

# The Analysis of Water Project Bid Rigging Behavior Based on Complex Network

Tiexin Cheng<sup>1</sup>, Ting Liu<sup>1,\*</sup>, Lingzhi Meng<sup>1</sup> and Chaoyang Wang<sup>2</sup>

<sup>1</sup>College of Management, Tianjin Polytechnic University, Tianjin 300387, China;

<sup>2</sup>Tianjin Water Conservancy Transaction Management Center, Tianjin 300204, China

\*Corresponding author

**Abstract**—There are a series of issues in the development of water project bidding areas, the illegal act of bid rigging and colluding is particularly worth attention, and how to prevent the rigging and colluding behavior in the bidding process has already become a current hot topic. The community structure characteristics of complex networks was applied to analyze the bid rigging and colluding behavior of tenderers, the community detection model of bid rigging and colluding based on complex network was established, and the rationality and feasibility of the model were illustrated by analyzing the winning rate of the tenderers in the community, which showed that the community detection model of bid rigging and colluding based on complex network could provide some reference for the bidding supervision departments to identify the rigging colluding behavior.

**Keywords**—project bidding; complex network; bid rigging; community detection; newman fast algorithm

## I. INTRODUCTION

In recent years, bid rigging and colluding for water projects in the bidding activities happens occasionally because of their regional and technical features, which seriously disrupts the normal order of market competition and hinders the healthy development of water project bidding market. How to control this illegal behavior has become a major difficulty in bid management of water projects. The traditional water project bidding analysis is mainly based on experience, competent judgment and perceptual knowledge. As a result, the results may tend to be arbitrary and uncertain. In addition, the subjective operations of the tenderers and the bidders are the main factors in the bid evaluation and supervision process. Therefore, illicit competition problems, such as bid rigging and colluding, are easy to occur. With the rise and development of big data in recent years, there are new ideas and methods for the behavior analysis of bid rigging and colluding. At the same time, the advent of complex networks and community structures has provided new tools and techniques for solving complex problems in many fields. Using suitable community detection algorithm to detect the community structure of a complex network makes the behavior analysis of the bid rigging more practical and objectivity. With this method, the potential illegal bidding behaviors can be more quickly found out so as to prevent and remedy bid rigging and colluding. Above all, the analysis of water project bid rigging and colluding behavior based on complex network demonstrates an

important practical significance for the water project bidding market.

## II. LITERATURE REVIEW

### A. Literature Review of Community Detection Algorithms in Complex Networks

The complex network can be understood as a network with a high degree of complexity, which consists of a large number of nodes and intricate connections. But there is not a unified definition for complex networks. Tsien Hsueshen gives a descriptive definition: a network with some or all properties of self-organization, self-similarity, attractor, small world and scale-free is called a complex network.

The WS small-world network model (proposed by Watts and Strogatz in 1998) and BA scale-free network model (proposed by Barabási and Albert in 1999) raise the research trend in complex networks [1-2]. As an important structural feature of complex networks, the community structure has been paid the same attention in recent years. The study of community structure testing can be traced back to 1970 as a probing optimization algorithm proposed by BW Kernighan et al. They conducted a study on the segmentation problem of the graph and gave a heuristic optimization algorithm called Kernighan-Lin method, which is a typical greedy algorithm based on the dichotomy of the typical community [3]. Other representative community discovery algorithms are as follows: the GN algorithm was given in 2002 by Girvan and Newman, which is a typical splitting algorithm based on the edge intermediary [5]. Although the algorithm is simple in thought, it has a high complexity; In 2004, Radicchi et al. proposed an improved GN algorithm called edge clustering coefficient method, which is a new splitting algorithm that can improve the speed of computation with less precision drop [6]; based on the GN algorithm, Newman in 2004 also provided a new agglomeration algorithm called Newman fast algorithm. The algorithm carries out hierarchical clustering according to the increment of module degree. After the clustering is completed, the community structure is divided by finding a local maximum module value[7-8]; In 2009, Dress discovered that unstable social structures in complex networks can be adequately detected by relaxing the modularity function (Q function) of Newman's algorithm[9]; in 2012, Bennett L proposed a hybrid integer non-linear programming (MINLP) based on the Newman algorithm for the classification of weighted networks

