

Research and Analysis on Software Project Risk Based on Bibliometrics

Ying Qu

School of Economics and Management
Hebei University of Science and Technology
Shijiazhuang, Hebei, China
Quying1973@126.com

Moran Fan

School of Economics and Management
Hebei University of Science and Technology
Shijiazhuang, Hebei, China
1870212671@qq.com

Abstract—The research on software project risk is the key problem in software project. A correct understanding of software project risk plays a decisive role in the success or failure of software project. In this paper, we use the related technology of crawler to crawl out the risk-related documents of software projects and analyze the word frequency of the information such as title, abstract and keywords of the documents. The main purpose is to analyze the source of the documents, discuss and summarize the hot research issues and research methods, and summarize and analyze the software projects in China. The purpose of this paper is to discuss the characteristics of software project risk research and the problems that need to be paid attention to, and to provide scientific basis and ideas for China's software project risk research.

Keywords—Software project risk; Crawler technology; Bibliometrics; Research trends

I. INTRODUCTION

In the 1980s, Boehm's monograph "Software Risk Management" was published, which laid a theoretical foundation in this field [1]. His outstanding contribution is to lead people into the research field of software project risk management and put forward the risk identification idea of TOP10 risk list. At the same time, Charette constructed similar risk management structure from different angles. SEI, as the largest base for research and practice of software risk management, considers that software risk assessment, continuous risk management and team risk management constitute the basic framework of software project risk management. Maryland University put forward the Riskit model, which aims to fully reflect and manage the causes of risk, trigger events and their impacts, and to assess risks in a certain step. Petra M. Bosch-Sijtsema et al. proposed how to share and generate practice-based and distributed project management through inter-organizational project interactions[2]. Hall's six-discipline model, which considers the combination of risk management and project management and pays attention to risk measurement and control, is a useful attempt to combine theory with practice. Microsoft believes that risk identification and management should be paid attention to at all stages of the project life cycle. It divides risk management into five steps: risk identification, risk analysis, risk activity planning, risk tracking and risk control. It emphasizes that this is a sequence of activities that are repeated and continuous in the process of project development.

Based on the above classical theories, Chinese scholars have gradually carried out the research on software project risk since 2003, and have achieved many results, which has brought tremendous impetus to the development of software industry in China. Documents in this study from theory to practice, from broad to concrete, from shallow to in-depth. This paper crawls out the relevant literature on software project risk from 2003 to 2017, and systematically analyses the changes in the number of documents, main research contents and main research methods in China and foreign countries. It also focuses on the research specialty of Chinese literature. The purpose is to find out the research characteristics of software project risk and the questions needing attention. It also provides reference for software project risk research.

II. ANALYSIS ON THE LITERATURE SOURCE OF SOFTWARE PROJECT RISK RESEARCH

The source analysis of literature is to analyze the basic situation of literature, including its author, publishing time, author's specialty, unit and so on. This paper mainly analyses the quantity of literature, unit and author's specialty. The purpose of this paper is to find out the changing rule of the number of publications of software project risk literature, the situation of the main supporting units and the basic situation of the main research fields in this field.

A. The Source and Acquisition Method of Document Data

In this paper, the self-defined acquisition process of Octopus is adopted. The crawling platform of Chinese literature is the knowledge retrieval platform (<http://search.cnki.net/>) of China Academic Journal Full-Text Database (CNKI). The subjects are "software project risk" and "IT project risk factors". Among them, the contents obtained by climbing include title, abstract, keyword, author, tutor, author's unit, author's specialty, publishing year, author's specialty and citation times. The total number of documents climbed out is 1030. All the Chinese literature years are from 2003 to 2017. The types of documents include: academic journals, doctoral dissertations, Master's thesis, conference paper, scientific and technological achievements, Chinese patent.

