflows and outflows of knowledge [7]. Neyer further pointed out that open innovation is consists of three types of participants; they are the internal core innovators which is usually a research and development department, internal edge innovators which includes marketing, production and other departments, external innovators which includes users, distributors, suppliers and competitors [8]. Toivonen put forward that as a typical knowledge-intensive business services, users are major participants in technological innovation in software industry [9].

In recent years, technological innovation in software industry has gradually attracted the close attention of domestic scholars. Technological innovation and application innovation are necessary to achieve leap-forward development of China's software industry [10]. Wang and Gao prove that the expected future income of the innovation is the fundamental reason for enterprises to participate in open source movement based on the open knowledge disclosure theory [11]. Liu and Chang analyzed characteristics of technological innovation of all kinds of different scale software companies using contingency table and bring forward that the innovation of enterprises should adapt to the regional innovation environment and expanding business scale blindly is not benefit to technological innovation [12]. According to the innovation process analysis of structured and object-oriented software development methods, Gao and Feng proposed four stages in innovation of software development method [13].

3. Methodology

3.1. Knowledge mapping method

The software industry is a typical knowledge-intensive business services (KIBS) and the innovation process involves a number of organizations and individuals. Freel put forward that highly skilled employees is the typical characteristic of KIBS, which usually include scientists, engineers while other users, suppliers and universities are the important external source of the KIBS [14]. Due to the diversity of the innovators in the software industry, it is difficult to statistic the R&D and patent data related to the software product innovation. The empirical research of KIBS accomplished by Hipp demonstrated that journals, conferences and other external explicit knowledge are the important information source in technological innovation [15]. Journals and other scientific literature help enterprises to clear the direction of innovation and enhance the absorptive capacity [16]. There scientific literature was selected to measure the open innovation in software products.

For the inheritance of scientific literature, the scientific literature expresses their recognition of the stock of knowledge through references, which citations reflect the process

of innovation and diffusion of technical knowledge [17]. Therefore, citation analysis methods have been an important quantitative measure method in technology innovation and diffusion. However, with the accelerated development of science and technology, it is hard to analyze the huge amounts of data with the traditional statistical analysis tools.

Citespace is developed by Professor Chen Chaomei of Drakes University using the Java language, which can be applied to diverse, time-sharing and dynamic literature co-citation network analysis. Citespace realized depth analysis of literature. Chen and Hu reveal the emerging field of regenerative medicine through Citespace analysis [18]. Zhao and Wang explored the status quo and hot topic in ubiquitous computing using Citespace [19]. Wu and Huang put forward the method of tracking and analyzing the changes of technology applications through analyzing the changes of the scientific literature over time [20]. Through text mining, Citespace clarify the intrinsic link between literatures and visually display citations between literatures, thus providing analytical tools for the innovation process of open innovation in software products.

3.2. Analysis tools

The most obvious difference between traditional innovation and open innovation is that enterprises should take advantage of the internal and external innovation information when they develop new technologies [21]. Open innovation is no longer confined within the enterprise and the innovation information can also be obtained from outside of the enterprise. Internal and external innovators in open innovation process also make open innovation process presents different characteristics. Therefore, it is necessary to select the appropriate indicators of knowledge mapping to measure the innovators and the innovation process so as to explore key factors that play important roles in the open innovation process.

Citespace reflects the duality between a research front and its intellectual base. If we define a research front as the state of the art of a specialty, what is cited by the research front forms its intellectual base [22]. When the scientific literature become the indicators of open innovation, the scientific literature stands for the outcome of open innovation and the citations represent the innovative information obtained in the open innovation. In the timelines view generated by Citespace, research front is the output of open innovation in different phrases and the citations are the innovators and their contribution in open innovation. The evolution of research fronts over time reflects the open innovation process.

Betweenness centrality metric and burst detection algorithm are two algorithms adopted in Citespace to identify in terms of research front concept and highlight potential pivotal point of paradigm shift over time [22]. Time zone view is generated to display research fronts and the shift process of different research fronts. According to the research goals, time zone of co-cited articles on Linux is generated to analyze the process of open source software innovation. The Time zone of co-cited authors on Linux is generated to discuss characteristics of participants of open source software innovation. Then the characteristics of process for open source software innovation are summarized.

4. Data Analysis

4.1. Sample selection

Linux is a Unix-like computer operating system assembled under the model of free and open source software development and distribution. The defining component of Linux is the Linux kernel, an operating system kernel first released 5 October 1991 by Linus Torvalds. The source code of the Linux operating system is completely open to the outside world.

