Use of Social Network Analysis in Bibliometric Researches[1]

Umut Al[*], Zehra Taşkın[**] and Güleda Düzyol[***]

[*] *umutal@hacettepe.edu.tr*
Hacettepe University, Department of Information Management, Turkey,
http://yunus.hacettepe.edu.tr/~umutal

[**] *ztaskin@hacettepe.edu.tr*
Hacettepe University, Department of Information Management, Turkey,
http://yunus.hacettepe.edu.tr/~ztaskin

[***] *gduzyol@hacettepe.edu.tr*
Hacettepe University, Department of Information Management, Turkey,
http://www.bby.hacettepe.edu.tr/akademik/guledaduzyol/

Curriculum Vitae

Umut Al is currently working as an associate professor at the Department of Information Management of Hacettepe University in Ankara, Turkey. He received his Ph.D. in Information Management from Hacettepe University (2008). His research interests include information seeking behaviours on the Internet, university libraries, and bibliometrics.

Zehra Taşkın is an M.A. student at Hacettepe University. She is working on a master thesis about standardization problems of Turkish addresses on citation indexes. Her main research areas are; bibliometrics, information seeking habits, databases and online information retrieval. She is working as a research assistant at the same university.

Güleda Düzyol is a doctoral student at Hacettepe University. She completed her master thesis about co-author analysis of librarianship in Turkey. Her research areas are; bibliometrics and statistical research methods of social sciences. She is also working as research assistant at Hacettepe University.

Abstract

The impact of published works is measured and evaluated quantitatively and qualitatively by bibliometric researches. In addition to this, various findings related to scientific communication are also obtained from bibliometric studies by analyzing certain characteristics of documents or publications. In recent years, as a result of increased interest to bibliometric studies, different scientific methods and techniques are observed for bibliometrics. Social network analysis (SNA) has also taken its place in the literature as a method which is frequently used in bibliometric studies, especially in the past few years. SNA can simply be defined as examining the social structure and its impacts. This method comprehends the social structure as a network, which connects the actors and the knots of the relationships between these actor pairs. These actors can be people, as well as groups, institutions and even countries. SNA is frequently used by several disciplines such as sociology, anthropology, social psychology, communications and economics. Using SNA, researchers examine the structure of communities, try to describe the network structures, and model the existing connections by visualizing the relationships between the communities. The aim of this study is to review the use of SNA in bibliometric studies.

[1] This study is supported by a research grant (no: 110K044) of the Turkish Scientific and Technological Research Center (TÜBİTAK).

Introduction

The evaluation of scientific publications is a very popular research area. The efficiency of scientific publications is usually investigated through bibliometric researches. Pritchard (1969, p. 348) defined the term bibliometrics as "the application of mathematical and statistical methods to books and other media of communication". Then, the bibliometricians began to deal with some specific research questions. Consequently, SNA has become one of the contemporary research methods which is frequently used in bibliometric studies in recent years.

As a research method, SNA has a history stretching back at least half a century (Newman, 2001, p. 016131-1). SNA can simply be defined as examining the social structure and its impacts. This method comprehends the social structure as a network, which connects the actors (e.g., people, countries) and the knots of the relationships between these actor pairs. Using SNA, researchers examine the structure of communities, try to describe the network structures, and model the existing connections by visualizing the relationships between the communities. In this study the SNA literature was reviewed and a few SNA applications based on bibliometric data were presented.

Method

We used Thomson Reuters' *Web of Knowledge* (*WoK*) as a data source for identifying Turkish scholars' publications. *WoK* covers five different databases, namely, *Science Citation Index Expanded*, *Social Sciences Citation Index*, *Arts & Humanities Citation Index*, *Conference Proceedings Citation Index (Science)*, *Conference Proceedings Citation Index (Social Science & Humanities)*. However *Conference Proceedings Citation Index (Science)* and C*onference Proceedings Citation Index (Social Science & Humanities)* are excluded in our study. The term "publication" is defined, unless otherwise indicated, as journal articles, meeting abstracts, notes, and etc. which were authored by the scholars affiliated with Turkish institutions and included in the citation indexes. To identify the publications within these databases, an online search was performed on March 10, 2011, by using the "address" field. To obtain reliable data, different forms of Turkish addresses in different languages (e.g., Turkey, Turkiye, Turkei, Turquie) were entered in the address field. At the end of the data cleaning process, a total of 198,687 publications, covering years from 1928 to 2009, were identified.

