

# *The Research on Co-authorship Network in Medical Information Education in China*

Bian Li

College of Information Engineering  
Hubei University of Chinese Medicine  
Wuhan, China

Wang Xiaoli

College of Information Engineering  
Hubei University of Chinese Medicine  
Wuhan, China

Zhao Zhen

College of Information Engineering  
Hubei University of Chinese Medicine  
Wuhan, China

**Abstract**—In this paper, CNKI data is the main sources, Chongqing VIP data and Wanfang data are the supplementary sources. It uses the method of social network analysis and visualization technology to analyze the characteristics about overall and local structure on co-authorship network in medical information education in China. The study found that the scale of integrated co-authorship network is large, the average degree of network is larger, the density of network is smaller, the average clustering coefficient is larger, the centrality of network is smaller; in most cases the co-authorship is in the internal agency cooperation in early stage, the partnership is undecided, in recent years, inter-agency co-authorship is widespread. The tightness of co-authorship network needs to be improved; the cooperation and exchange between the institutions need to be strengthened.

**Keywords**—medical information education; co-authorship network; social network analysis

## I. INTRODUCTION

Scientific cooperation has now become a major trend in promoting the development of Science, and the research on scientific cooperation is always based on the analysis of the phenomenon of co-authorship. In a certain period, the number and the cooperation of the authors of a research area, reflect the development speed and quality of scientific research cooperation and academic exchanges in this field.[1] The essence of co-authorship phenomenon study is to study the relationship between authors, and social network analysis method is specializing in relationship study. The research on the scientific co-authorship network abroad is earlier. Newman in 2001 in the “structure of scientific collaboration networks”, did a comparative study on the cooperation of scientists in the fields of biomedicine, physics and computer science for the first time[2]; Barabasi in 2002 in the “Evolution of the social network of scientific collaboration” studied the time variation of the network in the field of mathematicians[3]. In 2004, the Yoshikane. F. analyzed the 4 areas of electrical engineering, information processing, polymer and biochemical changes in the areas of cooperation mode of growth and change[4]. In

2007, Sun.F. studied the status and trend of the cooperation between industry and university in Japan from the perspective of co-authorship of academic publications.[5]. Co-authorship network research is mainly divided into the following aspects at home. First, theoretical research, studying the structural characteristics of the network, and finding a co-network model. Such as Liu Jie et al uses complex network in the network analysis of co-authorship in the field of physical chaos theory, revealing that the co-authorship of the field is of scaleless property and local small-world characteristic.[6] Xu Ling et al, studied the co-authorship of “Science Bulletin” based on the theory of complex network, and found that the network is a power law distribution.[7]second, case study, analyzing the characteristics and evolution of the network structure of a certain organization or a field of paper. Such as Liu Shengbo et al. who constructed the co-authorship network in the field of science and technology management, and revealed some regularity problems in the field of science and technology management in China.[8] Li Gang et al. studied the characteristics of the network structure, which is the research area of computer science, Wuhan University.[9] Chen Shaolong took school of management, Shanghai University as an example, and studied the co-authorship network characteristics in the college.[10]

At present, there are few papers on the subject of cross disciplinary research in China, and the research data is from important journals in the research field, or from a comprehensive database included in the relevant literature. Through literature research, we have not found the literature on the structure and characteristics of the co-authorship network structure and characteristics in the field of medical information education. In this paper, CNKI data is the main sources, Chongqing VIP data and Wanfang data are the supplementary sources, including retrieval of all the journals published in the field of medical information education in 1980~2014, using social network analysis method to study the medical information education in our country, in order to find out the pattern of the field and the evolution of the field.

