A Study on the Scientific Research Collaboration Network of "985 Project" Universities in China

Qiu Junping[1] and Wen Fangfang[2]

[1]*jpqiu@whu.edu.cn*
Wuhan University, Research Center of Chinese Science Evaluation (RCCSE), Hubei Wuhan, (P.R.China)

[2]*wenfangfang_zzu@yahoo.com.cn*
Wuhan University, School of Information Management (SIM), Hubei Wuhan, (P.R.China)

Abstract
Universities have become the main component in the national scientific innovation system. As the collaboration between different universities strengthens, the study on the scientific research collaboration is of increasingly vital value and importance. This paper chooses the "985 Project" universities as the sample to study on their scientific research collaboration relationship from the perspective of co-authorship based on the methods and tools of Social Network Analysis. The research results illustrate that the 39 universities has already established the primary scientific research collaboration, but the intensity still waits for enhancement. Besides, there exists significant correlation between scientific output and scientific collaboration. It means that to some extent to strengthen collaboration intensity can improve the quantity of universities' scientific output.

Keywords
Scientific collaboration; Collaboration relationship; Social Network Analysis; Co-authorship

Introduction
As early as the "Invisible College", there were universal collaboration in the academic community. After the 2nd World War, science presents the tendency of socialization. The organized scientific research obtained the huge success while the collaboration between researchers also strengthened. Especially under the background of "Big Science era", scientific research work became much more complicated and overlapping. Many projects, particular some major projects, can't be completed by one person or one organization. Therefore, no matter in the field of Science & Technology science or Humanities & Social science, scientific research collaboration strengthens unceasingly, with the collaboration ratio rises steadily to the direction of high level, deep level and multi-position[1]. Just like the founder of Cybernetics, Norbert Wiener said that the era of Edison individual invention had passed by and we had entered the era of scientific research collaboration.

As the unit crowded with knowledge and talents, university is shouldering the dual missions of scientific research and personnel training [2]. Since the reform and open policy, universities' quantity and size in China are expanding and their education quality and R&D level are also growing. Universities have become the essential component of national innovation system. With the collaboration tendency enhancing, scientific research collaboration between universities increases day by day. The collaboration between universities has many kinds of forms, such as undertaking R&D programs collaboratively, promoting R&D achievements collaboratively, publishing co-author papers and so on [3]. And co-author paper is the most direct way to collaboration and it is also the main manifestation of collaboration achievement. In this study we choose the 39 "985 project" universities in China as the sample. By means of the Social Network Analysis tools and methods, we discuss the scientific research collaboration between those 39 universities from the aspect of co-author papers statistic, in order to find some characteristics and rules to assistant the decision-making of the scientific research management departments.

Research Methods

2.1 Research Sample

This study chooses the "985 project" universities as research sample. Since those 39 universities play the critical role in the higher education of China. No matter personal training or scientific research, those 39 "985 project" represent the first-class level in China. Their names and location are listed in the Table 1.

Table 1 The list of "985 project" universities

Province	*Name*	*Province*	*Name*
Beijing	Tsinghua Univ	Shannxi	Xi'an Jiaotong Univ
	Beijing Univ		Northwestern Polytechnical Univ
	Beijing Univ of Aero & Astr		Northwest A&F Univ
	Renmin Univ of China	Shanghai	Fudan Univ
	Beijing Normal Univ		Shanghai Jiaotong Univ
	Beijing Institute of Technology		Tongji Univ
	Minzu Univ of China		East China Normal Univ
	China Agricultural Univ	Anhui	Univ of S&T of China
Fujian	Xiamen Univ	Sichuan	Sichuan Univ
Gansu	Lanzhou Univ		Univ of Electronic S&T of China
Guangdong	South China Univ of Technology	Tianjin	Tianjin Univ
	Sun Yat-sen Univ		Nankai Univ
Heilongjiang	Harbin Institute of Technology	Chongqing	Chongqing Univ
Hubei	Huazhong Univ of S&T	Jiangsu	Southeast Univ
	Wuhan Univ		Nanjing Univ
Hunan	Hunan Univ	Liaoning	Northeast Univ
	South central Univ		Dalian Univ of Technology
	National Univ of Defense Technology	Shandong	Shandong Univ
Jilin	Jilin Univ		Ocean Univ of China
Zhejiang	Zhejiang Univ		