According to the General Public License (GPL) agreement, anyone can be free to modify and redistribute the Linux. Linux Internationalization Initiative advocated by Sun, IBM and Linux Standard Base is responsible for providing a common, international applications platform for the different Linux distributions to facilitate the development and operation of Linux.

Comparing with Windows operating system developed by Microsoft, the Linux operating system has a more stable structure system, a higher level of security and low price. Therefore, the Linux operating system do not win the support of database software developers such as Oracle, Informix, Computer Associates, Sybase and IBM, but also win the favor of the hardware vendors such as DEC, Apple, Intel and so on. In addition, the Linux operating system also play important roles in the field of embedded system, telecommunications, cluster computing, and network security. The innovation of the Linux operating system is driven by organizations and individuals all over the world. The Linux operating system is a typical representative of the software products, which provide valuable samples for in-depth analysis of open innovation process in software products.

4.2. Data collection

The innovation of the Linux operating system was first initiated by Linus Torvalds. In 1991, Linus Torvalds studied at the University of Helsinki began to develop its own operating system which was called the Linux kernel later and the first version of Linux was released in September, which is widely regarded as the beginning of the Linux operating system. Since then, the organizations and individuals from different parts of the world continue to promote the innovation activities of the Linux operating system. Therefore, 1991 was selected as the start for the data collection in the literature. For that the information of scientific literature citation of 2012 not yet to statistics completely, 2011 was selected as the end for the data collection in the literature.

Web of Science developed by Thomson Reuters is one of the most important databases to access global academic information. The database contains the references cited in papers and makes an indicator for them according to the author, source and publication. A typical representative of the Science Citation Index-Expanded (SCI-EXPANDED), Social Science Citation Index (SSCI) and Arts and Humanities Citation Index (A&HCI) are three typical citation databases. For the purpose of the study is to measure the innovative features of Linux, SCI-EXPANDED database was selected as the download platform.

According to the innovation process of Linux operating system, we selected the "Topic = Linux" as search strategy and search the data on the platform of SCI-EXPANDED database. The time range is limited from January 1, 1991 to December 31, 2011. The document type is English article. Finally, 2399 literatures were collected as research data. The date of download is August 18, 2012.

4.3. Knowledge mapping analysis

Based on the above research, research data was imported into Citespace V.2.2.R1. We specify 3 as the length of a single time slice and the data from 1991 to 2011 were divided into 7 time slices. Top 30 cited articles of each time slice were selected as analysis objects and selected articles in each time slice were no more than 100. Then we operated Citespace and draw the Timelines of co-cited articles on Linux. After that, the nodes in the knowledge mapping were divided into 17 clusters through the function of clustering in Citespace. There are 143 nodes and 313 links in the figure of timelines of co-cited articles on Linux and the value of Mean Silhouette is 0.8659, which indicate that the clusters are divided properly and articles in each cluster are closely related. The value of Modularity Q is 0.7356, which

indicate that the knowledge mapping well reflect the clusters in the scientific literature and easy to identify the pivotal points.

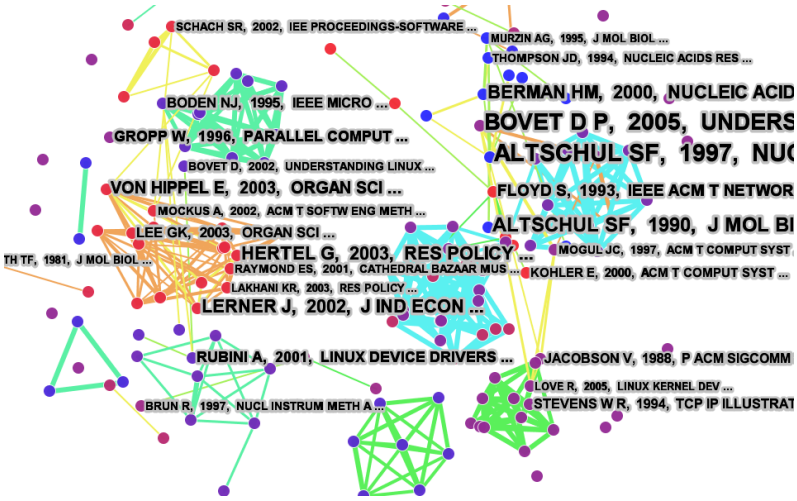

Figure 1. Time zone of co-cited articles on Linux

The analysis of clusters list from Citespace shows the field of flexible graphical user interface and nuclear microprobe were the research focus before 1990. Then the hot topic shift to the field of computer application, software environment and monitor between 1990 and 1995. After that, the field of asymmetric network and differentiated services router become the focus. Since 2000, business and modeling of Linux has become hot research field. The shift of research focus in different time slice is coincided with the actual innovation process in Linux. The initial innovation of the Linux operating system is mainly driven by a number of professional users and the innovation mainly focused on solving the technical problems of the graphical user interface and does not pay attention to the market demand. Then with the support of vendors such as IBM, Oracle, and Sybase, business become the innovation focus of the Linux operating system.

