International Journal of Machine Learning and Cybernetics (2020) 11:973–987
https://doi.org/10.1007/s13042-020-01082-x

ORIGINAL ARTICLE

Complex network analysis of three-way decision researches

Bo Yang[1,2] · Jinhai Li[1,2]

Received: 12 January 2020 / Accepted: 4 February 2020 / Published online: 15 February 2020

Abstract

In this paper, complex networks are used to analyze the dataset of three-way decision articles published before December 18, 2019 and downloaded from ISI Web of Science. The scientific collaboration network, university collaboration network, networks of scientific papers (i.e., citation network, bibliographic coupling network, co-citation network) and keywords network are constructed to reveal the relationships between authors, affiliations, papers and keywords, respectively. Some interesting results are obtained and used to answer the following questions: (1) which authors play a key role in developing three-way decision; (2) which affiliations actively promote the development of three-way decision; (3) which papers are important or influential in the field of three-way decision; (4) what are the closely related research issues around three-way decision.

Keywords Three-way decision · Granular computing · Rough set · Complex networks · Knowledge discovery

1 Introduction

Three-way decision was proposed by Professor Yao Yiyu [81, 82], and it is an effective mathematical tool to make decisions based on trisection idea [15, 18, 43, 45, 54, 92]. In recent years, the theory of three-way decision has been developed by incorporating granular computing, cognitive computing, rough set, formal concept analysis, fuzzy set, multi-attribute decision making, and so on. In the development of three-way decision, there have been appearing many hot and promising research issues such as decision-theoretic rough set [23, 38–40, 46, 75, 113], probabilistic rough set [48, 72, 83, 109], three-way granular computing [44, 89, 91], attribute reduction based on three-way decision [56, 105–107], fuzzy set oriented three-way decision [12, 22, 24, 25, 104], multi-granulation three-way decision [6, 36, 57, 59], cost-sensitive three-way decision [47, 102, 112], sequential three-way decision [19, 31, 83, 100, 103], three-way concept analysis [33, 35, 55, 60, 61, 63, 64, 88, 96, 110, 111], three-way concept learning [13, 34], three-way conflict analysis [21, 67, 90], clustering with three-way decision [1, 70, 93–95, 108], orthopairs [4] and shadowed sets [53]. In the meanwhile, a large number of useful and interesting results have been obtained. So it is natural and important to analyze the researches in three-way decision for the purpose of providing a reference for interested readers to read, study, understand and develop the theory of three-way decision.

The dataset of three-way decision articles we collected for complex network analysis in this paper were searched from ISI Web of Science by using the keywords: "three-way decision", "three-way decisions", "decision-theoretic rough sets" or "probabilistic rough sets" for those published before December 18, 2019. The dataset covers 549 papers and each of them was saved as a CIW (customer information warehouse) file for easy access. Number of the collected papers as a function of dates is shown in Fig. 1. It should be pointed out that the number of those papers written by only one author was 43 which accounts for 7.8% of total papers. We also found that the dataset of three-way decision articles includes 11 ESI highly cited papers [14, 29, 31, 34, 41, 59, 66, 82, 84, 94, 98].

The rest of this paper is organized as follows. In Sect. 2, we introduce the measurement indices and algorithms used in this paper to make data analysis based on complex networks. In Sect. 3, scientific collaboration network is

✉ Jinhai Li
jhlixjtu@163.com

Bo Yang
yangbo@kust.edu.cn

[1] Data Science Research Center, Kunming University of Science and Technology, Kunming 650500, Yunnan, People's Republic of China

[2] Faculty of Science, Kunming University of Science and Technology, Kunming 650500, Yunnan, People's Republic of China

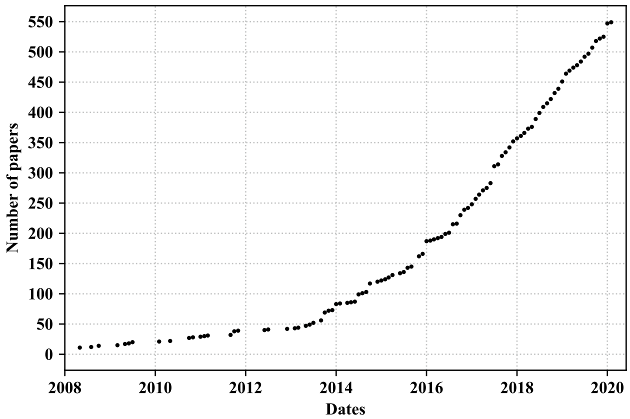

Fig. 1 Number of the collected papers as a function of dates

constructed to reveal the relationship between authors. In Sect. 4, university collaboration network is built to show the relationship between affiliations. In Sect. 5, citation network, bibliographic coupling network and co-citation network are investigated to obtain the relationship between three-way decision articles. In Sect. 6, keywords network is discussed to find the hottest research issues closely related to three-way decision. Finally, some useful conclusions are given in Sect. 7. The main content of our work can be shown in Fig. 2.

2 Models and algorithms

A network G can be presented by a set V of nodes and a set E of edges [50, 71], i.e., $G = (V, E)$. Generally speaking, networks can be divided into directed networks and undirected networks based on whether the edges have directions or not. Like undirected networks, directed networks were also encountered frequently in the real world. For instance, in citation network of papers, the nodes are papers and there is a directed edge from paper M to paper N if M cites N in its bibliography.

A network G can simply be represented by an adjacency matrix A no matter whether it is directed or undirected. For example, let G be a network with n nodes, and A be the adjacency matrix of G. In [50], if G is an undirected network, its adjacency matrix A can be defined as

$$A_{ij} = \begin{cases} 1, & \text{if there is an edge between the nodes } i \text{ and } j, \\ 0, & \text{otherwise;} \end{cases} \tag{1}$$

if G is a directed netwrok, the elements of the adjacency matrix A are given as follows:

$$A_{ij} = \begin{cases} 1, & \text{if there is an edge from the node } j \text{ to the node } i, \\ 0, & \text{otherwise.} \end{cases} \tag{2}$$

It should be pointed out that sometimes the importance of the connections between nodes of a network may be different from each other. If it happens, then we call such a network as a weighted network whose elements A_{ij} are viewed as the weights w_{ij} of the corresponding connections and they can be computed in a certain way in real instances. For example, in scientific collaboration network, weights often present frequency of cooperation between authors.

The degree of a node is an important notion in a network. Let G be a network with n nodes, and A be the adjacency matrix of G. In [50], if G is an undirected network, then the degree k_i of the node i is defined as the number of the edges directly connecting it to the other neighbors, i.e.,

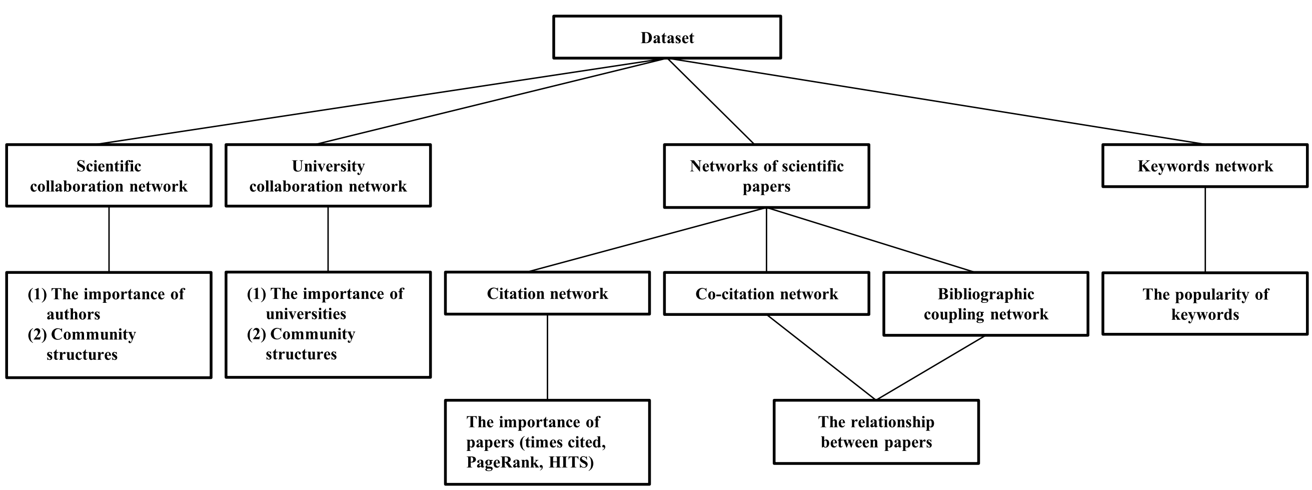

Fig. 2 The main content of our work

$$k_i = \sum_{j=1}^{n} A_{ij}. \tag{3}$$

If G is a directed network, the case may be more complicated and it needs an additional distinction between in-degree and out-degree [50]. The former means the number of incoming edges, while the latter means the number of outgoing edges. In this paper, we denote the in-degree and out-degree of the node i by k_i^{in} and k_i^{out}, respectively. Then we have

$$k_i^{\text{in}} = \sum_{j=1}^{n} A_{ij}, \quad k_j^{\text{out}} = \sum_{i=1}^{n} A_{ij}. \tag{4}$$

Another important notion in a network is network centrality. This index was used to find the most important or central nodes in a network. Different people may have different ideas about this issue. To the best of our knowledge, there have been many network centrality measures. In this paper, degree centrality, eigenvector centrality and PageRank value will be used to search for important authors and influential papers from the dataset of three-way decision articles.