After obtaining Turkish researchers' contributions, the SNA method is used to understand the relations between the actors. For the SNA analysis many SNA software platforms (such as *CiteSpace, Bibexcel*, *Pajek*, *UCINET*) were examined and it is decided to use the *CiteSpace. CiteSpace* is a freely available Java application which is designed as a tool for SNA and analyzes and visualizes co-citation networks (Chen 2004). Chaomei Chen (the creator of CiteSpace) underlined that "CiteSpace supports structural and temporal analyses of a variety of networks derived from scientific publications, including collaboration networks, author co-citation networks, and document co-citation networks" (Chen 2011). It also identifies scientific trends, significant publications, authors and journals.

Publications on SNA

The study primarily investigated the development of the research on SNA. *Scopus* and five citation indexes of *Web of Knowledge* were examined on 26 June 2011 in order to determine the number of SNA studies that were indexed by afore mentioned sources. Searches were conducted via "topic" field of citation indexes and "article title, abstract and keywords" field of *Scopus*. As a result, 1,523 publications and 2,169 publications were identified in citation indexes and in *Scopus*, respectively, between the years 1975-2010. Since some additional

time will be needed in order for the publications that were belong to 2011 to take part within the indexes completely, 2010 is determined as a cut-off point. It was found out that the article which was published in 1975, titled "An algorithm for clustering relational data with applications to social network analysis and comparison with multidimensional scaling" was the first study related to the topic in both two platforms. Therefore, 1975 was taken as a starting point for the searches. Figure 1 presents the distribution of SNA studies by years that were retrieved from the searches. It was observed that there was a very small number of studies on the topic before 1990 (19 publications in citation indexes, 29 publications in *Scopus*). Thus, only the development of literature from 1990 to 2010 was displayed in Figure 1. It was observed that the number of SNA studies have increased since 2003, and particularly after 2006, when compared to previous years. In citation indexes the number of SNA studies that were published between 2006-2010 were 77% (1,177 publications) of all studies and in Scopus it was 83% (1,795 publications). These results showed that the SNA topic has gained importance in recent years in the related literature.

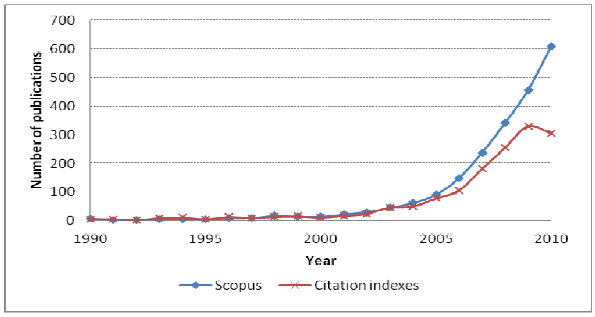

Figure 1: Number of publications on SNA (1990-2010)

Many studies combined the methods of SNA and bibliometrics in the literature (Larivière, Gingras & Archambault, 2006; Newman, 2001; Perianes-Rodríguez, Olmeda-Gómez & Moya-Anegón, 2010). In these researches, along with various bibliometric analyses, differences in collaboration patterns between the actors (such as authors, institutions, countries) were also studied.

In social network analysis studies which use bibliometric data, variables such as articles, citations, co-citation networks, collaborating authors or institutions are examined and some concepts are widely used. One of these concepts is centrality (Otte & Rousseau 2002, p. 441). There are different measures for centrality. Degree centrality, closeness centrality and betweenness centrality are among the most frequently used centrality measures.[2]

Findings

Turkey's contribution to the world's scientific literature has increased significantly during the last fifteen years. According to Thomson Reuters' classification, Turkey addressed researchers have made publications that belong to 247 different fields. Surgery is the most productive field of Turkish scholars (14,365 publications). Therefore, in this paper all the examples were based on this discipline's data.

