

Association for Information Systems AIS Electronic Library (AISeL)

WHICEB 2017 Proceedings

Wuhan International Conference on e-Business

Summer 5-26-2017

Big Data-driven Technology Innovation: Concept and Key Problems

Meng Li

School of Management, Wuhan University of Science and Technology, China

Zhiqing Zhang

School of Management, Wuhan University of Science and Technology, China, zhiqing.zhang@qq.com

Zhuqing Hu

School of Management, Wuhan University of Science and Technology, China

Follow this and additional works at: <http://aisel.aisnet.org/whiceb2017>

Recommended Citation

Li, Meng; Zhang, Zhiqing; and Hu, Zhuqing, "Big Data-driven Technology Innovation: Concept and Key Problems" (2017).
WHICEB 2017 Proceedings. 17.

<http://aisel.aisnet.org/whiceb2017/17>

This material is brought to you by the Wuhan International Conference on e-Business at AIS Electronic Library (AISeL). It has been accepted for inclusion in WHICEB 2017 Proceedings by an authorized administrator of AIS Electronic Library (AISeL). For more information, please contact elibrary@aisnet.org.

Big Data-driven Technology Innovation: Concept and Key Problems

Meng Li¹, Zhiqing Zhang², Zhuqing Hu^{3*}

¹School of Management, Wuhan University of Science and Technology, China

²School of Management, Wuhan University of Science and Technology, China

³School of Management, Wuhan University of Science and Technology, China

Abstract: In the background of big data, technological innovation has met some new opportunities and challenges. Based on expounding the concept and key technologies of big data, the concept, main data resources and characteristics of data driven technological innovation are analyzed. And some key problems of data driven technological innovation are discussed from technological and management perspective. From technological perspective, the key processes of data-driven technological innovation such as data acquisition, data processing, technology opportunity discovery and identification technology etc. are discussed and a technological framework is proposed based on Hadoop ecosystem. From management view, the idea of using big data to carry on operation and decision, matched decision-making culture and appropriate process, overall planning of large data applications and the stage of the target are analyzed as the main factors to affect the data-driven mode realization. This study has good values to the enterprises to build data driven innovative mode.

Keywords: big data; data driven technological innovation; multi-sources data; concept; key problems

1. INTRODUCTION

As the basic issue of technological innovation, the research on the motivation and driving factors of technological innovation has attracted much attention. Early innovative driving force theory mainly focused on the role played by technology supply and demand in the market for innovation. Classical innovative driving force theory is a theory that explains the theory of the relationship between the technology market and innovation, and it mainly contains technology-push theory, market-pull theory and dual power theory with technology and market^[1]. Domestic scholars divided technological innovation driven factors into internal drivers and external drivers, and they thought that technological innovation is the result driven by internal and external factors^[2]. In recent years, with the development of the theory of technology innovation driven factors, some new drivers and models are derived, such as design-driven technology innovation^[3], data-driven technology innovation^[4], etc.

As an important strategic resource, data implies great economic value. The study found that the management activities under the guidance of the data are spreading in the domestic and international business community, and the value of the data on the enterprise innovation is also widely noted by the theoretical circles, the industry, and even the government^[5-8]. For example, the U.S. government actively encourage the academic community and enterprises to pay attention to the analysis and use of massive data, and promote the research and application of development of new products and services^[5]. Domestically, much scholars are in the positive attention and discussion of opportunities and challenges for the enterprise innovation brought by big data, such as the report named "exploring the data-driven innovation as a new source of growth" by OECD^[6], the academician Roundtable meeting about "innovation driven by big data" held in Shanghai Expo in 2013 November^[7], the special topic seminar about "data driven enterprises" held by UFIDA company^[8] etc. In addition, Jay Lee and other scholars researched 179 large companies and found that those company who use data-driven decision-making model can increase productivity 5% or 6%, which can't be explained by other factors^[9]. O'Reilly company asserted that^[10]: "the data is the next' inside Intel', the future belongs to the

* Corresponding author. Email: zhiqing.zhang@qq.com (Zhang Zhiqing)

company and people who convert the data into the product."

As a "panning technique", Big data technology can extract value in multi-source, low value and low quality data through the integration and quality control of large data^[11]. Therefore, technological application of big data in the field of technological innovation can contribute to enhancing the intelligence, scientific and initiative in corporate technological innovation opportunity recognition and eliminating the subjectivity and preference in technology innovation opportunities found in a certain extent, and have good research value.