and the detection of overlapping communities[11]; In 2013, Meo put forward an algorithm to enhance the discovery of existing communities by using network weighting strategy. The algorithm enhances the community discovery by adding the edge center weight to the network topology [10]; In 2014, Li put forward a vertex similarity probability (VSP) model, which can find the community structure when the type of complex network is unknown[12]; Ferreira presented a complex network community detection algorithm based on time series clustering in 2016[14]; In 2017, Guerrero proposed a complex network adaptive community detection algorithm GGA based on genetic algorithm, which is guided by the modular index [14]; Soundarajan et al. put forward a HICODE meta-method in 2017 to find hidden communities that could find existing hidden communities[15]; Kumar et al. developed a complex network community detection algorithm based on rough set, which constructed the rough set using the regional connectivity around the nodes[16]. In the study of bid rigging of water projects, Newman's fast algorithm was chosen to divide the complex network community structure because the community structure of the complex network composed of bidding enterprises was undiscovered.

### B. Literature Review of Bid Rigging

Water project construction is inseparable from the bidding. However, there are various irregularities in the bidding process. Among them, the illegal activities of bid, such as bid rigging and colluding, are difficult to be prevented, which are greatly harmful to the society. Bid rigging is a kind of collusive bidding behaviour. It is a series of behaviors that make profits through the group, which is formed by several bidders. After that, by limiting competition or other ways, some stakeholders in the group win the bid. At present, there are three main aspects of bid rigging research. Firstly, by analyzing the organizational behavior of the bidding participants, the bid rigging behavior in the bidding process can be discovered. Marshall, at Penn State University in the United States, discovered the bid rigging behavior by conducting a cartel analysis of bidding participants [17]. The second point is to find out the evidence of bid rigging by analyzing the bidding price. In this area, Ballesteros-Pérez et al. found unusual bid prices by using tools to predict the bidding price cap and floor, which can provide effective support for the discovery of bid rigging behavior [18]. The third point is the behavior analysis of the participants in the bidding process. Lengwiler conducted a model analysis of the manipulation of bidders by corrupt auctioneers and proposed major solutions to the bid rigging behaviors [19].

Bid rigging analyses of projects are as follows: Reeves-Latour proposed a core-edge model to analyze the bidding network in the construction industry. The actual case analysis showed the evolution of the bidding network with bid rigging behaviors [20]; Li Qian used the evolutionary theory in game theory to establish an asymmetric evolution model for the behaviors of the bid rigging; Compared with the research on bid rigging behavior of the all kinds of projects, there are few researches on bid rigging behavior for water projects. Therefore, the research in this paper has some theoretical and practical significance.

## III. BID RIGGING COMMUNITY DETECTION MODEL BASED ON COMPLEX NETWORK

### A. Water Projects Bid Rigging Causes

Bid rigging behaviors seriously do harm to the fairness and justice of the bidding market. How to punish and prevent those problems have attracted extensive attentions. The main reasons for water projects bid rigging are as following three points:

- Firstly, the regional impacts are great. As water projects are more complex, longer-term, investment-intensive and more stringent than other projects, each bidding company that locates in the region of the water project has great advantages. Compared with non-local enterprises, local enterprises are more aware of water projects in the areas where they operate and can better meet the requirements of bidding and construction. Therefore, local companies are already dominated the market. When the water projects need bidding, a certain number of bidding companies are still needed for bid according to the regulations, which makes some enterprises bid rigging each other in order to win the bidding in their own area.
- Secondly, there are some difficulties in identifying and analyzing of bid rigging, and the punishments are not enough, therefore, it makes bid rigging unbridled. Although the government has promulgated the relevant laws to control the making of the bid rigging, due to the actual situation of the water projects, it is difficult to implement the relevant laws.
- Thirdly, the emergence of many bid rigging is often financially compensated by bidders who organize bid rigging, in the meanwhile other bidders often do not tip-off these illegal enterprises. It is very difficult for government officials to find relevant clues, which increases the difficulties of the investigation. Although the law, such as Bidding Law, stipulates the handling of bid rigging, the relevant provisions of law are not perfect, which affects the cognizance and handling of bid rigging behavior.