---

Education Department of Hubei Province in Humanities and social sciences research project “the research on the social needs and employment status of undergraduate students in the information science of TCM”

Education Department of Hubei Province in 2013 the teaching quality project “the comprehensive reform of Chinese medicine information management specialty”

## II. RESEARCH PROCESS

### A. Data Sources and Processing

The medical information education in China began with the in-service education and training of medical library personnel in 1980s. In the mid-1980s, library and information science was set up by the former State Education Commission and the Ministry of health, the name of which was then changed to "information science (medicine, pharmacy)", "information management and information system (medical direction)" and "medical informatics", etc. Considering its connotation and extension keeps expanding, only using words like "medical", "information", "intelligence" to search the literature of the field may miss part of the relevant literature. Therefore, combining the characteristics of medical information education, the author determined a number of main search terms and limited search terms and developed following retrieval types:

Search type 1: the main search terms {(medical information) (medical information) (pharmaceutical information) (TCM) (TCM) (clinical information) (medical information) (medical information) (medical information)} and limited search terms (training) (Education) (subject) (Professional) (course)}

Search type 2: the main search terms {(biological information technology) (Biomedical Engineering) (Biomedical Engineering) and limited search terms (training) (Education) (Teaching) (subject) (Professional) (course)}

Search type 3: the main search terms {(computer) (Information Management) (Library Information)} (And) limited search terms (1) {(training) (Education) (Teaching) (subject) (subject) (course)} (And) limited search terms (2) {(Medicine)}.

Search type 4: the main search terms (Information Literacy)} and limited search terms {(Medicine)}

CNKI data is the main sources, Chongqing VIP data and Wanfang data are the supplementary sources. The paper uses the method of accurate retrieval of the subject, and the retrieval

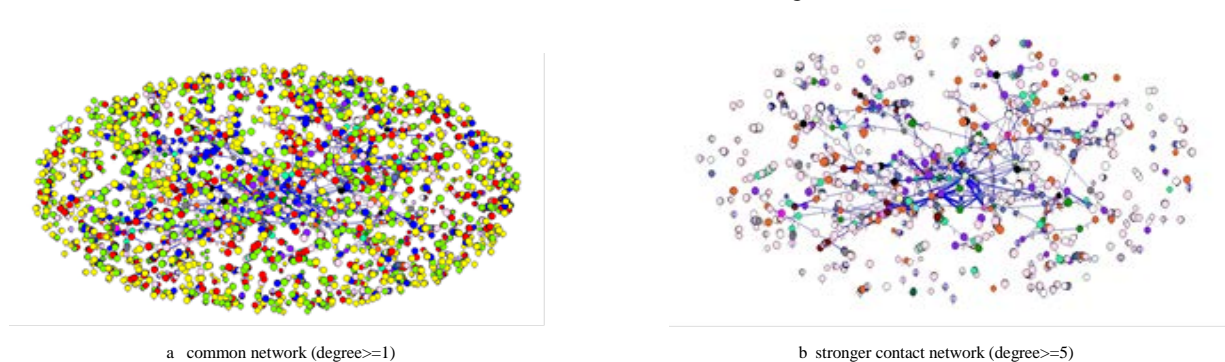
time is from 1979 to 2014. A total of 6097 papers were retrieved. The retrieved bibliographic information was imported into NoteExpress3.0 to construct the literature information management database, and then used software and manual ways to catalogue for retrieval, excluded notice, manuscripts and other non-primary papers, and finally got 5181 relevant literature. According to the mechanism of the location of the same name of the author to distinguish, the same name of the same institution has not been found. We carried out a statistical analysis on the number and author of co-authored papers, only counting the first six authors of the co-authored paper.

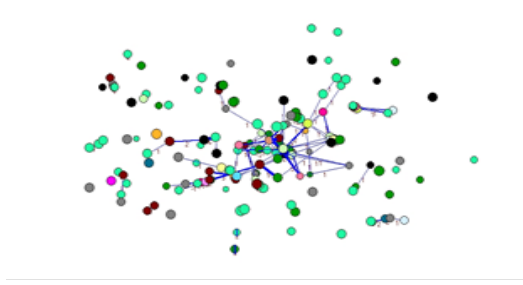
### B. Research Methods and Tools

The social network analysis method can comprehensively analyze the relationship between the individual and the whole, the local and the whole, and the relationship between the different networks[10]. Social network analysis method is the main method for the analysis of co-authorship network [13]. Take the author as the nodes, and co-authorship relationship as the edges to construct co-authorship network. The co-authorship of the article is equal to the reciprocal relationship. In this paper, the author uses the method of social network analysis to analyze the overall and local structural characteristics of the co-authorship in the field of medical information education in China. The overall network structure is analyzed from four aspects: the network degree distribution, the network density, the aggregation coefficient and the network center potential; and is analyzed from the average path length, clustering coefficient, degree centrality analysis of the three aspects of the largest connected sub graph chart structure. Use the NoteExpress3.0 managing bibliographic information, and Pajek drawing co-authorship network topology mapping, and then analyze the structure characteristics.