2. THE CONCEPT AND CHARACTERISTICS OF DATA-DRIVEN TECHNOLOGY INNOVATION

From the point of view of origin, data-driven is not a new concept, which has been widely applied in the field of computer, linguistics, control engineering, manufacturing industry, power and equipment fault diagnosis^[12]. And the data-driven technology innovation is developed with the propose of the concept of big data. The data-driven technology innovation mainly refers a process that the data that is changing in the environment or management leads to demand for new technology and evolutes along the supply chain evolution, thus promotes the enterprise to conduct technology innovation^[13]. Data-driven technology innovation is different in content and paradigm, and it is a new model of technological innovation from intuition to scientific decision-making, which has the following characteristics:

(1) Through searching, comparing, clustering and classifying huge amounts of data, we can greatly enhance the intelligence and initiative of the management of technological innovation and opportunity discovery and identification and improve the ability to judge on the market.

(2) The data from the network and other media industry statistics contains information about products and technological innovation. The analysis of such information can help enterprises to find the point of technical innovation and make the right decisions.

(3) Enterprise technological innovation includes product innovation, elements innovation and elements combination innovation. As for most driven models, data is behind the design pattern. Data-driven is on the opposing side, which can also conduct technology innovation by comprehensive method based on business-driven, relationship-driven and data-driven.

3. THE MAIN SOURCES OF DATA THAT CAN DRIVE TECHNOLOGY INNOVATION

The data that can drive the technology innovation has the characteristic of multisource and heterogeneity, it mainly contains six data sources as follows:

(1) The data from the policy changes in the business-operating environment. Such as the five-year plan of national level and long-term planning involved the development of the industry. This kind of data have significant long-standing, directivity and strategic influence on the enterprise decision-making of technological innovation, which is mainly applied in the aspect of supporting the decision-making in the key technological innovation of enterprises.

(2) The competitive intelligence from competitors' product innovation and technological innovation. The competitive intelligence is a kind of information that companies are more concerned about. Through the analysis in view of the main competitor intelligence, companies can be more clearly in the forefront of technology innovation, and then combine with the opponent and its structure and technical features information to make integrated decision, so as to make the right choice.

(3) The data from customer complaints and expectations. Such kind of data forms complex and occurs mainly in blogs, forums and message system. The content of this kind of data mainly discusses about a aspect of a kind of product or a product, such as the complaints in automobile gearbox and engine. Enterprises can find the lack of products through statistical clustering the different user ID complaints. In addition, through the

sentiment analysis of social media, enterprises can conform the brand view of the potential customers and the reaction to products and the behavior from different customer groups and stakeholders, and recognize the characteristics, situation or the response to activity of emotional expressing to guide enterprises in technological innovation.

(4) The tendency descriptive data in the development of technology. Generally, this kind of data is aggregated data. The form is mainly professional analysis reports. Through statistical computations of representative company from different industries, we can find out the trend of technology in industry, so as to guide the technology development of the whole industry and guide enterprises in technological innovation decision-making.

(5) The new product development data from the lower reaches of the supply chain. Such information has significantly characteristics that evolves along the supply chain. For example, through analyzing the data, we can find that the new models in automobile industry generally shows a reduction trend, which gives the iron and steel industry in upstream a suggestion that they should increase investment in high strength steel R & D to meet the technical requirements of the downstream market.

(6) The operating data of changes in the sales from the organization. Internal data is both directly and indirectly for enterprise technological innovation. When the sales change, you can find its incentive through the market analysis. If it is caused by technical backwardness, it should carry out technological innovation or improve other elements.

From the above analysis, it is not difficult to find that: The data used in driving technological innovation mainly comes from external and social fields and has larger inaccessibility and uncertainty. Internal data are mainly used for supporting decisions of technology innovation of external data driven, it can also conduct enterprise technology innovation in a locally driving way. However, as for driving ability, Internal data is weaker than external data.

4. ANALYSIS ON THE TECHNOLOGICAL FACTORS OF DATA-DRIVEN TECHNOLOGY INNOVATION

How to make use of the data to drive the enterprise to carry on the technical innovation is a difficult problem that involves many dimensions such as technology and management. From a technical point of view, it is mainly about how to obtain, process and apply the data to carry out the key problems such as technical opportunity discovery.