Different from degree centrality, eigenvector centrality further considered the importance of the nodes that they are connected to [2, 50, 71]. In fact, this is very reasonable because if all the nodes in a network are assigned with different scores in advance, the connections to high-scoring nodes will naturally contribute more than those to low-scoring nodes.

For an undirected network G with n nodes, let A be the adjacency matrix of G, λ be a constant, and $x = (x_1, x_2, \ldots, x_n)$ be a vector of centralities of all nodes. If x_i is defined as (see e.g. [69] for details)

$$x_i = \frac{1}{\lambda} \sum_{j=1}^{n} A_{ij} x_j, \tag{5}$$

then the above equation can be rewritten in the matrix form as

$$\lambda x = Ax. \tag{6}$$

It can be observed that λ is an eigenvalue of the adjacency matrix A with the corresponding eigenvector x. In the eigenvector centrality measure, when an appropriate eigenvalue λ of the adjacency matrix A is obtained, we find the required eigenvector x.

Google's PageRank algorithm can be considered as a generalization of the eigenvector centrality measure. Compared with the eigenvector centrality measure, PageRank algorithm further considered the importance of the nodes that they are pointed to as a relative value [3]. In other words, the contribution of the importance of a node is also affected by the out-degree of this node. The more the number of the edges directly connecting one node to the other neighbors is, the less the contribution of the importance of this node is. In [50], the PageRank value was defined by

$$x_i = \alpha \sum_{j=1}^{n} A_{ij} \frac{x_j}{k_j^{\text{out}}} + \beta_i, \tag{7}$$

where x_i is the centrality measure of the web page i, k_j^{out} is the out-degree of the web page j, and α and β_i are two positive constants. Note that α is used to keep a balance between the first item and the second item in the equation, and it was often set to be 0.85 for the purpose of accelerating the convergence speed. The parameter β_i can be determined by the text similarity between the web page and query condition. In PageRank algorithm, the PageRank value will be iterated until the computational results are unchanged or only slightly changed. In this case, the PageRank algorithm will be terminated and the required vector x of centralities can be obtained.

Hyperlink-induced topic search (HITS) algorithm can provide more information than the other network centrality algorithms [20, 50]. It gives each node two values: authority centrality and hub centrality. The former shows the authority of a node which can be quantified by the number of nodes with high hub centrality connecting to it, and the latter shows the importance of a node which can be quantified by the number of nodes with high authority centrality that it is pointing to. For example, in citation network of papers, nodes with high authority centrality mean the important or influential papers in the research field, and nodes with high hub centrality mean the ordinary papers which cite many important or influential papers.

The concept of a connected component of a network comes from graph theory. It is a maximal subnetwork in which each pair of nodes is connected by a path. In real networks, there often exist some large components which include most of nodes and the rest of nodes are divided into many small components disconnected to each other [50, 68]. For example, scientific collaboration network has some large cooperative groups and many small cooperative groups.

Detecting community structures of a network is an important issue in the domain of complex network analysis. If a network has community structures, it means that the network can be easily grouped into sets of nodes within which connections are dense but between which they are sparse [49, 50]. For instance, the scientific collaboration network and university collaboration network in this paper are divided naturally into communities.

3 Scientific collaboration network

Scientific collaboration network is a social network where nodes are scientists and links are co-authorships. The 549 papers collected in this paper with the topic of three-way decision involves 709 authors. In order to avoid too many authors' influence on the final analytical results, this paper only considered the papers where the number of authors is less than or equal to 5.

The scientific collaboration network visualization was realized by Gephi which is an open source network analysis and visualization software package, and it is shown in Fig. 3. This weighted network has 644 nodes and 1235 edges and it includes 68 connected components. The number of connections repeated between two nodes was converted into the weight of a link. The node size is proportional to its degree (the number of coauthors). The node color is based on connected components. The edge width is based on its weight (the number of cooperation in pairs). Top 10 components colored the nodes with color and the remain components colored the nodes with gray. The first largest component represented by Professor Yao Yiyu has 366 nodes and 830 edges accounted for 56.83%. The second largest component represented by Professor Xue Zhan'ao has 13 nodes and 37 edges accounted for 2.02%. The third largest component represented by Professor Zhu Yanhui has 12 nodes and 29 edges accounted for 1.86%.

In Fig. 4, the community structures in the largest components of the scientific collaboration network are shown. The network was divided into 18 communities. The node color is based on community cluster membership. The node size is proportional to its degree (the number of coauthors). Top 10 communities colored the nodes with color and the remain colored the nodes with gray.

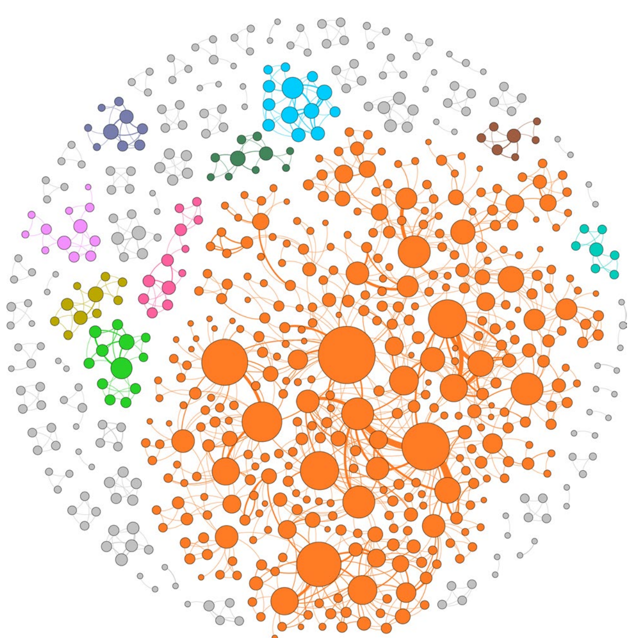

Fig. 3 Scientific collaboration network

Table 1 shows the ranking of communities in the investigation of three-way decision model. Members of influential communities and their proportion of total researchers are listed in the table.

In order to obtain the influence of authors in the three-way decision field, author rankings are listed with different evaluating measures and benchmarks in Table 2. The evaluation indices are the number of papers published, the amount of coauthors, the frequency of collaboration with other authors, and eigenvector centrality (considering both quantity and quality of coauthors). In the 2nd, 3rd, 4th columns of Table 2, each of them includes two parts: author name and the value of certain index.

4 University collaboration network

In this section, we construct a weighted network in which nodes represent universities and undirected edges indicate co-occurrence pairs of universities. The number of connections repeated between two nodes was converted into the weight of a link.

The university collaboration network visualization was realized by Gephi, and it is shown in Fig. 5. This weighted network has 156 nodes and 268 edges and it includes 15 connected components. The node size is proportional to its degree (the number of co-occurrence pairs between this node and other universities). The node color is based on connected components. Top 10 connected components colored the nodes with color and the remain colored the nodes with gray. The edge width is based on its weight (the number of connections repeated between the two nodes of this edge). The largest component represented by University of Regina has 135 nodes and 245 edges accounted for 77.56%.

In Fig. 6, the community structures in the largest components of the university collaboration network are shown. The network was divided into nine communities. The node color is based on community cluster membership. The node size is proportional to its degree.

Table 3 shows the ranking of communities. Members of influential communities and their proportion of total universities are listed in the table.