B. Document time distribution

There are 709 Chinese documents annotating the publication time, and the number of Chinese documents increases first and then decreases. The reasons for this phenomenon can be summarized as follows: First, with the deepening and development of research, the risk of software projects in China has changed from vague unknown to enough understanding. Full recognition of software project risk makes it possible to identify, control and even manage software project risk effectively, and then avoid software project risk effectively, and ultimately reduce the risk of software project, leading to the gradual cooling of research hotspots. Therefore, there is a decreasing trend in the number of Chinese literatures. Secondly, there are many homogenization studies, and it is not difficult to find more evaluation articles in the crawling results of Chinese literature. This shows that the research of software project risk mainly focuses on the evaluation after identifying the existing risks. The research of unknown risks is relatively deficient and difficult to discover, which leads to the research of existing risks. Relative saturation has been reached, and the research on uncertain risks is not timely enough.

C. Analysis on Distribution of Literature Authors and Professionals

The author's units in Chinese literature are analyzed for word frequency (as shown in Figure 2), and the frequency of different units (universities) is analyzed. Based on high frequency words, the hot research units (universities) related to software project risk are obtained. The higher frequency units are: Beijing University of Posts and Telecommunications,

ShanghaiJiaotong University, and Telecommunications. Zi University of Science and Technology, Nanjing University of Technology, Tianjin University, Huazhong University of Science and Technology, etc. Most of the literature of Beijing University of Posts and Telecommunications is mainly about the theoretical research and evaluation of software project risk management, among which Chen Shuangquan, Wu Jinyun and others should start from enterprises or government departments to evaluate software project risk [3][4]; Nanjing University of Technology also tends to evaluate risk in the research of software project risk. In terms of evaluation, Chen Yue et al. used evaluation methods to study software project risk [5].

Author's Professional Analysis: From the display of the author's professional word clouds of Chinese literature (as shown in Figure 3), the major occupying a larger proportion are: project management, management science and engineering, business management, project management (professional degree), software engineering, business management (professional degree), business management, computer application technology, software. Engineering (professional degree), computer software and theory, Master of Engineering (professional degree), technology, economy and management, etc. Among the majors with more than 10 publications, the management majors accounted for 66.7%, and the computer majors accounted for 33.3%. It can be explained that in the research of software project risk, management specialty has a large proportion of research on software project risk management, while computer technology-based research is relatively weak.



