Knowledge-Based Systems 193 (2020) 105464

Contents lists available at ScienceDirect

Knowledge-Based Systems

journal homepage: www.elsevier.com/locate/knosys

GMM: A generalized mechanics model for identifying the importance of nodes in complex networks[☆]

Fan Liu [a,b,c], Zhen Wang [d], Yong Deng [a,*]

[a] Institute of Fundamental, University of Electronic Science and Technology of China, Chengdu, 610054, China
[b] Yingcai Honors of school , University of Electronic Science and Technology of China, Chengdu, 610054, China
[c] School of Mathematical Sciences , University of Electronic Science and Technology of China, Chengdu, 610054, China
[d] Center for Optical Imagery Analysis and Learning, School of Mechanical Engineering, Northwestern Polytechnical University, Xi'an 710072, Shanxi, China

ARTICLE INFO

Article history:
Received 22 August 2019
Received in revised form 26 December 2019
Accepted 29 December 2019
Available online 10 January 2020

Keywords:
Influential nodes
Complex networks
Gravity model

ABSTRACT

How to assess the importance of nodes in the network is an open question. There are many ways to identify the importance of nodes in complex networks. However, these methods have their own shortcomings and advantages. In particular, some methods based on the importance of nodes between interactions between nodes have been proposed. These methods utilize local information or path information. How to combine local and global information is still a problem. In this paper, a generalized mechanical model is proposed that uses global information and local information. To verify the effectiveness of the method, some experiments were performed on a total of ten real networks. In particular, an innovative experimental network-based quality assessment was proposed to validate the method of identifying the importance of nodes.

1. Introduction

The real world is very complex with uncertainty. Complex networks theory plays a key role in many fields [1], such as physics [2–4], finance [5], game theory [6], disease propagation [7], society [8,9], time series [10,11], network design [12], adaptive control [13,14], uncertainty [15], bionic algorithm [16,17] and so on [18–21]. For example, fractals are common in the real networks. How to qualitatively quantify fractal models is an important issue [22]. Network model has been widely used in knowledge-based system or other intelligent systems. For example, deep learning of neural networks is a hot topic. To deal with some problems in knowledge-based systems, complex networks model is more and more used in recent years due to its efficiency to model different components or factors in complex system. Identification of the vital nodes in the networks is a very important topic. As a result, ranking the influential nodes in knowledge based system is paid great attention. In particular, how to evaluate the importance of nodes in complex networks is very meaningful [23]. The key nodes are more influential in the structure and function of the network than other nodes. The identification of key nodes can be applied in many fields [24], such as disease [25], resource allocation [26], society [27,28], social networking services [29–31], biological information [32] and so on [33]. For instance, important nodes play a key role in communities [34]. Different communities are overlapping in some real cases [35]. Li et al. [36] proposed a novel method to detect the overlapping communities by seed community in weighted-networks. In addition, important nodes have a significant impact on the network robustness [37]. Node activity also can influence the immunization, in which most disease propagation models can be established by utilizing different initial nodes [38]. Different nodes also have different influences in information dissemination. Wang et al. [39] proposed a new information dissemination model .

There are many methods to assess the importance of nodes in complex networks including degree centrality (DC) [40], closeness centrality (CC) [41], betweenness centrality (BC) [42], PageRank (PC), H-index [43], K-order [44] so on [45]. DC is a simple method of measuring the centrality by calculating the number of connections between the surrounding nodes and the node [40]. For DC [40], the greater the degree of the node, the more important it is. The CC calculates the reciprocal of the sum of the shortest path of the node and other nodes as the centrality value, and the closer the node is to the center, the closer it is to other nodes [41]. CC is sensitive to the structure of the network, and even small fluctuations can change the ordering of

☆ No author associated with this paper has disclosed any potential or pertinent conflicts which may be perceived to have impending conflict with this work. For full disclosure statements refer to https://doi.org/10.1016/j.knosys.2019.105464.

* Corresponding author.
E-mail address: dengentropy@uestc.edu.cn (Y. Deng).

https://doi.org/10.1016/j.knosys.2019.105464
0950-7051/

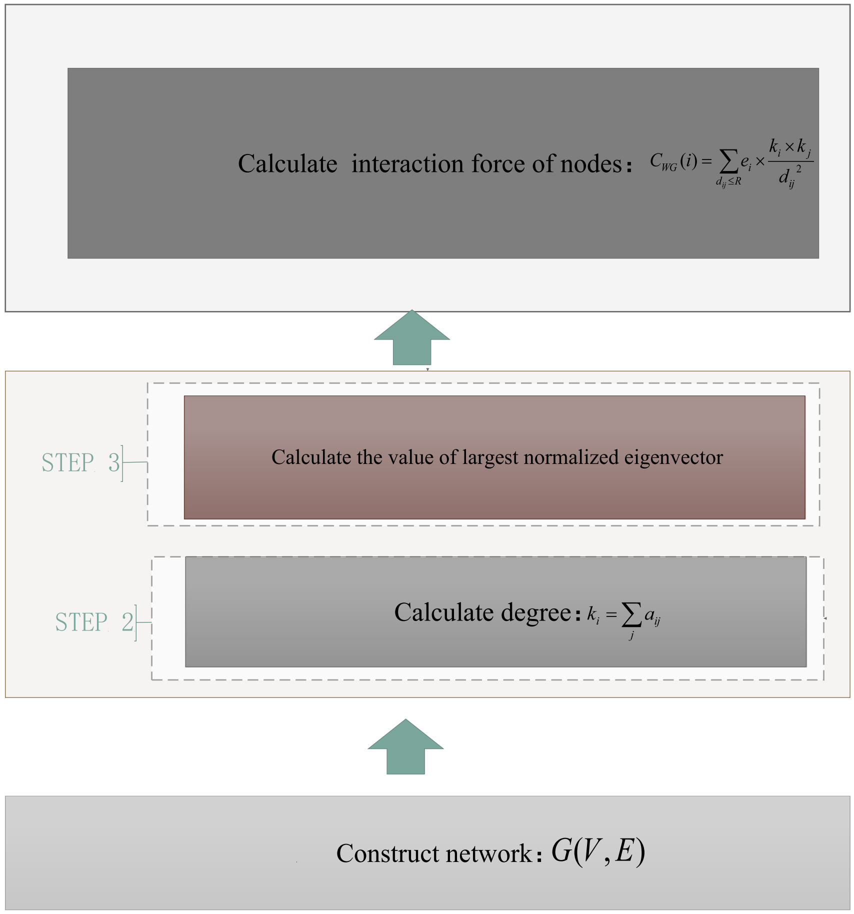

Fig. 1. The proposed method is divided into four steps. The first step is to build a network. The second step calculates the node degree of each node as the quality of each node interaction. The third step assigns each node a first weight. The fourth step calculates the centrality scores of each node interaction.

nodes [41]. BC calculates the extent to which a node is located at other nodes' path [42]. CC [41] and BC [42] use the degree of concentration of the path to evaluate the importance of the node and utilize the global information. The PC [46] uses the information about the neighbors to iterate to assess the influence of the node, and it is stable on a scale-free network, but very sensitive to random networks. In addition, evidence theory as a tool for information fusion can be applied in many areas, such as decision making, uncertainty, data fusion, and complex networks. The method of node importance evaluation based on evidence theory has also been proposed. Based on the information fusion algorithm with evidence theory [47–50] and fuzzy theory [51], Mo and Deng [52] proposed an evidence-based approach to assess the influence of nodes on network, which combines methods of central measurement such as DC [40] and BC [42].

