ISSN 2616-5783 Vol.4, Issue 3: 35-41, DOI: 10.25236/AJHSS.2021.040308

Research on Co-authorship Network of Sports Culture in China Based on Social Network Analysis (SNA)

Xin Lu[1], Wentao Zhang[2*]

[1]School of Mathematics and Information Science, Nanchang Normal University, Nanchang 330032, China
[2]Physical Education College of Jiangxi Normal University, Nanchang 330022, China
[*]E-mail: wentaozhang1978@163.com
[*]Corresponding Author

Abstract: *The co-authorship network of Chinese sports culture field is constructed while the co-authored papers published in CNKI database from 2000 to 2019. SNA methods and complex network theory are employed to analyze the structural characteristics, groups of core authors, and distribution of the network resources based on overall density, average distance, clustering coefficient, and centrality of co-authorship network. The result shows that the information liquidity is poorer, the connections are relatively loose, and there are some groups that authors are familiar with each other in the co-authorship network of sports culture. The scientists' productivity and co-author groups' size follow the power law distributions. The field has formed a large scale and influential groups and core authors, and mainly in small-scale cooperation. Moreover, most of the authors who published amounts of papers have limited co-authored papers, and a few core authors have high resource control. What's more, prolific authors should expand the scope of cooperation to improve the connectivity of network. And the same time, more attention should be pay to these core authors and hub authors.*

Keywords: *sports culture, co-authorship network, social network analysis, complex network*

1. Introduction

The development of information technology has made it possible for scholars from different regions and fields to share and exchange knowledge and experience. At the same time, the comprehensive, in-depth and intersecting research of various disciplines require the cooperation of personnel from different disciplines and directions to solve the bottleneck of scientific research. The phenomenon of scientific research cooperation is becoming more and more common and normal, and plays a role that cannot be ignored. A detailed discussion of the network structure of the cooperative relationships is of great significance to the study of resource sharing, experience exchange and information dissemination, etc. in the research network and has become a hot topic at present.

In the co-authored network, the researchers are treated as nodes, and the partnership as an edge, and the sheer number of papers and authors determines the complex network characteristics of the co-authored network. Under the requirements of the times of building "sports power" and "cultural power", the research of sports culture in China has become an important subject of great concern and research in the current cultural circles. Based on the research of previous people, this paper comprehensively uses social network analysis method and complex network knowledge to carry out a systematic analysis of the author's co-author network in the field of domestic sports culture, with a view to discovering the characteristics of the co-author network in the field of sports culture in China.

2. Materials and methods

2.1 Data sources and preprocessing

This study was conducted with the CNKI Journal Full Text Database as the data source, the name "Sports Culture", the subject category "Sports", the time range 2000 to 2019, and the data source category selected from all Journals, SCI Source Journals, EI source journals, core journals, and CSSCI

ISSN 2616-5783 Vol.4, Issue 3: 35-41, DOI: 10.25236/AJHSS.2021.040308

source journals, and a total of 5986 literature data were obtained. Excluding invalid data such as no authors, editorial writings, etc. A total of 5,937 valid literature data were finally obtained, including 6131 authors. Co-authored 2485 papers, accounting for 41.86% of the valid data, and co-authored papers involving 4382 authors. Therefore, the cooperation rate in the field of domestic sports culture research in China (the ratio of the total number of co-authored papers to the total number of papers) is 41.86 percent, the degree of cooperation (the ratio of the total number of authors to the total number of papers) is 1.03, indicating the overall cooperation situation is relatively low and needs to be strengthened.

2.2 Research Methods

Social Network Analysis (SNA) is a research method to quantitatively analyze the relationship structure and attributes of social networks [1]. By analyzing the various attributes of the network, such as network density, average path length, clustering coefficient and centrality, core/edge structure, etc., SNA explains the overall structure of the network and the interrelations between nodes in many ways, which can be used not only to analyze individuals, but also to Institutional partnerships can also visualize the evolution paths and trends of an area through keyword co-presentation.

3. Analysis of Authors' Co-authorship Network in Domestic Sports Culture

3.1 Co-authorship network construction

The construction of a common matrix that reflects the author's cooperative relationship is the basis for analyzing the co-authored network in the field of domestic sports culture. On the basis of full consideration of sample data characteristics and data processing capacity, this paper assumes that : (1) the authors in the co-author's paper are equal, and that the order of signatures does not affect their cooperative relationship; (2) There are only two states of co-authorship: "with (1)" and "without (1)".