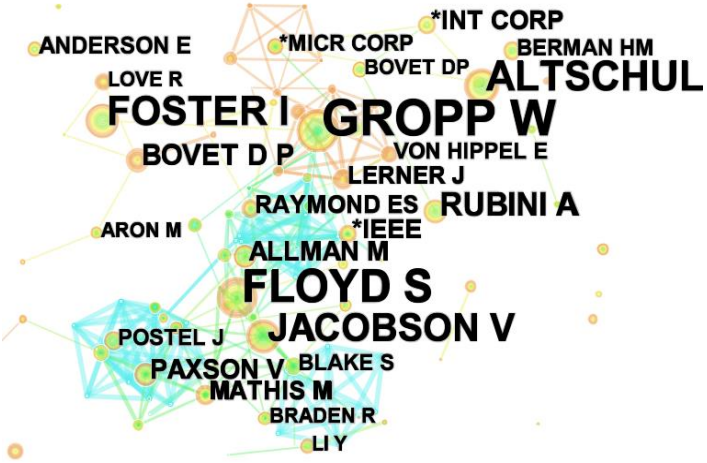

Figure 2. Time zone of co-cited authors on Linux

Figure 2 shows that there are three pivotal points with the highest centrality value. As table 1 shows, core articles are An Architecture for Differentiated Services written by Blake and published in 1998, Jini Technology Core Platform Specification written by SUN Micr Inc and published in 2000, and RRC Protocol Specification written by ETSI and published in 2001. The centrality value of Linux Myths written by Micr Corp and published in 1999, Java Remote Method Invocation Specification written by Sun Micr Inc and published in 1998, and Running Linux written by Welsh and published in 1995 are 0.03. The top 6 pivotal points includes individuals, enterprises as well as Standardization Organizations. And besides Linux Myths, the other points which play important roles in Linux innovation are all technical standards or technical specifications.

Table 1. Core articles of time zone of co-cited articles on Linux

Author	Year	Articles
Blake S.	1998	An Architecture for Differentiated Services
SUN Micr Inc	2000	Jini Technology Core Platform Specification
ETSI	2001	RRC Protocol Specification
Micr Corp	1999	Linux Myths
Sun Micr Inc	1998	Java Remote Method Invocation Specification
Welsh M.	1995	Running Linux

There are 22 burst points in the timelines of co-cited articles on Linux. The highest burst value is 8.19, which is the Understanding the Linux Kenrnel written by Bovet and published in 2005. The book plays a particularly prominent role in the turning point of the Linux innovation. Besides it, four authors' burst values are between 5.00 and 5.99; eight authors' burst values are between 4.00 and 4.99. And the other authors' values are between 3.00 and 3.99. Above all, all the burst point are individuals which means individuals play more important roles in the turning point of innovation in Linux operating system.

5. Discussion

5.1. Technical standards are technology platforms of open source software innovation

Time zone of co-cited articles on Linux shows that technical standards provide technology platforms for open innovation in software products. For instance, An Architecture for Differentiated Services defines architecture for implementing scalable service differentiation in the Internet [23]. Jini Technology Core Platform Specification provides the source code and specifications for Jini technology [24]. RRC Protocol Specification is the radio resource allocation protocol for the access network (UTRAN) and user equipment (UE) in the third generation mobile communication system [25]. These technical standards and specifications ensure the compatible between open innovations of Linux operating system and other software or hardware technology or product, so as to promote the open innovation in software products.

Technical standard refers to a series of files with mandatory requirements or guidance, the contents of which include the technical details or requirements, and its purpose is to allow the product or service meeting certain safety standards or market access requirements. Technical standards codified accumulated technical experience and formed the basis of new technology generated. When innovation is regarded as generation and diffusion process of new knowledge, technical standards become a strong support for innovation [26] . In the process of open source software innovation, technical standards reduce the adverse effects of path

dependence. Meanwhile it provides a public technology integration platform for enterprises to integrate dispersed technical resources to achieve innovation.

