Visualization of the Knowledge Base and Research Front of Scientific Collaboration Research

Jianhua Hou[1] and Chaomei Chen[2]

[1] *hqzhixing@gmail.com*
WISE Lab S&T Management and Ethic research centre, Dalian University of Technology, Dalian-116085 (China) , The college of Information Science and Technology, Drexel University, Philadelphia PA-19104 (U. S.)

[2] *chaomei.chen@cis.drexel.edu*
The college of Information Science and Technology, Drexel University, Philadelphia PA-19104 (U. S.)

Introduction

Collaboration has become an integral part of research in the world of 'big science' (Galison & Hevly, 1992; Weinberg, 1961). An individual scientist can seldom provide all of the expertise and resources necessary to address complex research problems (Hara et. al. 2003). The trend toward specialization in science has brought a need for multidisciplinary collaboration to bring together the knowledge, skills, and abilities required for the advancement of research (Stevens & Campion, 1994). Co-authorship of a paper can be thought of as documenting collaboration between two or more authors (Newman, 2004). Collaborative scientific research, formally acknowledged by co-authorships of scientific papers, originated, developed, and continues to be practiced as a response to the professionalization of science (Beaver & Rosen, 1978). Katz and Martin (1997) distinguish between collaboration at different levels.

The goal of our study is to survey and analyze the body of literature on scientific collaboration with the aid of knowledge domain visualization techniques by co-citation analysis. We aim to highlight central research topics and the research front concerning scientific collaboration.

Data Source

We retrieved documents by entering the topic words *scientific collaboration* or *scientific coope ration* into the Web of Science. The search was limited to the period 1993–2007 and to the documental typology "articles" so that only original research papers would be analyzed. There are 1455 document records that contain 28355 citations in total. The data distribution in terms of time (annual) shows that the articles on scientific collaboration are on the rise, specifically from 46 in 1993 to 191 in 2007. Then the data was processed by using CiteSpace software (Chen, 2006).

Co-citation linking refers to the frequency of the same article being cited by another article, while co-citation strength refers to the mutual citation of articles among themselves. The centrality of the citation nodes and the frequency of citation are used to achieve maximum analysis of the clusters.

CiteSpace, a visualized information analyzing system developed by Chaomei Chen of Drexel University, is an effective instrument for the detection and analysis of changes in the knowledge base and research front.

Visual Analysis of the Scientific Collaboration Domain

We ran CiteSpace to obtain the document co-citation network shown in Fig.1. We then ran "Find Cluster" to note clusters of documents and used TF*IDF to analyze the clusters (Fig. 2). There are 47 document clusters in the scientific collaboration network. The different colours in the cluster show the evolutionary relationship between knowledge base and research front in the field of scientific collaboration; different clusters also typify the evolution of research bases in different fields. We used TF*IDF to analyze and label the clusters of documents by title words.

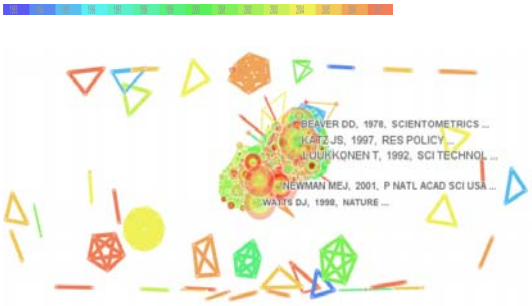

Figure 1□Map of the document co-citation network of the scientific collaboration domain

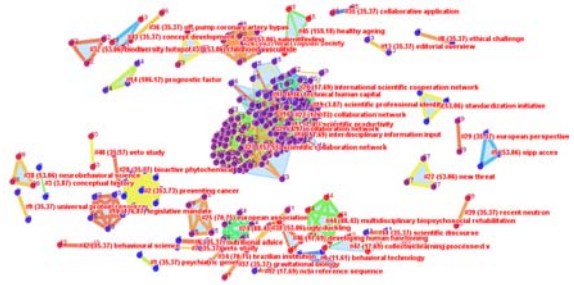

Figure 2. Labeled clusters (from using TF*IDF)

The knowledge base itself changes with the research front over time. Research front, the way

we use it here to highlight the emergence or bursting forth of a new trend, is the definition domain of time as a variable and its knowledge base as its co-domain. Therefore, the knowledge base of a research front may thus be defined as the citation route of the research front in scientific documents.

