Scientometrics
DOI 10.1007/s11192-013-1072-y

An investigation of collaborations between top Chinese universities: a new quantitative approach

Weichao Wang · Yishan Wu · Yuntao Pan

Received: 14 May 2013

Abstract A new quantitative method is introduced to analyze the collaboration among different organizations. The method defines the collaboration score based on the number of people involved in collaboration, and then the collaboration strength is obtained by summing up the collaboration scores with this method. We choose "Project 985" universities, which represent the top universities in China, as an example to study the collaboration network, strength in leading collaboration and strength in participating collaboration. Results based on Scopus show some characteristics of such collaboration and verify the feasibility of the new approach.

Keywords Quantitative method · Collaboration score · Collaboration strength · "Project 985" university

Introduction

With the increase of interdisciplinary communication and the acceleration of innovation, completing complex research projects by a single organization or even a single country is a "mission impossible", especially for big-science projects. In order to adapt to the new situation, scientific collaboration is an inevitable trend. New collaboration patterns are changing the global balance of science, so established superpowers need to keep up or be left behind, says Jonathan Adams (2012). Consequently, collaboration is of interest to more and more researchers. Newman analyzed the collaboration networks of scientists in biology and medicine as well as various sub-disciplines of physics, and found that scientific

W. Wang (✉)
Peking University, Beijing 100871, China
e-mail: wangweichao526@gmail.com

Y. Wu · Y. Pan
Institute of Scientific and Technical Information of China, Beijing 100038, China

Published online: 22 June 2013

communities seem to constitute a "small world", in which merely five or six steps were needed to get from one randomly chosen scientist in a community to another (Newman 2001a). Then he used networks in biology, physics and mathematics to answer a broad variety of questions about collaboration patterns, such as the numbers of papers authors write, how many collaborators they have, what the typical distance is between scientists, and how patterns of collaboration vary between subjects and over time (Newman 2004). Jarneving carried out a comprehensive study on regional-foreign research collaboration and applied various multivariate methods for the description of collaborative networks of various compositions and at various levels of aggregation (Jarneving 2010). Qiu and Ma (2011) choose the Chinese top universities as the example to study their scientific research collaboration relationship from the perspective of co-authorship based on the methods and tools of social network analysis.

To the best of our knowledge, the scientific collaboration among organizations or people has been all studied quantitatively by frequency or time, which means if two organizations appear in the same paper, one collaboration is counted simply (Bordons et al. 1996; Luukkonen et al. 1993; Wang et al. 2012; Newman 2001a, b, c; Newman 2004; Jarneving 2010). However, the fact is that the number of people involved in a scientific paper usually reflects the difficulty and complexity of the project. As to the process of collaboration, a two-person work is usually not on the same scale as a ten-person work no matter the depth, width or academic contribution of the work. Therefore, considering all the collaborations equally is not fair and rational. In addition, first affiliation included collaborations and first affiliation excluded collaborations need to be distinguished as well. Because the first affiliation often dominates the collaboration, the strength of collaboration in the case of first affiliation excluded collaborations, like the collaboration between second affiliation and third affiliation, appears weak comparing to first affiliation included collaborations.

In this study, we introduced the "collaboration score" concept based on the number of people involved in collaboration. The collaboration strength was calculated by additively, and strength in leading collaboration and strength in participating collaboration were discussed respectively. We use this new approach to analyze collaborations among top Chinese universities. Data are taken from Scopus.

Data and methods

Data sources

In this study, data is collected from Scopus, which is a product of Elsevier publishing corporation. Scopus has already covered 738 high quality Chinese journals (ISTIC 2012). The broad coverage of local and international journals by Scopus will make our research more reliable, which is one of the most important reasons we choose Scopus as our data source. Data refer to publications from the year 2011.