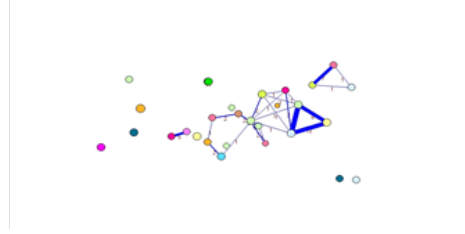
### C. Global Network Topology Change

The author of the research field of medical information education in our country is working with the network topology structure in Fig. 1.





c more stronger contact network (degree $\geq$ 10)



d most stronger contact network (degree $\geq$ 15)

Fig. 1. Cooperative network topology graph

In order to facilitate the study of the network topology graph of network topology, we call these 4 networks a, b, c and d. The structural characteristics of these four networks are analyzed from four aspects average degree, network density, average aggregation coefficient and center potential, as shown in Table 1.

TABLE I. THE BASIC CHARACTERISTICS OF THE OVERALL NETWORK DIAGRAM IN DIFFERENT SITUATIONS

|   | node<br>number | average<br>degree | network<br>density | average<br>aggregation<br>coefficient | Central<br>potential |
|---|----------------|-------------------|--------------------|---------------------------------------|----------------------|
| a | 7242           | 3.3571            | 0.0005             | 0.6317                                | 0.0082               |
| b | 1844           | 4.8416            | 0.0026             | 0.7697                                | 0.0137               |
| c | 164            | 2.2927            | 0.0141             | 0.5984                                | 0.0541               |
| d | 30             | 1.7333            | 0.0598             | 0.7119                                | 0.1207               |

Degree is a basic and important statistical feature of complex networks. The degree  $K_i$  of the  $i$  node refers to the number of other nodes connected to the node, which can reflect the importance of the nodes in a certain extent. The arithmetic average of the degree of all nodes is called the average degree of the network. Network density is the ratio of the number of lines (or edges) that are actually present in the network, or the number of lines (or edges) that are most likely to be generated in the network. The density of a network is a measure of the completeness of the network graph, and to a certain extent, the quantity and complexity of the relationship between the network and the network is characterized. The average probability of an aggregate factor represents the average of two nodes connected to the same node in the network. In co-authorship network, the average clustering coefficient reflects the close degree of co-authorship. Central potential describes the center of the whole network.

From Fig. 1 and table 1 we can see that with the increase of the degree of network nodes, the nodes in the network decreases significantly. The average degree of network increases firstly and then decreases with the increase of the degree of nodes. That means the average degree of the network increases when co-authoring times is less than 5, and it decreases when co-authoring times is more than 5. The whole network density increases with the increase of the degree of nodes, the average aggregation coefficient of the 4 networks has not changed much, the network center potential increases

with the increase of the degree of nodes, and the center potential of the whole network is relatively small.

The foregoing analysis shows that, for the co-authorship network of Chinese medical information education research fields, its overall density is not high, centrality is smaller, average degree and average clustering coefficient is larger, indicating the number of relationship and complexity of the co-authorship network of the field are not large, and there is no obvious central tendency. With the increasing of the nodes in the network, the average degree of the network decreases, the density and the center of the network are increasing, which shows that although the overall co-authorship degree is not high, there is a group of co-authorship in this field. The number of nodes and nodes in the 4 networks is large, but the average clustering coefficient is small, which shows that the "triangle" is higher in the network.

#### D. Analysis of Network Evolution

In order to have an in-depth understanding of the structural characteristics of co-authorship network in Chinese Medical Information Education, and find out authors or groups who plays an important role in the co-authorship network evolution process, we have to analyze the evolution of the co-authorship network. We take 5 years as a stage and select the most of the co-authorship of each phase of the network as the research object. The connected sub graph is a kind of special stator diagram of the reachable path between any two nodes in the sub graph. The local characteristics of the whole network can be reflected by the nature of the largest connected sub graph. Next, we analyze the average path length and the average clustering coefficient of the largest connected sub graph, and evaluate the effect of the authors' node in the network topology by 4 kinds of central index[11]. Network1~ network7 of the largest connected sub graph are shown in Fig. 2.

