

Quantitative Analysis of Papers on Chinese Citation Analysis Research in Recent Years

-Based on data of CNKI from 2005 to 2014

Ping Zhang¹, Xindi Wang¹

¹ Beijing Jiaotong University, Beijing 100044, China

Keywords: Citation analysis; Bibliometrics; CNKI

Abstract. In this paper, the author uses EXCEL to make classification statistic and quantitative analysis of the distributed situation of the years, periodicals, funds and subjects and author situation of 1,629 relevant papers in Chinese citation analysis research from 2005 to 2014 included in CNKI's Chinese Academic Journals Network Publishing Database, and makes analysis using Bibliometrics related laws such as the document growth law, Plath index, Brad Quantitative, Lokat's Law, Zipf's law and h index to explore the status and development trend of the Chinese citation analysis research, and reveal the pattern and characteristics of development to provide a reference for further research.

Citation analysis belongs to the category of bibliometrics, and it is one of the most active research fields in current bibliometrics, scientometrics, webometrics, library and information science. It is increasingly widely used in the evaluation of journal, document, scientific research and talent, information service and knowledge management. The development of citation analysis was initiated by P. LK Gross et al.'s first research in 1927 and developed rapidly with the advent of the Science Citation Index. In 1981, Wang Chongde published the first citation analysis article in the China's Journal of Information Science, which marked the start of the research on the Chinese periodical citation^[1]. During the transition period between the 20th century and the 21st century, Liu Yanping et al. and Zhou Yunping et al. statistically analyzed the 90's Chinese citation analysis papers respectively^{[2][3][4]}. In 2006, Ye Xiejie et al. and Guo Sang et al. made a quantitative analysis of the Chinese citation analysis research papers from 2000 to 2004 and 2007 to 2011 respectively^[5]. This paper analyzes the status quo and development trend of the Chinese citation analysis research during the last ten years by using bibliometrics method. Compared with the existing documents, the research range is longer and more conducive to reveal the patent and characteristics of development of this field in China, thus providing the useful reference for the research in the future^[1].

1 Data and Methods

1.1 Data sources

The author uses the CNKI's Chinese Academic Journals Network Publishing Database as the source, which is the largest full-text database of continuous dynamic update of Chinese academic periodicals in the world. Its full collected rate of core journal and source journal of the important evaluation database is 99%, and the document overall rate is 99.9%. So far, its collected kinds of academic journals reaches 8,035, and the total amount of full-text documents more than 42,380,000.

1.2 Statistical method

The author selected subject term as the search path to ensure the integrity of the search results, and used the "Citation Analysis" as the subject term of the search to retrieve the relevant documents of citation analysis from 2005 to 2014 in the Chinese Academic Journals Network Publishing

Database, 3,450 documents were retrieved and EXCEL was employed to make classification statistic and quantitative analysis of the distributed situations of the years, periodicals, funds and subjects and authors situation of all papers to find out the development condition and patterns of the Chinese citation analysis in recent years.

2 Results Analysis

2.1 Analysis of years distribution

According to the statistical data, there were 3,450 research papers regarding to citation analysis in China between 2005 and 2014, of which 1,816 were core journals, accounting for 52.64% of the total. In Fig. 1, we could see that the annual total research volume of citation analysis of 2005-2007 was about 280. After that, the document quantity increased from 300 in 2008 to 417 in 2010, a substantial growth was found. It basically maintained at 400 in 2011-2013. We could see from Fig. 2 that the proportion of core journals was around 40%-50% from 2005 to 2007, and the proportion greatly increased in the period of 2008-2010, and stabled at about 55% in 2011-2013. The development pattern of the two and the model of growth of the document amount was basically consistent. Based on the pattern found by the researchers and Price's view, we can know that the development of the subject and the growth of the document amount will go through the stage of slow development, rapid development and stable development^[6]. It was pointed out in research of Liu Yanping et al. and Zhou Yunping et al. that the Chinese citation analysis research experienced the stages from the slow development to the rapid development in the 1990s, and by the late 1990s, it gradually entered a relatively stable period. Since the 21st century, as the international community popularized the application of the network databases such as digital library and CNKI, Vip and Spring. The Chinese citation analysis research began to jump out of the limitations of traditional citation analysis, and started to seek the networked, digital intelligent and diversified development, and shifted from the traditional approaches to the network citation analysis methods. This transition was a leap in the process of informationization, so the document amount had a great boost. While it tended to become stable in recent years. The whole process of the decade was also a microcosm of the document growth pattern.