Fig. 1 Chinese literature source word cloud

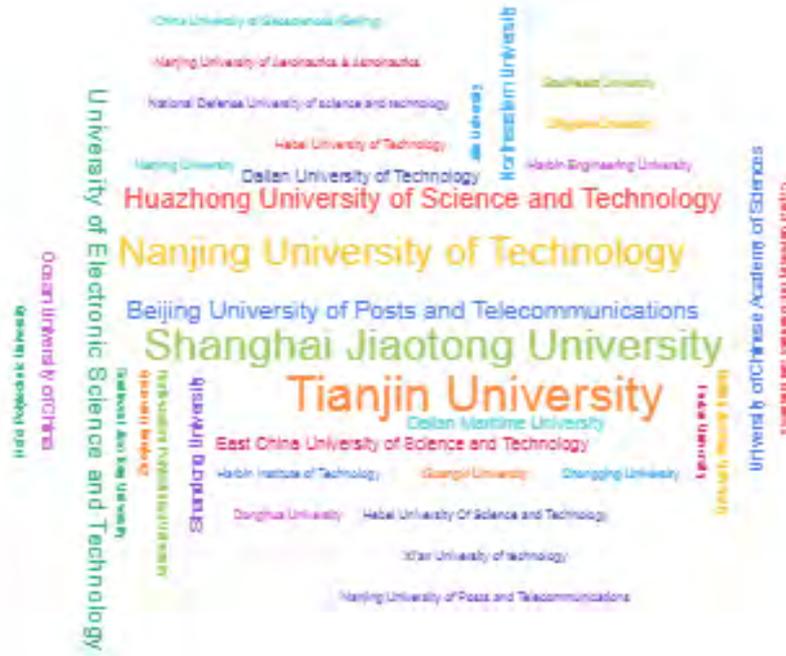


Fig. 2 Chinese literature author's professional word cloud

III. HOTSPOTS AND METHODS OF DOCUMENT RESEARCH

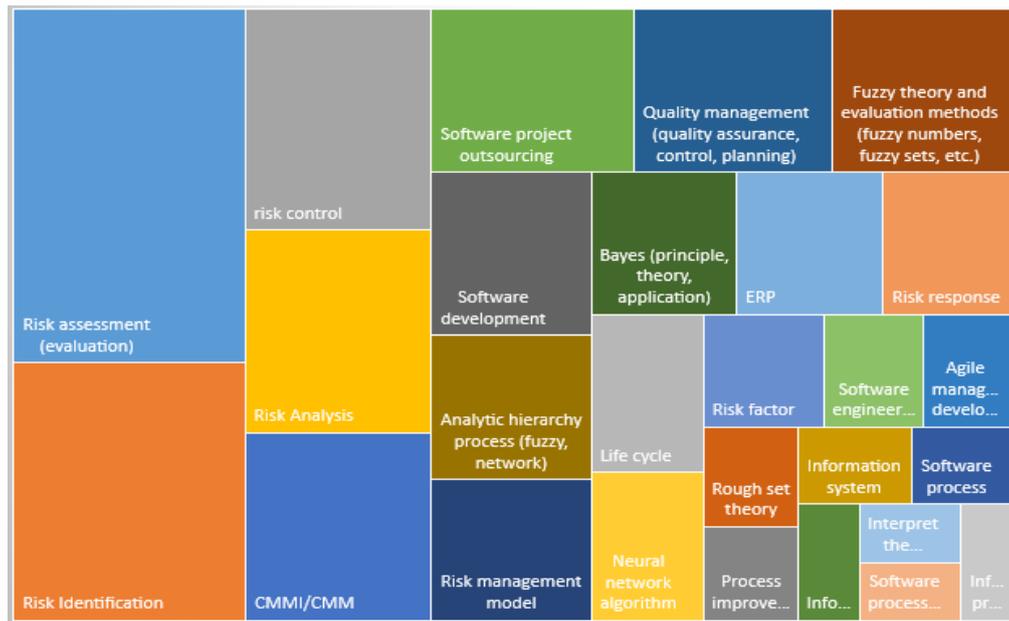


Fig. 3 Chinese literature keyword tree diagram

All keywords, abstracts and titles in Chinese literature are sorted out, and word frequency analysis of keywords, abstracts and titles is carried out by using tree graph in ECXEL365 (as shown in Figure 4). In the tree diagram, we can clearly get the research hotspots and main methods of Chinese literature related to software project risk from 2003 to 2017. Based on the frequency of keyword, title and summary segmentation, the research hotspots and research methods of software project

risk are obtained. Higher frequency words are risk assessment (evaluation), risk identification, fuzzy theory, software development, analytic hierarchy process, Bayesian, neural network, risk factors, etc.

A. Hotspot Analysis of Literature Research

The research hotspots of Chinese and foreign literatures based on high-frequency words are almost the same, which mainly include: identification of software project risk, evaluation of software project risk, control of software project risk, response of software project risk, management of software project quality, process and process improvement of software project, risk. Factors, construction of software projects, agile development projects, future development prospects, etc.

1) Research on Risk Identification of Software Projects

In the literature of risk identification, Cao Xianye and others mentioned that the identification of risk factors is the first step in risk management. The author proposed a framework for identifying risk factors in IT outsourcing. The application of this framework found 10 new risk factors and formed Chinese conditions. List of risk factors under the circumstances, and rank 17 software project risk factors [6]. Gaoyi et al. designed a set of reasonable and practical risk identification and response system, which combines with interviewing experts and organizing special conferences, to scientifically and systematically identify and respond to software project risks [7]. The overwhelming majority of the literature studies in the form of software project risk assessment, in which the identification of software project risk as part of the evaluation study, lays the foundation for the risk assessment of software project.

2) Risk Assessment of Software Projects

There are a lot of Chinese documents to evaluate the risk of software project, 106 articles to identify the risk clearly before the risk evaluation of software project, 144 articles to control and analyze after the risk evaluation of software project, among which there are many unspecified documents to illustrate the risk identification. Risk assessment and final risk control and analysis are mutually reinforcing. With the accurate identification of risk, evaluation is meaningful. The result of evaluation is to control and analyze the risk of software project better. If there is no final risk control in the analysis, evaluation will lose its value.