TABLE II. THE AVERAGE PATH LENGTH AND CLUSTERING COEFFICIENT OF THE LARGEST CONNECTED SUB GRAPH

|                                   | average path length | average clustering coefficient |
|-----------------------------------|---------------------|--------------------------------|
| the largest connected sub graph 1 | 0.500               | 1.000                          |
| the largest connected sub graph 2 | 0.607               | 0.829                          |
| the largest connected sub graph 3 | 0.700               | 0.706                          |
| the largest connected sub graph4  | 1.135               | 0.600                          |
| the largest connected sub graph5  | 0.961               | 0.583                          |
| the largest connected sub graph6  | 1.550               | 0.558                          |
| the largest connected sub graph7  | 2.399               | 0.464                          |

From table 2, with the development of medical information education in China, the average path length of the co-network is increasing, which shows that the network information transmission speed decreases with the increase of network size. The average clustering coefficient decreases, which shows that the rate of "triangle" structure in the network gradually decreases, and closeness of co-authorship decreases with the increase of network scale.

From Fig. 2, the number of nodes in the connected sub graph 1~4 is 6, 8, 10, 13, and the number of papers is 1, 2, 3 and 8. Further analysis shows that the nodes in the connected sub graph 1~4 are mostly come from the same organization and only several nodes are from other organizations. For example, all authors of connected sub Fig. 1 are from Shanghai first medical college, authors of Fig. 2 are from West China Medical University and Luzhou Medical College, authors of Fig. 3 are from the Guangdong General Hospital of Guangdong military area command, most of the authors of Fig. 4 belong to the Jilin University School of public health, only one of the authors is from the Military Medical Science Academy of the PLA. Thus we can see that in the first 4 stages, the largest connected sub graph is of small size and simple co-authorship, most of which is internal co-authorship. So we only analyze the last 3 connected sub graphs.