5 Networks of scientific papers

In this section, we construct a directed network (citation network) and two undirected networks (bibliographic coupling network and co-citation network) by bibliography to

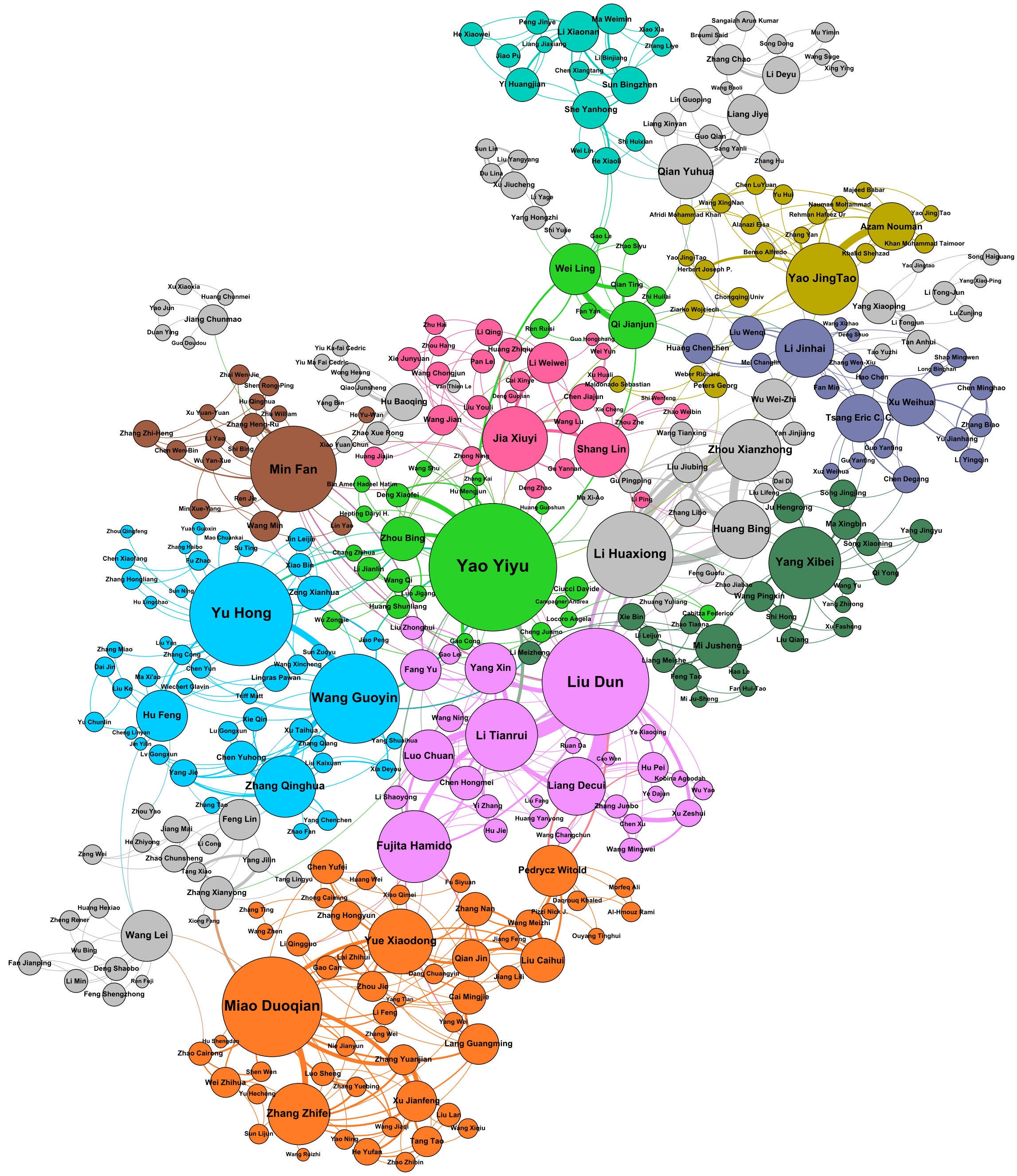

Fig. 4 Community detection of scientific collaboration network

reveal the relationship between papers. In the analysis of the dataset of three-way decision papers, we used digital object identifier (DOI) which is a string of numbers, letters and symbols as the unique label to identify a paper. Moreover, the PageRank and HITS algorithms will be used to search for the most important or influential papers in the networks.

5.1 Citation network

Citation network is a directed network in which nodes are papers and there is a directed edge from paper M to paper N if M cites N in its bibliography [50, 71].

The citation network visualization was realized by Gephi, and it is shown in Fig. 7. This directed network

Table 1 Community ranking of modularity class

No.	Members of community	Percentage %
1	Miao Duoqian, Zhang Zhifei, Yue Xiaodong, Xu Jianfeng, Liu Caihui, Pedrycz Witold, Zhang Yuanjian, Qian Jin, Lang Guangming, Wang Meizhi, Zhang Nan, Cai Mingjie, Zhou Jie et al.	14.48
2	Yu Hong, Wang Guoyin, Zhang Qinghua, Hu Feng, Chen Yuhong et al.	12.57
3	Yao Yiyu, Wei Ling, Qi Jianjun, Zhou Bing, Qian Ting, Deng Xiaofei et al.	7.92
4	Liu Dun, Li Tianrui, Liang Decui, Fujita Hamido, Luo Chuan, Yang Xin, Xu Zeshui, Fang Yu, Chen Hongmei, Hu Pei et al.	7.65
5	Jia Xiuyi, Shang Lin, Li Weiwei et al.	7.65
6	Yang Xibei, Mi Jusheng, Wang Pingxin et al.	6.28
7	Yao JingTao, Azam Nouman et al.	6.01
8	Li Jinhai, Xu Weihua, Tsang ECC, Liu Wenqi et al.	5.74
9	Min Fan, Wang Min et al.	4.64
10	Li Xiaonan, Sun Bingzhen, She Yanhong et al.	4.37
11	Qian Yuhua, Liang Jiye, Li Deyu, Zhang Chao et al.	4.37
12	Li Huaxiong, Zhou Xianzhong, Huang Bing, Zhang Libo, Liu Jiubing, Wu Weizhi et al.	3.83
13	Zhang Xianyong, Yang Jilin et al.	3.01
14	Hu Bao Qing, Zhao Xue Rong et al.	2.73

Table 2 Authors ranking

No.	Number of papers	Number of coauthors	Number of cooperation	Eigenvector centrality
1	Yao Yiyu 55	Yao Yiyu 34	Liu Dun 100	Liu Dun
2	Liu Dun 45	Liu Dun 28	Li Tianrui 83	Yao Yiyu
3	Li Tianrui 34	Yu Hong 27	Li Huaxiong 68	Li Tianrui
4	Liang Decui 30	Miao Duoqian 26	Yao Yiyu 67	Fujita Hamido
5	Yu Hong 29	Wang Guoyin 23	Miao Duoqian 67	Li Huaxiong
6	Miao Duoqian 29	Min Fan 22	Liang Decui 56	Yang Xin
7	Li Huaxiong 25	Li Huaxiong 22	Zhou Xianzhong 52	Luo Chuan
8	Yao JingTao 23	Fujita Hamido 18	Huang Bing 49	Zhou Xianzhong
9	Wang Guoyin 23	Li Tianrui 18	Yu Hong 48	Fang Yu
10	Zhou Xianzhong 17	Yao JingTao 18	Wang Guoyin 47	Huang Bing
11	Hu Bao Qing 16	Yang Xibei 18	Fujita Hamido 36	Chen Hongmei
12	Azam Nouman 15	Jia Xiuyi 16	Yao JingTao 36	Liang Decui
13	Huang Bing 15	Yue Xiaodong 16	Luo Chuan 33	Min Fan
14	Jia Xiuyi 14	Zhang Zhifei 15	Yue Xiaodong 31	Yu Hong
15	Wei Ling 14	Zhou Xianzhong 15	Zhang Zhifei 31	Wang Guoyin
16	Min Fan 14	Zhang Qinghua 15	Yang Xin 31	Wang Ning
17	Zhang Qinghua 13	Li Jinhai 14	Min Fan 30	Miao Duoqian
18	Zhang Yanping 12	Liang Decui 14	Zhang Qinghua 30	Hu Pei
19	Shang Lin 12	Huang Bing 14	Wei Ling 28	Yang Xibei
20	Fujita Hamido 11	Shang Lin 13	Jia Xiuyi 27	Jia Xiuyi
21	Qi Jianjun 11	Qian Yuhua 13	Xu Jianfeng 27	Yue Xiaodong

has 4578 nodes and 14,627 edges. The node size is proportional to its in-degree (times cited by the papers in our dataset). This investigation shows that about 210 papers were never cited at all accounted for 4.59%. For the remainder, 2874 papers have one citation accounted for 62.78%, 554 papers have two citations accounted for 12.1%, and 201 papers have three citations accounted for 4.4%. Only 194 papers have 10 or more citations accounted for 4.24%, and just 7 papers have 100 or more citations accounted for 0.15%.

