The Evolution Analysis of the Hot Spots and Trends of Visualization Research

Yong-yue CHEN*, Lei-lei BU and Xue-yang LI
Department of Management of Wuhan Textile University
*Corresponding author

Keywords: Visualization, Hot spot and trend analysis, Mapping knowledge domain.

Abstract. With the continuous development of large data and cloud computing, the analysis of mass data, the methods and tools of information visualization analysis have become a new trend and hot spot. Based on the background of visualization, from the two dimensions of space and time, through the analysis of a large number of literature data, this paper combs the current research topics of researchers, and provides a reference for comprehensive and effective analysis and tracking of hot topics and trends in thematic areas research.

Introduction

With the development of information technology and the relevant industries such as big data and cloud computing, the field of scientific research is continuously expanding and mining. The data in different disciplines also show an explosive growth. Massive data analysis, as well as a more intuitive and comprehensive presentation of massive data and information has become an urgent reality, the visual analysis methods and tools in the big data environment has become the current hot spots and trends [1, 2, 3]. Among them, the use of visual methods to track the front hot spots, predict the development trend is an important research method which current researchers pay attention to [4].

The essence is that researchers use visual tools and methods to analyze and process a large amount of relevant literature data to form visual diagrams and tables of information of concern to researchers and rely on the related tools of computers to draw the relevant image. It can discover the research hot spots and development trends, find out the topics of their research, search for the literature of interest, and effectively help the researchers to better understand the hidden information behind the knowledge. Based on this point, this paper analyzes and analyzes the visualization method of hot spots and trend analysis based on keywords.

Visualization Tool Selection and Analysis Method Introduced

Visualization Tools to Choose

At present, there are many kinds of visual software, such as Prefuse, Ucinet and Citerspace, and they have their own advantages and characteristics [1]. In the data processing, Prefuse more emphasis on the corresponding changes in the number of figures, and Ucinet and Citerspace can only handle the relevant software has been converted data source. In support of the function, Prefuse is relatively more powerful, more able to support research workers do dynamic queries and searches, the scalability and extensibility are more powerful than the other two tools, while Ucinet and Citerspace mainly analyze and generate visual co-citing networks. Compared to Citerspace used to study trends and cutting-edge research and other emergencies, Ucinet more is a simpler, easy-to-use tool for the simple analysis of certain characteristics of a collaborative network has a very fast and simple advantage, and Citerspace is a good visual tool in tracking the growth of new disciplines, showing real-time hot topics in research and research trends in real time. In this paper, we mainly use the advantages and characteristics of Citerspace for data visualization analysis [4].
Analysis Methods Introduced

Research Hot Spot Analysis

The word frequency analysis method can identify the research hotspot and development trend of the subject field through the statistical analysis of the key words of the core content of the document and the occurrence frequency of the keyword in the literature of a certain scientific research field. Generally we think that the keyword article is the core content of refining and essence of the article [5, 6, 7, 10]. Therefore, when some keywords appear constantly in the literature of this field, we think that it represents the research hotspot in this field. Also formally based on this idea, Citespace II to statistical and analysis of the frequency of co-occurrence between different keywords, and use the visual method to display the frequency and size of the keywords in the form of a graph in front of the readers, so as to provide a basis for the research workers to analyze a certain area.

Research Frontier Analysis

Tracking the hot front of the discipline is conducive to scientific researchers to better grasp the scientific trends and trends. The frontiers of the subject are determined according to the mutational terms extracted from the literature titles, abstracts and the high frequency words of the article [8, 9]. The mutation algorithm is better for us to detect the occurrence of mutation words, so as to better apply the time series analysis and discover the frontier of the subject. In this paper, we mainly extract the terminology of catastrophe from the key words of the literature, and use the “time zone” of Citespace II to analyze the frontier of research in the field.

Sources and Processing of Data

Sources of Data

The relevant data used in this article are from CNKI. Using "visual analysis" as the key word, the journal database of CNKI was retrieved. The selected time period is all the journals in the 15 years from "2002 to 2016", which include SCI source journals, EI source journals, core journals and CSSCI data, which retrieved a total of 2710 articles, and use these basic data to build the original data required for this study.

Data Processing

According to the requirements of Citespace II software, the experimental data are required to be stored in the same file in the form of text files. During the process of analysis, when creating a project, the folder path of the previously saved data can be directly selected according to the needs of scientific researchers, which will be more conducive to the development of research workers related work. At the same time, the data were standardized, such as: data de-duplication, the author of the same name, similar keywords so as to ensure the objectivity and accuracy of the analysis in the experimental process.