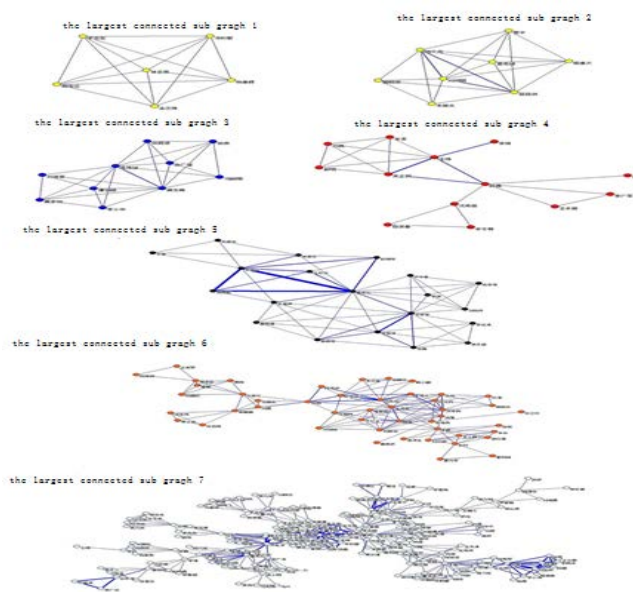


Fig. 2. Network 1~ network 7 of the largest connected sub graph

TABLE III. THE CENTER OF THE LARGEST CONNECTED SUBGRAPH NODE IN NETWORK 5

| NO. | author        | nDegree | author        | nBetweenness | author        | nCloseness | author        | nEigenvector |
|-----|---------------|---------|---------------|--------------|---------------|------------|---------------|--------------|
| 1   | Dong Xiuzhen  | 18.421  | Dong Xiuzhen  | 44.347       | Dong Xiuzhen  | 79.167     | Dong Xiuzhen  | 81.455       |
| 2   | Zou Huiling   | 14.737  | Qi Jiaxue     | 19.883       | Qi Jiaxue     | 67.857     | Zou Huiling   | 78.975       |
| 3   | Qi Jiaxue     | 9.474   | Zou Huiling   | 13.938       | Wang Jianqi   | 61.29      | Yang Guosheng | 63.557       |
| 4   | Yang Guosheng | 8.947   | Li Xueqi      | 8.187        | Li Xueqi      | 59.375     | Wang Songjun  | 33.472       |
| 5   | Li Xueqi      | 6.842   | Wang Jianqi   | 6.823        | Zou Huiling   | 57.576     | Zhao Ruigang  | 27.849       |
| 6   | Wang Songjun  | 4.737   | Lin Aihua     | 6.238        | Lu Quncang    | 55.882     | Wang Jianqi   | 18.178       |
| 7   | Lu Quncang    | 4.211   | Lu Quncang    | 1.949        | Yan Huachun   | 54.286     | Qi Jiaxue     | 14.66        |
| 8   | Zhao Ruigang  | 4.211   | Zhang Jing    | 0.39         | Dong Xiaojian | 54.286     | Lin Aihua     | 14.606       |
| 9   | Wang Jianqi   | 4.211   | Zhao Ruigang  | 0.195        | Liu Ruigang   | 54.286     | Li Xueqi      | 10.076       |
| 10  | Zhang Jing    | 3.684   | Yang Guosheng | 0.195        | Li Bing       | 54.286     | Lu Quncang    | 8.38         |

From table 3, we can see that the most important part of the network in the 2000-2004, which is the most important part of the network, has increased in the first ten years, and the central index of the top authors, Zou Huiling, Dong Xiuzhen, Qi Jiaxue, played an important role in the field. In the number of centers, the difference degree of the center is relatively large between the number of the first Dong Xiuzhen and the second Qi Jiaxue. It shows that Dong Xiuzhen is not only the network of co-authorship or co-authorship of the largest number of authors, but also plays an important role in the communication of the entire network. Further study finds out that, among the top ten authors of various centrality indicators, the vast majority of them belong to biological engineering college, Fourth Military Medical University, only one of them belongs to information science and engineering college, Central South University.

TABLE IV. THE CENTER OF THE LARGEST CONNECTED SUBGRAPH NODE IN NETWORK 6

| NO. | author       | nDegree | author       | nBetweenness | author       | nCloseness | author       | nEigenvector |
|-----|--------------|---------|--------------|--------------|--------------|------------|--------------|--------------|
| 1   | Liu Zhicheng | 9.565   | Zhang Xu     | 41.971       | Quan Haiying | 50.549     | Liu Zhicheng | 63.958       |
| 2   | Quan Haiying | 7.826   | Quan Haiying | 35.725       | Zhang Xu     | 46         | Ye Yalin     | 49.379       |
| 3   | Deng Junming | 7.391   | Li Xia       | 33.527       | Liu Zhicheng | 44.231     | Quan Haiying | 49.349       |
| 4   | Ye Yalin     | 6.522   | Niu Lihong   | 23.188       | Deng Junming | 42.593     | Li Haiyun    | 42.987       |
| 5   | Tao Qiang    | 6.087   | Liu Zhicheng | 13.259       | Ye Yalin     | 42.202     | Deng Junming | 42.089       |
| 6   | Zhang Xu     | 5.652   | Shan Nana    | 12.464       | Ren Zhaohui  | 42.202     | Duan Yuanmin | 34.479       |
| 7   | Duan Yuanmin | 5.652   | Zhang Haiyan | 8.599        | Li Xia       | 41.818     | Zhang Xu     | 34.446       |
| 8   | Li Xia       | 5.217   | Zhou Maoxin  | 8.502        | Tao Qiang    | 41.441     | Tao Qiang    | 28.659       |
| 9   | Li Haiyun    | 5.217   | Zhao Ling    | 8.502        | Liu Jinghua  | 40         | Tian Xin     | 26.923       |
| 10  | Liu Jinghua  | 4.348   | Deng Junming | 8.007        | Duan Yuanmin | 39.316     | Ren Zhaohui  | 23.857       |