5.2. User participation is an effective way to knowledge transfer of open source software innovation

The research sample involves enterprises, standardization organizations and individuals and so on. Burst of articles on Linux show that the burst nodes in the open innovation of Linux operating system are all individuals. Knowledge transfer of open source software innovation achieve through user participation. For example, Understanding the Linux Kernel authored by Bovet enable users to know better of the data structures, algorithms and programming of the Linux Kernel which promote the user innovation of Linux. It indicates that the burst in the open innovation of Linux operating system is mainly caused by the individual, which means that user innovation dominant the open innovation in software products.

In the context of open innovation, global liquidity of technical resources has greatly promoted the rapid growth of new knowledge, which makes it increasingly difficult for a single enterprise entirely on their own to complete the whole product innovation. In the process of open source software innovation, due to different types of users with different application background and strong technology transfer capabilities, they could quickly realize the information search and matching and achieve knowledge creation and transfer. Therefore, User participation is becoming an effective way to knowledge transfer of open source software innovation. Lakhani and Wolf put forward that the feeling of creativity of individual users is the most powerful power in the open innovation [27]. The original innovator of the Linux operating system is experienced and professional technical users, and they tend to be the leading users of the Linux operating system which have fanatical dedication to improving the Linux operating system. They often constantly seek new solutions to satisfy individual special needs. Schreier and Prügl suggest that leading users have more knowledge and consumption experience, so they always with stronger internal control capabilities and innovative spirit, and can adopt more new products [28]. Therefore, user innovation becomes important innovation form of open innovation in software products.

5.3. Innovative society is an important intermediary of open source software innovation

Time zone of co-cited authors on Linux shows that the innovator of the Linux operating system realizes the virtualization cooperation. And the analysis of the time zones of co-cited articles on Linux demonstrates that the virtualization cooperation promote the open innovation of Linux operating system. von Hippel point out that the innovative society refers to a network of interconnected personal or corporate, and the personal or corporate in the network exchange information through face-to-face, electronic or other means [26]. Therefore, the open innovation process of the Linux operating system indicates that innovative society is an intermediary of open source software innovation.

Participants in open innovation process will face many problems such as Arrow Information Paradox, unequal status of technology buyers and sellers, confirming the effectiveness of information source and fostering bilateral market. In the open innovation process of Linux operating system, the innovative societies are mainly composed of users. Baldwin and Clark point out that the efficiency of user innovation societies is higher than that of the single manufacturer-centered innovation model [29]. Uses with different technological background in innovation can improve the efficiency of innovation [30]. As an important intermediary of software product innovation, innovative society establish contacts among

open source software innovators and provide technical support services, thus solving above problems in the open innovation process.

6. Conclusion

With the open innovation has gradually become an important form in software products, the paper measure the process and innovators of open innovation with the indicator of the scientific literature and the open innovation process in software products are studied. The study shows that technical standards are technology platforms of open source software innovation. Open innovation in software products should be supported by technical standards. User participation is an effective way to knowledge transfer of open source software innovation. User innovation is an important form of open innovation in software products. Innovative society is an important intermediary of open source software innovation and the driving force of open innovation in software products.

Acknowledgements

This work is supported by National Natural Science Foundation of China (71002094), National Social Science Foundation of China (Grant No.12BJY013), Ministry of education of Humanities and Social Science project of China (Grant No.12YJA790151）.