Table 4 shows the ranking of papers by in-degree (times cited by the papers in our dataset) based on citation network. Top 12 influential papers [26, 32, 34, 41, 66, 74, 82, 87, 89, 95, 99, 115] and their cited times are listed. Note that the citations in Table 4 contain two parts: the former is the in-degree of a paper within the citation network, while the

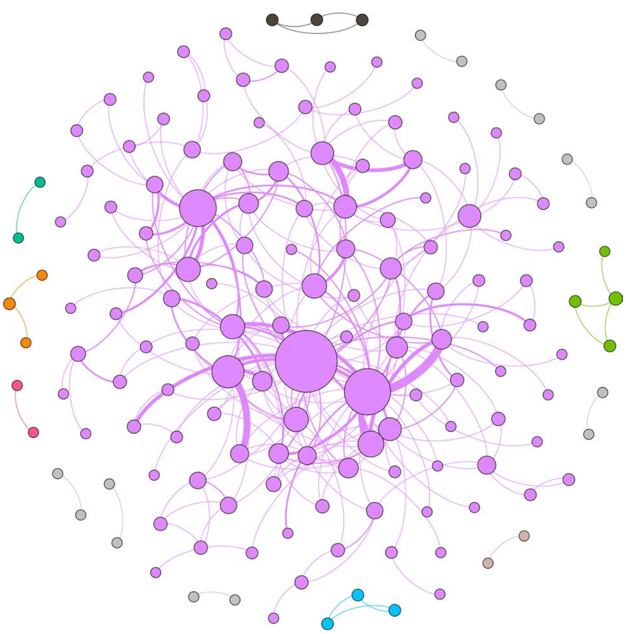

Fig. 5 University collaboration network

latter quotes from ISI Web of Science (their topics are not limited to three-way decision). By the way, we list in Table 4 only the papers from our downloaded dataset. However, the papers [8, 16, 17, 30, 51, 52, 58, 65, 77–80, 85, 86, 93, 97, 114] beyond our dataset are also given in Table 5 for the convenience of interested readers' reference. In other words, the topics of the papers listed in Table 5 may not be three-way decision. These additional papers are recommended to readers because they were cited very frequently together with those in Table 4. Generally speaking, if readers want to well understand the ideas of the influential papers in Table 4, it is necessary to read those in Table 5 at the same time.

By Table 5, we can obtain more information on papers outside our downloaded dataset. There are two reasons: (1) keywords of some papers do not cover three-way decision, such as Nos. 6, 7, 8, 10, 11, 12, 13, 14, 16, 17 in Table 5; (2) some papers have not been included in ISI Web of Science (of course, they do not have any citation in ISI Web of Science, but still have citations in our citation network), which are presented in Table 5 with "–". In other words, these important papers outside our dataset can be successfully traced by complex network analysis although they were not included in our dataset.

Table 6 shows the ranking of papers by PageRank algorithm on citation network. Top 12 influential papers [7, 26, 28, 32, 34, 73, 74, 82, 87, 95, 99, 115] and their cited times are listed. To compare with Table 4, there are significant differences for these two ranking methods in terms of top 12 papers. The reasons are as follows: the ranking results

Fig. 6 Community detection of university collaboration network

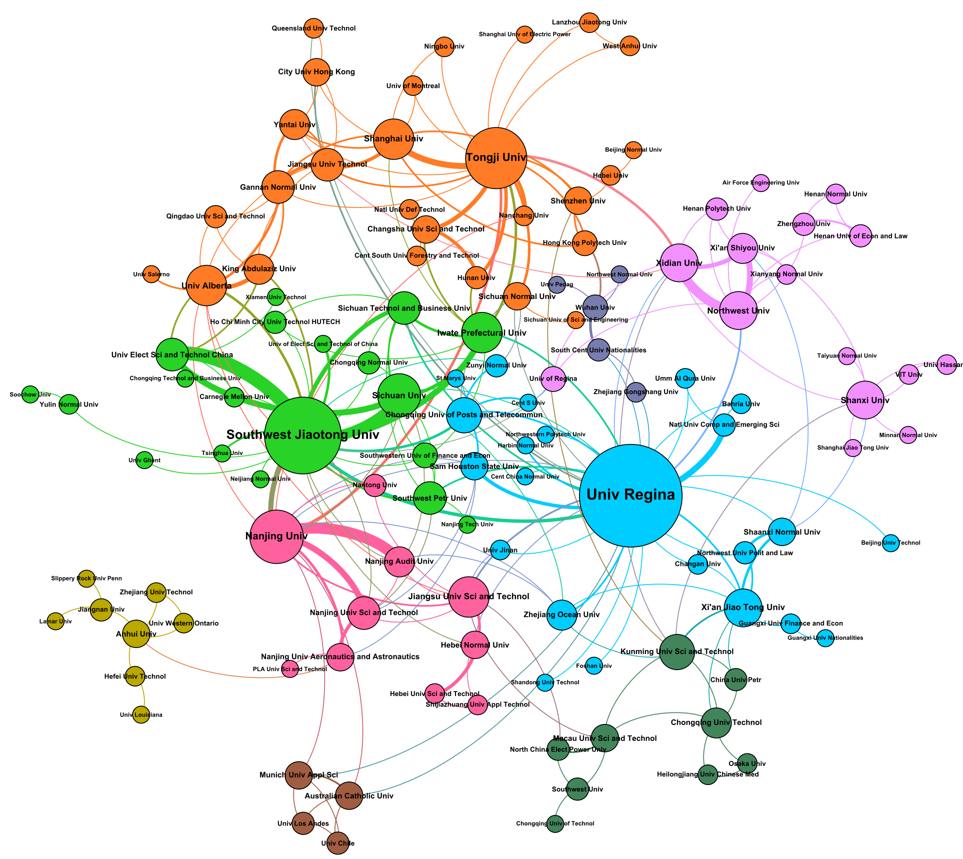

Fig. 6 Community detection of university collaboration network

Table 3 Community ranking of modularity class

No.	Members of community	Percentage %
1	Tongji Univ, Shanghai Univ, Univ Alberta, Gannan Normal Univ, Jiangsu Univ Technol, Changsha Univ Sci and Technol, King Abdulaziz Univ, Yantai Univ, Hunan Univ, etc.	22.31
2	Univ Regina, Chongqing Univ of Posts and Telecommun, Xi'an Jiao Tong Univ, Natl Univ Comp and Emerging Sci, Shaanxi Normal Univ, Sam Houston State Univ, etc.	19.01
3	Southwest Jiaotong Univ, Sichuan Univ, Univ Elect Sci and Technol China, Iwate Prefectural Univ, Sichuan Technol and Business Univ, Southwest Petr Univ, etc.	15.7
4	Northwest Univ, Xidian Univ, Xi'an Shiyou Univ, Shanxi Univ, etc.	13.22
5	Nanjing Univ, Nanjing Audit Univ, Nanjing Univ Sci and Technol, Jiangsu Univ Sci and Technol, Hebei Normal Univ, Nanjing Univ Aeronautics and Astronautics, etc.	8.26
6	Kunming Univ Sci and Technol, Macau Univ Sci and Technol, Chongqing Univ Technol, etc.	7.44
7	Anhui Univ, Jiangnan Univ, etc.	6.61
8	Wuhan Univ, South Cent Univ Nationalities, etc.	4.13
9	Munich Univ Appl Sci, Australian Catholic Univ, etc.	3.31

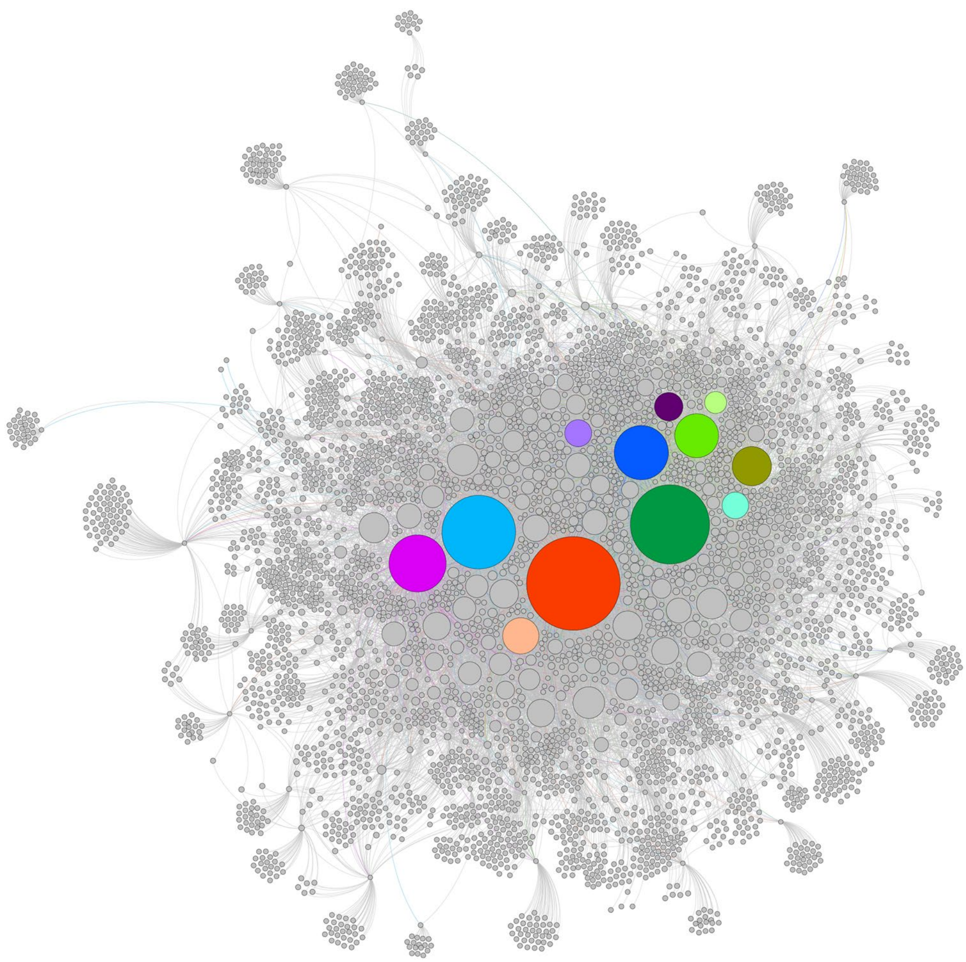

Fig. 7 Citation network

shown in Table 4 completely depend on the quantity of citations, while those in Table 6 depend on both the quantity and quality of citations.