### B. Bid Rigging Community Detection Model Based on Complex Network

#### 1) Design complex networks based on the relationship among bidding enterprises

The specific definition of the graph is described as follows: a graph is an ordered pair  $G = (V, E)$  comprising a set  $V = \{v_i\}$  of vertices or nodes or points together with a set  $E = \{e_i\}$  of edges or arcs or lines, which are 2-element subsets of  $V$  and  $E$ . The order of a graph is  $v_i$ , its number of vertices. The size of a graph is  $e_i$ , its number of edges. According to the relationship between nodes and edges in the graph, the graph can be transformed into matrix form. Adjacency matrix is a matrix that can be used to represent the relationship between nodes in the graph. The definition of adjacency matrix in undirected weighted graph is as follows :  $G = (V, E)$  is an undirected graph,  $V(G) = \{v_1, v_2, \dots, v_n\}$  is the node in the graph,  $E(G) = \{e_1, e_2, \dots, e_m\}$  is the edge in the graph, and the square matrix  $A = (a_{ij})_{n \times n}$  is the adjacency matrix. Let

$$a_{ij} = \begin{cases} w_{ij}, v_i \text{ and } v_j \text{ are adjacent} \\ 0, v_i \text{ and } v_j \text{ are not adjacent} \end{cases} \quad (1)$$

Among which,  $w_{ij}$  is the weight of the edge connecting node  $v_i$  and  $v_j$ .

Due to the complexity of the relationship between the bidding companies and the bid rigging behaviors, it is more difficult to identify bid rigging behaviors among enterprises. The concept of big data and its continuous development provide new ideas and methods for solving the problems of bid rigging. The application of complex networks can well illustrate complex models in the fields of social sciences, management sciences and engineering techniques. Because of the complexity of the bid rigging, we set up a complex network based on the complex relationships between bidding enterprises. Assuming that the total number of firms participating in the bidding project is  $n$  (the same firm only counts once), the adjacency matrix  $A=(a_{ij})_{n \times n}$  of the bidding firm can be defined as follow:

$$a_{ij} = \begin{cases} w, v_i \text{ and } v_j \text{ participated in } w \text{ bidding projects together} \\ 0, v_i \text{ and } v_j \text{ did not participate in same bidding project} \end{cases} \quad (2)$$

From the adjacency matrix, it can be seen how frequently companies participate in same bidding projects. Converting the adjacency matrix to a complex network graph could show the relationship between the enterprises more intuitively. When drawing a network diagram, each enterprise participating in the bidding is regarded as a node. If two enterprises jointly participate in the same bidding project, an edge can be connected between the two enterprises, and the weight of the edge is considered as 1. If two firms participate in  $w$  projects together, the weight between two firms is  $w$ . According to the network diagram, we can get the total number of edges and the number of edges connecting to each node, that is the degree of nodes.

## 2) Detection of community structures in complex networks using Newman fast algorithm

Through the analysis of the structures of the complex networks formed by enterprises participating in the bidding, it is possible to detect whether there is a corporation in the bidding process. If there is obvious community structure in the complex network, bid rigging behavior among enterprises can be considered. In this paper, the Newman fast algorithm for complex networks is used to find and determine community structures.

Newman fast algorithm is a classic community detection algorithm, and the Newman fast algorithm can effectively partition the community structure of complex networks. Newman fast algorithm is a clustering analysis method based on greedy algorithm. The specific steps of Newman fast algorithm are as follows:

- Firstly, initializing the network with  $n$  nodes for  $n$  communities, that is to say each node is an independent community. Initialize the module degree function  $Q = 0$ , the initialized  $e_{ij}$  and  $a_i$  are satisfied with the following conditions.

$$e_{ij} = \begin{cases} 1/(2m), \text{ if there is a edge between node } i \text{ and } j \\ 0, \text{ other} \end{cases} \quad (3)$$

$$a_j = k_j / (2m) \quad (4)$$

Among them,  $k_j$  is the degree of node  $i$ , and  $m$  is the total number of edges in the network.

- Secondly, if there is a connecting edge between the two communities, the two communities are merged into a single one. Then calculates the combined module degree increments.

$$\Delta Q = e_{ij} + e_{ij} - 2a_i a_j = 2(e_{ij} - a_i a_j) \quad (5)$$

According to the principle of greedy algorithm, each merge should be in the direction of maximizing or decreasing  $Q$ -measure.

- Thirdly, repeat the second step. Continue to merge the communities until the entire networks are merged into a community so far. This process should be performed at most  $n-1$  merges.