Figure 2a and Figure 2b illustrate the collaboration network of Turkey addressed surgery publications. The most collaborative partner of Turkey is the USA in surgery field (Figure 2b). Czech Republic, Japan and Lebanon have high betweenness centrality ratios. This means

[2] The definitions of centrality measures and detailed information can be obtained from Otte & Rousseau 2002, pp. 442-443.

these three countries are core nodes that make connections to other nodes (countries). For example, Lebanon supply linkage for some countries such as Saudi Arabia, Syria and Oman.

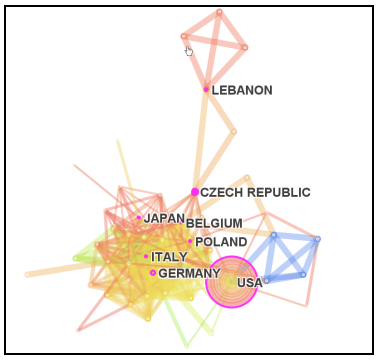

Figure 2a: Turkey's collaboration network (centrality)

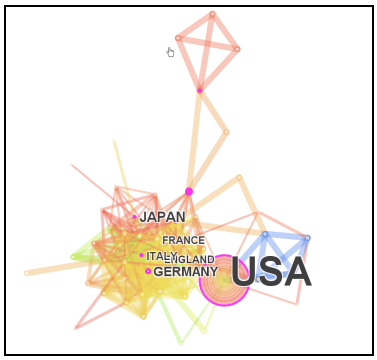

Figure 2b: Turkey's collaboration network (frequency)

In Figure 3, the network structure of keywords' (descriptors and identifiers) co-occurrences in surgery field has been revealed. As a keyword, "brain" has a pivotal node in the network with the highest betweenness centrality (0.16). In the figure, the colours show the first connections in time.

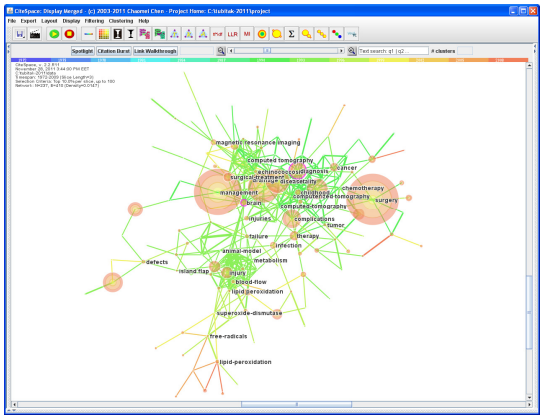

Figure 3: The keyword network of Turkey addressed surgery publications

Conclusion

SNA as research method has been used by many scholars of various scientific disciplines in the last years. Our literature search supported that SNA is an emerging research field. In summary, this study aimed to create awareness about the use of SNA method in bibliometric studies.

References

Chen, C. (2004). Searching for intellectual turning points: Progressive knowledge domain visualization. *Proceedings of the National Academy of Sciences of the USA*, 101(Suppl. 1): 5303-5310. Retrieved November 24, 2011 from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC387312/pdf/1015303.pdf.

Chen, C. (2011). CiteSpace: Visualizing patterns and trends in scientific literature. Retrieved November 24, 2011 from: http://cluster.cis.drexel.edu/~cchen/citespace/.

Larivière, V., Gingras, Y. & Archambault, E. (2006). Canadian collaboration networks: A comparative analysis of the natural sciences, social sciences and the humanities. *Scientometrics*, 68, 519-533.

Newman, M.E.J. (2001). Scientific collaboration networks. I. Network construction and fundamental results. *Physical Review E*, 64, 016131.

Otte, E. & Rousseau, R. (2002). Social network analysis: a powerful strategy, also for the information sciences. *Journal of Information Science*, 28, 441-453.

Perianes-Rodríguez, A., Olmeda-Gómez, C. & Moya-Anegón, F. (2010). Detecting, identifying and visualizing research groups in co-authorship networks. *Scientometrics*, 82, 307-319.

Pritchard, A. (1969). Statistical bibliography or bibliometrics? *Journal of Documentation*, 25, 348-349.