Due to the large number of authors involved in the collaboration, the overall co-authoring network will be extremely complex and will not be able to conduct targeted analysis. Therefore, 989 authors whose publishing frequency ≥2 were selected as the research objects, accounting for 22.57 percent of the total. Subsequently, the built author collaboration in the field of sports culture, the Inverse Network Matrix, was imported into Gephi and a reasonable threshold was set to map the 865-bit author collaboration (124 orphaned nodes have been removed), as shown in Figure 1. Discover that the co-author of this build is a highly unconnected network with a total of 865 nodes and 2038 connections. Each node represents an author, and its size represents the number of papers co-authored; The line between nodes represents the existence of cooperation, and its thickness indicates the number of cooperation times. According to the analysis, Bai Jinxiang of Jisukong University participated in the largest number of co-authored papers, with 20 papers, and Zhong Haiping of Jisukong University collaborated the most times, with 8 papers.

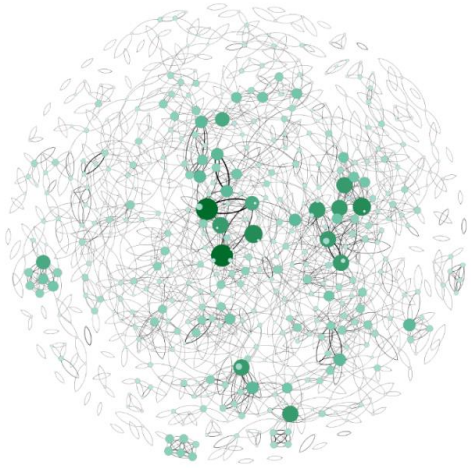

Figure. 1 Map of Co-authorship Network

ISSN 2616-5783 Vol.4, Issue 3: 35-41, DOI: 10.25236/AJHSS.2021.040308

3.2 Network characteristic parameter analysis

In the analysis of social network, the structural characteristics of a network can be described by means of average distance, density overall and clustering coefficient. The authors of this study in the field of domestic sports culture co-authored the network characteristic parameters see Table 1. In order to compare the structural differences of the co-authors' network in different fields in detail, the author also selects the social networks in the literature[2-4].

The average path length is the average of the number of edges between any two nodes and is an important measure of the smooth flow of information across the network, which can be expressed as:

$$L = \frac{1}{\frac{1}{2}N(N+1)} \sum_{i \geq j} d_{ij} \tag{1}$$

The smaller the L value, the faster the information in the network propagates, and the closer the nodes are connected[5]. According to formula (1), the average network distance of the co-authored network in the field of domestic sports culture is calculated to be 11.72. At least two scholars from 11-12 people are required to establish a cooperative relationship. Compared to other co-authored networks, this path is very long and the dissemination of information is less efficient. As can be seen, the co-author network in the field of sports culture does not have the characteristics of the small world and is a co-author network with poor information flow and difficult communication between researchers. In this network, information dissemination efficiency is low, and the author needs a longer time to respond to the emergence of new information.

Network density is the ratio of the number of existing cooperative links of all nodes in the network to the maximum number of connections that could theoretically exist, i.e. $d=2L/k(k-1)(0\leq d\leq 1)$, L for the actual number of cooperative links, k is the number of nodes[6] Network density is evaluated for the degree of looseness of the relationship between authors. The higher the density is, the closer the network node is, the smoother the information propagation, and the more productive it is. Compared with co-authored networks in other fields, the network density of co-authored networks in the field of sports culture is 0.0027 and the density is smaller. Because of the large scale of the network co-authored in this paper, the connection between its nodes is looser. Although some nodes are very closely connected, the overall network density is low due to sparse connections between groups. Another angle shows that there is a lot of room for improvement in the cooperation of authors in the field of sports culture in China, and it is necessary to continuously strengthen the cooperation between authors in the field of sports culture, and to enhance the connectivity of the overall network, so as to promote the exchange of knowledge and the sharing of experience in the field of sports culture.

Clustering coefficient refers to the possibility that the neighboring nodes of a node are still neighbors. The clustering coefficient of the whole network is equal to the average value of the clustering coefficient of all nodes, and its size reflects the universality of small and medium-sized groups in the network [7]. By running Gephi software to get a clustering factor in the field of domestic sports culture is 0.362, the network has a clear aggregation effect, indicating that the field of sports culture co-authors in the network there are some internal familiar with each other, it is easy to form cooperative relations of the group.