The object of our research in this study is the so-called "Project 985" universities (hereinafter referred to as 985UNI), whose total number is 39. "Project 985" was initiated by the Chinese ministry of education on May 4, 1998 (the fifth month of 1998, therefore called "Project 985" for short), aiming to promote the development and reputation of the Chinese higher education system. The project involves both national and local governments, which allocate large amounts of funds to certain selected universities in order to build them into new research centers, improve their facilities, help them hold international

conferences in China or send Chinese faculty abroad to attend international conferences, and attract world-renowned faculty and visiting scholars (Fu et al. 2011; Zhang et al. 2013). It goes without saying that 985UNIs represent the top-level of the pyramid in China's higher education system. With strong teaching and research capacity, they always attract the attention from researchers as well as the general public. Table 1 lists all the 39 "985UNIs". In the following sections, the 985UNI-985UNI collaborated papers are defined as those co-authored papers that have more than one institutional affiliation, the first one of which is a 985UNI and at least one of the other ones is also a 985UNI; the 985UNI-University collaborated papers are defined as those co-authored papers that have more than one institutional affiliations as well, but the first one of which is a 985 UNI and at least one of the rest is any university (rather than a research institute, a company, a hospital, etc.), no matter 985UNI or not.

The number of scientific papers from Mainland China recorded by Scopus in the year 2011 is 234,600 (excluding papers by Hong Kong, Macao and Taiwan), contributing almost 12.26 % of the whole world's production. 985UNIs published 90,800 papers among the total, including 50,400 collaborated papers as defined above.

Methods

"The total number of scientists is growing largely by the square of the number of outstanding scientists", says Derek J. de Solla Price in his famous book "Little science, big science". As is well known, the "square root law" was first proposed in this book, which states that half of the scientific papers are contributed by the top square root of the total number of scientific authors. We could follow his suit to define the collaboration score. Consider a 100-author paper, applying the "square root law" means that we assume that there are about ten authors that have a major contribution. These ten authors can be considered to be outstanding contributors.

The number of authors in a paper can reflect the degree of collaboration. The more authors involved, the bigger the collaboration scale. Considering the predominant contribution of relatively outstanding authors to collaboration in one paper, the collaboration score can be expressed as

$$CollaborationScore = \sqrt{a} \tag{1}$$

by assuming the collaboration score of a single author paper is 1 point, where *CollaborationScore* is the collaboration score of one paper, a is the number of authors, the square root of a is the number of outstanding authors. It is more reasonable to use "square root law" to define the collaboration score than simply use collaboration authors only.

In this study, the collaboration strength was calculated additively. Assuming S_{AB} is the collaboration strength between Organization A (the first affiliation) and Organization B (the participating affiliation), than S_{AB} is calculated as

$$S_{AB} = \sum_{\mathrm{i}=1}^{n} \left(\frac{CollaborationScore_i}{Org_i} \times t_i \right) \tag{2}$$

where i is one of the Organization A's papers whose collaboration partners include Organization B, Org_i is the number of organizations in paper i, $CollaborationScore_{\mathrm{i}}$ is the collaboration score of paper i, n is the total number of Organization A's papers whose collaboration partners include Organization B, while t_i is the number of occurrence of

Table 1 List of all the 39 "985UNIs"

Location	Full name	Abbreviation
Beijing	Tsinghua university	THU
Beijing	Peking university	PKU
Beijing	Beijing university of aeronautics and astronautics (Beihang)	BUAA
Beijing	Renmin university of China	RUC
Beijing	Beijing normal university	BNU
Beijing	Beijing institute of technology	BIT
Beijing	Minzu university of China	MUC
Beijing	China agricultural university	CAU
Shanghai	Fudan university	FDU
Shanghai	Shanghai Jiao Tong university	SJTU
Shanghai	Tongji university	TJU
Shanghai	East China normal university	ECNU
Shaanxi	Xi'an Jiao Tong university	XJTU
Shaanxi	Northwestern polytechnical university	NPU
Shaanxi	Northwest A and F university	NAFU
Hunan	Hunan university	HNU
Hunan	Central south university	CSU
Hunan	National university of defense technology	NUDT
Hubei	Huazhong university of science and technology	HUST
Hubei	Wuhan university	WHU
Guangdong	Sun Yat-Sen university	SYSU
Guangdong	South China university of technology	SCUT
Jiangsu	Nanjing university	NJU
Jiangsu	Southeast university	SEU
Sichuan	Sichuan university	SCU
Sichuan	University of electronic science and technology of China	UESTC
Tianjin	Tianjin university	TU
Tianjin	Nankai university	NKU
Shandong	Shandong university	SDU
Shandong	Ocean university of China	OUC
Liaoning	Northeastern university	NEU
Liaoning	Dalian university of technology	DUT
Fujian	Xiamen university	XMU
Gansu	Lanzhou university	LZU
Heilongjiang	Harbin institute of technology	HIT
Jilin	Jilin university	JLU
Zhejiang	Zhejiang university	ZJU
Anhui	University of science and technology of China	USTC
Chongqing	Chongqing university	CQU