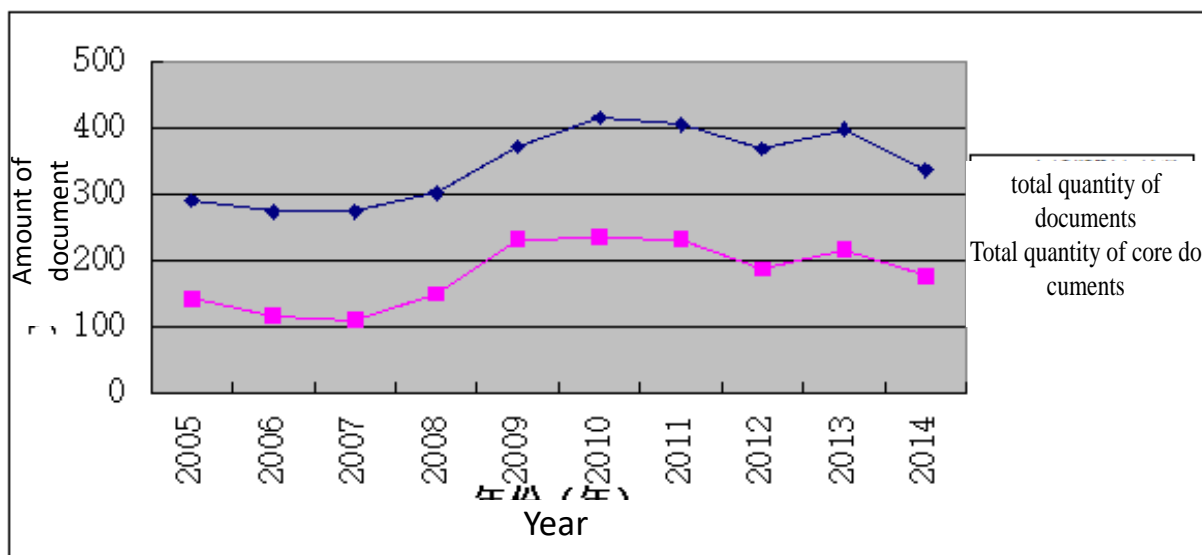


Fig.1 Annual distribution of document amount

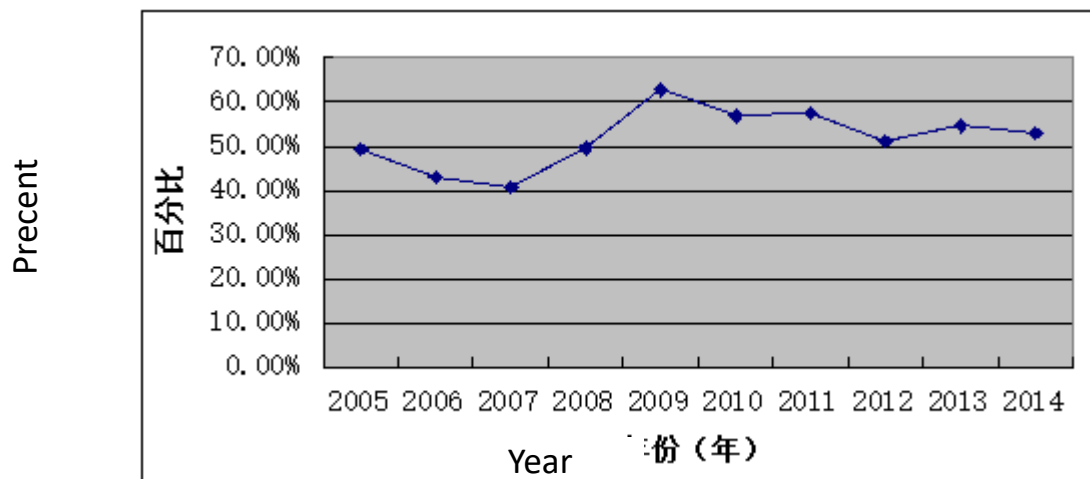


Fig. 2 The proportion of document amount of core journals accounting for the annual total document quantity

2.2 Distribution of journals and disciplines

In the distribution of the disciplines in Fig. 3, it could be seen that the documents of citation analysis research mainly focused on the field of the library information and library discipline, the document amount accounted for 73.0% of proportion of the total papers, and the medical related disciplines accounted for 9.0%. And the document amount of the library information and library fields had a significant increasement compared with the statistical results of the citation analysis research of Ye Xiejie et al., while it decreased compared with the research results of Guo Sang et al. in 2013, the document amount of the non-library information and library fields have declined and increased compared with the previous researches respectively. Thus, the citation analysis of the 21st century have been widely concerned in various fields, the library information and library fields are still the main research fields of citation analysis, library information periodical remain a main source of citation analysis research. The medical-related subjects are also increasing emphasis on citation analysis.