Recently, some methods based on interaction between nodes have been proposed, such as the inverse-square law model [53], Gravity model [54]. The inverse-square law model considers the interaction between nodes to assess the importance of nodes. The Gravity models [54] uses neighbors and path information to assess the influence of the node. Fei et al. [53] proposed a method for measuring the influence of a node based on the inverse square law, which means that if the degree of a node is larger and the closer to other nodes, the influence of the node is greater. However, in some applications it is necessary to consider the global influence of the nodes. For example, in disease transmission models, it is necessary to study the influence of different objects in the network, and the strategy is also a key factor for dynamic networks [55,56]. Important nodes also play a key role in information dissemination, and Gao et al. [57] propose a non-Markov information propagation model. Adding edges to the network is also an important way to optimize information propagation model [58]. In addition, some Gravity models based on node interactions are time-consuming for large networks [53, 54]. Li et al. [54] solved this problem by introducing a truncation radius. Moreover, how to better use local and global information to assess the influence of nodes remains a problem. For example, Wen and Deng [59] proposed a method based on local dimension to identify influential nodes in the network. To better solve the above problems, it is necessary to propose a more generalized model. Inspired by the laws of gravity, the generalized mechanics model (GMM) for identifying the importance of nodes in complex networks was proposed, which takes into account local and global information. If the degree of one node is larger (local information), the shorter distance and the greater weight (global information) with other nodes, the influence of the node is greater. In addition, the GMM inherits the truncation radius of the gravity model proposed by Li et al. [54] to deal with the time-consuming problem of large-scale networks.

The contributions of this paper are summarized as follows.

- A generalized mechanical model used to evaluate the influence of the node was proposed. Based on this model, a weighted gravity model (WGravity) is applied to the evaluation of node importance. The WGravity model combines global and local information to assess the impact of a node.
- A novel method for verifying the importance assessment of nodes has been proposed. Based on the connectivity of complex networks, complex network quality has been proposed. Then, by attacking the nodes in the network, the quality of the final complex network of the network is tested.

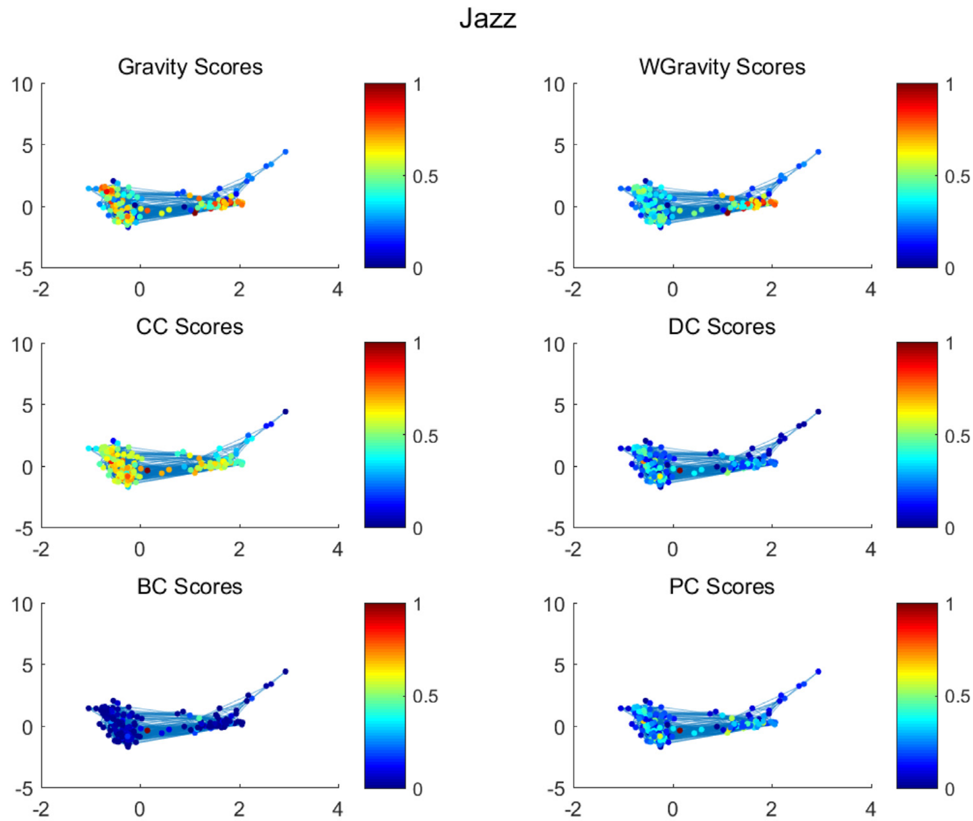

Fig. 2. This figure describes the central scores of the six methods in Jazz. Gravity and WGravity, DC performance is basically the same, the value of each node in CC is high.

The structure of this paper is as follows. Section 2 presents some methods for nodes influence assessment. The generalized mechanics model and its application are introduced in Section 3. Section 4 analyzes the effectiveness of the proposed method. Section 5 summarizes the work of this paper.

2. Preliminaries

Given a graph $G = (V, E)$, where V represents the nodes, E represents the edges, degree centrality [40] (DC) , closeness centrality [41] (CC) , betweenness centrality [42] (BC), PageRank [46] (PC), Gravity model [60], and eigenvector centrality [61] (EC) will be introduced in this section.

2.1. Degree centrality (DC)

DC is a simple method of measuring the centrality by calculating the number of connections between the surrounding nodes and the node [40]. For DC [40], the greater the degree of the node, the more important it is. DC measure is defined as follows [40].

$$DC(i) = k_i = \sum_j a_{ij} \tag{1}$$

where the k_i is the degree of the node i, a_{ji} is the connection between node j and i. The $DC(i)$ is centrality score of node i.

2.2. Closeness centrality (CC)

The CC calculates the reciprocal of the sum of the shortest path of the node and other nodes as the centrality value. CC measure is defined as follows [41].

$$CC(j) = \sum_i (\frac{a_{ji}}{|V| - 1})^2 (\frac{1}{d_{ji}}) \tag{2}$$

where a_{ji} represents the connection between node j and i, d_{ji} represents the shortest distance from node j to node i. The $CC(j)$ is the centrality score of node j.

2.3. Betweenness centrality (BC)

BC [42] calculates the degree of concentration of the path to evaluate the importance of the nodes. BC measure is defined as follows [42].

$$BC(j) = \sum_{i,k \neq j} (\frac{n_{ik}(j)}{N_{ik}}) \tag{3}$$

where d_{ik} is the number of paths from node i to k through node j and N_{ik} represents the total number paths through node j. The $BC(j)$ is the centrality score of node j

2.4. PageRank (PC)

The PC [46] calculates the information of the neighbors to iterate to assess the influence of the node. PC measure is defined as follows [46].

$$PC(j)^t = \sum_{i=1}^{|V|} (a_{ji} \frac{PC(i)^{t-1}}{k_i}) \tag{4}$$

where k_j is the out degree of node j and a_{ji} represents the connection between node j and node i, and $PC(j)^t$ represents the importance of node j at t steps. The $PC(j)$ is the centrality score of node j.