### 3) Use modularity function to evaluate the community structure partition

Generally speaking, the nodes within the communities are more closely related and the connections among the different communities are sparse. For the evaluation criterion of the complex network community structure, the Newman and Girvan modularity function is currently recognized by people. It is called  $Q$ -measure that guides the search and helps to evaluate the fit of community structures in networks. Random networks do not have community structure. Modularity can be understood as the difference between a complex network and a random network in a certain community structure. The greater the difference is, and the greater the corresponding  $Q$ -measure is. The larger the  $Q$ -measure is, and the better the result of partition of community structure is.

The Newman fast algorithm was used to get the hierarchical tree structure chart of complex network community structure. Disconnecting at different locations in the tree chart can result in different network community structures. Each community structure corresponds to a  $Q$ -measure of a local modularity function. The community structure with the largest local  $Q$ -measure is the best network community structure. The range of  $Q$ -measure is between 0 and 1. A large number of studies found that  $Q$ -measures in the actual network are generally between 0.3 and 0.7. If the  $Q$ -measure is greater than 0.3, we could think that the community structure of the network is obvious. If the bidding enterprise's complex network has an obvious community structure, it can be considered that there is an obvious cooperative behavior among the enterprises. Therefore, it is possible to initially determine that there is a bid rigging suspicion among the enterprises in the community. The relevant departments can find out the existence of bid rigging behavior through further investigation and analysis of suspected enterprises.

IV. EMPIRICAL ANALYSIS

A. Set up a complex network

The data of this paper came from Tianjin Water Conservancy Transaction Management Center. Firstly, the collected data were reorganized, some unrepresentative data were excluded and the finally 23 companies were involved in the study of the community testing model. By analyzing the bidding situation of 23 enterprises, we got the adjacency matrix (23 order matrix) of complex network. The data of the adjacency matrix is shown in FIGURE I .

0	5	3	4	4	0	2	0	0	2	2	1	2	1	1	0	1	1	0	0	0	0	1
5	0	1	3	1	0	1	1	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0
3	1	0	1	3	5	3	1	5	0	0	1	1	1	0	0	0	0	0	0	0	0	1
4	3	1	0	0	0	0	0	0	2	1	0	0	0	0	0	0	0	0	0	0	0	0
4	1	3	0	0	1	5	1	2	0	1	1	1	1	0	0	0	0	0	0	0	0	0
0	0	5	0	1	0	4	2	9	0	0	0	0	0	0	1	0	0	1	1	0	0	0
2	1	3	0	5	4	0	2	5	0	0	1	1	1	0	1	0	0	0	0	0	0	0
0	1	1	0	1	2	2	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	5	0	2	9	5	2	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
2	0	0	2	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	1	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
1	0	1	0	1	0	1	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0
2	1	0	1	0	1	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0
1	0	1	0	1	0	1	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0
1	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	1	0
0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0
1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0

FIGURE I. ADJACENCY MATRIX OF BIDDING ENTERPRISE

From the adjacency matrix, it can be seen how often companies participated in same bidding projects. If the value is small, such as 0 or 1, it shows that the behaviors of bidding between enterprises are relatively normal; On the contrary, the larger value means that more companies participated in the same bidding projects more frequently. In the process of bid rigging behavior analysis, particular attention should be paid to these parts of enterprises.

As the adjacency matrix of complex network is determined, the adjacency matrix can be visualized by using Ucient software, which can express the relationship between the bidding enterprises more visible. The initial complex network diagram is shown in FIGURE II, FIGURE II contains a total of 23 nodes and 65 edges.

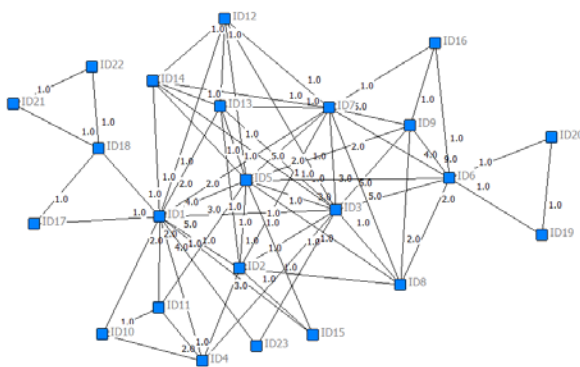


FIGURE II. INITIAL COMPLEX NETWORK DIAGRAM

B. Community Detection

Newman fast algorithm can be implemented by MATLAB analysis software. The tree structure diagram of the results of community division is shown in FIGURE III.