Thematic Research Hot Spots and Trend Analysis

Research Hot Spot Analysis

In general, the key words can express the core content and essence of the literature to a certain extent. Therefore, word frequency statistics of key words in a certain field can reflect the research hotspot and development in this field [5, 7]. In this paper, Citespace II is used to count and analyze the frequency of occurrence of different keywords, and visualize the frequency and size of keywords by using visual methods to provide a reference for researchers to discover hotspots in a particular field. Table 1 shows some high-frequency keywords and high-level keywords by Citespace II. Figure 1 shows the visualization of research hotspots presented by Citespace II.
Table 1. High-frequency keywords in the network and high center keyword rankings.

| ranking | High-frequency keywords | Frequency | High central keywords | Centrality |
|---------|-------------------------|-----------|-----------------------|------------|
| 1       | Visualization           | 400       | Visual Analysis       | 0.71       |
| 2       | Visual Analysis         | 328       | Information System    | 0.48       |
| 3       | Knowledge map           | 237       | Three-dimensional model| 0.45       |
| 4       | Research Hotspot        | 102       | Collnet               | 0.45       |
| 5       | Citespace               | 88        | Citation Analysis     | 0.44       |
| 6       | Information Visualization| 81       | 3D Visualization      | 0.41       |
| 7       | Bibliometric            | 77        | Data Mining           | 0.25       |
| 8       | Social Network Analysis | 68        | Opengl                | 0.25       |
| 9       | GIS                     | 67        | Csci                  | 0.24       |
| 10      | Co-word analysis        | 60        | Content Analysis      | 0.24       |
| 11      | Geographic Information System| 51 | Geographic Information System | 0.23       |
| 12      | 3D Visualization        | 41        | Data warehouse        | 0.23       |
| 13      | Big Data                | 39        | Co-citation analysis  | 0.22       |
| 14      | Data Mining             | 38        | Visualization         | 0.18       |
| 15      | Data Visualization      | 38        | Clustering Analysis   | 0.15       |
| 16      | Clustering Analysis     | 37        | GIS                   | 0.14       |
| 17      | Research Frontier       | 35        | Information Visualization| 0.13       |
| 18      | Scientific knowledge map| 33        | Co-word analysis      | 0.1        |
| 19      | Csci                    | 29        | Idl                   | 0.1        |
| 20      | Bibliometrics           | 28        | Algorithm             | 0.09       |

In the visualization network, Citespace II can generate the keyword co-occurrence network, calculate the centrality of each keyword, and cluster some high-frequency keywords. High-center keywords are the key hot spots in this research field\(^{[11,12]}\).

![Figure 1. Key words co-occurrence network.](image)

Each node in the visual network of Figure 1 symbolizes a keyword, and the size of a circle in a node means that the frequency of the keyword appears in the network, the connection between nodes in a network indicates the co-occurrence relationship between keywords in the network. When the connection between two nodes is thicker, the frequency of two keywords co-occurrence is higher. There is a close relationship with other nodes, which is the key point to connect different fields or topics. The keywords, the co-occurrence relationship between keywords, and the clustering relationship of keywords are clearly and accurately displayed in the form of knowledge map, a clear
view of the effect allows us to quickly identify a few key points with a strong sense of interpretation. Through the analysis of key points, we can not only find out the research hotspots, but also find some new phenomena and rules which are not reflected in traditional word frequency analysis.

From Figure 1 we can see that the entire view is mainly divided into two larger clusters, the lower left corner of the main visualization as the center, including: data visualization, visual analysis, etc., the upper right corner mostly relates to knowledge management, including: knowledge management, bibliometrics, citation analysis, co-word analysis, knowledge map and so on.

Research Frontier Analysis

Tracking the hot front of the discipline is conducive to researchers to better grasp the scientific trends and trends, mutation algorithm can be used to detect the emergence of mutant words, In order to make better use of time series related literature analysis and retrieval [8, 9], In Citespace II, research frontiers are based on abrupt terminology extracted from identifiers of titles, abstracts, keywords, and documentation. In this study, we selected the node type as the key word, detected a total of 2171 abrupt terminology, and select the display type as "time zone" to generate a research frontier glossaries and frontier view of the 15 years of visual analytics.