Table 7 shows the ranking of papers by HITS algorithm on citation network. That is, the third ranking method was used to rank papers and the ranking results are given in Table 7. It can be observed from Tables 4 and 7 that the parameters times cited and authority seem to have similar effects on the evaluation of academic papers for our dataset since only one paper is different between them. This is not

Table 4 Paper ranking of citation network

No.	Paper	Times cited	ESI
1	Yao [82]	307/541	Yes
2	Yao [87]	79/123	No
3	Li et al. [34]	63/125	Yes
4	Li et al. [32]	60/76	No
5	Ziarko [115]	42/156	No
6	Liang et al. [26]	38/49	No
7	Zhang et al. [99]	38/55	No
8	Yang et al. [74]	29/46	No
9	Yao [89]	29/47	No
10	Sun et al. [66]	28/66	Yes
11	Lang et al. [21]	27/35	No
12	Yu et al. [95]	26/34	No

Table 5 Paper ranking beyond our dataset

No.	Paper	Times cited	ESI
1	Pawlak [52]	257/–	No
2	Yao [85]	181/–	No
3	Yao and Wong [77]	171/–	No
4	Yao [79]	135/–	No
5	Ziarko [114]	119/–	No
6	Li and Zhou [30]	109/183	No
7	Herbert and Yao [8]	85/133	No
8	Yao [78]	82/215	No
9	Zadeh [97]	77/–	No
10	Yao and Zhao [80]	77/391	No
11	Simiński [65]	76/15	No
12	Jia et al. [16]	68/170	Yes
13	Yao [86]	63/80	No
14	Yu et al. [93]	62/123	No
15	Pawlak et al. [51]	60/–	No
16	Qian et al. [58]	60/451	Yes
17	Jia et al. [17]	57/100	No

surprising because the ranking methods based on times cited and authority have the same characteristic of depending on the quantity only. To be more concrete, the former depends on the quantity of citations, and the latter depends on the quantity of hubs connecting to it. In other words, if we view citations and hubs as the same, then the ranking methods based on times cited and authority will be similar.

Note that HITS algorithm also output top 12 hub papers when it obtained top 12 authority papers. In Table 8, top 12 hub papers [10, 11, 26–28, 37, 42, 56, 57, 62, 66, 101] and their in/out-degrees (times cited by the papers in our dataset/the number of bibliographies) are listed. According to the discussion in Sect. 2, we know that top 12 hub papers in Table 8 were used to help the generation of top

Table 6 Paper ranking of citation network on PageRank

No.	Paper	Times cited	ESI
1	Yao [82]	307/541	Yes
2	Yao [87]	79/123	No
3	Ziarko [115]	42/156	No
4	Herbert and Yao [7]	24/48	No
5	Li et al. [32]	60/76	No
6	Li et al. [34]	63/125	Yes
7	Zhang et al. [99]	38/55	No
8	Liang et al. [26]	38/49	No
9	Liang et al. [28]	24/45	No
10	Yu et al. [95]	26/34	No
11	Xu and Guo [73]	24/60	No
12	Yang et al. [74]	29/46	No

Table 7 Paper ranking of citation network on authorities

No.	Paper	Times cited	ESI
1	Yao [82]	307/541	Yes
2	Yao [87]	79/123	No
3	Li et al. [34]	63/125	Yes
4	Li et al. [32]	60/76	No
5	Ziarko [115]	42/156	No
6	Liang et al. [26]	38/49	No
7	Zhang et al. [99]	38/55	No
8	Sun et al. [66]	28/66	Yes
9	Yang et al. [74]	29/46	No
10	Lang et al. [21]	27/35	No
11	Herbert and Yao [7]	24/48	No
12	Yao [89]	29/47	No

Table 8 Paper ranking of citation network on hubs

No.	Paper	In/out-degree
1	Liu and Liang [42]	12/63
2	Qian et al. [57]	5/83
3	Qian et al. [56]	25/58
4	Hu et al. [11]	19/47
5	Liang et al. [26]	38/49
6	Liang et al. [27]	2/50
7	Liu et al. [37]	1/38
8	Qiao and Hu [62]	5/42
9	Sun et al. [66]	28/61
10	Liang et al. [28]	24/44
11	Hu et al. [10]	12/44
12	Zhang et al. [101]	6/54

12 authority papers in Table 7. In fact, the papers with high hub centrality and the papers with high authority centrality are beneficial to each other in HITS algorithm because they need to search for each other. In other words, readers can easily find influential papers from the papers with high hub centrality.

5.2 Bibliographic coupling network

Citation network provides a simple and intuitive way to show citation patterns but not the only one. An alternative representation is the bibliographic coupling network. Two papers are said to be bibliographic coupling if they cite the same other papers [50, 71]. In this section, we construct a weighted network in which nodes are papers and undirected edges indicate strengths of coupling which are the number of common citations between two papers.

The bibliographic coupling network visualization was realized by Gephi, and it is shown in Fig. 8. This weighted network has 532 nodes and 100,518 edges. The node size is proportional to its degree (the number of common citations). We only drew the nodes and ignored the huge number of edges. There are a slight difference among the sizes of nodes. The reason is that all the papers belong to the same research field. Here, we do not give the ranking of papers based on bibliographic coupling network because it lacks of convincing due to only slight difference among nodes.

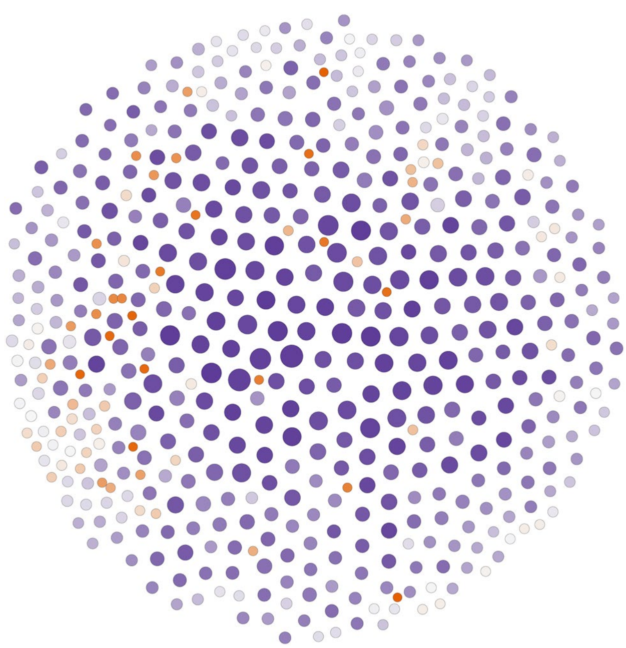

Fig. 8 Bibliographic coupling network

5.3 Co-citation network

Another undirected network of citation patterns is co-citation network. Two papers are said to be co-cited if they are both cited by the same third paper [50, 71]. In this section, we construct a weighted network in which nodes represent papers and undirected edges indicate strengths of co-citation which equal to the number of other papers that cite both of them.

The co-citation network visualization was realized by Gephi, and it is shown in Fig. 9. This weighted network has 247 nodes and 9744 edges. The node size is proportional to its degree (the number of co-citations).

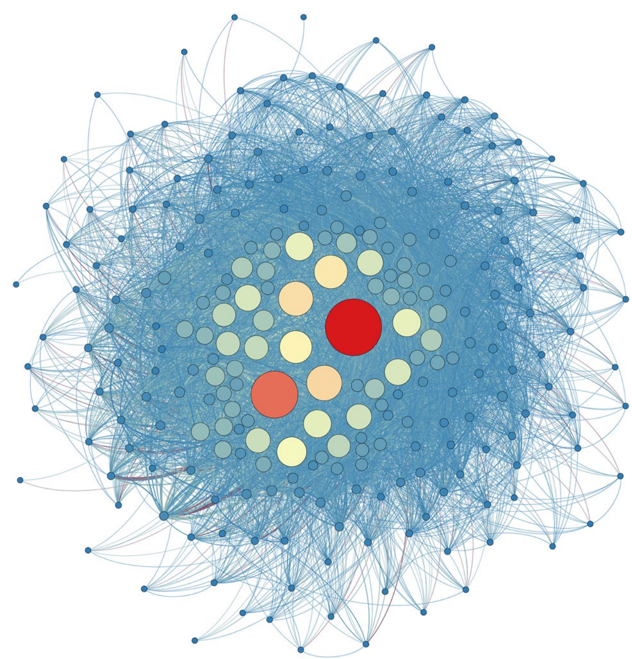

Fig. 9 Co-citation network

Table 9 Paper ranking of co-citation network

No.	Paper	Times cited	ESI
1	Yao [82]	541	Yes
2	Yao [84]	386	Yes
3	Li et al. [31]	126	Yes
4	Hu [9]	122	No
5	Liu et al. [41]	118	Yes
6	Yu et al. [94]	113	Yes
7	Qian et al. [59]	239	Yes
8	Yao and Azam [76]	93	No
9	Li et al. [32]	76	No
10	Deng and Yao [5]	104	No
11	Li et al. [34]	125	Yes
12	Zhang and Min [98]	102	Yes

Table 9 shows the ranking of papers by degree on co-citation network. Top 12 influential papers [5, 9, 31, 32, 34, 41, 59, 76, 82, 84, 94, 98] and their cited times are listed.