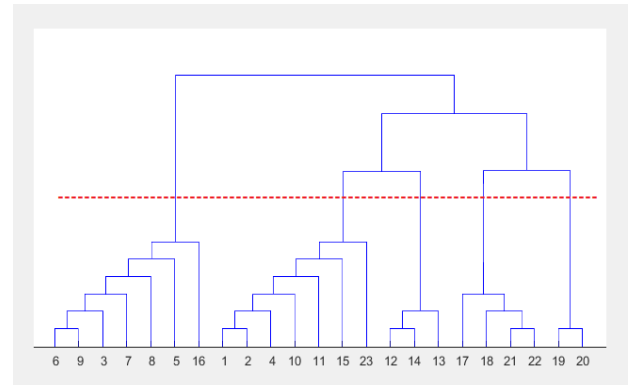


FIGURE III. TREE STRUCTURE DIAGRAM OF COMMUNITY DIVISION RESULTS

In FIGURE III, the complex network can be divided into five communities with the red dotted line. In this case, the Q-measure of the module degree function is 0.3732 which is greater than 0.3. Therefore, it can be considered that the community structure of the network is reasonable. This result showed that there was a clear community structure among these companies. The results of community division are shown in TABLE I, and the corresponding community structure is shown in FIGURE IV.

TABLE I. THE RESULTS OF THE COMMUNITY DIVISION

Community ID	Company ID
S1	ID3, ID5, ID6, ID7, ID8, ID9, ID16
S2	ID1, ID2, ID4, ID10, ID11, ID15, ID23
S3	ID19, ID20
S4	ID17, ID18, ID21, ID22
S5	ID12, ID13, ID14

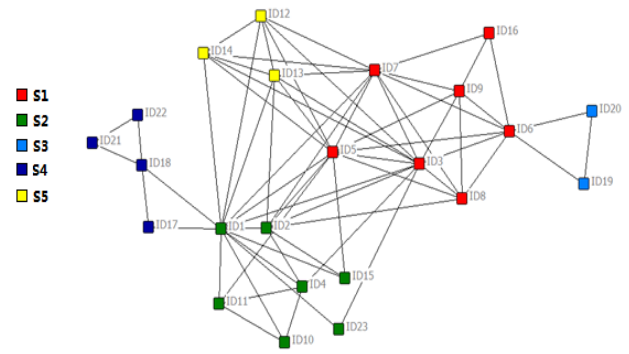


FIGURE IV. COMMUNITY STRUCTURE DIAGRAM

It can be seen from FIGURE IV that in the community S1 and the community S2, the nodes are closely connected and the weights of the sides are relatively large, which indicates that the enterprises in the community S1 and the community S2 have obvious cooperative behaviors. These companies are likely to bid rigging. And the other three communities contain fewer



nodes, and the weight values of the side are small, hence it can basically excluded the cooperative behavior of enterprises.

### C. Verify the validity of the community detection model

We selected a part of those enterprises which have obvious cooperative behavior and analyze their bid winning rate. We defined the bid winning rate with cooperative behavior and bid winning rate of non-cooperative behavior as follows:

$$C_1 = a_1/b \quad (6)$$

$$C_2 = a_2/b \quad (7)$$

Among them,  $C_1$  is the bid winning rate with cooperative behavior;  $C_2$  is the bid winning rate of non-cooperative behavior;  $a_1$  is the bid winning times with cooperative behavior;  $a_2$  is the bid winning times of non-cooperative behavior;  $b$  is the total number of participating bids.

After the statistical calculation of the winning rate of 14 enterprises in community S1 and community S2, it is found that ID6, ID9, ID3 and ID1 and ID2 in community S1 are more representative. The bid winning rates of these five companies are shown in TABLE II

TABLE II. THE BID WINNING RATE OF SOME ENTERPRISES WITH SIGNIFICANT COOPERATIVE BEHAVIOR

Company ID	ID1	ID2	ID3	ID6	ID9
<b>bid winning rate with cooperative behavior</b>	42.86 %	28.57 %	35.71 %	18.19 %	13.33 %
<b>bid winning rate of non-cooperative behavior</b>	35.71 %	14.29 %	28.57 %	10.23 %	8.69%

It can be seen from TABLE II, when the enterprise are cooperated to participate in the bidding, the winning rates are significantly improved, which illustrates the rationality of the community model. According to the analysis of the above results, the community detection model can be applied to detect the bid rigging behavior in the water project bidding.

