Journal of Informetrics 9 (2015) 118–134

Contents lists available at ScienceDirect

Journal of Informetrics

journal homepage: www.elsevier.com/locate/joi

Visualization and quantitative study in bibliographic databases: A case in the field of university–industry cooperation

Feng Feng[a], Leiyong Zhang[a], Yuneng Du[b,*], Weiguang Wang[c]

[a] School of Public Affairs, University of Science and Technology of China, China
[b] College of Economics & Management, Anhui Agricultural University, Hefei 230036, China
[c] Chinese Academy of Sciences, China

ARTICLE INFO

Article history:
Received 23 November 2013
Received in revised form
24 November 2014
Accepted 26 November 2014
Available online 24 December 2014

Keywords:
Bibliometrics
Visualizing
Quantitative
CiteSpace
University–industry collaboration

ABSTRACT

CiteSpace is a visual document analysis software, by which performances and trends of certain disciplines can be displayed for a given period. Moreover, the evolution of a frontier research can be explored by such software as well. This research focuses on the visualization and quantitative study in bibliographic databases by taking the university–industry collaboration studies as an example. Using the Web of Science (WOS), 587 publications and over 30,000 references were selected for analysis, which produced the following results: (1) Our method can clearly reveal the key elements of certain disciplines, such as the largest share of publications, the most frequently cited authors and journals in the university–industry cooperation research field; (2) The relationships among the frequently cited authors, references, journals and keywords can be explained visually in the university–industry cooperation research field; (3) Of special note is that the potential problems and evolutionary trends of certain research fields such as university–industry cooperation can also be ascertained via our method; (4) In general, according to the case study, our visualization and quantitative method evolved a new research framework to evaluate the performance of some research areas.

1. Introduction

Quantitative research of literatures appeared in the early 20th century but had not formed an independent discipline until Pritchard (1969) first proposed that the terminology Statistical Bibliography should be replaced by Bibliometrics. Since then the theory and practice of literature studies on the basis of bibliometrics has become widespread in academic research (Diem & Wolter, 2013). The metamorphosis of Bibliometrics generated by the development of networks and computer technology has made graphical study and visualization research of literatures possible. That images contain much more information than digits or words of the same size is a foregone truism (Ma & Xi, 1992). This becomes evident when looking at all kinds of visualization software. Exploration and prediction of frontier science has increasingly become more popular over the years, and analyzing text messages and citation information of literatures by computer provides a new perspective of Bibliometrics. Complicated phenomenon and analysis results can be obtained in the process of visualization research. Using the images

* Corresponding author at: 130 West Changjiang Road, Anhui Agricultural University, Hefei, Anhui, China. Tel.: +86 13083062933.
E-mail address: duyuneng@gmail.com (Y. Du).

http://dx.doi.org/10.1016/j.joi.2014.11.009
1751-1577/

generated by computer software, researchers can quickly understand the research status and be able to forecast the possible research directions in the future (Chen, 2006). Thus new study areas will be identified through visualization study, yet it must be noted that the basis of visualization study is co-citation analysis (Ma & Xi, 1992).

Co-citation analysis theory suggests that any new knowledge is derived from existing ones. If two papers are cited by one or more papers, then they have co-citation relationship, and the number of citations is called the co-citation degree. A high co-citation degree means that the two papers are closely related to each other (Small, 2003). White and McCain (1998) tried to describe the subject structures of information science with the method of Author Co-citation Analysis (ACA), which were developed by them since 1981 (White, 1990). Cluster analysis, multi-dimensional scaling and factor analysis were applied in their study with the help of SPSS. After that, lots of studies on the description and analysis of different areas appeared, and some retrieval systems and analysis software were developed based on ACA. Traditional ACA method gained huge success, but some researchers still questioned it. The arguments were focused on the following points: (1) whether Pearson correlation coefficient r was suitable for the measure of correlation between the authors (See Ahlgren, Jarneving, & Rousseau, 2003; Bensman, 2004; Leydesdorff & Zaal, 1988; White, 2003); (2) which matrix should be used to generate the similarity coefficient matrix (See Kruskal & Wish, 1978; Leydesdorff & Vaughan, 2006; Miguel, Moya-Anegón, & Herrero-Solana, 2008; Waltman & Eck, 2007). To solve this dilemma, the optimization of the ACA was developed into two directions. Some researchers tried to choose proper similarity measurement according to the specific data, or better ways to generate a similarity matrix, by which visual maps can be created with the help of SNA (Social Network Analysis) software such as Pajek, Ucinet or VxOrd. Other researchers tried to use the co-citation matrix directly, for example, White (2003) introduced Pathfinder Network Scaling (PFNETs) into ACA, and the original co-citation data in the co-citation matrix was used directly. In the visual maps which were created with PFNETs, the nodes represented the authors, the lines between nodes represented weighted path, and the co-citation number was the weight (White, 2003).