Organization B in paper *i*. In other words, if different schools or departments of the same organization appear more than once in a single paper, the collaboration score contributed by this paper needs to be multiplied by the number of occurrences. For example, a paper

whose first affiliation is College of Engineering in PKU has four collaboration partners, which are respectively department of physics in THU, department of chemistry in THU, department of physics in SJTU and school of optical and electronic information in HUST. If the total collaboration score of this paper is 1 point, then according to the above description, this paper's contribution to the collaboration strength between PKU and THU is 0.4 point (two-fifth of total collaboration score), its contribution to PKU-SJTU is 0.2 point (one-fifth of total collaboration score), while its contribution to PKU-HUST is also 0.2 point (one-fifth of total collaboration score).

For the same reason, the mutual collaboration strength is defined as

$$M_{AB} = M_{BA} = S_{AB} + S_{BA} \tag{3}$$

where S_{BA} is the collaboration strength between Organization B(the first affiliation) and Organization A (the participating affiliation).

Because of the significant place of first affiliation in collaboration, if the first affiliation is not a 985 UNI, then we regard this paper as representing only a weak collaboration among 985 UNIs and drop it. For example, if a paper's first affiliation is neither Organization A nor B, it will not be counted for the collaboration strength between Organization A and Organization B.

In the above study, we consider first author's affiliation as the first affiliation of a paper. In many cases, corresponding author usually takes the leading position in scientific collaborations. However, in most cases for Chinese publications, first author and corresponding author belong to the same country or to the same organization. For example, among the 234,600 papers form Mainland China we mentioned above which are counted by first author, over 98 % have the Chinese corresponding author. PKU has 3232 papers in the first author case while 3199 in the corresponding author case. For this reason, the collaboration strength does not appear to be much different between these two cases. In the following sections, we only take first author papers into consideration to reduce the workload.

Results and discussion

Collaboration strength network

According to Eq. (2), the collaboration strength matrix for 985UNI was obtained using traversal method for all the papers indexed by Scopus, as shown in Table 2. We see from the matrix in Table 2 that 71.01 % of the cells is non-empty (excluding the diagonal). If we set a threshold of 5 points to the matrix, this ratio would decrease to 21.63 % significantly. It may be inferred that, a 985UNI collaboration network is developing, though the intensity is still weak. The cell representing the strongest collaboration is SJTU-FDU with the value of 111.81 points, and that the value of FDU-SJTU is as high as 72.41 points. Consequently, the mutual collaboration strength between these two universities is strongest as well by adding up the values of the above two cells.

Based on Eq. (3), a network in terms of the mutual collaboration strength among 985UNIs was built. Here we use Pajek software for network analysis and visualization, as shown in Fig. 1. A threshold of 10 points was set in the figure to reduce the low-point lines so as to give a clearer picture. In Fig. 1, the width and grayness of lines are proportional to the value. The wider and the blacker are the lines, the stronger is the collaboration. Table 3 shows the top 15 pairs of 985UNIs in terms of mutual collaboration strength.

Table 2 Collaboration strength matrix of 985UNIs

First affiliation	Collaborating affiliation										
	PKU	BUAA	BIT	BNU	DUT	UESTC	NEU	SEU	FDU	NUD T	H I T
PKU	–	15.91	5.89	21.26	2.71	1.53	0.00	3.31	17.43	2.75	0.00
BUAA	25.29	–	3.16	3.10	3.15	0.00	0.00	0.00	4.54	1.67	3.16
BIT	10.81	9.87	–	1.00	0.00	1.90	0.00	0.60	0.00	2.12	6.44
BNU	29.18	2.80	2.61	–	1.17	2.81	0.58	0.93	1.37	0.00	1.45
DUT	5.47	0.00	1.73	1.82	–	0.00	11.29	2.41	2.79	0.00	32.01
UESTC	6.18	1.21	1.17	1.78	3.24	–	0.50	5.08	0.60	0.00	1.98
NEU	9.68	0.00	2.56	3.02	8.53	1.86	–	2.00	1.32	0.58	3.87
SEU	4.68	2.91	0.00	1.44	1.53	1.26	1.30	–	3.41	0.00	8.01
FDU	15.30	0.00	0.00	4.71	1.37	0.00	6.00	0.00	–	0.41	0.75
NUDT	0.00	4.75	0.00	0.00	0.00	3.10	0.00	1.53	2.66	–	0.00
HIT	12.14	8.96	15.54	0.00	38.86	2.04	6.41	2.57	3.52	1.30	–
											.....