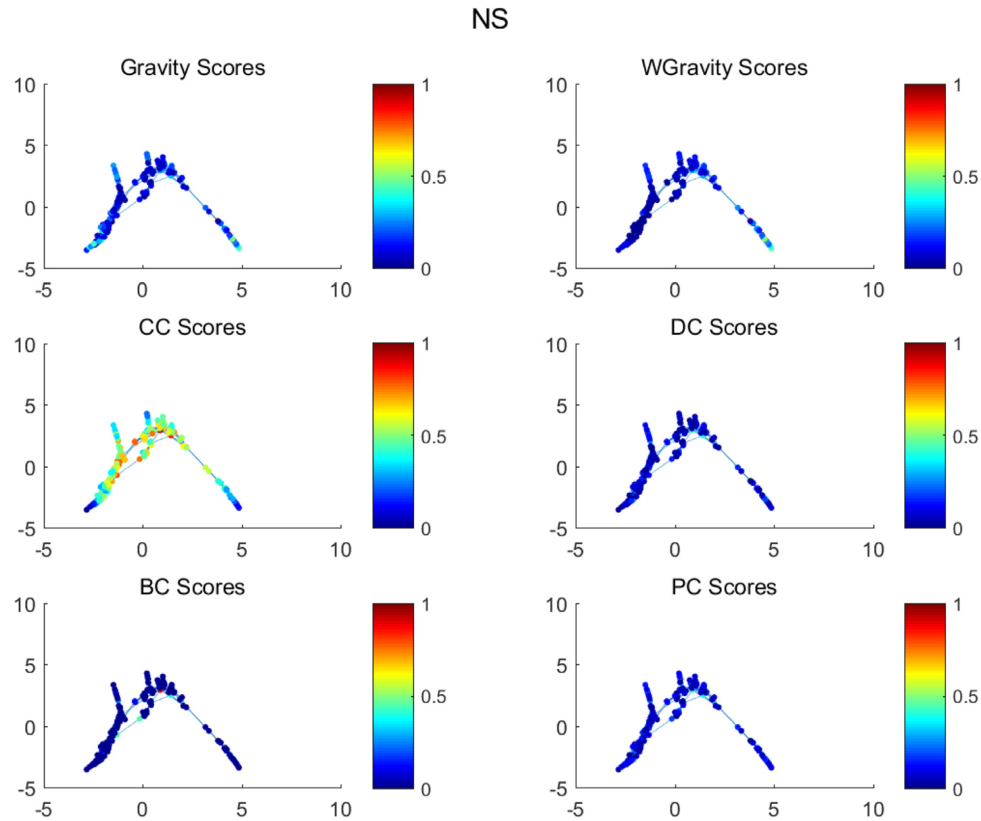

Fig. 3. This figure describes the central scores of the six methods in NS. PC, BC, WGravity, and Gravity are difficult to distinguish the differences in node importance, but CC gives a significant difference globally.

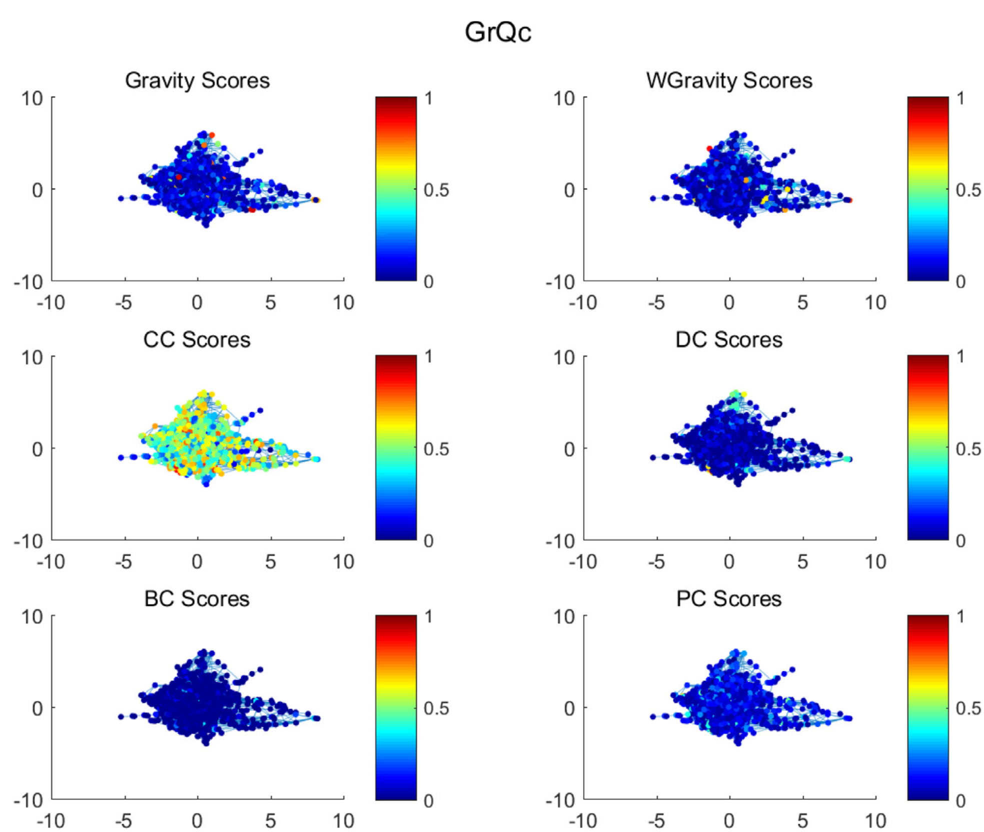

Fig. 4. This figure describes the central scores of the six methods in GrQc. There is a slight difference between WGravity and Gravity, but the order of importance of each node globally is not much different.

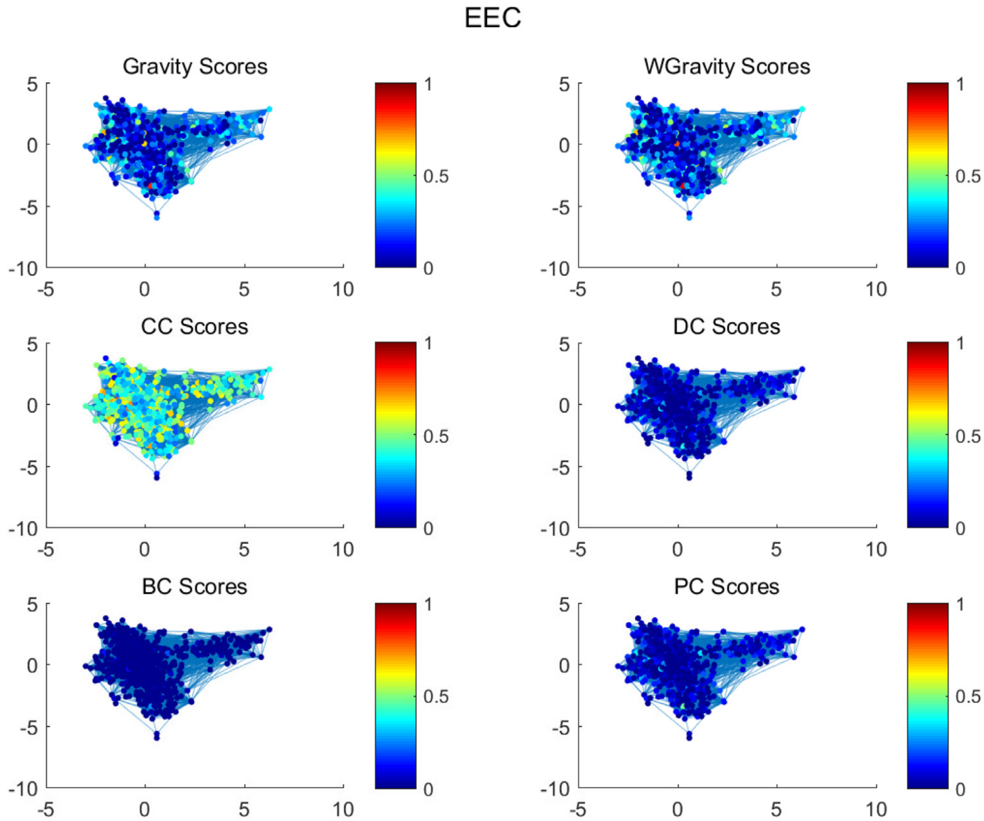

Fig. 5. This figure describes the central scores of the six methods in EEC. BC, PC, and DC are difficult to distinguish the differences of each node, but WGravity and Gravity can better judge the importance of the nodes.