From table 4 we can see that, in 2005-2009 the scale of the largest connected sub graph in the network is larger than the previous one, The centrality of the center of the degree of the center of the top ten is not obvious, and the maximum value is only 9.565, indicating that the network is distributed evenly, and the aggregation of nodes is not obvious. Liu Zhicheng's degree centrality and feature vector centrality are the largest, but betweenness centrality and closeness centrality slightly decrease, indicating that he may be the node who has published

the most papers or has the most co-authorship papers in the network.. Quan Haiying and Zhang Xu's proximity to the center of the degree and the number of the center are higher, indicating that these two authors are not only close to the other authors in the network, but also the network of the bridge node. In the further study, the network in the vast majority of nodes belongs to the Capital Medical University biomedical engineering college; the mechanism of the internal co-authorship is more common.

TABLE V. THE CENTER OF THE LARGEST CONNECTED SUBGRAPH NODE IN NETWORK 7

| NO. | author         | nDegree | author        | nBetweenness | author         | nCloseness | author         | nEigenvector |
|-----|----------------|---------|---------------|--------------|----------------|------------|----------------|--------------|
| 1   | Wang Wei       | 4.949   | Wang Wei      | 68.221       | Wang Wei       | 36.617     | Wang Wei       | 72.114       |
| 2   | Liu Yan        | 2.284   | Lu Hongbing   | 33.901       | Cui Lei        | 32.242     | Cao Jindan     | 41.137       |
| 3   | Lu Hongbing    | 2.157   | Cui Lei       | 28.204       | Xu Peiyang     | 32.137     | Lu Hongbing    | 37.802       |
| 4   | Liu Yamin      | 1.904   | Xu Peiyang    | 26.101       | hang Shijing   | 31.723     | Jiao Chun      | 35.721       |
| 5   | Jiao Chun      | 1.586   | Zhang Shijing | 20.828       | Lu Hongbing    | 31.369     | Zhang Guopeng  | 34.148       |
| 6   | Wang Liwei     | 1.523   | Wang Min      | 20.693       | Cao Jindan     | 30.926     | Chang Xiaohong | 34.148       |
| 7   | Dong Xiuzhen   | 1.523   | Ji Zhenyu     | 15.362       | Li Houqing     | 30.733     | Mou Dongmei    | 30.809       |
| 8   | Zhang Guopeng  | 1.459   | Zhang Yanwu   | 11.447       | Jiao Chun      | 29.58      | Jian Weiping   | 23.812       |
| 9   | Chang Xiaohong | 1.459   | Zhang Bin     | 11.447       | Zhang Guopeng  | 29.535     | Duan Lei       | 22.309       |
| 10  | Cao Jindan     | 1.459   | Zhang Hao     | 10.141       | Chang Xiaohong | 29.535     | Wu Xiaoling    | 22.309       |

From the above Fig. 2, the scale and complexity of the network in 2010-2014 have changed a lot compared with the previous stage. From the table 5, the centrality of the degree of Wang Wei, Liu Yan, Lu Hongbing is the largest, and the cumulative number of times is as high as 78, 36, 34 times, it can be inferred that they are the high yield authors in the network. The betweenness centrality of Wang Wei, Lu Hongbing and Cui Lei is the largest, indicating that they play a role in communicating and connecting other authors of the bridge junction in the network; top 10 authors of closeness centrality have similar closeness centrality, indicating that there are many nodes of high closeness centrality, which are of high continuous degree with other nodes. Wang Wei, Cao Jindan, Lu Hongbing, their central character of the feature vector is the highest. Wang Wei's feature vector center is much higher than other authors and the degree of the center is the largest, not only he has published a large number of papers, and the co-author also have published a number of papers. From Fig. 2 we can clearly see that Wang Wei is currently the core of the research field of medical education in China, and he is a bridge to communicate with other authors in the field.

Further analysis revealed that the largest connected sub graph of the co-authorship network contains a number of agencies, such as Jilin University School of public health represented by Wang Wei, Fourth Military Medical University School of Biomedical Engineering represented by Xiu Zhen Dong, Shandong Medicine & Health Science and Technology Information Institute, Medical Information Research Center represented by Liu Yamin, to China Medical University Medical Information Department represented by Cui Lei, Central South University Xiangya School of medicine represented by Houqing Li, Huazhong University of Science and Technology, Tongji Medical College, Medicine Health

Management College represented by Zhang Huazhong, Nanjing Medical University School of Basic Medical Sciences, Department of Biomedical Engineering represented by Duan Lei, and so on. It indicates that at this stage, the co-authorship has been extended from within organization to co-authorship of many organizations and agencies.