In these literature on risk assessment of software projects, there are many corresponding risk analysis, control and suggestions after evaluation. The ultimate purpose of risk assessment of software projects is to analyze and control the risk of software projects. The source of risk of software projects is the premise of evaluation, and only when the source of risk of software projects is clear can the risk of software projects be built. Establish an accurate index system framework.

3) Research on Risk Control of Software Projects

There are 24 separate studies on risk control in Chinese literature. Among them, Feng Zhen et al. established a fuzzy linear programming model for risk control scheme design. The goal is to establish a risk control scheme, and to minimize the risk of software projects under the constraints of risk control cost [8]. Xuan Chuncheng and others have integrated the four dimensions of IT governance mechanism and control methods, including risk control objectives, risk control subjects, risk control activities and risk control procedures, to build a risk control system for IT projects, aiming at achieving good implementation results for IT projects [9]. The risk control of software project is mostly a part of the risk research of software project. Correct understanding of risk is the premise of software project control, and all of them are for better operation effect of software project.

B. Analysis on Literature Research Methods

The research methods based on high frequency words mainly include: fuzzy theory and evaluation methods (fuzzy numbers, fuzzy sets, etc.), analytic hierarchy process, network analysis, risk management model establishment, Bayesian theory, ERP, neural network algorithm, agile management and development, rough set theory, interpretive structure model, etc.

Among the above methods, fuzzy theory, Bayesian, neural network, analytic hierarchy process, rough set theory and other methods are mostly used for software project risk assessment. Rao Zhengzhao et al. have assessed the risk of software projects, mainly focusing on the research of software requirement risk analysis relying on measurement technology. Therefore, the improved grey clustering method is used to assess the risk, and the feasibility of its assessment is studied [10]. Xu Jiazhen and Fang Anru used Bayesian theory to evaluate software project risk [11]. Huang Cunxin et al. used the method of neural network to evaluate the risk of software project [12].

IV. CONCLUSION

As an industry bred in the information age, software project will face another challenge with the rapid development of information. The existing research has made great progress, and the research on software project risk has gradually become mature. However, from the typical analysis above, there are still some problems such as high complexity, long cycle and large scale of software project risk, some of which are determined by the characteristics of software project and software project risk, such as the high complexity of software project. But there are still some unsolved problems. At the same time, under the new mode of big data and cloud computing, there is no doubt that it brings new challenges to software users and developers.

ACKNOWLEDGMENT

It gets help from the project of Hebei Education Department of China (ZD2017027) and is also supported by the project from Social Science Foundation of Hebei Province of China (HB18GL045).

REFERENCES

- [1] Boehm, B.W. Software risk management. Washington D.C: IEEE Computer Society Press, 1989.
- [2] Petra M. Bosch-Sijtsema, Lars-Henrik Henriksson. Managing Projects with Distributed and Embedded Knowledge through Interactions. *International Journal of Project Management*. 2014(32) :1432-1444.
- [3] Chen Shuangquan. Risk management of hotel informatization construction in Guangdong Mobile Training Institute. Beijing University of Posts and Telecommunications, 2012.
- [4] Wu Jinyun. Software project risk management. Beijing University of Posts and Telecommunications, 2012.
- [5] Chen Yue. GJB5000A based H enterprise software project management process improvement. Nanjing University of Science and Technology, 2013.
- [6] Cao Xian Ye. Research on the risk factors of IT outsourcing based on Chinese situation. Central South University, 2009.
- [7] Gao Yi. BN company YD_OA project risk identification and response research. Shandong University, 2013.
- [8] Feng Zhen, Zhang Suidi. Software project risk control scheme design. *Computer system application*, 2010, 19 (06): 49-51.
- [9] Qin Chun Cheng. Research on risk control of IT project in construction enterprise group. Guangxi University, 2016.
- [10] Rao Zhengzhao, Pu Tianyin. Application of Grey Cluster Evaluation in Software Demand Risk Assessment. *Coal Technology*, 2012, 31(12): 202-203.
- [11] Xu Jiazhen. Research on project family risk assessment based on Bayesian network. Hunan University, 2014.
- [12] Fang Anru. Empirical study on ERP system application for Chinese enterprise. Harbin University of Technology, 2010.