As one can see from Table 3, 12 of 15 pairs are universities located in the same city, and two of the remaining three pairs are in the same region, with one being Northeast China and another being East China. Obviously, the collaboration among 985UNIs is intensely region-dependent. This regional bounding conforms to the minimum distance principle for collaboration and the economic rule of minimum input for maximum output. Among these relationships, three universities in Shanghai, SJTU, FDU and TJU, form a firm triangle. THU, PKU and ZJU spread most widely in collaboration with the other 985UNIs, while NPU, RUC and MUC are comparatively isolated in the network.

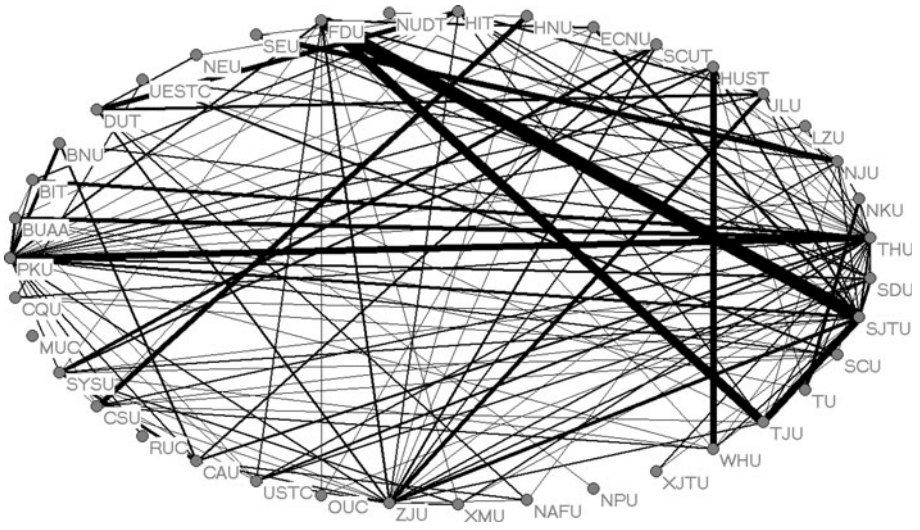

Fig. 1 Mutual collaboration strength network of 985UNIs

Table 3 Top 15 pairs of 985UNIs in terms of mutual collaboration strength

Rank	985UNI A	985UNI B	Mutual collaboration strength	City or cities involved
1	SJTU	FDU	184.22	Shanghai
2	FDU	TJU	115.10	Shanghai
3	SJTU	TJU	101.75	Shanghai
4	HUST	WHU	94.77	Wuhan
5	PKU	THU	89.76	Beijing
6	SEU	NJU	75.16	Nanjing
7	HNU	CSU	74.20	Changsha
8	DUT	HIT	70.87	Dalian-Harbin
9	BUAA	THU	63.25	Beijing
10	ZJU	JLU	60.25	Hangzhou-Changchun
11	NKU	TU	52.79	Tianjin
12	PKU	BNU	50.44	Beijing
13	BIT	THU	49.91	Beijing
14	SJTU	ZJU	48.36	Shanghai-Hangzhou
15	SCUT	SYSU	46.65	Guangzhou

Strength in leading collaboration (SLC) and strength in participating collaboration (SPC)

SLC was calculated by summing up any 985UNI's collaboration strength with all the other 985UNIs, while SPC was calculated by summing up all the other 985UNIs' collaboration strength with the given 985UNI. We ranked the 985UNIs by SLC and SPC in Table 4.