### III. CONCLUSION

This paper is based on the previous research results; we did a research of co-authorship network in medical information education in China. Through the analysis of the overall structure and local structure characteristics of the co-authorship, we find that the cooperation in the field of medical information education in China has the following characteristics:

a) The overall network size is larger (including 7242 nodes, 12157 edges), the average degree of the network is large, but the network density is small, the average clustering coefficient is relatively large, but the overall network centrality is small. Comparison of two groups of seemingly contradictory data shows that China's medical information education and research field of the overall cooperation is of large scale, the co-authorship network contains a lot of small co-authors, their internal cooperation is more stable and tends to belong to the same institution, and there is no significant difference in the production of the article.

b) The co-authors are gradually expanded to the outside of the organization, and the scale of the co-authors is growing rapidly, and in recent years, the phenomenon of cross agency is common. From the analysis of the paper, we can know that, during the period of 1980~2009, the nodes in each stage of the largest connected sub graph are part of the same organization. The author of the largest connected sub graph of 2010~2014 is from a dozen of different organizations. This shows that in recent years, in the field of scientific research, enhancing cooperation, expanding cooperation scope, exchange and cooperation between agencies are increasing.

c) Wang Wei as the representative of the Jilin University School of public health (formerly Bethune Medical University), Dong Xiuzhen, as the representative of the Biomedical Engineering College of The Fourth Military Medical University, Liu Yamin as the representative of the Shandong provincial medical and Scientific Information Research Institute of Medical Information Research Center, Cui Lei as the representative of the Department of medical informatics, Li Houqing as the representative of Central South University Xiangya Medical School, Zhang Shijing as the representative of the Tongji Medical College, Huazhong University of Science and Technology medical and Health Management School. The researchers in core research institutes are the important nodes of the whole network. Some of them are the academic leaders, some of them have high yield in the field, some of them are of both two aspects. They not only attracted a large number of co-authors, but also played the role of "connecting people" in the network.

## REFERENCES

- [1] Lin Li. Research on the cooperation network organization and cooperation of scientific research papers [D] .Changchun: Jilin University,2010.
- [2] Newman M.E.J.The Structure of Scientific Collaboration Networks[J].Proceedings of the National Academy of sciences,2001,98(02):404-409.
- [3] A.L.Barabasi,H.Jeong,et.Evolution of the social network of scientific collaborations[J].Physica A 311,2002:590-614.
- [4] Yoshikane,F. And K.Kageura.Comparative Analysis Of Coauthorship Networks of Different Domains:The Growth and Change of Networks[J].Scientometrics,2004,60(3):433-444.
- [5] Sun,Y.,M.Negishi,et al.Coauthorship linkages between universities and industry in Japan[J].Research Evaluation,2007,16(4):299-309.
- [6] Liu Jie,Lu Junan. Complex network of small scientific research cooperation and its analysis[J]. Complex systems and complexity science,2004,1(03):56-61.
- [7] Xu Ling,Hu Haibo,Wang Xiaofan. Empirical analysis of a Chinese scientist working on the network [J]. Complex systems and complexity science,2009,6(01):20-28.
- [8] Liu Shengbo,Yang Ying et. Research on the co authors of the paper in the field of science and technology management in China [J]. Science and technology management, 2010,(03):228-230.
- [9] Li Gang,Li Ang. The study of academic groups based on social network analysis--Taking the Computer College of Wuhan University as an example [J]. Journal of information resources management, 2011,(03):43-47.
- [10] Chen Shaolong. The application of the social network analysis method to the Management College of Shanghai University[J]. Modern information, 2012,32(4): 165-168, 177.
- [11] Huang Kaimu,Fan Zhenjia,et. Research on the co authors of journal papers in the field of competitive intelligence in China[J]. Information journal, 2015,34(2):142-147.