The situation of 17 journals with more than 40 published papers were listed in Table 1, with a total of 1,469 papers, accounting for 42.58% of the total quantity of papers. Among them, the published paper amount of information periodicals for Journal of Intelligence was the largest with a quantity of 176, and that of Library and Information Service was 160. Bradford's law states that there was relation of $1: a: a^2 \dots (a > 1)$ between the amount of journals in the core area with the largest number of papers and that of the contained journals in the subsequent area containing the same number of papers in the core^[7]. Table 1 showed that the published paper amount of these journals was close to 1/3 of the total volume, and located in the core area, which was in line with the definition of the journal of the core area for the Bradford's law, indicating that 17 journals of Table 1 was a source of the important journals of citation analysis research currently engaged in China.

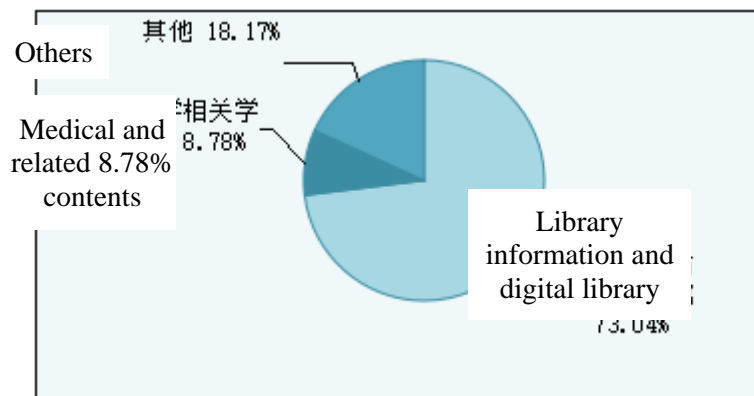


Fig. 3 Discipline category distribution

Table 1 Journals that contain more than 40 papers

| Journals | Published paper amount | Proportion | Cumulative proportion |
|--|------------------------|------------|-----------------------|
| Journal of Intelligence | 176 | 5.10% | 5.10% |
| Library and Information Service | 160 | 4.64% | 9.74% |
| Information Science | 131 | 3.80% | 13.53% |
| Modern Information | 130 | 3.77% | 17.30% |
| Journal of Library and Information Sciences in Agriculture | 113 | 3.28% | 20.58% |
| Science and Technology Information Development and Economy | 105 | 3.04% | 23.62% |
| Chinese Journal of Scientific and Technical Periodicals | 99 | 2.87% | 26.49% |
| Intelligence Theory and Practice | 94 | 2.72% | 29.22% |
| Journal of Southwest University for Nationalities (Humanities and Social Sciences) | 77 | 2.23% | 31.45% |
| Intelligence Exploration | 74 | 2.14% | 33.59% |
| Chinese Journal of Medical Library and Information Science | 48 | 1.39% | 34.98% |
| Information and Documentation Services | 47 | 1.36% | 36.35% |
| Journal of Medical Informatics | 45 | 1.30% | 37.65% |
| Library and Information Knowledge | 44 | 1.28% | 38.93% |
| Chinese Journal of Library Science | 43 | 1.25% | 40.17% |
| Library and Information | 42 | 1.22% | 41.39% |
| Research in Science | 41 | 1.19% | 42.58% |
| Other | 1981 | 57.42% | 100.00% |

2.3 Author analysis

In Table 2, 13 authors with more than 10 published papers were listed, the document amount totaled 236, accounting for 6.8% of the total amount, so the core author groups in the field of citation analysis haven't been formed. The highest yield of the author was Qiu Junping from Wuhan University, with the amount of 57 papers and 5.7 papers annually. Followed by Jiang Chunlin from Dalian University of Technology, with the amount of 35 papers. We can see that these authors have

made contributions that can not be ignored for the field of citation analysis research. It conforms to Lakat's law, that means a small number of writers wrote a large number of documents, and most of the authors wrote few documents.