(1) Data acquisition technology. From the point of view of data sources, there are two aspects. On the one hand, enterprises can produce "super size" or "massive data" in the process of production and service. On the other hand, the Internet contains rich social value and commercial value and scientific research value of big data^[14]. Regardless of what kind of sources, the data need to be converted into a form that can be used by decision-making layer through data mining, screening, and finishing techniques. In specific applications, the main use of technology for data acquisition is web crawler. But due to the different structure of the different data sources, we need to carry out targeted crawler design.

(2) Data processing technology. The data that can drive technological innovation have many characteristics such as multi-source heterogeneity, interactivity, timeliness, sociality, sudden, dynamics, high noise and so on^[15]. In addition, the unstructured and qualitative data is too much and has higher real-time and lower value density. Those characteristics lead such kind of large data more difficult to analyze. Based on this, there is an urgent need to do some research for the data that can drive technological innovation such as simplification, storage and so on. The research is mainly contains carrying out data simplifying and cleaning by the subject oriented data analysis and enhancing data storage and invoked flexibility by using the current popular NoSQL database to

store relevant data.

(3) Technology opportunity discovery and identification technology. When we have a large amount of data, you need to unearth the trends of technological innovation, opportunity and other relevant information through text analysis and cluster method^[16]. But this is very challenging work, which need to be supported by natural language processing, text analysis and other methods. In specific applications, we can consider the use of the subject-oriented large data analysis and conduct the recognition and identification of technological innovation opportunities through the feature extraction, dimension segmentation and algorithm improvement^[17]. In addition, the visualization of the results is also a key issue in the applications.

5. THE KEY PROCESSES AND TECHNIQUES OF DATA DRIVEN TECHNICAL INNOVATION

How to use the multisource data to drive technological innovation is the key problem of this research. The main processes and techniques of data-driven technological innovation is shown as Figure 1.