6 Keywords network

In this section, we construct a weighted network in which nodes represent keywords and undirected edges indicate co-occurrence pairs of keywords. The number of connections repeated between two nodes was converted into the weight of a link.

In Fig. 10, the keywords network visualization is shown in which weighted degree is not less than 10. This weighted network has 109 nodes and 544 edges. The node size is proportional to its degree (the number of co-occurrence pairs of keywords). The edge width is based on its weight. Top 10 keywords are three-way decision, decision-theoretic rough set, rough set, probabilistic rough set, granular computing, attribute reduction, fuzzy set, multi-granulation, cost-sensitive and loss function.

Furthermore, we trisected pairs of keywords according to occurring frequently together with three-way decision, and the trisection results are shown in Fig. 11. This star network has 100 nodes and 99 edges. In the figure, there are high frequency keywords (the frequency is greater than or equal to 11), middle frequency keywords, and low frequency keywords (the frequency is less than or equal to 3). These three types of keywords excluding three-way decision are colored with different colors from the inner to the outer. The trisection results of keywords provide useful hints for future researches of three-way decision since high frequency keywords often mean well-established researches, middle frequency keywords probably present emerging researches, and low frequency keywords may be new novel researches.

7 Conclusions

We have studied the dataset of three-way decision articles downloaded from ISI Web of Science. Concretely, the scientific collaboration network, university collaboration network, networks of scientific papers (i.e., citation network, bibliographic coupling network, co-citation network) and keywords network have been constructed to show the relationships between authors, affiliations, papers and keywords,

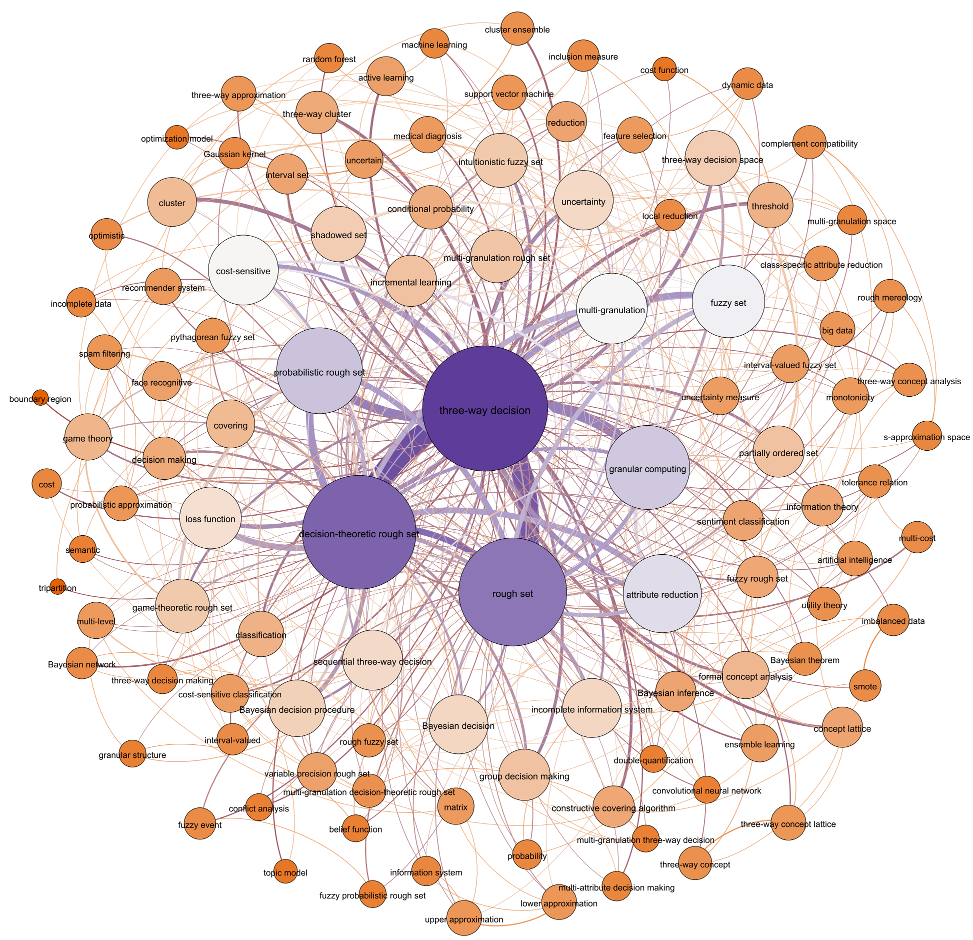

Fig. 10 Keywords network

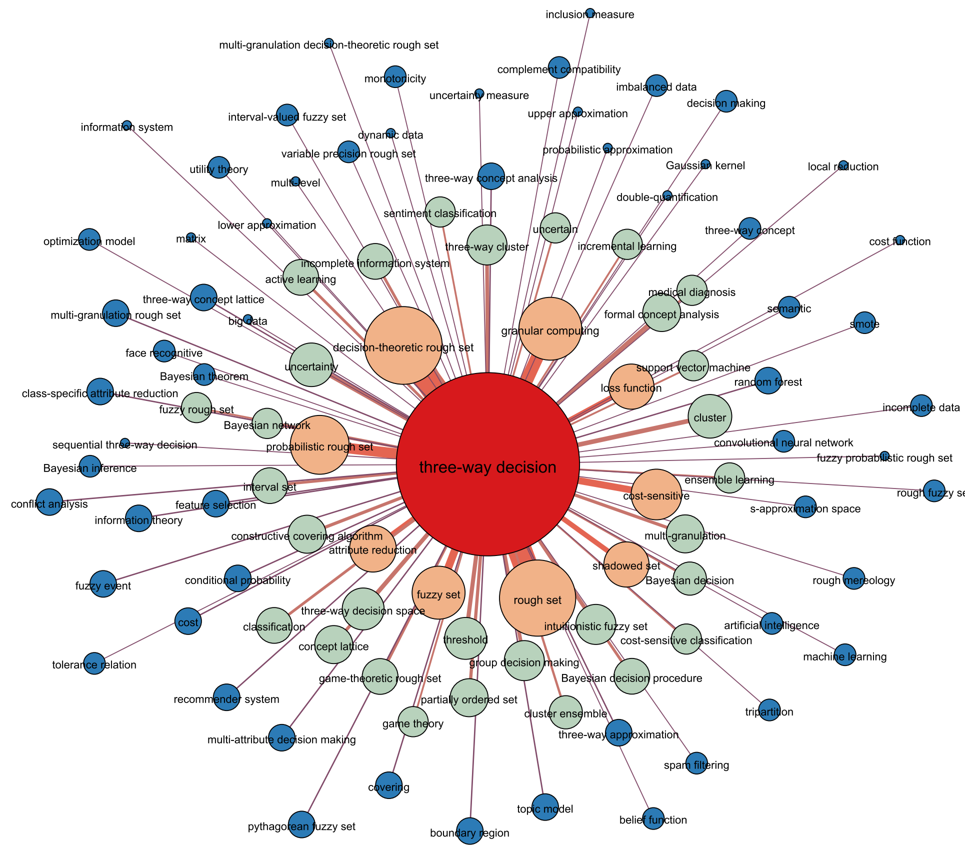

Fig. 11 Star network of keywords on three-way decision

respectively. Some interesting results are summarized as follows:

(1) The scientific collaboration network has shown that most of papers were completed by different authors with cooperation. Yao Yiyu, Liu Dun, Li Tianrui, Liang Decui, Yu Hong, Miao Duoqian, Li Huaxiong, Yao JingTao, Wang Guoyin, Min Fan, et al. played a huge role in the development of three-way decision. Moreover, the community structure has presented the detailed collaboration among them and other authors.

(2) University collaboration network has shown the cooperation between different research affiliations. A clear cooperative relationship among universities has been established with close geographical location, especially those in the same province. University of Regina, Southwest Jiaotong University, Tongji University, Nanjing University, Sichuan University, and University of Electronic Science and Technology of China actively promoted the development of three-way decision.

(3) Networks of scientific papers include citation network, bibliographic coupling network and co-citation network. It has been shown that Professor Yao Yiyu's two pioneering papers occupied the core position. Some important papers outside our dataset can be successfully traced by complex network analysis. It is interesting that there is a strong correlation between the weighted degree of papers in the co-citation network and ESI highly cited papers.