H index is an important indicator to measure the impact of the author in scientific research. Higher H index indicating that the author's scientific research influence is greater [8] [9]. H index of the authors' paper influence in the field of citation analysis research was obtained in this paper by counting the number of papers published in a single citation analysis research and the citation frequency of the papers. It can be seen that Qiu Junping and Zhao Xing had the highest H index, which showed that they had the greatest influence in the field of citation analysis, and also showed that the author's H index and the number of published papers were not necessarily proportional.

Fig. 4 showed the distribution of the author institutions, of which Nanjing University was ranked first with a total of 249 papers, accounting for 7.22%; followed by Wuhan University and Dalian University of Technology, the proportions were 4.52% and 3.07% respectively. Thus, Nanjing University, Wuhan University and Dalian University of Technology are of significance in the field of citation analysis research. Fig. 4 showed that the author institutions were colleges and universities and their research institutes, indicating that the state and its universities attached importance to the citation analysis research.

Table 2 Author distribution

| Author | Author Institution | Literature | H Index |
|---------------|--|------------|---------|
| Qiu Junping | Wuhan University | 57 | 15 |
| Jiang Chunlin | Dalian University of Technology | 35 | 8 |
| Liu Zeyuan | Dalian University of Technology | 22 | 8 |
| Yang Hua | The Second Affiliated Hospital of China Medical University | 18 | 5 |
| Yang Siluo | Xiangtan University | 16 | 7 |
| Zhao Rongying | Wuhan University | 11 | 5 |
| Zhao Xing | Chongqing University | 12 | 9 |
| Cui Lei | China Medical University | 11 | 4 |
| Hua Weina | Nanjing University | 11 | 4 |
| Deng Sanhong | Nanjing University | 11 | 5 |
| Yuan Peiguo | Nanjing University | 11 | 5 |
| Ye Jiyuan | Nanjing University | 11 | 6 |
| Liu Zeyuan | Dalian Institute of Technology | 10 | 6 |

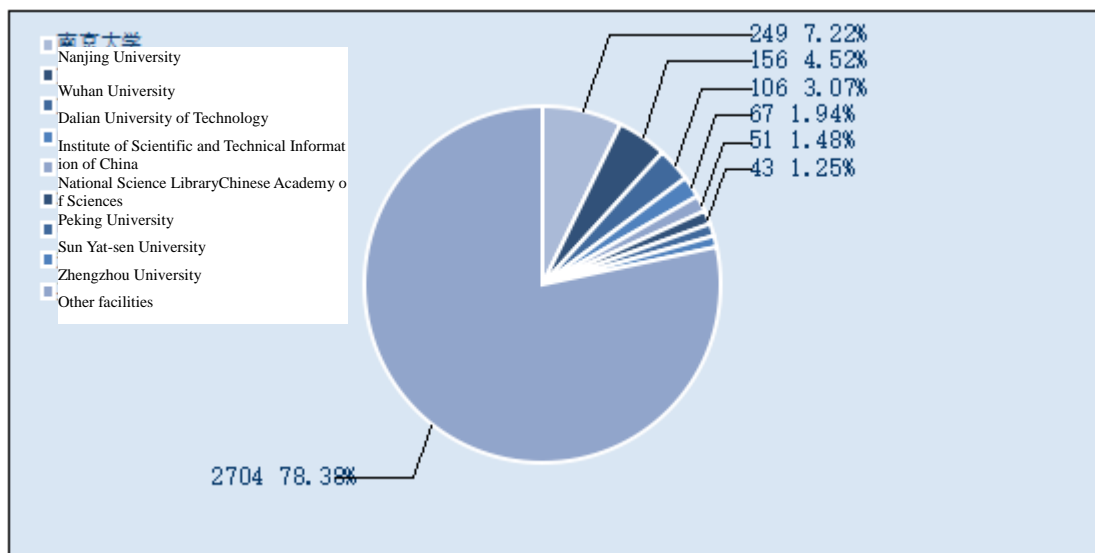


Fig. 4 Author Institutional Distribution

2.4 Hot topic analysis

In this paper, the frequency of the key words used in the paper that very year was analyzed, and the research of the citation research field was analyzed by the word frequency distribution of Zipf's law^[10]. In Table 3, the top five keywords with frequency of use of that very year in the citation analysis field were given. It could be seen that in the field of citation analysis, the annual research hot topics were different. It was observed that the usage of "citation analysis" and "bibliometrics" decreased markedly with time, but still occupied the first and second positions. The reason of this one is that the field that we made statistics and researched was citation analysis filed, and citation analysis belongs to the bibliometrics category, it must be our focus of research. The "literature metrology" and "bibliometrics" are the same and this is because they were produced at the same time void of standardized treatment. It was worth great attention that CSSCI had higher utilization rate in 2009 -2011, indicating that it was research hot spot in those three years. Knowledge Facebook had high frequency of use in nearly three years of 2012-2014 , and has become a new research hotspot. It can be also obtained that the frequency of use of periodical evaluation, published paper analysis variable, but it has been the focus of research that favored by the author.