In the two directions above, visual maps are important concepts. In some sense, visualization is the future direction for bibliographic studies. As mentioned above, since images contain much more information than words or data, this area can be called mapping knowledge domains. In studies, co-citation analysis of authors, titles, keywords and institutions can be explored. The performance of certain study areas can be shown to the readers intuitively by maps with authors, literatures, journals and institutes as nodes in it. CiteSpace is a typical tool for co-citation network analysis and visualization. Based on some concepts in information science (research front, intellectual base and time-variant duality), CiteSpace can generate two complementary views, viz., cluster views, and time-zone views (Synnestvedt, Chen, & Holmes, 2005). One of the functions of CiteSpace is drawing visual co-citation maps. Separate co-citation networks are generated first and then a whole map is combined with the separate ones. Important literatures can be recognized in the map due to their prominent features. Thus, the finding of Turning Points can be simplified to the searching of key nodes in the visualization maps, and the evolution of the area can also be detected and monitored with the key nodes (Chen, 2003). Another function of CiteSpace is in detecting the hottest topics and predicting future research trends. Hybrid network of cited papers and their citing papers are mapped by the software and burst terms of the area are detected by an algorithm called burst detection (Chen, 2006), which reveal the research trends.

This paper tried to use CiteSpace to develop visualization and quantitative study in bibliographic databases. University–industry collaboration was selected to be the target area, because this field has become a hot issue in recent years. For a long time, the university–industry collaboration was an important issue related to management studies, science and technology policy studies, innovation studies, industrial organization or network, and science of science studies (Agrawal, 2001; Alice, 2007; Hancock & McCurry, 1993; McMillan & Hamilton, 2003; Thune, 2007). University–industry relationship is not just a link or a connection, it is rich in meanings. Universities not only draw R&D funds from enterprises to develop science and technology, but also cultivate more talents through the talent exchange program with firms. At the same time, industry may increase profits and market share via the technical progress of universities (Cao, Zhang, Feng, & Du, 2013; Feng, Zhang, Du, Ma, & Fu, 2012). Generally, cooperation of universities and industry must go through an interface or media which can be tangible such as diffusion of technology transfer offices, technology transfer center, or intangible such as technology, knowledge, patents or even information (Nelson, 2001; Shane, 2005; Siegel, Waldman, & Link, 2003; Thursby, Jensen, & Thursby, 2001). Heterogeneity and complementarity are the two key reasons in the formation process of university–industry cooperative relation (Kumar, 1994). Upon upgradation of cooperation, university and industry may build a certain economic entity together, which is common in practice (Jonas, Göran, Inger, John, & Mats, 2006; Pickard, 1944). Previous studies have given lots of details of collaborations between universities and industry, but there is little concern for bibliometric analysis of the university–industry collaborations studies. Although several studies have emphasized the importance of literature analysis on university–industry links (Aurora & Luisa, 2012), to our knowledge, a more quantitative and graphical approach based on bibliometric techniques has yet to be undertaken. As a consequence, the university–industry cooperation is a perfect example to construct our visualization and quantitative research framework in bibliographic databases, since the visual images with graphical techniques are really needed in the university–industry cooperation study (Aurora & Luisa, 2012).

In this paper, visualization and quantitative research framework was created through the following regimen: the history of the university–industry collaboration studies was mapped and hot topics in recent years were found; using bibliometrics analysis method with the help of Web of Science and the program of CiteSpace II, the trend of university–industry collaboration studies was depicted quantitatively (Chen, 2004, 2006); Panoramic information was revealed and the key literatures and key authors of the university–industry collaboration areas were identified in our work (Hou & Hu, 2013). The ultimate