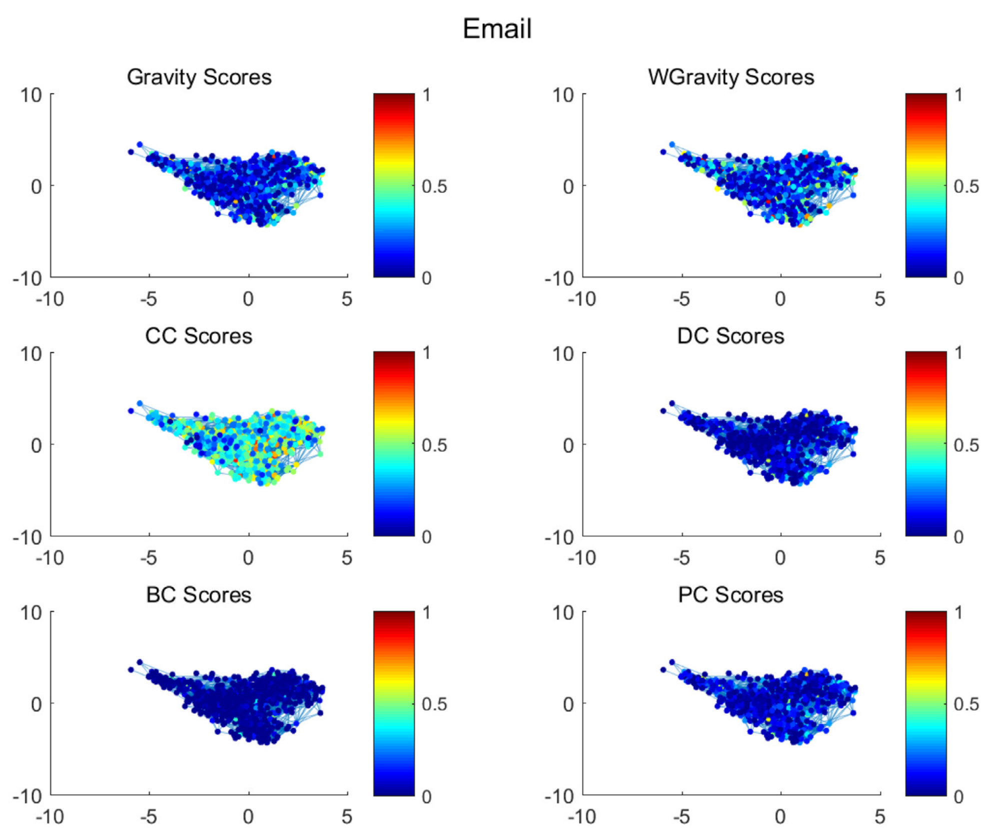

Fig. 6. This figure describes the central scores of the six methods in Email. CC assigns a higher value to the importance of nodes, but can distinguish the differences of each node. However, BC and DC are difficult to distinguish the difference between nodes.

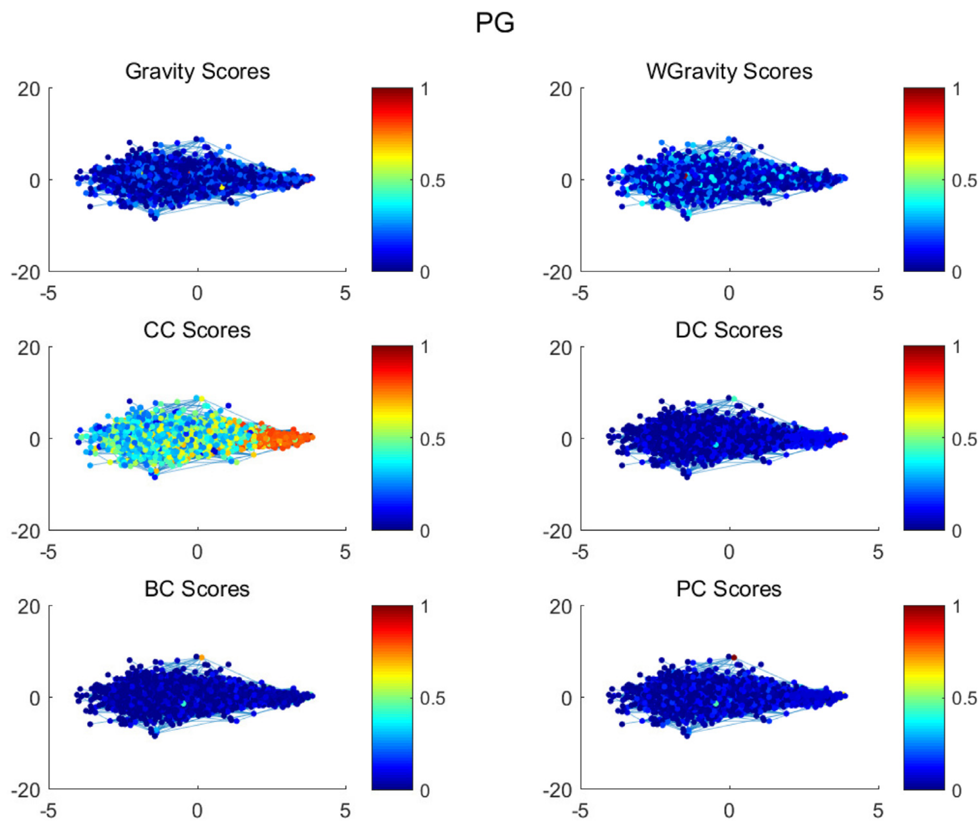

Fig. 7. This figure describes the central scores of the six methods in PG. CC has an advantage in judging the importance of global nodes. Gravity and WGravity are more consistent, especially the distribution of node importance.