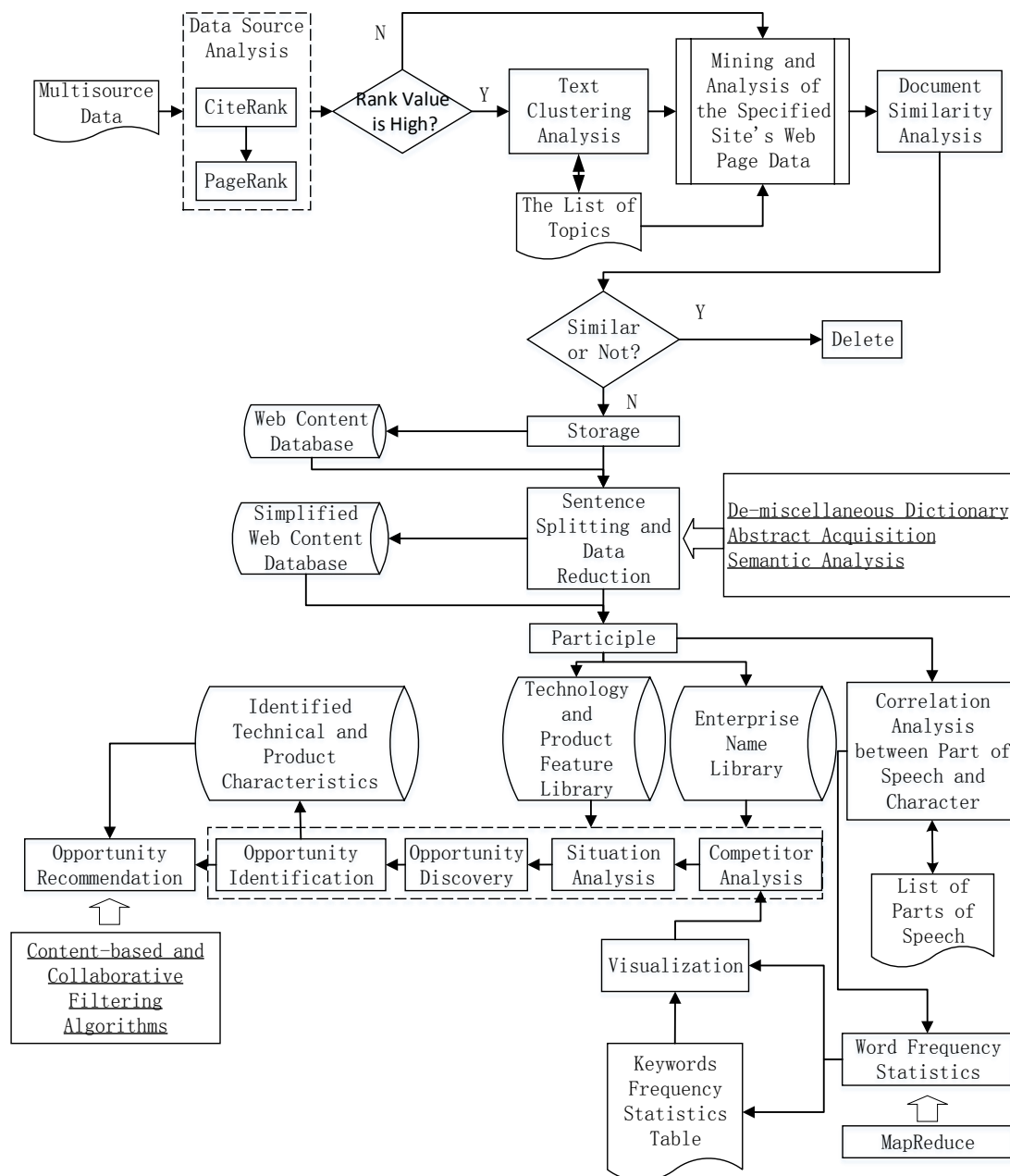


Figure 1. The main processes and techniques of data-driven technological innovation

Firstly, aiming at the dynamic characteristics of the large data of competitive intelligence, this paper will focus on the web information traversal algorithm, and enhance the pertinence and accuracy of the analysis through the theme identified of the document. Though selecting a representative website (page) as the data source and classifying the corresponding web pages and websites based on the relevance of the topic, this paper uses the higher rank value as the training sample for cluster analysis and divide it into several topics. And then, on this basis, the analysis of other sites (web) content is analyzed, the "website - articles - sentence - word" four cascade management method is adopted, and the keyword is attached as an attribute to the appropriate website, sentence, sentence, or vocabulary.

Secondly, in recent years, according to the practice of data mining, it is found that the density of valuable information is very low and there are a lot of redundant information, which leads to that the mining data must be simplified and conducted noise reducing to improve the efficiency of data processing. In the document and sentence level, this paper intends to use a similarity calculation method based on frequency statistics combined with the abstract acquisition and semantic analysis to simplify data and remove redundant data. At the lexical level, noise data is removed through self-learning and expert method to set up complex dictionary. MapReduce computing framework will make full used and Hadoop large data processing technology will be adopted in data processing. The framework of opportunity discovery and recognition system of technology innovation based on big data analysis as shown as Figure 2.

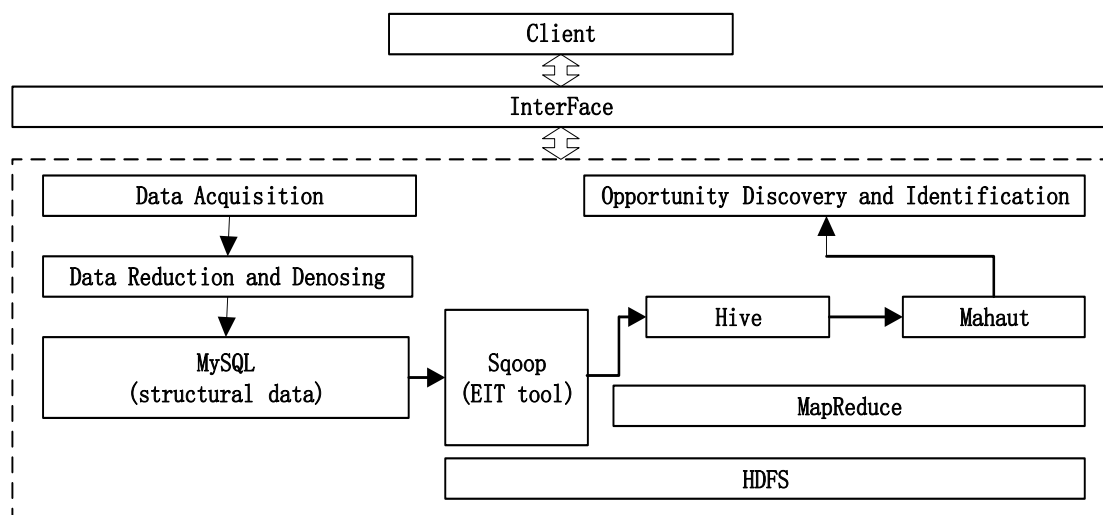


Figure 2. Framework of opportunity discovery and recognition system of technology innovation based on big data analysis