## V. CONCLUSION

In this paper, the community structure model based on complex networks is applied to bid rigging behavior analysis. The results demonstrate that the Newman fast algorithm can reasonably divide the community structure of complex networks. Through the community detection of the complex network of the enterprises involved in bidding, we can find the enterprises with obvious cooperate behaviors. This analysis provides a certain reference for the relevant departments to identify whether there is bid rigging behavior in enterprises. However, due to the limited amount of data, there still are some shortcomings. We hope more data to test the model can be collected in the future.

## ACKNOWLEDGMENT

This paper was funded by Tianjin philosophy and social science research project (Project No. TJGL16-020).

Then, I would like to express my heartfelt gratitude to Tianjin Water Conservancy Transaction Management Center for the raw data they provided.

## REFERENCES

- [1] Watts D J, Strogatz S H. Collective dynamics of 'small-world' networks [J]. *Nature*, 1998, 393 (6684): 440-442.
- [2] Barabási A-L, Albert R. Emergence of scaling in random networks [J]. *Science*, 1999, 286 (5439): 509-512.
- [3] B. W. Kernighan, S. Lin. An efficient heuristic procedure for partitioning graphs [J]. *Bell System Technical Journal*, 49(2):291-307, 1970.
- [4] G. Palla, I. Derenyi, I. Farkas, et al. Uncovering the overlapping community structure of complex networks in nature and society. *Nature*, 2005, 435(7043):814-818.
- [5] M. Girvan, M. E. J. Newman. Community structure in social and biological networks. *Proceedings of the National Academy of Sciences*, 2002,99(12):7821-7826.
- [6] F. Radicchi, C. Castellano, F. Cecconi, et al. Defining and identifying communities in networks. *PNAS*, 2004,101(9):2658-2663.
- [7] [10] M. E. J. Newman. Fast algorithm for detecting community structure in networks [J]. *Phys. Rev. E*, 2004, 69:066133
- [8] M. E. J. Newman. Modularity and communities structure in networks [J]. *Proc Natl. Acad. Sci*, 2006, 103(23): 8577-8582
- [9] .Jiang J Q, Dress A, Yang G. A spectral clustering-based framework for detecting community structures in complex networks[J]. *Applied Mathematics Letters*, 2009, 22(9):1479-1482.
- [10] Meo P D, Ferrara E, Fiumara G, et al. Enhancing community detection using a network weighting strategy[J]. *Information Sciences An International Journal*, 2013, 222(3):648-668.
- [11] Bennett L, Liu S, Papageorgiou L G, et al. A Mathematical Programming Approach to Community Structure Detection in Complex Networks[J]. *Computer Aided Chemical Engineering*, 2012, 30(4):1387-1391.
- [12] Li K, Pang Y. A unified community detection algorithm in complex network[J]. *Neurocomputing*, 2014, 130(3):36-43.
- [13] Ferreira L N, Zhao L. Time series clustering via community detection in networks[J]. *Information Sciences*, 2015, 326(C):227-242.
- [14] Guerrero M, Montoya F G, Baños R, et al. Adaptive community detection in complex networks using genetic algorithms[J]. *Neurocomputing*, 2017.
- [15] He K, Li Y, Soundarajan S, et al. Hidden Community Detection in Social Networks[J]. 2017.
- [16] Kumar P, Gupta S, Bhasker B. An upper approximation based community detection algorithm for complex networks[M]. Elsevier Science Publishers B. V. 2017.
- [17] Marshall R C, Marx L M. Bidder collusion[J]. *Journal of Economic Theory*, 2007, 133(1):374-402.
- [18] Ballesteros-Pérez, Pablo, González-Cruz, et al. Detecting abnormal and collusive bids in capped tendering[J]. *Automation in Construction*, 2013, 31(5):215-229.
- [19] Lengwiler Y, Wolfstetter E. Auctions and corruption: An analysis of bid rigging by a corrupt auctioneer[J]. *Journal of Economic Dynamics & Control*, 2010, 34(10):1872-1892.
- [20] Reeves-Latour M, Morselli C. Bid-rigging networks and state-corporate crime in the construction industry[J]. *Social Networks*, 2016.