The top 3 SPC are THU, PKU and FDU. Besides, their SPC are all greater than SLC. It can be easily learned that other universities in the 985UNI network are willing to collaborate with them. This is because they have not only powerful scientific research capabilities, but also a disciplinary advantage, which was built by years of significant investment from the government and their long-term development goal of comprehensive university. On the other hand, the highest SLC is SJTU, showing its expertise in seeking opportunities to collaborate with others.

Both SLC and SPC show the activity and vigor of related universities in the collaboration network. As we can see from Table 4, five of the bottom ten in SLC ranking is located in the remote areas such as Northwest, Southwest and Northeast of China. There are also two liberal arts universities, one marine-oriented university and one with military background in the remaining five of the bottom ten. A similar phenomenon is observed in SPC ranking. The bottom ten according of the SPC consist of six remote universities, two liberal arts universities, one marine-oriented university and one with military background. Obviously, the geographical factor is of the most importance that impinges upon the ranking. Most of the remote universities have difficulties in communication, resource sharing, traffic and funding, dramatically falling behind those of universities located in Central China, East China and/or the coastal areas. Secondly, in such a 985UNI network dominated by comprehensive universities or science and engineering universities, the liberal arts universities and marine-oriented universities perform poorly because of their

Table 4 SLC and SPC of 985UNIs and their ranking

SLC rank	98UNI	SLC	SPC rank	985UNI	SPC
1	SJTU	396.07	1	THU	406.28
2	THU	287.57	2	PKU	355.41
3	PKU	258.79	3	FDU	354.64
4	FDU	253.69	4	SJTU	324.99
5	ZJU	243.19	5	ZJU	274.09
6	TJU	201.85	6	TJU	200.72
7	HIT	198.38	7	NJU	178.15
8	HUST	187.63	8	SYSU	175.23
9	SDU	176.96	9	HUST	160.05
10	JLU	171.31	10	USTC	152.03
11	CSU	162.87	11	DUT	150.19
12	SYSU	159.72	12	HIT	148.66
13	NJU	159.61	13	WHU	143.58
14	DUT	146.03	14	SDU	141.37
15	WHU	139.42	15	JLU	133.17
16	CAU	139.21	16	BUAA	132.22
17	BUAA	130.08	17	SCUT	126.95
18	SEU	124.58	18	NKU	124.11
19	SCUT	115.41	19	BNU	111.43
20	SCU	109.96	20	CSU	110.38
21	UESTC	105.96	21	CAU	107.43
22	USTC	105.46	22	SCU	95.48
23	HNU	102.36	23	XMU	94.73
24	XMU	99.40	24	HNU	94.71
25	TU	97.70	25	SEU	92.93
26	BNU	96.88	26	BIT	92.30
27	ECNU	95.51	27	ECNU	86.92
28	CQU	88.48	28	TU	85.19
29	BIT	86.08	29	LZU	80.07
30	NKU	75.17	30	CQU	71.20
31	XJTU	72.61	31	UESTC	61.44
32	LZU	70.29	32	NEU	55.76
33	NEU	61.02	33	XJTU	49.36
34	NAFU	52.78	34	NAFU	37.80
35	NUDT	49.37	35	NPU	33.51
36	OUC	47.53	36	NUDT	31.55
37	NPU	39.71	37	OUC	30.71
38	RUC	25.33	38	RUC	28.39
39	MUC	12.16	39	MUC	12.96

limited number of subjects as well as weak tradition in collaboration for scholars in social sciences and humanities. Moreover, for reasons of confidentiality and security, universities with military background usually lock the door against the others.