Table 5 showed the distribution of funds supporting scientific research. It can be seen that the National Natural Science Foundation and the National Social Science Fund have given absolute support, and further proved the country shared great attention to the field of citation analysis research, which provided a useful reference for us to study the application fund.

Table 3 Key word distribution

| Year | Keyword 1/Frequency of use that very year (%) | Keyword 2/Frequency of use that very year (%) | Keyword 3/Frequency of use that very year (%) | Keyword 4/Frequency of use that very year (%) | Keyword 5/Frequency of use that very year (%) |
|------|--|---|--|--|--|
| 2005 | Citation analysis/42.47 | Bibliometrics/14.38 | Journals/6.51 | Core Journals/6.16 | Citation/6.16 |
| 2006 | Citation analysis/44.00 | Bibliometrics/13.82 | Papers analysis/8.00 | Author analysis/6.18 | Journals/5.82 |
| 2007 | Citation analysis/42.18 | Bibliometrics/ 14.55 | Papers analysis/9.45 | Bibliometrics/6 .91 | Journals/6.55 |
| 2008 | Citation analysis/34.54 | Bibliometrics/ 13.82 | Papers analysis/6.25 | Citation/5.92 | Bibliometrics/5.9 2 |
| 2009 | Citation analysis/42.63 | Bibliometrics/ 13.40 | CSSCI/8.31 | Journal evaluation/6.17 | Papers analysis/5.09 |
| 2010 | Citation analysis/37.65 | Bibliometrics/ 11.51 | CSSCI/9.11 | Academic impact/4.80 | Citation/4.08 |
| 2011 | Citation analysis/36.36 | Bibliometrics/ 9.58 | CSSCI/7.62 | Journal evaluation/6.88 | Papers analysis/4.91 |
| 2012 | Citation analysis/35.95 | Bibliometrics/ 10.00 | Knowledge of Facebook/5.14 | Bibliometrics/4 .86 | CSSCI/4.32 |
| 2013 | Citation analysis/26.57 | Bibliometrics/ 6.77 | Knowledge of Facebook/6.27 | Bibliometrics/5 .26 | Cited frequency/4.01 |
| 2014 | Citation analysis/30.03 | Bibliometrics/ 6.61 | Bibliometrics/6. 31 | Journal evaluation/5.71 | Knowledge of Facebook/5.11 |

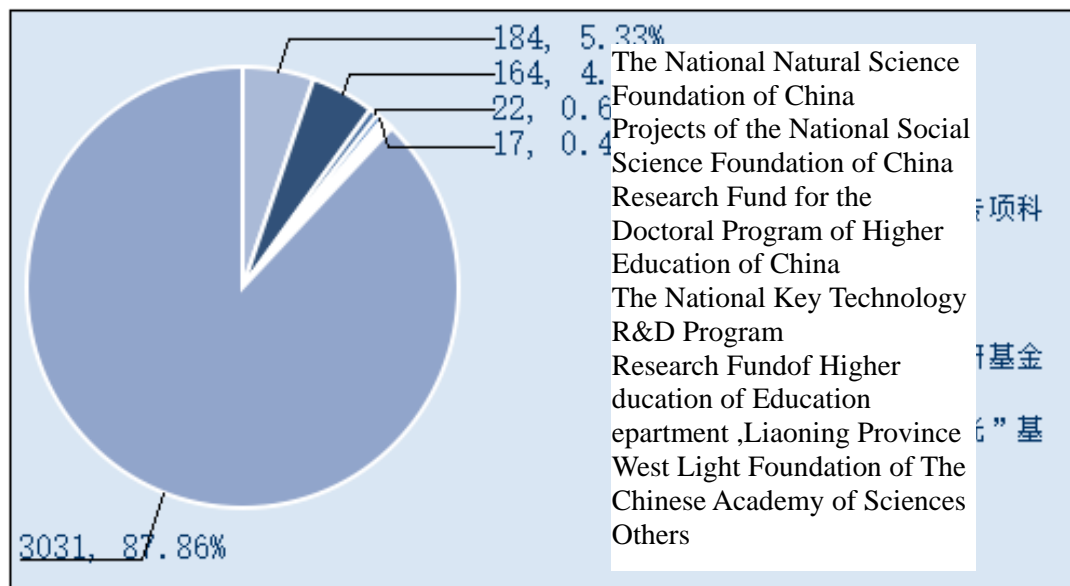


Fig. 5 Fund distribution

3 Trend Analysis

3. 1 Diversification of hot spots of research topics

With the development of science, the informatization infiltrates various fields, the rapidly growth research fields of various disciplines also promotes the interdisciplinary research of various disciplines. The mutual use of documents is an important way to promote the inheritance, utilization, infiltration and development of knowledge. Citation analysis can reveal the characteristics of the