Table 1 Comparison of Different Co-authorship Network Features

Co-authorship network type	Node	Average path length	Network density	Clustering coefficient
Author cooperation in the field of sports culture	865 authors	11.719	0.0027	0.362
Cooperative networks in the field of competitive intelligence	1079 authors	2.999	0.006	0.804
Sports Research Co-authorship Network	935 authors	——	0.2059	——
Physical Education Curriculum Standards Authors Network	200 authors	1.002	0.0024	0.606

3.3 Degree distribution

The number of edges directly connected to a node is the degree of that node, and the degree of the author's co-author network represents the number of times of cooperation with that author. Statistical

ISSN 2616-5783 Vol.4, Issue 3: 35-41, DOI: 10.25236/AJHSS.2021.040308

results can be found in Table 2, due to the frequency of ≥2, resulting in the author co-author network there are 124 isolated nodes. The number of authors with node degrees of 1, 2, and 3 was the largest, accounting for 82.54 percent of the total, while the number of authors with larger degrees was very small, with only 12 nodes greater than 8, accounting for 1.39 percent of the total. The scope of cooperation among authors in the field of sports culture in China is small, and most authors send essays and cooperate less frequently.

Table 2 Degree Distribution of Co-authorship Network

Node Degree	Node number	Proportion	Node Degree	Node number	Proportion	Node Degree	Node number	Proportion
1	314	36.30%	6	14	1.62%	12	2	0.23%
2	249	28.79%	7	8	0.92%			
3	151	17.46%	8	3	0.35%			
4	77	8.90%	9	8	0.92%			
5	37	4.28%	10	2	0.23%			

The complex network degree distribution that obeys the power law distribution is also known as a scaleless network, which refers to a network that has a large number of nodes with a small number of degrees and a few nodes with a large number of degrees, and a large number of nodes called hubs. Degree distribution function:

$$P(k) = c \bullet k^{-\gamma} \tag{2}$$

It is the percentage of *nodes that measure the* number of nodes with k in the whole network (k is the number of nodes, *c is* the scale factor, *γ is* the scaleless network coefficient, *c*>0 and *γ*>0). Taking the logarithm of P(k) function, we can get ln(P(k))=ln(c) -γ·ln(k) , that is, under the condition of double logarithm, the power law distribution is a straight line with negative slope, so as to judge whether the co-authorship network satisfies the power law. The linear fit of the coefficient is calculated γ about 2.419. Thus, it can be determined that the author co-author network in the field of domestic sports culture is a scale-free network with a scale coefficient of 2.419.

3.4 Scale analysis of connected subgraphs

A connected subgraph is a subgraph that has a reach path between any node, but does not contain loops. The scale of the connectivity submap refers to the number of nodes contained in the connectivity submap, and its scale distribution reflects the number of connected submaps of different sizes in a complex network. The author co-author of the network, the number of authors of the collaboration group is the size of the connected submap.

Table 3 Connected Subgraph of Co-authorship Network

authors	connectivity sub-maps	proportion (%)	authors	connectivity sub-maps	proportion (%)	authors	connectivity sub-maps	proportion (%)
2	51	10.31	7	2	1.42	20	1	2.02
3	18	5.46	9	1	0.91	543	1	54.90
4	9	3.64	13	1	1.31			
5	6	3.03	14	1	1.42			
6	2	1.21	18	1	1.82			

After statistics, the domestic sports culture field authors co-authored a network of 94 connectivity sub-maps (removal of isolated nodes), detailed statistics can be found in Table 3. Statistics The scale distribution of the connectivity subpic in the field of discovery is still subject to the power law distribution, and the fit γ is about 0.623. The largest co-authoring network in the field of sports culture in China consists of 543 nodes, accounting for 54.90% of the total; the number of authors with two or three co-authors accounted for 15.77%. This is can be seen that the field of domestic sports culture has formed a large-scale cooperative groups, with small-scale cooperation as the main.

3.5 Core/edge structure analysis

A core/edge structure is a node that is in a relative core and edge position that is screened based on the closeness of the connections between nodes in the network. Using a continuity model, this article calculates the core values(coreness)of 865 nodes through multiple iterations and determines that 58 authors, including Bai Jinxiang, Tian Zugu, and Zhang Dongxiu, are the core authors of the domestic sports culture co-author network, accounting for all of them 6.71% of the authors.