(4) Keywords network has shown that decision-theoretic rough set, rough set, probabilistic rough set, granular computing, attribute reduction, fuzzy set, multi-granulation, cost-sensitive, loss function and sequential three-way decision are the hottest research issues around three-way decision.

Acknowledgements The authors would like to thank Professor Yao Yiyu for his valuable comments and suggestions on the preliminary draft. This work was supported by the National Natural Science Foundation of China (Nos. 11947041 and 11971211).

References

1. Afridi MK, Azam N, Yao JT, Alanazi E (2018) A three-way clustering approach for handling missing data using GTRS. Int J Approx Reason 98:11–24
2. Borgatti SP (2005) Centrality and network flow. Soc Netw 27(1):55–71
3. Brin S, Page L (1998) The anatomy of a large-scale hypertextual web search engine. Comput Netw ISDN Syst 30(1–7):107–117
4. Ciucci D (2011) Orthopairs: A simple and widely used way to model uncertainty. Fundamenta Informaticae 108(3–4):287–304

5. Deng XF, Yao YY (2014) Decision-theoretic three-way approximations of fuzzy sets. Inf Sci 279:702–715
6. Hao C, Li JH, Fan M, Liu WQ, Tsang ECC (2017) Optimal scale selection in dynamic multi-scale decision tables based on sequential three-way decisions. Inf Sci 415:213–232
7. Herbert JP, Yao JT (2009) Criteria for choosing a rough set model. Comput Math Appl 57(6):908–918
8. Herbert JP, Yao JT (2011) Game-theoretic rough sets. Fundam Inf 108(3–4):267–286
9. Hu BQ (2014) Three-way decisions space and three-way decisions. Inf Sci 281:21–52
10. Hu BQ, Wong H, Yiu MFC (2016) The aggregation of multiple three-way decision spaces. Knowl Based Syst 98:241–249
11. Hu BQ, Wong H, Yiu KFC (2017) On two novel types of three-way decisions in three-way decision spaces. Int J Approx Reason 82:285–306
12. Hu BQ (2016) Three-way decision spaces based on partially ordered sets and three-way decisions based on hesitant fuzzy sets. Knowl Based Syst 91:16–31
13. Huang CC, Li JH, Mei CL, Wu WZ (2017) Three-way concept learning based on cognitive operators: an information fusion viewpoint. Int J Approx Reason 83:218–242
14. Hu JH, Yang Y, Chen XH (2018) A novel TODIM method-based three-way decision model for medical treatment selection. Int J Fuzzy Syst 20(4):1240–1255
15. Jia XY, Zheng K, Li WW, Liu TT, Shang L (2012) Three-way decisions solution to filter spam email: an empirical study. In: International conference on rough sets and current trends in computing, pp 287–296
16. Jia XY, Liao WH, Tang ZM, Shang L (2013) Minimum cost attribute reduction in decision-theoretic rough set models. Inf Sci 219:151–167
17. Jia XY, Tang ZM, Liao WH, Shang L (2014) On an optimization representation of decision-theoretic rough set model. Int J Approx Reason 55(1):156–166
18. Jia XY, Deng Z, Min F, Dun Liu D (2019) Three-way decisions based feature fusion for Chinese irony detection. Int J Approx Reason 113:324–335
19. Ju HR, Pedrycz W, Li HX, Ding WP, Yang XB, Zhou XZ (2019) Sequential three-way classifier with justifiable granularity. Knowl Based Syst 163:103–119
20. Kleinberg JM (1999) Authoritative sources in a hyperlinked environment. J ACM 46(5):604–632
21. Lang GM, Miao DQ, Cai MJ (2017) Three-way decision approaches to conflict analysis using decision-theoretic rough set theory. Inf Sci 406:185–207
22. Liang DC, Liu D, Pedrycz Witold, Hu Pei (2013) Triangular fuzzy decision-theoretic rough sets. Int J Approx Reason 54(8):1087–1106
23. Liang DC, Pedrycz W, Liu D, Hu P (2015) Three-way decisions based on decision-theoretic rough sets under linguistic assessment with the aid of group decision making. Appl Soft Comput 29:256–269
24. Liang DC, Liu D (2015) A novel risk decision making based on decision-theoretic rough sets under hesitant fuzzy information. IEEE Trans Fuzzy Syst 23(2):237–247
25. Liang DC, Liu D (2015) Deriving three-way decisions from intuitionistic fuzzy decision-theoretic rough sets. Inf Sci 300:28–48
26. Liang DC, Liu D, Kobina A (2016) Three-way group decisions with decision-theoretic rough sets. Inf Sci 345:46–64
27. Liang DC, Xu ZS, Liu D (2017) A new aggregation method-based error analysis for decision-theoretic rough sets and its application in hesitant fuzzy information systems. IEEE Trans Fuzzy Syst 25(6):1685–1697
28. Liang DC, Xu ZS, Liu D (2017) Three-way decisions with intuitionistic fuzzy decision-theoretic rough sets based on point operators. Inf Sci 375:183–201
29. Liang DC, Xu ZS, Liu D, Wu Y (2018) Method for three-way decisions using ideal TOPSIS solutions at Pythagorean fuzzy information. Inf Sci 435:282–295
30. Li HX, Zhou XZ (2011) Risk decision making based on decision-theoretic rough set: a three-way view decision model. Int J Comput Intell Syst 4(1):1–11
31. Li HX, Zhang LB, Huang B, Zhou XZ (2016) Sequential three-way decision and granulation for cost-sensitive face recognition. Knowl Based Syst 91:241–251
32. Li HX, Zhang LB, Zhou XZ, Huang B (2017) Cost-sensitive sequential three-way decision modeling using a deep neural network. Int J Approx Reason 85:68–78
33. Li JH, Mei CL, Lv YJ (2013) Incomplete decision contexts: approximate concept construction, rule acquisition and knowledge reduction. Int J Approx Reason 54(1):149–165
34. Li JH, Huang CC, Qi JJ, Qian YH, Liu WQ (2017) Three-way cognitive concept learning via multi-granularity. Inf Sci 378:244–263
35. Li MZ, Wang GY (2016) Approximate concept construction with three-way decisions and attribute reduction in incomplete contexts. Knowl Based Syst 91:165–178
36. Lingras P, Chen M, Miao DQ (2008) Rough cluster quality index based on decision theory. IEEE Trans Knowl Data Eng 21(7):1014–1026
37. Liu CH, Pedrycz W, Jiang F, Wang MZ (2018) Decision-theoretic rough set approaches to multi-covering approximation spaces based on fuzzy probability measure. J Intell Fuzzy Syst 34(3):1917–1931
38. Liu D, Li TR, Ruan D (2011) Probabilistic model criteria with decision-theoretic rough sets. Inf Sci 181(17):3709–3722
39. Liu D, Li TR, Liang DC (2012) Three-way government decision analysis with decision-theoretic rough sets. Int J Uncertain Fuzziness Knowl Based Syst 20:119–132
40. Liu D, Li TR, Liang DC (2014) Incorporating logistic regression to decision-theoretic rough sets for classifications. Int J Approx Reason 55(1):197–210
41. Liu D, Liang DC, Wang CC (2016) A novel three-way decision model based on incomplete information system. Knowl Based Syst 91:32–45
42. Liu D, Liang DC (2017) Three-way decisions in ordered decision system. Knowl Based Syst 137:182–195
43. Liu JB, Li HX, Zhou XZ, Huang B, Wang TX (2019) An optimization-based formulation for three-way decisions. Inf Sci 495:185–214
44. Li XN (2019) Three-way fuzzy matroids and granular computing. Int J Approx Reason 114:44–50
45. Li YF, Zhang CQ, Swan Jason R (2000) An information filtering model on the web and its application in JobAgent. Knowl Based Syst 13(5):285–296
46. Li W, Miao DQ, Wang WL, Zhang N (2010) Hierarchical rough decision theoretic framework for text classification. In: IEEE international conference on cognitive informatics, pp 484–489
47. Min F, He HP, Qian YH, Zhu W (2011) Test-cost-sensitive attribute reduction. Inf Sci 181(22):4928–4942
48. Nauman M, Azam N, Jingtao Yao JT (2016) A three-way decision making approach to malware analysis using probabilistic rough sets. Inf Sci 374:193–209
49. Newman M (2006) Modularity and community structure in networks. Proc Natl Acad Sci 103(23):8577–8582
50. Newman M (2018) Networks. Oxford University Press, New York
51. Pawlak Z, Wong SK, Ziarko W (1988) Rough sets: probabilistic versus deterministic approach. Int J Man Mach Stud 29(1):81–95