| Term Frequency           | Term Frequency           | Term Frequency           |
|--------------------------|--------------------------|--------------------------|
| 2002—2006                | 2007—2011                | 2011—2016                |
| Intelligent substation (20) | Visualization (365)    | Knowledge Map (220)     |
| City 3D Geological Information System (50) | OpenGL (53) | Co-word analysis (58)    |
| Visualization (48)       | Random linear element (150) | h index (20)            |
| Residential area planning (26) | Data mining (35)     | Biosafety (23)           |
| Non-point source pollution (15) | Content analysis (16)  | Citespace II (30)        |
|                          | Cognitive Neuroscience (13) | Community (25)          |
|                          | Citespace (45)           |                          |
|                          | Co-citation analysis (25) |                          |
|                          | Information Science (15) |                          |
|                          | Finite element software (9) |                          |

Table 2 lists the research terms and their frequency in the field of visual analysis from 2002 to 2016. Figure 2 shows the time-zone view of the research frontier in this 15-year field of visual analysis. This view contains 56 nodes (keywords), size of the node represents the same keyword frequency, from which we can detect the forefront of research and development trend in recent years in the field of visual analysis, from 2002 to 2006, the research on visual analysis is relatively scarce in the past five years. From 2007 to 2011, the five years is a boom period of visual analysis and research. Its research and analysis methods are mainly qualitative analysis and visualization analysis. According to the comparative analysis of key words and documents such as "Data Mining", "Content Analysis" and "Cognitive Neuroscience", we can find that the five-year visual analysis method has been developed rapidly in other disciplines.
Moreover, the qualitative research rate gradually decreased from 2011 to 2016, and the quantitative analysis increased rapidly.

**Conclusion**

To track the hot spot of academic frontier and predict the trend of its development has always been an important part of the attention of the researchers. In this paper, the key words of the subject area of "Visual Analysis" are selected, and the key words are analyzed. Visual method is applied to grasp the hot topics and trends of academic frontiers.

Through the cluster analysis of high-frequency keywords, we find that the visual analysis from 2002 to 2015 has expanded from the simple data and content analysis to a wider area of social sciences and industries. In recent years, research on bibliometrics, co-citation analysis and knowledge map has become a major hotspot. At the same time, the trend of visual analysis combined with new fields such as big data and cloud computing is getting closer and closer. Moreover, the visual analysis in the past fifteen years shows a shift from qualitative analysis to quantitative analysis, and with the updating and strengthening of technologies such as big data, the research on visualization has gradually become a hot spot and trend.

**Acknowledgement**

This research is supported by National Social Science Foundation of China under Grant 14BTQ057. It is also supported by the Humanities and Social Science General Project of Hubei Province Department of Education under Grant 14Y023.

**References**

[1] Huo Liang, Zhao Lemen. Visualization and Its Application in Information Analysis [J]. Information studies: Theory and Practice & Application, 2017, 40(4): 111-116.

[2] Shen Jiujiu, Yang Xiaoyue. A Visual Analysis of the Current Situation of Quantitative Research on Library and Information Science in the Background of Big Data [J]. Library, 2017(6):46-52.

[3] Miao Rui-sheng; Ma Hai-qun. Literature Measurement Analysis of "Big Data" Research in the Field of Library and Information Science in China [J]. Information Science, 2017(3):93-97.
[4] Chen Yongyue, Tian Wenfang. Visual Analysis Method Research on Hotspot Track and Trend Prediction in Subject Area [J]. Information studies: Theory and Practice & Application, 2017, 40(6):117-121.

[5] Su Nan. Visualization Analysis of Big Data in China based on CSSCI [J]. Information Science, 2016, V36 (3):112-116.

[6] Zhao Bin, Dong Ying, Yang Xiao-jie. Knowledge Mapping and Hot Topics in the Study of Information Ecology in China—Based on the Perspective of Bibliometric Analysis [J]. Information Science, 2017(9):61-66.

[7] Sun Z, Soochow University Library. Distribution and Frontier Evolution of University Library Service in Recent Decade: Based on Bibliometric CNKI Core Journals and Visual Analysis [J]. Library Journal, 2017.

[8] Bernard J, Daberkow D, Fellner D, et al. VisInfo: a digital library system for time series research data based on exploratory search—a user-centered design approach [J]. International Journal on Digital Libraries, 2015, 16(1):37-59.

[9] Nazemi K, Retz R, Burkhardt D, et al. Visual trend analysis with digital libraries[J]. 2015:1-8.

[10] Tang J, Hechi University Library. Multidimensional Visual Analysis of Studies on Library's Subject Service in China [J]. Information Research, 2014.

[11] Zhang J S, Humanities F O, University D. Visual Analysis of Domestic Self-confidence Document Information [J]. Journal of Dalian University, 2016.

[12] Chen C, Wu H. Visual Analysis on Graphene Patent Information of China [J]. Journal of Modern Information, 2014.