2.2 Data Source

This analyzes the scientific collaboration network of "985 project" from the aspect of co-authorship papers. Two universities' names appear together in the identical paper can be considered as collaboration while the number of co-authorship papers can be considered as the collaboration frequency. We choose the CNKI database as the data source which is the largest Chinese Journals Full-text database in the world. The CNKI database contains 6968 academic journals covering all kinds of fields such as science, engineering, technology, agriculture, philosophy, medicine, humanities and social science. We use each two "985 project" universities' names as retrieval terms to retrieval the number of their co-authorship papers which can be seen as the collaboration frequency of the two universities. Time span is ten years from 2000 to 2009. Based on the retrieval results we construct the initial co-authorship matrix for the 39 "985 project" universities with the value in the matrix are the number of co-authorship papers as well as the collaboration frequency of each pair of universities.

2.3 Analysis Tools

This study makes use of some Social Network Analysis tools to give statistic analysis and visualization presentation to the collaboration relation of those 39 "985 project" universities. The whole study mainly uses three kinds of tools: Ucinet, Pajek and SPSS.

（1）Ucinet

Ucinet is one of the most popular Social Network Analysis tools which can process many forms of documents such as .txt, .excel and so on. Meanwhile Ucinet has strong function of matrix analysis and can compute the density and coreness of social network, as well as Core/Periphery analysis and role analysis etc. This study mainly uses the functions of density and coreness calculation and Core/Periphery structure analysis.

（2）Pajek

Pajek is a visualization procedure especially for the large scale network and its strong point is graph function. Pajek only contains a few statistic functions. Thus most statistic work is completed by Ucinet and SPSS. We just make use of the graph function of Pajek to produce the collaboration network figure.

（3）SPSS

SPSS（Statistical Product and Service Solutions）is a statistic software integrating data calculation and analysis, with the basic functions of data management, statistic analysis, graph analysis etc. In this study we mainly use SPSS to carry on curve fitting, correlation analysis and regression analysis.

Results Analysis

3.1 Entire Network Analysis

Those 39 universities construct 741 collaboration pairs and 725 of them have collaboration relation. There are only 16 structural holes and the network is almost complete graph. It means that those “985 project” universities have established broad collaboration relations initially. But from the aspect of collaboration frequency, there is great difference between each pair, with the max value is 1434 while the minimum is only 1. Entire network analysis can demonstrate the overall appearance of the scientific collaboration network of “985 project” universities at the macroscopic level. From which we can find that some universities have frequent collaboration, such as Wuhan University & Huazhong University of S&T, Fudan University and Shanghai Jiaotong University, Hunan University & South Central University, Tsinghua University & Beijing University, with the frequencies are 1434, 1271, 1051, 913, 836 respectively.

3.2 Density Analysis

From the network density analysis, we have found that the overall density of the collaboration network of “985 project” universities is as high as 0.9771 regardless of collaboration intensity. When we take the collaboration intensity into account and set threshold value to the collaboration frequency. As the collaboration relations below some threshold value are deleted, the density of whole network will decrease exponentially along with the growth of threshold value (see as the figure 1).