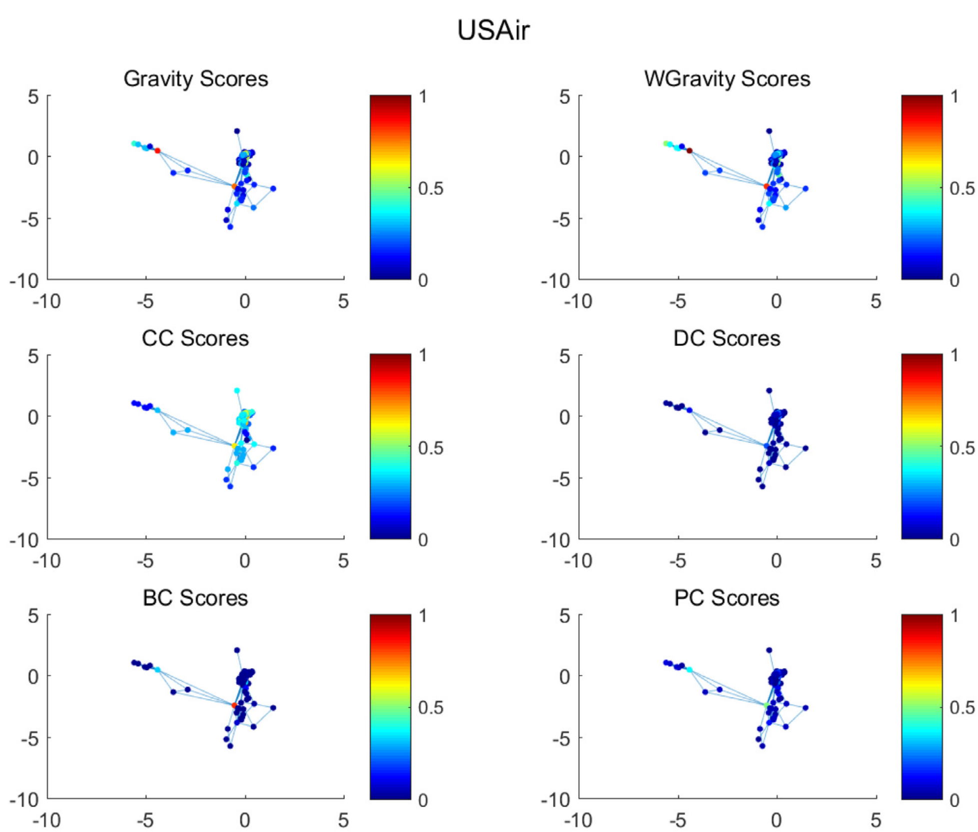

Fig. 8. This figure describes the central scores of the six methods in USAir. WGraviy is slightly different from Gravity.

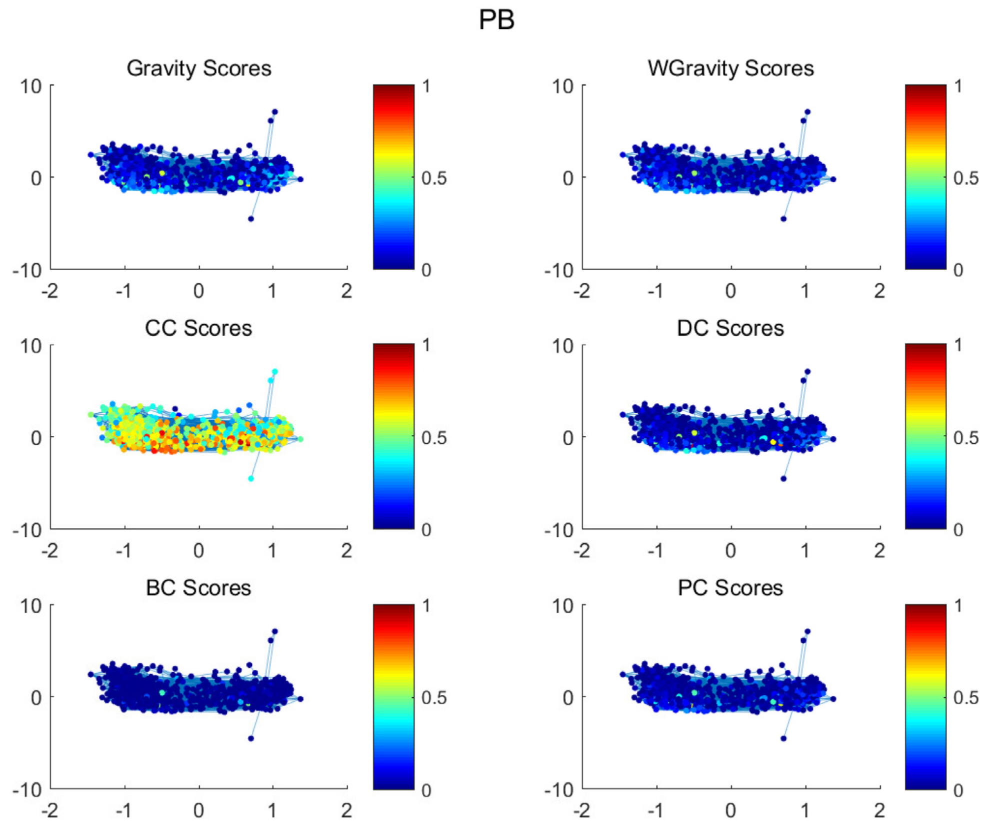

Fig. 9. This figure describes the central scores of the six methods in PB. CC can clearly see the difference in global importance of nodes.

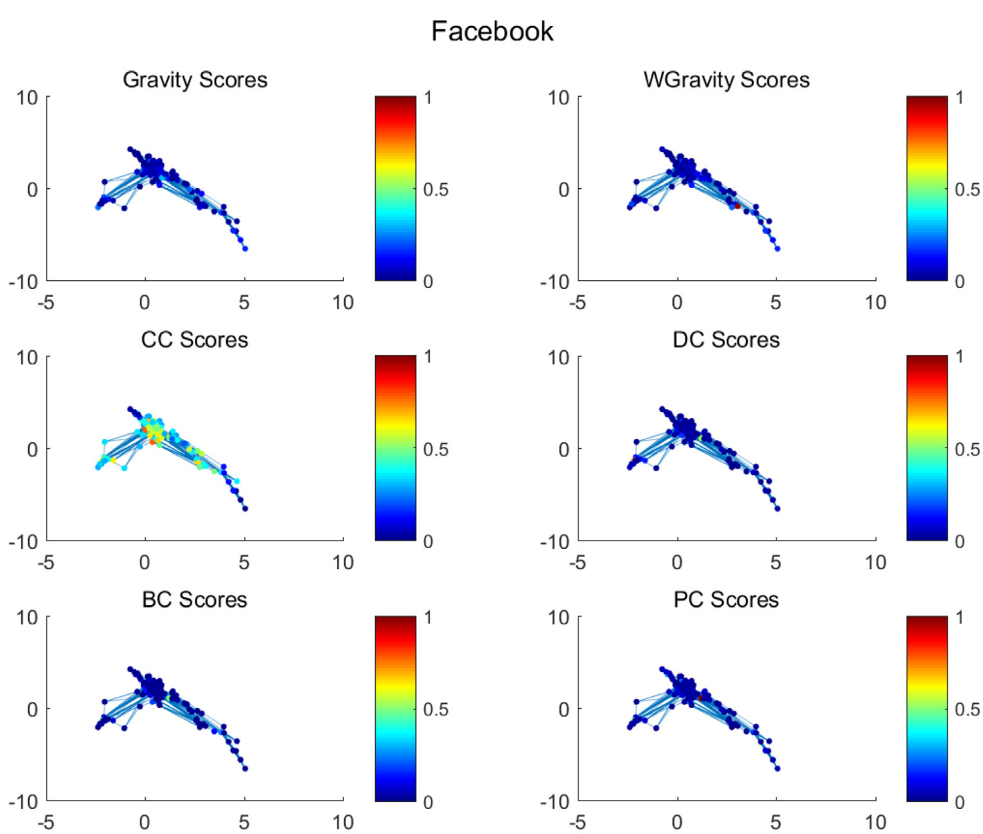

Fig. 10. This figure describes the central scores of the six methods in Facebook. The darker the node, the greater the importance of the node.