Thirdly, this study intends to use the utility matrix to predict the degree of membership of an enterprise for a technical feature, and to achieve the recommendation of technology opportunity based on the content and collaborative recommendation algorithm. The basic idea is: Firstly, identifying the enterprise (C_1, C_2, \dots, C_n) and technical features (T_1, T_2, \dots, T_m) frequency through large data analysis, constructing relational data pair and utility matrix at different time coordinates, obtaining the values (b_{ij}) by dividing the numerical value of each column with the maximum value of a column, and converting it into an integer value between [0,5] according to a certain rule to express the degree of membership and correlation between technical characteristic and the enterprise's. Then, analyzing the similarity between different enterprises based on the attributes of product structure and scale, and calculating the similarity between different features by clustering analysis of technical characteristics to predict the degree of membership of an enterprise for a technical feature (transform value of 0). And finally, automatically setting the threshold or sorting from large too small to achieve automatic recommendation of technology opportunity.

6. ANALYSIS ON THE MANAGEMENT FACTORS OF DATA-DRIVEN TECHNOLOGY OVATION

From a management point of view, doing a good job of data-driven technology innovation requires a number of synergies:

(1) The idea of using big data to carry on operation and decision is needed. In addition to the attraction from technology itself, the application of a new technology depends to a greater degree in the cognition and acceptance of management layer, as well as the use of large data. Due to a lot of uncertainty and high risk in the large data applications, the attention and guidance from management layer, especially the leadership, becomes more important.

(2) Matched decision-making culture and appropriate process are needed for support. Through the study, McAfee and Brynjolfsson (2012) found that effective application of big data in the management need managers to change the decision-making culture such as organizational leadership, management talent, technology, decision-making and enterprise culture and so on, which is because the organizational decision-making culture has very important value for innovation performance^[18]. Gandomi and Haider (2015) found that only by the effective flow can a large number of fast moving and diverse data be transformed into meaningful insights. Therefore, if companies want to realize data-driven technology innovation, changing the organization's decision-making culture and building a reasonable and feasible decision-making process are needed^[19].

(3) Developing the overall planning of large data applications and determining the stage of the target and the key areas of application are needed. Based on the research of more than 1000 experts and technical personnel in the Global, IBM Institute for Business Value and Sa ïl Business School of University of Oxford proposes five suggestions for big data applications. The core is regarding customers as the center to develop the strategic planning and comprehensive enterprise blueprint of big data, and starting from the existing data to set and complete the periodical strategy target, and then gradually establish a system analysis and increase the analytical ability of big data according to business priorities^[20].

(4) Necessary performance management is needed to realize the model of data-driven technology innovation. The biggest feature of big data applications is the big uncertainty. Therefore, it is necessary to develop measurable indicators to analyze investment yield of big data, analysis and control the risk, and strengthen the decision-making management. Only by possessing these abilities can big data applications take effect, provide comprehensive, unified and accurate information for decision-maker and help enterprises derive insights and explore business opportunities and value in the explosion of data.

7. CONCLUSIONS

Enterprises have an urgent desire in how to improve the ability of technological innovation, especially the ability of discovering the innovation opportunity actively. Big data provides a good opportunity for technology innovation. But how to use big data to carry out technology innovation is the difficult and key point of the research and application. This paper has conducted the preliminary discussion on the origin, concept, characteristics and key issues of data-driven technology innovation to open the research door of data-driven technology innovation and lay the foundation for the enterprises to realize the model of data-driven technology innovation. However, there are too many questions need to be explored for data-driven technology innovation under the background of big data. More in-depth research in the contents, ways, methods and models of construction are needed to clear acceptance characteristics of different types of enterprises for big data and the critical elements of realizing data-driven technological innovation, and then to better guide the practice of technological innovation.