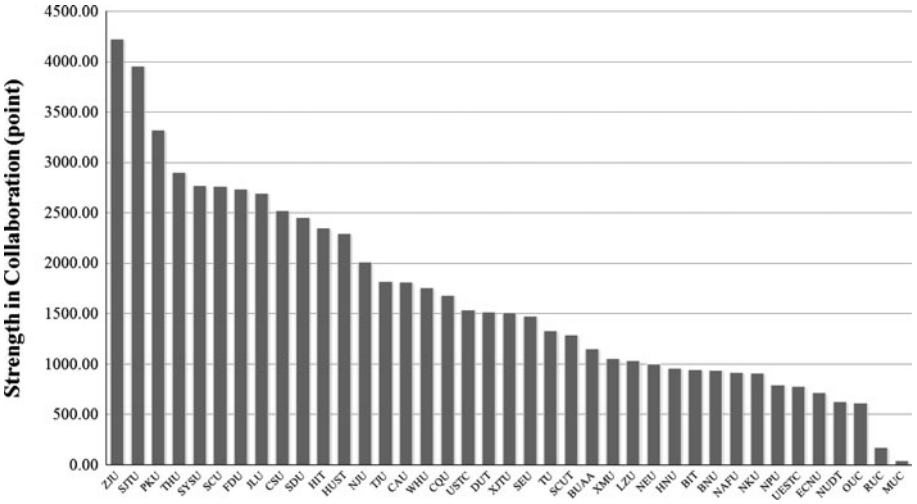

Fig. 2 Total collaboration scores of 985UNIs

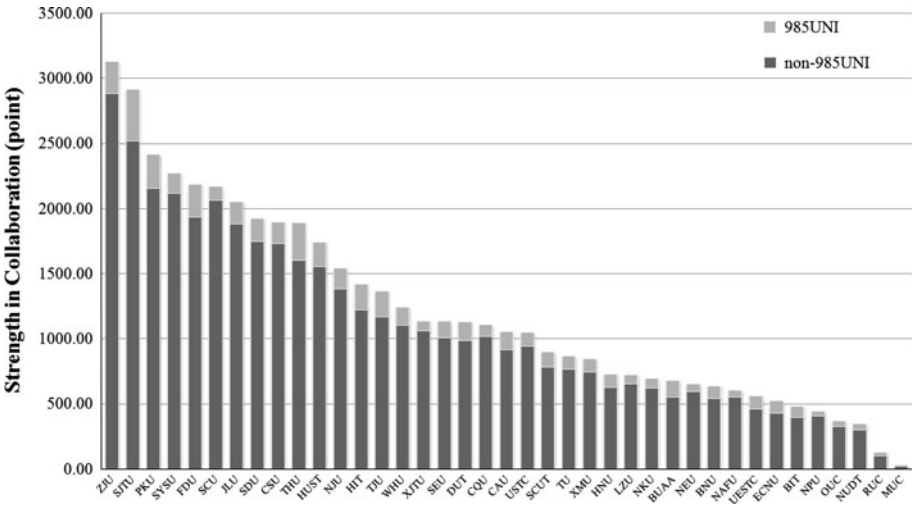

Fig. 3 Collaboration strength of 985UNI-university

985UNI-university collaboration

Breaking out of the world of 985UNI, we still used the above methods to perform statistical analysis on 985UNI-University collaboration. Total collaboration scores of 985UNIs, including their collaboration with universities, institutions, enterprises, hospitals, government agencies and so on, were obtained and ranked in Fig. 2. In this section, we only focused on their collaboration with universities, as shown in Fig. 3. Here, the grey part in columns represents collaboration with 985UNIs, while the dark part represents collaboration with non-985UNIs (including local and international universities). For 39

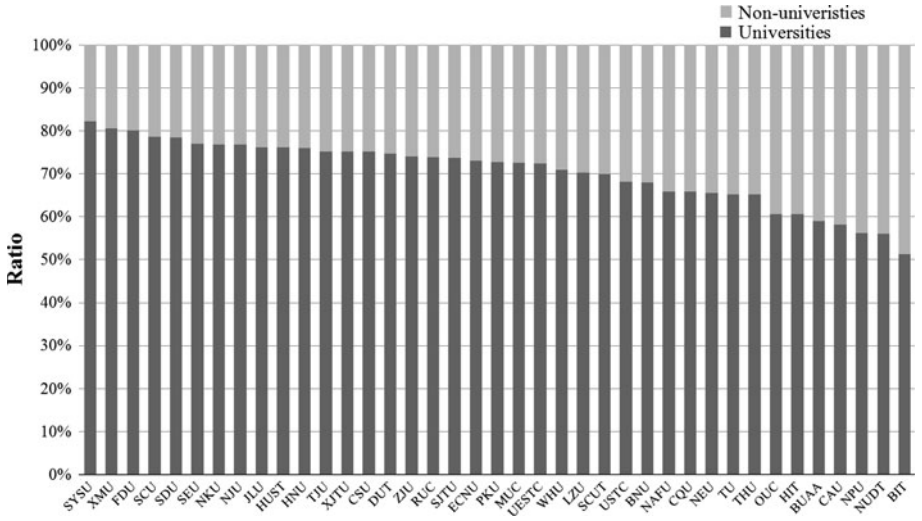

Fig. 4 Collaboration ratios of 985UNI-university

"985UNIs", the average ratio corresponding to the grey part is 10.96 %. SCU, XJTU and SYSU are the lowest three, with the ratio being 5.07, 6.41 and 7.03 %, respectively.