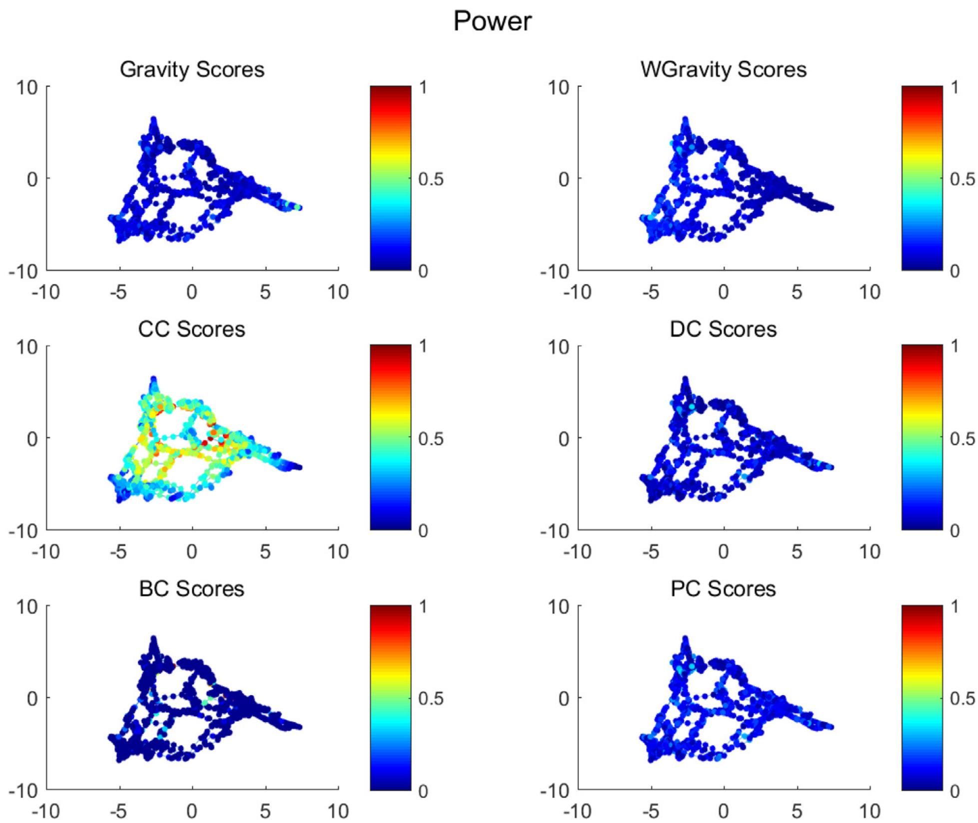

Fig. 11. This figure describes the central scores of the six methods in Power. The darker the node, the greater the importance of the node.

2.5. Eigenvector centrality (EC)

The EC assigns relative scores to all nodes in a network [61]. The EC is a complex central measurement method. The connection of a node has few nodes. If its EC value is high, it is also important [62]. Given a $n \times n$ matrix A, x_j is the value of the jth entry of the normalized largest eigenvector. Eigenvector centrality (EC) measure is defined as follows [61].

$$EC(j) = \frac{1}{\lambda}\sum_{i=1}^{|V|}(a_{ji}x_i) \tag{5}$$

where λ is the largest eigenvalue of the A. The $EC(j)$ is the centrality score of node j.

2.6. Gravity centrality

Gravity centrality measure is defined as follows [60].

$$CG(j) = \sum_{i \neq j}\frac{k_j k_i}{d_{ji}^2} \tag{6}$$

where k_i is the degree of the node i and d_{ji} is the shortest distance between node j and node i. The $CG(j)$ is the centrality score of node j.

3. Proposed method

Complex networks come from the real world, such as transportation networks, social networks. Based on the complex network also follows the laws of the real world, some methods based on the mechanical model to measure the importance of nodes have been proposed. However, how to make better use of global and local information and lower time complexity is still a problem to be solved. Inspired by these methods, a weighted gravity (WGravity) model was proposed in this paper, which combines both global and local information with low time complexity compared to Gravity model.

STEP 1: Construct network.

Given a real network, the graph $G = (V, E)$ is constructed to represent the network, where V represents the nodes and E represents the edges.

STEP 2: Calculate the degree of nodes.

In the mechanical model, the interaction of two objects is proportional to the mass and inversely proportional to the square of the distance. In the proposed method, the degree of the node is compared to the mass of the object in the real world.

$$k_i = \sum_j a_{ij} \tag{7}$$

where k_i is the degree of node i and a_{ij} is the connection between node i and node j.

STEP 3: Calculate the value of the largest normalized eigenvector.

In the real world, the influence of each object is different. If the influence of the object is greater, the force it interacts with other objects is also greater. How to measure the influence of each object has also become a key issue. Inspired by eigenvector, the value of the largest eigenvector is used as the weight of each node.

$$AX = \lambda X \quad e_i = X_i \tag{8}$$

where λ and X is the largest eigenvalue and normalized eigenvector, and e_i is the ith value of X.

In previous studies [53,54], the Gravity model did not take into account the weight of each node, meaning that the ability

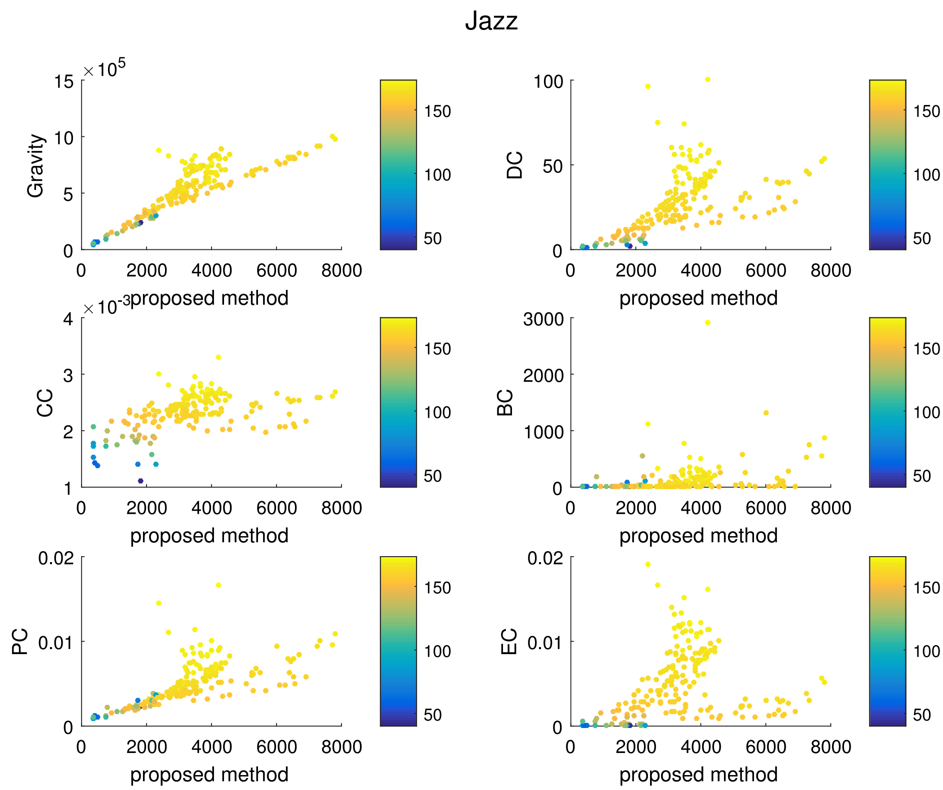

Fig. 12. The figure depicts the relationship between the proposed method and other methods in Jazz. Both the abscissa and the ordinate are the centrality scores of the nodes. The proposed method shows a good linear correlation with PC and Gravity. The darker the color, the greater the influence of the node in the SI model. If the slope trend of the scatter plot indicates the correlation between the proposed method and the method of comparison. If the slope is positive, it indicates a positive correlation.