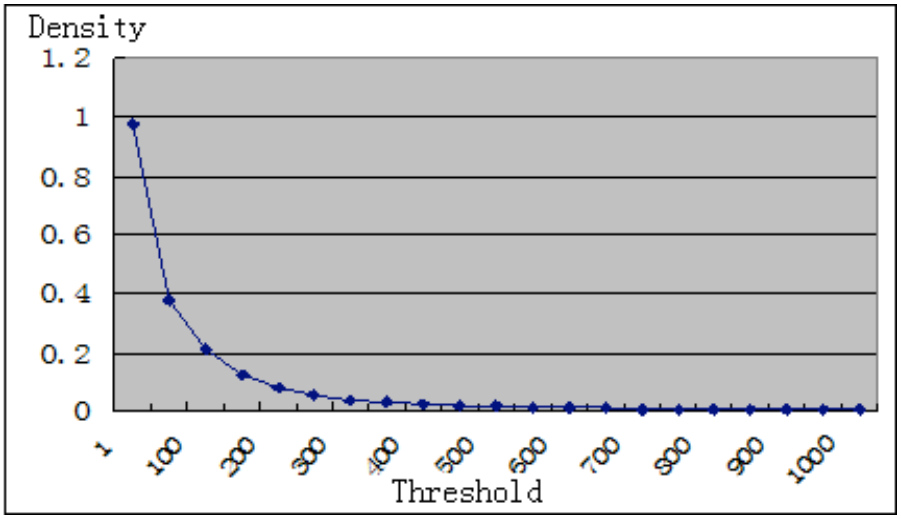

Figure 1 Density under different threshold

It is important to take the collaboration intensity into account when we calculate the network density. The scientific collaboration network based on co-authorship papers is an assignment network. It means that there is a value between each pair of universities which represents the collaboration intensity. According to the theories and methods of Social Network Analysis, the study of evaluation network must be based on the collaboration intensity. Especially for entire network analysis, the result can not illustrate the realistic situation of the network without considering intensity. For example, at the beginning we get the overall density without considering intensity, the overall density of those 39 universities is as high as 0.9771. But when we take intensity into account and delete the collaboration relation under some threshold we have found that the density is decreasing exponentially along with the growth of threshold of intensity. It means that the collaboration relation of those "985 project" universities is not very close. Therefore, it is much more reasonable and objective to introduce the index of intensity into the calculation and analysis of network density.

When we introduce the index of intensity into the density analysis we have found that most universities stay at the level of low-frequent collaboration. The collaboration relations below the 50 frequency account for nearly 2/3. It means that many universities' collaboration relations are established by several authors. And most of those collaborators are supervisors and graduate students. Namely an in-service teacher of A University studies at B University. He publishes a co-author paper with his supervisor. Then paper will have two affiliations: A University and B University. That is the two universities have scientific collaboration relation. Usually an in-service graduate student may have more than one co-author papers with supervisor during the undergraduate period. Thus we find that many co-author papers between two universities are completed by graduate student and his supervisor, especially for those low-frequent collaboration relations. In fact such kind of collaboration relation is very fragile. If the in-service student graduates, then the collaboration relation between A University and B University will be broken. From density analysis we have found that most "985 project" universities just have low-frequent collaboration relation which is very fragile. Based on such a discovery we give the suggestion that Chinese universities should establish much more broad, comprehensive and intensive collaboration with each other.

3.3 Core/Periphery Structure Analysis

Through Core/Periphery Structure Analysis we have get the coreness of each "985 project" university. Coreness is an index representing the node's position and influence in the scientific collaboration network [4]. There are 8 universities have higher coreness which means they are much more active and influential in the scientific collaboration network. They

are Beijing University, Shanghai Jiaotong University, Fudan University, Tsinghua University, Tongji University, Huazhong University of S&T, Zhejiang University, Wuhan University. Core/Periphery model is deduced from the realistic world which can reflect the abstract and simple relationship scheme of realistic social phenomenon. A.Bavelas [5] had confirmed a supposition in his pioneering research. That is the closer to the core the more influential of the actor. Such a discovery can explain the knowledge dissemination, information transmission, resources sharing and organization efficiency of social network. Generally speaking, those nodes located in the central position always play critical role in the entire network. They are much more active in the network's activities and have great influence to other universities and to the efficiency of the whole network.

The essence of scientific collaboration is resources sharing and knowledge dissemination and its goal is to realize the result of "1+1>2" to improve the efficiency of the whole organization. From the result of Core/Periphery Structure Analysis, we have a quantitative understanding of the position of each "985 project" university. In the scientific collaboration network of "985 project" universities, there are 8 universities with higher coreness are located in the core position in the network which means that those 8 universities have the stronger openness and influence. They play a critical role in the scientific research field and even the whole higher education system. This discovery is useful to assistant the decision-making of scientific research management department. Since these universities play a key role to enhance the collaboration intensity and scientific efficiency, if governments give some investment priority to these universities with high coreness, the overall scientific efficiency of the collaboration network of Chinese universities will be enhanced greatly.