We list all nodes of the network ranked by frequency and the top 5 clusters by size using CiteSpace. The 10 most frequently cited articles indicate the research focus that attracts the most attention and some of the basic topics in this field. Based on the recognition of the keywords in the cluster, it is possible to draw out the topics in the knowledge base of the scientific collaboration domain and its research focus. We may infer from the statistics that the following are the most heatedly discussed topics:

1) Concerning scientific collaboration:

What is scientific collaboration; How to measure scientific collaboration; How was scientific collaboration originated; What is the relationship between scientific collaboration and productivity;

2) Related studies in the scientific cooperative framework:

The "Small Worlds"; Dynamics of collaboration of small world framework□The vertex connectivity follows a scale-free Power-law distribution; The statistical mechanics of network topology and dynamics□The shortest route, framework priority, centrality;

3) Concerning international scientific collaboration:

Differences in co-authoring international collaboration; Structure of the international scientific cooperative framework; Influence of cognizance, history, society, geography and economy on international collaboration; Countries' orientations in international scientific collaboration;

After a detailed scrutiny of the burst articles that represent the research front, we find that, in the late 20th century, the articles with the highest burst rate are mostly concentrated in the origin and definition studies of international scientific collaboration and in its influence factors. In the first years of the 21st century, new trends appeared in scientific collaboration research: scientific collaboration networking, network structure and paradigms. Furthermore, we find that the techniques of social network and scientometrics as applied in international scientific collaboration are being studied to a considerable degree. Particularly, the quantitative study of the network structure of scientific collaboration and the application of paradigms in scientific collaboration are attracting increasing academic attention.

Conclusion and Discussion

Scientific collaboration is a product of the professionalization of scientific endeavours. The development of science fosters collaboration among scientific researchers, institutions and even those from different nations. International collaboration in science has itself long been a focus of research, the study of which involves large-scaled cross-disciplinary efforts, information science, computer science, and scientometrics. The new techniques of information science, social network and the Internet have sped up the application of a multi-layered technology in this field. The map of document co-citation network of the scientific collaboration shows the evolution of the research knowledge domain. From 1980s to 2000, the majority of the research foci are on the origin and the development of scientific collaboration including conception and the measurement. After 2000, with scientists from physics and mathematics adding to the research in the field, there are new research trends emerging such as those involving analyzing scientific collaboration networks using complex theory, bibliometrics and social network technology as represented by Newman, Albert and Glanzel.

References

Beaver, D. D. & Rosen, R. (1978). Studies in scientific collaboration. Part 1□The professional origins of scientific coauthorship. *Scientometrics,* 1, 65- 84.

Chen, C. (2006). CiteSpace II: Detecting and visualizing emerging trends and transient patterns in scientific literature. *Journal of the American Society for Information Science and Technology*, 57, 359 - 377.

Galison, P. & Hevly, B. (Eds.). (1992). *Big science: the growth of large-scale research. Stanford*, CA: Stanford University Press.

Noriko Hara，Paul Solomon, Seung-Lye Kim, and Diane H. Sonnenwald. (2003). An Emerging View of Scientific Collaboration: Scientists' Perspectives on Collaboration and Factors that Impact Collaboration. *Journal of the American Society for Information Science and Technology*, 54(10):952–965, 2003.

Katz, J. S. & Martin, B. R. (1997). What is research collaboration. *Research Policy,* 26, 1- 18.

Newman, M.E.J. (2004). Co-authorship networks and patterns of scientific collaboration. *Proceedings of the National Academy of Sciences*, 101, 5200–5205.

Stevens, M. J. & Campion, M. A. (1994). The knowledge, skill, and ability requirement for teamwork: implications for human resource management. *Journal of Management,* 20, 503–530.

Weinberg, A. M. (1961). Impact of large-scale science on the United States. *Science,* 134, 161–164.