52. Pawlak Z (1981) Information systems theoretical foundations. Inf Syst 6(3):205–218
53. Pedrycz W (1998) Shadowed sets: representing and processing fuzzy sets. IEEE Trans Syst Man Cybern Part B Cybern 28(1):103–109
54. Peters JF, Ramanna S (2016) Proximal three-way decisions: theory and applications in social networks. Knowl Based Syst 91:4–15
55. Qi JJ, Qian T, Wei L (2016) The connections between three-way and classical concept lattices. Knowl Based Syst 91:143–151
56. Qian J, Dang CY, Yue XD, Zhang N (2017) Attribute reduction for sequential three-way decisions under dynamic granulation. Int J Approx Reason 85:196–216
57. Qian J, Liu CH, Yue XD (2019) Multigranulation sequential three-way decisions based on multiple thresholds. Int J Approx Reason 105:396–416
58. Qian YH, Liang JY, Yao YY, Dang CY (2010) MGRS: a multi-granulation rough set. Inf Sci 180(6):949–970
59. Qian YH, Zhang H, Sang YL, Liang JY (2014) Multigranulation decision-theoretic rough sets. Int J Approx Reason 55(1):225–237
60. Qian T, Wei L, Qi JJ (2017) Constructing three-way concept lattices based on apposition and subposition of formal contexts. Knowl Based Syst 116:39–48
61. Qian T, Wei L, Qi JJ (2019) A theoretical study on the object (property) oriented concept lattices based on three-way decisions. Soft Comput 23(19):9477–9489
62. Qiao JS, Hu BQ (2018) On transformations from semi-three-way decision spaces to three-way decision spaces based on triangular norms and triangular conorms. Inf Sci 432:22–51
63. Ren RS, Wei L (2016) The attribute reductions of three-way concept lattices. Knowl Based Syst 99:92–102
64. Ren RS, Wei L, Yao YY (2018) An analysis of three types of partially-known formal concepts. Int J Mach Learn Cybern 9(11):1767–1783
65. Simiński K (2012) Neuro-rough-fuzzy approach for regression modelling from missing data. Int J Appl Math Comput Sci 22(2):461–476
66. Sun BZ, Ma WM, Xiao X (2017) Three-way group decision making based on multigranulation fuzzy decision-theoretic rough set over two universes. Int J Approx Reason 81:87–102
67. Sun BZ, Chen XT, Zhang LY, Ma WM (2020) Three-way decision making approach to conflict analysis and resolution using probabilistic rough set over two universes. Inf Sci 507:809–822
68. Tarjan R (1972) Depth-first search and linear graph algorithms. SIAM J Comput 1(2):146–160
69. Umadevi V (2013) Case study—centrality measure analysis on co-authorship network. J Glob Res Comput Sci 4(1):67–70
70. Wang PX, Shi H, Yang XB, Mi JS (2019) Three-way k-means: integrating k-means and three-way decision. Int J Mach Learn Cybern 10(10):2767–2777
71. Wang XF, Li X, Chen GR (2012) Network science: an introduction. Higher Education Press, Beijing
72. Wong SKM, Ziarko W (1987) Comparison of the probabilistic approximate classification and the fuzzy set model. Fuzzy Sets Syst 21(3):357–362
73. Xu WH, Guo YT (2016) Generalized multigranulation double-quantitative decision-theoretic rough set. Knowl Based Syst 105:190–205
74. Yang X, Li TR, Fujita H, Liu D, Yao YY (2017) A unified model of sequential three-way decisions and multilevel incremental processing. Knowl Based Syst 134:172–188
75. Yang XP, Yao JT (2012) Modelling multi-agent three-way decisions with decision-theoretic rough sets. Fundam Inf 115(2–3):157–171
76. Yao JT, Azam N (2015) Web-based medical decision support systems for three-way medical decision making with game-theoretic rough sets. IEEE Trans Fuzzy Syst 23(1):3–15
77. Yao YY, Wong SKM (1992) A decision theoretic framework for approximating concepts. Int J Man Mach Stud 37(6):793–809
78. Yao YY (2003) Probabilistic approaches to rough sets. Expert Syst 20(5):287–297
79. Yao YY (2008) Probabilistic approcah to rough set. Int J Approx Reason 49:255–271
80. Yao YY, Zhao Y (2008) Attribute reduction in decision-theoretic rough set models. Inf Sci 178(17):3356–3373
81. Yao YY (2009) Three-way decision: an interpretation of rules in rough set theory. International conference on rough sets and knowledge technology. Springer, Berlin, pp 642–649
82. Yao YY (2010) Three-way decisions with probabilistic rough sets. Inf Sci 180(3):341–353
83. Yao YY, Deng XF (2011) Sequential three-way decisions with probabilistic rough sets. In: IEEE 10th international conference on cognitive informatics and cognitive computing, pp 120–125
84. Yao YY (2011) The superiority of three-way decisions in probabilistic rough set models. Inf Sci 181(6):1080–1096
85. Yao YY (2012) An outline of a theory of three-way decisions. International conference on rough sets and current trends in computing. Springer, Berlin, pp 1–17
86. Yao YY (2013) Granular computing and sequential three-way decisions. In: International conference on rough sets and knowledge technology. Springer, Berlin, pp 16–27
87. Yao YY (2016) Three-way decisions and cognitive computing. Cogn Comput 8(4):543–554
88. Yao YY (2017) Interval sets and three-way concept analysis in incomplete contexts. Int J Mach Learn Cybern 8(1):3–20
89. Yao YY (2018) Three-way decision and granular computing. Int J Approx Reason 103:107–123
90. Yao YY (2019) Three-way conflict analysis: reformulations and extensions of the Pawlak model. Knowl Based Syst 180:26–37
91. Yao YY (2020) Three-way granular computing, rough sets, and formal concept analysis. Int J Approx Reason 116:106–125
92. Yao YY (2020) Tri-level thinking: models of three-way decision. Int J Mach Learn Cybern. https://doi.org/10.1007/s13042-019-01040-2
93. Yu H, Liu ZG, Wang GY (2014) An automatic method to determine the number of clusters using decision-theoretic rough set. Int J Approx Reason 55(1):101–115
94. Yu H, Zhang C, Wang GY (2016) A tree-based incremental overlapping clustering method using the three-way decision theory. Knowl Based Syst 91:189–203
95. Yu H, Jiao P, Yao YY, Wang GY (2016) Detecting and refining overlapping regions in complex networks with three-way decisions. Inf Sci 373:21–41
96. Yu HY, Li QG, Cai MJ (2018) Characteristics of three-way concept lattices and three-way rough concept lattices. Knowl Based Syst 146:181–189
97. Zadeh LA (1965) Fuzzy sets. Inf Control 8(3):338–353
98. Zhang HR, Min F (2016) Three-way recommender systems based on random forests. Knowl Based Syst 91:275–286
99. Zhang HR, Min F, Shi B (2017) Regression-based three-way recommendation. Inf Sci 378:444–461
100. Zhang LB, Li HX, Zhou XZ, Huang B (2020) Sequential three-way decision based on multi-granular autoencoder features. Inf Sci 507:630–643
101. Zhang QH, Zhang Q, Wang GY (2016) The uncertainty of probabilistic rough sets in multi-granulation spaces. Int J Approx Reason 77:38–54
102. Zhang QH, Xie Q, Wang GY (2018) A novel three-way decision model with decision-theoretic rough sets using utility theory. Knowl Based Syst 159:321–335

103. Zhang QH, Lv GX, Chen YH, Wang GY (2018) A dynamic three-way decision model based on the updating of attribute values. Knowl Based Syst 142:71–84
104. Zhang QH, Xia DY, Liu KX, Wang GY (2020) A general model of decision-theoretic three-way approximations of fuzzy sets based on a heuristic algorithm. Inf Sci 507:522–539
105. Zhang XY, Miao DQ (2017) Three-way attribute reducts. Int J Approx Reason 88:401–434
106. Zhang XY, Tang X, Yang JL, Lv ZY (2020) Quantitative three-way class-specific attribute reducts based on region preservations. Int J Approx Reason 117:96–121
107. Zhang XY, Yang JL, Tang LY (2020) Three-way class-specific attribute reducts from the information viewpoint. Inf Sci 507:840–872
108. Zhang Y, Yao JT (2017) Gini objective functions for three-way classifications. Int J Approx Reason 81:103–114
109. Zhao XR, Hu BQ (2016) Fuzzy probabilistic rough sets and their corresponding three-way decisions. Knowl Based Syst 91:126–142
110. Zhi HL, Jinhai Li JH (2019) Granule description based knowledge discovery from incomplete formal contexts via necessary attribute analysis. Inf Sci 485:347–361
111. Zhi HL, Qi JJ, Qian T, Wei L (2019) Three-way dual concept analysis. Int J Approx Reason 114:151–165
112. Zhou B, Yao YY, Luo JG (2014) Cost-sensitive three-way email spam filtering. J Intell Inf Syst 42(1):19–45
113. Zhou B (2014) Multi-class decision-theoretic rough sets. Int J Approx Reason 55(1):211–224
114. Ziarko W (1993) Variable precision rough set model. J Comput Syst Sci 46(1):39–59
115. Ziarko W (2008) Probabilistic approach to rough sets. Int J Approx Reason 49(2):272–284