References

[1] St. Laurent and M. Andrew, "Understanding Open Source and Free Software Licensing", O'Reilly Media, **(2008)**.
[2] E. von Hippel and J. Chen, "The Major Shift Towards User-Centred Innovation: Implications for China's Innovation Policy Making", Journal of Knowledge-based Innovation in China, vol. 1, no. 1, **(2009),** pp. 16-27.
[3] G. Krogh and E. Hippel, "The Promise of Research on Open Source Software", Management Science, vol. 52, no. 7, **(2006),** pp. 957-983.
[4] M. Osterloh and S. Rota, "Open Source Software Development-Just Another Case of Collective Invention?", Research Policy, vol. 36, **(2007),** pp. 157-171.
[5] G. Krogh and E. Hippel, "Special Issue on Open Source Software Development", Research Policy, vol. 32, **(2003),** pp. 1149-1157.
[6] D. Chatterji, "Accessing External Sources of Technology", Research Technology Management, vol. 39, no. 2, **(1996),** pp. 48-56.
[7] H. Chesbrough, W. Vanhaverbeke and J. West, "Open Innovation: Researching a New Paradigm", Oxford University Press, **(2006)**.
[8] A. K. Neyer, A. C. Bullinger and K. M. Moeslein, "Integrating Inside and Outside Innovators: A Sociotechnical Systems Perspective", R and D Management, vol. 39, no. 4, **(2009)**, pp. 410-419.
[9] M. Toivonen, "Expertise as Business", Long-Term Development and Future Prospects of Knowledge-Intensive Business Services (KIBS), Helsinki University of Technology, Laboratory of Industrial Management, Doctoral dissertation series, **(2004)**.
[10] Y. Qin, G. Zhang and K Wang, "The Innovation and Development of Chinese Software Industry Based on Software Outsouring", Science of Science and Management of S and T, vol. 1, **(2006),** pp. 126-131.
[11] Q. Wang and S. Gao, "Study on Dynamic Games Model of Open Source Innovation based on Open Knowledge Disclosure", Journal of Industrial Engineering/Engineering Management, vol. 24, no. 4, **(2010),** pp. 104-109.
[12] S. Liu and Y. Chang, "Characteristics of Technological Innovation of Different Scale Software Enterprises in China", Science of Science and Management of S and T, vol. 7, **(2004),** pp. 21-23.
[13] Y. Gao and X. Feng, "Research on the Innovative Evolution Process of the Software Development Method", Studies in Science of Science, vol. 27, no. 2, **(2009),** pp. 180-186.
[14] M. Freel, "Patterns of Technological Innovation in Knowledge-Intensive Business Services", Industry and Innovation, vol. 13, no. 3, **(2006),** pp. 335-358.
[15] C. Hipp, "Knowledge-intensive Business Services in the New Mode of Knowledge Production", AI & Soc, vol. 12, no. 1/2, **(1999),** pp. 88-106.

[16] J. W. Spencer, "Firms' Knowledge-sharing Strategies in the Global in Novation System: Empirical Evidence from the Global Flat Panel Display Industry", Strategic Management Journal, vol. 24, **(2003),** pp. 217-233.
[17] M. Weinstok, "Encyclopedia of Library and Information Science", Philadelphia: ISI Press, **(1977)**.
[18] C. Chen, Z. Hu, S. Liu and H. Tseng, "Emerging Trends in Regenerative Medicine: A Scientometric Analysis in CiteSpace", Expert Opinions on Biological Therapy, vol. 12, no. 5, **(2012),** pp. 593-608.
[19] R. Zhao and J. Wang, "Visualizing the Research on Pervasive and Ubiquitous Computing", Scientometrics, vol. 86, **(2011),** pp. 593–612.
[20] F. Wu and L. Huang, "Analysis on the Change of Technologies Domain of Application Based on Discipline Changed", Studies in Science of Science, vol. 30, no. 4, **(2012),** pp. 503-511.
[21] H. W. Chesbrough, "Open Innovation: The New Imperative for Creating and Profiting from Technology", Harvard Business School Publishing, **(2003)**.
[22] C. Chen, "Cite Space Ⅱ: Detecting and Visualizing Emerging Trends and Transient Patterns in Scientific Literature", Journal of the American Society for Information Science and Technology, vol. 57, no. 3, **(2006)**, pp. 359-377.
[23] S. Blake, "An Architecture for Differentiated Services", http://www. ietf.org/rfc/rfc 2475.txt, **(1998).**
[24] SUN Micr Inc., "Jini Technology Core Platform Specitication", http://www-csag.ucsd.edu/ teaching/ cse291s03/Readings/core1_2.pdf, **(2000).**
[25] ETSI, "RRC Protocol Specification", **(2001).**
[26] H. A. Robert and D. S. Ram, "The Role of Standards in Innovation", Technological Forecasting and Social Change, vol. 64, **(2000),** pp. 171–181.
[27] K. R. Lakhani and B. Wolf, "Why Hackers Do What They Do: Understanding Motivation and Effort in Free/Open Source Software Projects", In Feller J, Fitzgerald B, Hissam S, Lakhani KR, eds., Perspectives on Free and Open Source Software, MIT Press, **(2005)**.
[28] M. Schreier and R. Prügl, "Extending Lead-user Theory: Antecedents and Consequences of Consumers' Lead Userness", The Journal of Product Innovation Management, vol. 25, no. 4, **(2008),** pp. 331-346.
[29] E. Von Hippel, "Democratizing Innovation", The MIT Press, **(2005)**
[30] C. Y. Baldwin and K. B. Clark, "Does Code Architecture Mitigate Free Riding in the Open Source Development Model?" Management Science, vol. 52, no. 7, **(2006),** pp. 1116-1127.

Author

Dan Tian

She received her PhD in Management (2008) from Dalian University of Technology. Now she is an assistant professor of management science at school of Management Science and Engineering, Dongbei University of Finance and Economics. Her current research interests include technology management of software industry and risk management.