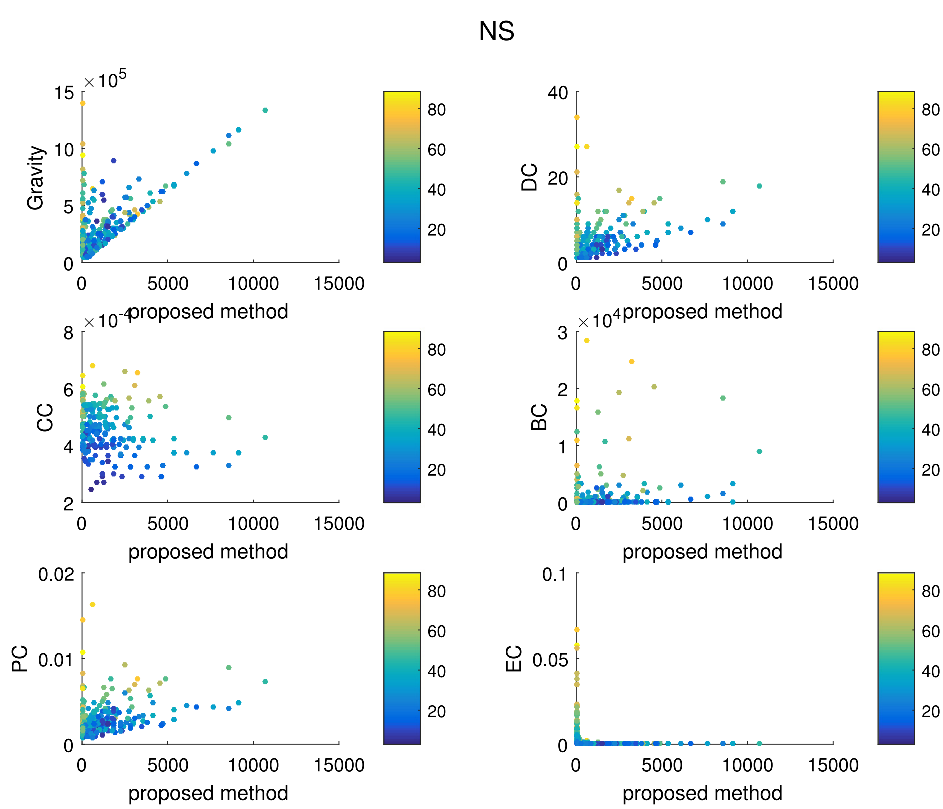

Fig. 13. The figure depicts the relationship between the proposed method and other methods in NS. The proposed method shows a strong correlation with Gravity.

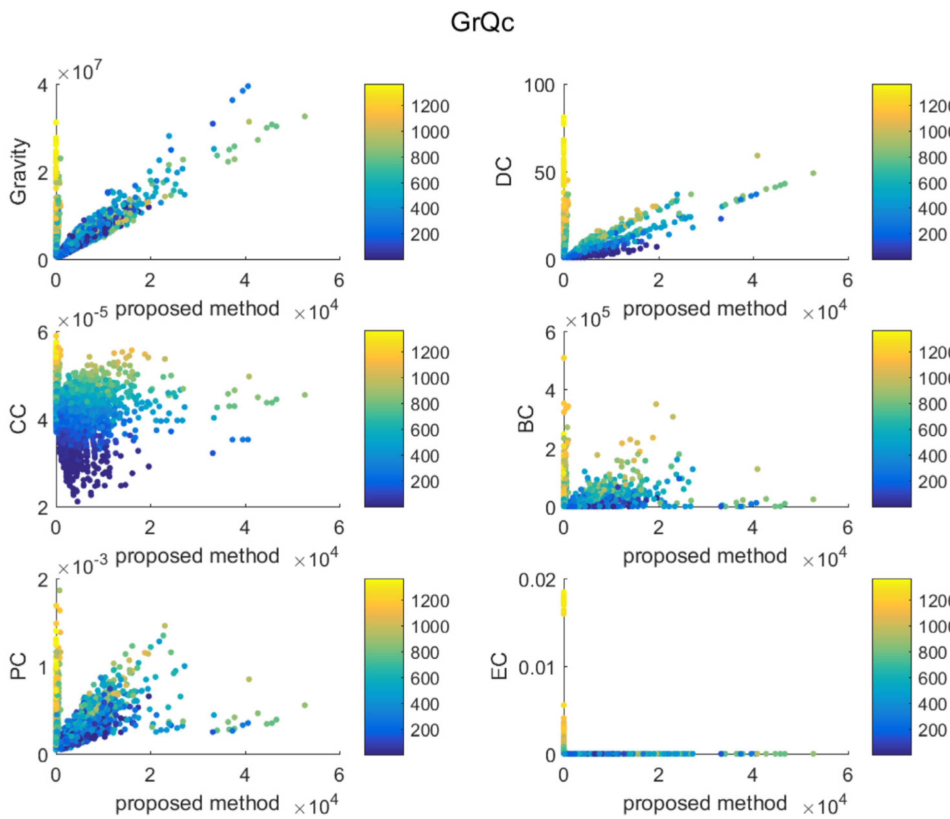

Fig. 14. The figure depicts the relationship between the proposed method and other methods in GrQc. The proposed method shows a good correlation with PC, with different differences from CC and BC.

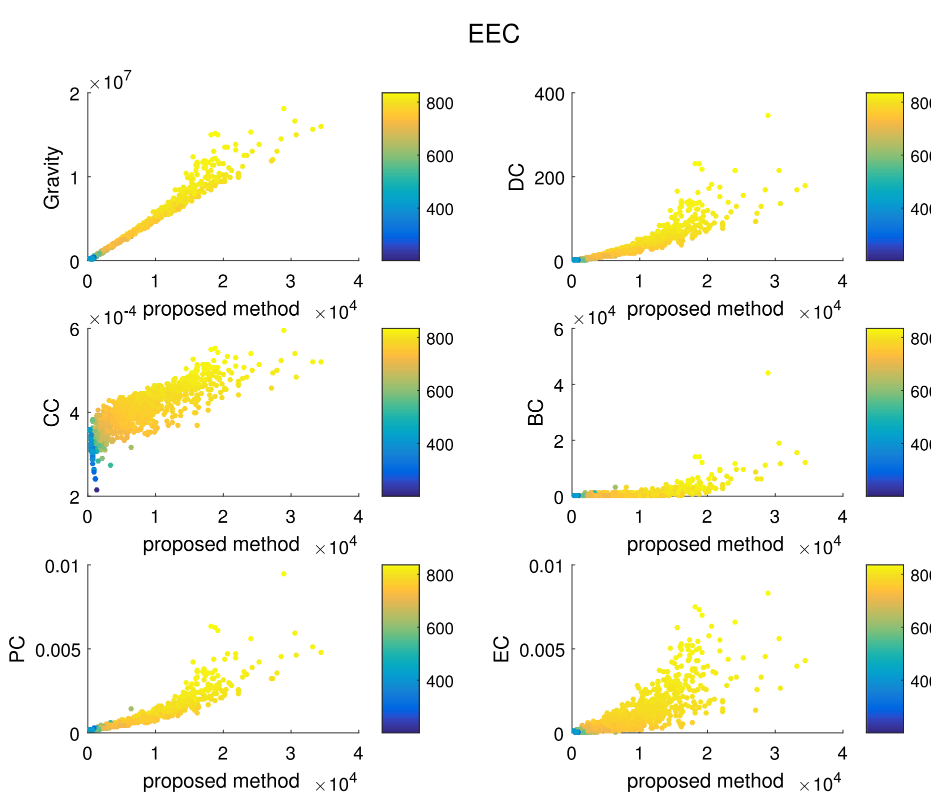

Fig. 15. The figure depicts the relationship between the proposed method and other methods in EEC. The proposed method shows a strong correlation with the other six methods.