ACKNOWLEDGEMENT

This research was supported by the Natural Science Foundation of Hubei Province under Grant 2015

CFB564 and the Humanities and Social Science Planning Foundation of Ministry of Education under Grant 15 YJA630103 and the Humanities and Social Science Key Program of Ministry of Education under Grant 17D008.

REFERENCES

- [1] Wang Haishan. (1992).Theoretical model of technological innovation motive mechanism[J]. *Science Technology and Dialectics*, 9(6): 22-27+64(In Chinese)
- [2] Zhang Guoqiang, Feng Taozhu, Wei Jujin. (2010).The theory discussion about the power of enterprise technology innovation [J]. *Technology and Innovation Management*, 31(1): 23-26(In Chinese)
- [3] Chen Xin, Chen Guodong. (2012).Three dimensional network of the innovation drive: a meta model——theoretical integration of design-driven innovation and technological innovation[J]. *Studies in Dialectics of Nature*, 28(5): 66-71(In Chinese)
- [4] Zhu Donghua, Zhang Yi, Wang Xuefeng et al. (2013).Research on technology innovation management under the environment of big data[J]. *Science of Science and Management of Science & Technology*, 34(4): 172-180(In Chinese)
- [5] Zhang Baoming. (2013-7-19).Big data: New introduction of technological innovation in the US[EB/OL]. <http://www.chinavalue.net/media/Article.aspx?ArticleId=106980> (In Chinese)
- [6] China Science News. (2013-11-26). Hot discussion about data-driven innovation at Academician round table conference in Shanghai [EB/OL]. <http://www.stcsm.gov.cn/xwpt/kjdt/335055.htm>(In Chinese)
- [7] Data-driven innovation. (2013-7-10)[EB/OL]. http://whb.news365.com.cn/ewenhui/whb/html/2013-07/10/content_91.htm(In Chinese)
- [8] UFIDA. (2013-12-06).How to realize management innovation by data-driven for enterprises[EB/OL]. <http://www.enet.com.cn/article/2013/1206/A20131206337177.html> (In Chinese)
- [9] Jay Lee, Edzel Lapira, Behrad Bagheri, et al. (2013).Recent advances and trends in predictive manufacturing systems in big data environment [J]. *Manufacturing Letters*, 1(1): 38-41.
- [10] Thomopoulos R, Destercke S, Charnomordic B. (2013).An iterative approach to build relevant ontology-aware data-driven models [J]. *Information Sciences*, 221(2): 452–472.
- [11] XU Jian-Xin, HOU Zhong-Sheng. (2009).Notes on Data-driven System Approaches[J]. *Acta automatica sinica*, 35(6): 668-675.
- [12] LEE J, LAPIRA E, BAGHERI B, et al. (2013).Recent advances and trends in predictive manufacturing systems in big data environment[J]. *Manufacturing Letters*, 1(1): 38-41.
- [13] KUSIAK A. (2009).Innovation: a data-driven approach[J]. *International Journal of Production Economics*, 122(1): 440-448.
- [14] He Fei, He Keqing. (2014).The discussion of large data and its scientific issues and methods[J]. *Journal of Wuhan University(SCIENCE EDITION)*, 60(1): 1-12(In Chinese)
- [15] Geum Y, Lee H J, Lee Y, et al. (2015).Development of data-driven technology roadmap considering dependency: An ARM-based technology roadmapping [J]. *Technological Forecasting & Social Change*, 91(1): 264–279.
- [16] Ohsawa Y. (2002).Introduction to Chance Discovery[J]. *Journal of Contingencies and Crisis Management*, 10(2): 61-62.
- [17] Yi Zhang, Douglas K.R. Robinson, Alan L. Porter, et al. (2015).Technology roadmapping for competitive technical intelligence[J]. *Technological Forecasting and Social Change*, 110: 175-186.
- [18] McAfee A, Brnjofsson E. (2012).Big data: the management revolution[J]. *Harvard Business Review*, 90(10): 61-68.
- [19] Gandomi A, Haider M. (2015).Beyond the hype: Big data concepts, methods, and analytics[J]. *International Journal of Information Management*, 35(2): 137-144.
- [20] Feng Zhiyan, Guo Xunhua, Zeng Dajun, etc. (2013).A number of preface topic of business management research under the background of big data[J]. *Journal of Management Science*, 16(1): 1-9(In Chinese)