Because the university part is the biggest share of collaboration scores, the rankings in Figs. 2 and 3 do not vary much. Then, we plotted and ranked these shares in Fig. 4. The university part contributes at least half of the total scores, and the highest three, namely SYSU, XMU and FDU, even reach more than 80 percent. However, high percentage is not necessarily good because we should always emphasize the diversity of development and the commercialization of research results.

Conclusions

In this study, we introduced a new method to study the collaboration among different organizations quantitatively, with Chinese 985UNIs as an example. Some characteristics of such collaboration were discovered by analyzing the network, SPC, SLC and total collaboration scores.

The collaboration score was defined only by the number of people involved in this research. But beyond that, the subject factor needs to be considered. There is a huge gap among different disciplines in terms of the number of collaborators. This situation is similar to the impact factor in journal assessment, useful but not comprehensive. In addition, for papers published in journals of different levels, their contribution to the collaboration should not be given equal treatment. More reasonable indicators and weights need to be considered and designed in the following study.

Nowadays, especially in biomedical papers, it is not rare to see the special authorship functions such as "equal first authors" and "equal corresponding authors" (Hu 2009). Such cases may somewhat affect the collaboration strength in our approach. Unfortunately, this information cannot be found in Scopus yet.

All the results above are based on scientific papers. In our future work, we plan to use more kinds of research outputs, such as patents, research grants, research reports, economic

and social benefits from research and so on to achieve a better understanding of collaboration.

Acknowledgments The study was supported by the project of "National Natural Science Foundation of China" (70973118). The authors wish to thank anonymous reviewers who gave valuable suggestion that has helped to improve the quality of the manuscript.

Reference

Adams, J. (2012). Collaborations: the rise of research networks. *Nature, 490*, 335–336.

Bordons, M., Gomez, I., Fernandez, M. T., Zulueta, M. A., & Mendez, A. (1996). Local, domestic and international scientific collaboration in biomedical research. *Scientometrics, 37*(2), 279–295.

Fu, H. Z., Chuang, K. Y., Wang, M. H., & Ho, Y. S. (2011). Characteristics of research in China assessed with essential science indicators. *Scientometrics, 88*(3), 841–862.

Hu, X. J. (2009). Loads of special authorship functions: Linear growth in the percentage of "equal first authors" and corresponding authors. *Journal of the American Society for Information Science and Technology, 60*(11), 2378–2381.

Institute of scientific and technical information of China. (2012). *Statistical data of Chinese S&T papers*. Beijing: Institute of scientific and technical information of China.

Jarneving, B. (2010). Regional research and foreign collaboration. *Scientometrics, 83*(1), 295–320.

Luukkonen, T., Tijssen, R. J. W., Persson, O., & Sivertsen, G. (1993). The measurement of international scientific collaboration. *Scientometrics, 28*(1), 15–36.

Newman, M. E. J. (2001a). The structure of scientific collaboration networks. *PNAS, 98*(2), 404–409.

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

Newman, M. E. J. (2001c). Scientific collaboration networks. II. Shortest paths, weighted networks, and centrality. *Physical Review E, 64*, 016132.

Newman, M. E. J. (2004). Coauthorship networks and patterns of scientific collaboration. *PNAS, 101*(s1), 5200–5205.

Qiu, J. P., & Ma, F. (2011). *A study on the relationships between citation and collaboration* (pp. 693–699). Durban: Proceedings of the ISSI.

Wang, X., Xu, S., Wang, Z., Peng L., Wang C. (2012). International scientific collaboration of China: collaborating countries, institutions and individuals. *Scientometrics*, doi: 10.1007/s11192-012-0877-4.

Zhang, H., Patton, D., & Kenney, M. (2013). Building global-class universities: assessing the impact of the 985 Project. *Research Policy, 42*(3), 765–775.