ISSN 2616-5783 Vol.4, Issue 3: 35-41, DOI: 10.25236/AJHSS.2021.040308

It should be noted that the core author determined according to the location of the node in the network is fundamentally different from the traditional core author based on the frequency of the text. To further compare the results, the author based on traditional measurement methods, according to Price's law and the highest frequency of 20, 151 core authors with the original number of articles higher than 3.35 were calculated.

In terms of quantity , there is a gap of nearly 100 places between the two , which is more obvious. There are also significant differences in author composition. Only 23 of the core authors based on SNA sent more than 3.35 articles and 60.34 percent sent fewer than 3.35. By comparison, it is found that most high-yielding people in the field of sports culture cooperate with others less often, partners are more fixed, and the scope of cooperation is narrower.

3.6 Centrality analysis

3.6.1 Degree center degree analysis

Degree center degree is closely related to node degree, Bai Jinxiang, Tian Zuguo participated in 20 and 17 paper cooperation, degree center degree ranked first. In the network of 865 authors of domestic sports culture, each author has a cooperative relationship with at least one author, so the minimum degree of center is 1. The statistics found that 314 authors had a degree center of 1, accounting for 36.30 percent of the 865 authors. At the same time, 74 authors (8.55 percent) had degree center degree ≥5. There were only 4 authors ≥10, accounting for only 0.46 percent. In other words, more than 74 percent of authors have worked with at least two authors, only slightly more than 8 percent have co-authored articles with more authors, and less than 1 percent have worked with a wide range of authors. Combined with the data, it can be learned that the domestic sports culture field authors cooperation is more extensive, most of the authors have cooperative relations, but limited to the cooperation of small groups, follow-up research should strengthen the exchange and sharing of knowledge and experience between different groups, enhance network connectivity.

3.6.2 Mediation centrality analysis

The degree of the node, the degree of other nodes connected to the node, and the location of the network in which it is located are closely related to the degree of mediation center. Liu Yong, a scholar from Hengyang Normal College, has nine documents on his cooperation, and Zhou Jiansheng and Chen Yonghui, who have a direct relationship with him , also have 6 and 4 co-authored papers respectively , and the shortest path of the other nodes is also more likely to pass through Zhou Jiansheng and Chen Yonghui. At the same time, Liu Yong is still in the "Bai Jinxiang as the core" and Zhang Ping as the core between the two groups , the shortest path between the two sub-groups through his probability is also higher , the joint role led to Liu Yong's highest degree of intermediary center .

Further statistics show that in a network of 865 authors, the node has a minimum centering value of 0, i.e. the node is not on the shortest path between any two authors in the network. The statistics found that 444 authors had a median center of 0, accounting for 51.33 percent of the total, and 122 authors had a center greater than 1. In addition, Liu Yong and three other authors have a center greater than 9, and they have very high control over the information resources in the network. It can be found that in the co-author network in the field of sports culture in China, the control of information resources is unbalanced, fewer authors have a large amount of information resources, and most authors lack effective resource control.

3.6.3 Analysis of proximity centrality and eigenvector centrality

Because proximity to the center, feature vector centrality is not often used, this article analyzes the two together. Proximity to the center is most relevant to the network location of the node. Zhou Jiansheng and Zhang Ping co-authored fewer papers (6 and 5, respectively), but they are located between the three sub-groups with Bai Jinxiang, Zhang Liangxiang and Zheng Daming as the core respectively, and are the only bridge between them, so their average distance to the rest of the network is shorter; The three sub-groups co-authored papers with a higher frequency, resulting in a higher proximity to the center of the two. In addition, Bai Jinxiang and Zhong Haiping's subnet contains the most author nodes, which also leads to the proximity of the two people side by side in the first place. Numerically, the difference between the two people's proximity to the center and the degree of center is small, indicating that the subnet does not have an absolute leadership core.

The degree of the node and the degree to which it is connected affects the size of the feature vector center. Because Bai Jinxiang and Tian Zuguo participated in the paper more times, so it obtained a

ISSN 2616-5783 Vol.4, Issue 3: 35-41, DOI: 10.25236/AJHSS.2021.040308

higher characteristic vector center, and at the same time with the two people more times of cooperation Zhong Haiping, Long Peilin, Wang Yaqiong, Zhang Dongxiu and other people also obtained a higher value.