3.4 Degree Centrality Analysis

Through degree centrality analysis and correlation analysis, we have found that there is significant correlation between the scientific output (the number of papers) and degree centrality of each university. See as the Figure 2, the correlation coefficient is as high as 0.903 which means that the correlation analysis result is significant. The higher the degree centrality is, the higher the scientific output is. That is those universities with higher degree centrality will have higher scientific output and publish more papers. Degree centrality is a very important term in the Social Network Analysis which is deduced from the concept of "star" in the Sociomentrics [6]. The degree centrality of one node depends on the number of other nodes connected to the node directly. Thus a node of higher degree centrality is located in the core of network and has close and direct relations with many other nodes. In this study one university's degree centrality is calculated by the number of its collaborators and the intensity of their collaboration relations. Thus the degree centrality can reflect a university's collaboration degree in the scientific collaboration network both from the quantity and intensity of collaboration relations. High degree centrality means that the university has close and intensive collaboration relations with many other universities.

We make use of SPSS to carry on the correlation analysis with the two indexes, the degree centrality and the scientific output (the number of papers). From a new aspect it confirms the controversial problem that is if collaboration can enhance scientific output. Then we get the conclusion that to enhance collaboration can promote scientific output. Such a discovery is consistent with Beaver and Rosen's [7] viewpoint. In fact, in the scientific collaboration network, one university's collaboration degree and scientific output have interactive relation. Namely collaboration can promote scientific output while the development of scientific research can also propel the development of scientific collaboration. The interactive relation between scientific output and collaboration is the interactive relation between productive force and production relation. Only by enhancing team collaboration and adjusting the

relation between scientific production and collaboration, can we promote the development of productive force and scientific output [8].

Correlations

			papers	degree cent
Spearman's rho	papers	Correlation Coefficient	1.000	.903**
		Sig. (2-tailed)	.	.000
		N	39	39
	degree centrality	Correlation Coefficient	.903**	1.000
		Sig. (2-tailed)	.000	.
		N	39	39

**. Correlation is significant at the 0.01 level (2-tailed).

Figure 2 Correlation Analysis of papers and degree centrality

Conclusion

However because of some disturbance factors, such as false co-author, fake collaboration and so on. Some people still question if it is scientific and reasonable to study academic collaboration by co-author papers. We also acknowledge that this method has some limitation. Besides, our sample is also very limited. So this study has some deficiencies as following. Firstly, paper is only a kind of scientific output. Other forms, such as patents, monographs, report and so on, are also the important component of scientific output. We analyze the scientific collaboration of universities only from the aspect of co-author papers. Our discoveries and conclusions are of great limitation. It can not reflect the comprehensive and realistic collaboration situation of those universities. Secondly, our study mainly concentrates on the macroscopic statistic and analysis and our conclusions are also stay at the preliminary level. Few essential rules and characters are detected. For the two reasons above, this paper may not precisely reflect the scientific collaboration relation of those “985 project” universities. We just intend to make use of Social Network Analysis methods and tools to illustrate the overall situation of scientific collaboration of “985 project” universities during the recent 10 years from the aspect of co-author papers statistic, in the hope that we can get a quantitative understanding of Chinese universities’ scientific collaboration. As for the problems and deficiencies of this study, we will supplement and better it in the following research.

Acknowledgments

This paper is supported by the Independent Fund of PhD Scientific Research of Wuhan University which is the Fundamental Research Funds for the Central Universities. We thank all the members of the project for their helpful comments on and insights into this work.

References

Zheng Haiyan. (2007). Statistical analysis of Chinese humanities & social sciences papers included in the SSCI and A&HCI(1995～2004). *Management and Review of Social Sciences*, 4, 47-55.

Qu Zhentao.(2000).The historical evolution of university’s social function. *Guangming Daily*, 2000-12-01.

Liang Liming.(2008). Strong geographical preferences in scientific collaboration between "985 Project" universities. *Science of science and management of S&T*, 11, 112-116.

Luo Jiade.(2005). Social Network Analysis. Beijing: Social Science Document Press.

Freeman. L. C. (1979). Centrality in social networks: Conceptual Clarification. *Social Networks*, 1,215-239

Wasserman Stanley, Faust K. (1994).Social Network Analysis: Methods and Applications. Cambridge:Cambridge University Press.

D. Beaver, R.Rosen. (1978). Studies in scientific collaboration. *Scientometrics*, 1, 65-84.
Ding Kun. (2009).The network character and college team status in coauthor network of science of science and management of science and technology. *Science of Science and Management of S&T*. 12,10-15