research object and the relationship between these elements, and the conclusions obtained are reliable and persuasive. Therefore, it is widely used in various disciplines and interdisciplinary fields. Statistics showed that for the last 10 years, the proportion of information and library science in the Chinese citation analysis has declined, non-information science research has increased, the keyword was no longer a single model, indicating that the research focus and application field began to show the trend of diversification.

3. 2 Diversification and integration of analysis methods

The credibility of the citation analysis research depends on the normative nature of the citations, the reliability of the authors' references, the precision ratio and recall ratio of the selected data, and the reasonableness and correctness of the researchers' statistical methods. In recent years, the higher requirements of papers published in the research ensured the reliability of the author's citations, a variety of search paths and improvement of the network database ensured the data precision ratio and recall ratio. In the 21st century, it could be seen that bibliometric methods or citation analysis methods were no longer used for the analytical methods in the high-level journals such as core journals. Many analysis methods such as clustering analysis, webometrics, neural networks, probability statistics and fuzzy mathematics were used, and a variety of analytical methods were even used in a paper for a few documents, making citation analysis research methods diversified and integrated.

3. 3 Intellectualization and networking of analytical means

Nowadays, how to reduce the work of manpower is the focus of our research and work, citation analysis field began to seek more convenient and effective analytical means and tools to help the researchers. In recent years, network data such as CNKI, Vip, Wan Fang and Spring has been widely used, developed and continuously improved, it was the the citation analysis research to form the network database as the main data source, and computer information processing software such as SPASS and Matlab as the intelligent, network-based development stage of the measurement statistics analysis tools.

3. 4 Statistical measurement indexation

The main tools of SCI journals such as *Citation Index of Chinese Sci-tech Journals* and *Chinese Scientometrics: Papers and Quotations* based on "Science Citation Index" and "Social Science Citation Index", etc. provide a reliable scientific basis for the purposes such as journal evaluation and citation research, which promotes in-depth research of the citation analysis field.

References

- [1] Guo Sang, Jia Jianrui. Quantitative analysis of papers on Chinese citation analysis research from 2007 to 2011 [J]. Journal of Library and Information Sciences in Agriculture, 2013,06: 57-60.
- [2] Liu Yansu, Zhang Qing. Summary of the Domestic Citation Analysis in the 1990s [J]. Information Science, 1998,05: 431-435.
- [3] Zhou Yunping, Sun Yuan. Study on the Current Situation of Chinese Citation Analysis and Its Development in the 21st Century [J]. Library and Information Service, 2001,02: 80-82 + 89.
- [4] Li Li. Research and development of the Chinese citation analysis in 1990s [J]. Library and Information Forum, 2001,02: 18-21.
- [5] Ye Xiejie, Cai Lizhu, Di Ming. Quantitative analysis of the papers on the Chinese citation analysis research in recent years [J]. Modern Information, 2006,02: 130-132.
- [6] Qiu Fengming, Chen Huilan. Paper Analysis of Chinese Citation Analysis Research Based on

Content Analysis [J]. *Information Science*, 2010,09: 25-28 + 66.

[7] Hu Lan. Citation Analysis on the Volleyball Scientific Research Papers in the Core Periodicals of Sports in China from 2004 to 2012 [D]. Nanjing Institute of Physical Education and Sports Science, 2007, 23 (4): 114-12.

[8].Liang Yongxia, Liu Zeyuan, Yang Zhongkai, Wang Xianwen. Knowledge Mapping of Frontiers and Evolution in Citation Analysis Field [J]. *Studies in Science of Science*, 2009,04: 516-522.

[9] Zhang Na. Bibliometrics analysis on the competitive aerobics papers of the core journals of the China's sports in 2003-2012 [D]. Shanghai Normal University, 2014.

[10] Zhaxi Yuzhen. A Summary of research and development for Bibliometrics Citation Analysis Field [J]. *Tibet Science and Technology*, 2011,10: 15-16.