3.6.4 Correlation analysis

The correlation analysis of the degree center, intermediary center, proximity center, characteristic vector center and number of documents issued by the 865 authors of the co-author network in the field of sports culture is analyzed, and the results are in Table 4. It can be seen that the number of authors and intermediary center degree of correlation between the highest, indicating that high-yielding authors in the control of resources is more obvious advantages, but they are limited to a small range of cooperation and exchanges, the lack of a broader co-author. The correlation between degree center degree and intermediary center degree is higher than its correlation with near center degree and characteristic vector center degree, which shows that some authors in the network, while co-authoring with a large number of authors, also provide a bridge for cooperation and communication among other authors, and promote the connection between different authors and groups. Proximity to the center is most relevant to the characteristic vector center, indicating that the more frequently the author cooperates, the higher the author's position in the sub-group, which also means that cooperation is an important way to increase the author's influence and solve complex problems.

4. Conclusion

(1) The average path length of the co-authored network is larger, it does not have the characteristics of the small world, the information flow in the network is poor, and the author communication is difficult, the network density is low, the overall connection of the co-author network is low, the internal connection is looser, and the clustering coefficient is 0.362, the network has a distinct aggregation effect, and the domain has some familiar groups.

(2) Network Nodes have the largest number of authors with low degrees, very few authors with high degrees, and the scope of cooperation among authors in the field of domestic sports culture is small. By statistical analysis and linear fitting of the node degrees of the co-authorship network in this field, it is found that the network is a scale-free network with a scaling coefficient of approximately 2.419.

(3) According to the analysis of the authors' co-author-network connectivity sub-chart statistics, the largest co-author network in the field of sports culture in China consists of 543 nodes, and the number of authors of the 2-3 co-author-network accounts for 15.77%, indicating that a large-scale cooperative group has been formed in this field, with small-scale cooperation as the main. By analyzing the subgraph size statistics and curve fitting of the coauthor network, it is found that the coauthor network follows a power law distribution with a fitting coefficient of 0.623 and has the characteristics of scale-free network.

(4) Core/Edge Structural Analysis shows that the core authors group has only 58 authors, most of the high-yielding people in the field of domestic sports culture have fewer cooperations with others, the node core is lower, the partners are fixed, and the scope of cooperation is narrow. High-yielding authors should consciously work with more people to expand their circle of collaboration to improve the overall connectivity and density of the network.

Acknowledgement

Funding: This work was supported by Science and Technology Project Founded by the Educational Department of Jiangxi Province(GJJ181076).

This work was supported by the College Social Sciences Foundation of Jiangxi Province(TY162002).

References

[1] HUANG Kaimu, FAN Zhenjia, LU Shengjun, et al. Research on the co authorship network of journal papers in the field of competitive intelligence in China [J]. Journal of Information Science and Technology, 2015, 34(2): 142-147.

[2] Zhang Huan, Dong Baolin, Cai Yujun. Current situation of sports scientific research cooperation

ISSN 2616-5783 Vol.4, Issue 3: 35-41, DOI: 10.25236/AJHSS.2021.040308

in China[J]. Journal of Wuhan Institute of Physical Education,2014, 48(3): 16-22.
[3] Zhang Yang, Liu Jinyuan. Research on the co-authorship network of papers in the field of competitive intelligence based on SNA [J]. Library and Information Knowledge, 2012, 2: 87-94.
[4] Lu Dongdong. Research on the co-authorship network of papers in the field of Physical Education and Health Curriculum Standards [J]. Physical Education Research and Education, 2014, 29(5): 79-82.
[5] Chen C, Ibekwe-SanJuan F, Hou J. The structure and dynamics of co-citation clusters: A multiple-perspective co-citation analysis[J]. Journal of the American Society for Information Science and Technology, 2010, 61(7): 1386-1409.
[6] Perianes-Rodríguez A, Olmeda-Gómez C, Moya-Anegón F. Detecting, identifying and visualizing research groups in co-authorship networks[J]. Scientometrics, 2010, 82(2): 307-319.
[7] Zhü K. Fuzzy analytic hierarchy process: Fallacy of the popular methods[J]. European Journal of Operational Research, 2014, 236(1): 209-217.