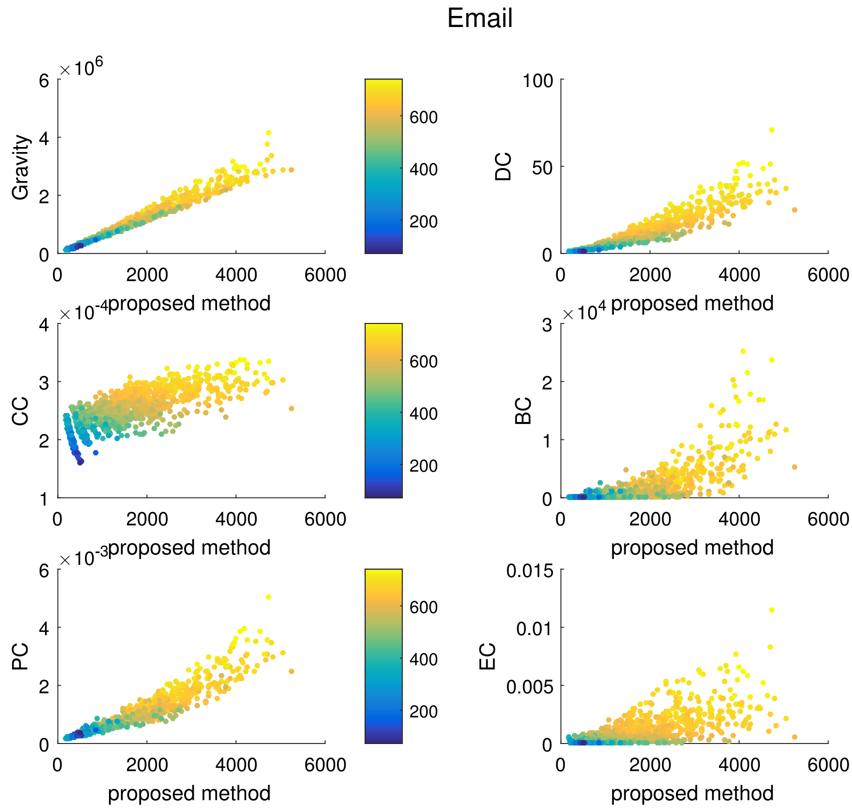

Fig. 16. The figure depicts the relationship between the proposed method and other methods in Email. The proposed method shows a strong correlation with its Gravity, DC, and PC.

to interact with nodes is the same. In fact, the power of each node is different in many real applications. For example, in social networks, the influence of some celebrities is obviously greater than that of common people. Therefore, it is necessary to consider the weight of each node. Inspired by EC [61], from a global perspective, the weights of different nodes are distinguished by assigning a relative score to each node. However, when all items in the feature vector are required to be non-negative, only the largest eigenvalue is satisfied.

STEP 4: Calculate interaction forces of nodes.

In the real world, the interaction between two objects is often proportional to mass and inversely proportional to the square of the distance. In this model, the degree of the node is taken as the quality, and the shortest distance between the nodes is taken as the distance between the nodes. WGravity centrality measure is defined as follows.

$$C_{WG}(i) = \sum_{d_{ij} \leq R_J} e_i \times \frac{k_i \times k_j}{d_{ij}^2} \tag{9}$$

where d_{ij} is the shortest distance between node i and node j. R is the radius of influence of each node, and the radius of influence is set to $0.5\langle d \rangle$ (we assume that the influence of each node does not exceed half of the radius of the network).

In conclusion, the GMM considers both the local and global information. For example, If the degree of one node is larger (local information), the shorter distance and the greater weight (global information) with other nodes, the importance of the node is greater. In addition, the GMM inherits the truncation radius of the Gravity model proposed by Li et al. [54] to tackle with the time-consuming problem in some large-scale networks. Moreover, inspired by EC [61], the relative score (global information) is assigned to each node to control the relative interaction of each node. To summarize, the whole steps of proposed method are shown in Fig. 1.

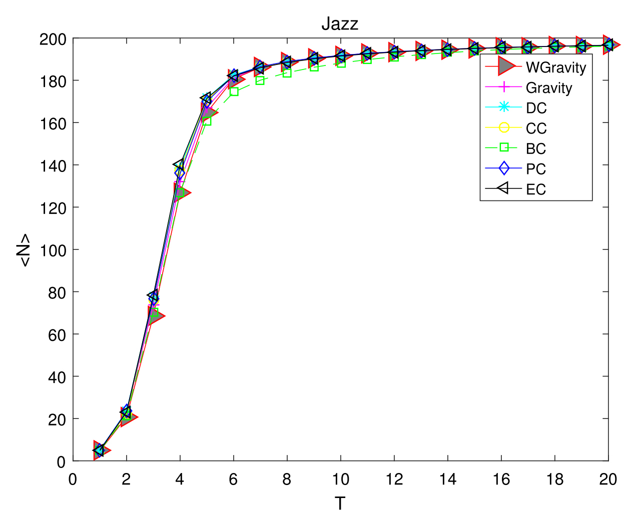

Fig. 17. This figure compares the number of average infected nodes in the top 100 nodes selected by Jazz in different methods. If the curve rises faster, it means that the selected top 100 nodes have a greater influence on the SI model. $F(t)$ represents the average number of nodes infected by the node at time t. The rising trend of all methods is basically the same, but the average number of influences of BC after time 4 is the least.

COMPLEXITY ANALYSE

Consider a network G with n nodes and m edges corroding with the adjacency matrix A. The time complexity of the proposed method is divided into three terms in total: the shortest path of the node is $O(m)$, the maximum eigenvector e corresponding to the adjacency matrix A is $O(n^2)$, and the degree of the computing node is $O(n^2)$. So its total time complexity is $O(n_k\langle 2n^2 + m\rangle)$,

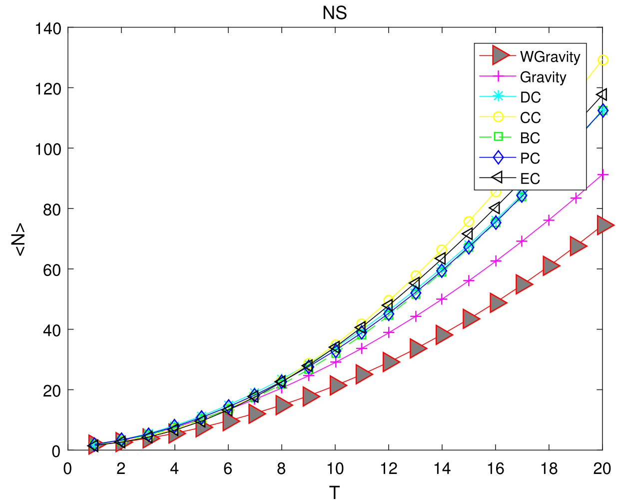

Fig. 18. This figure compares the number of average infected nodes in the top 100 nodes selected by NS in different methods. All methods performed significantly differently in NS. The top100 node of CC has the most influence, and the curve rises faster. WGravity's top100 nodes have the least number of impacts.

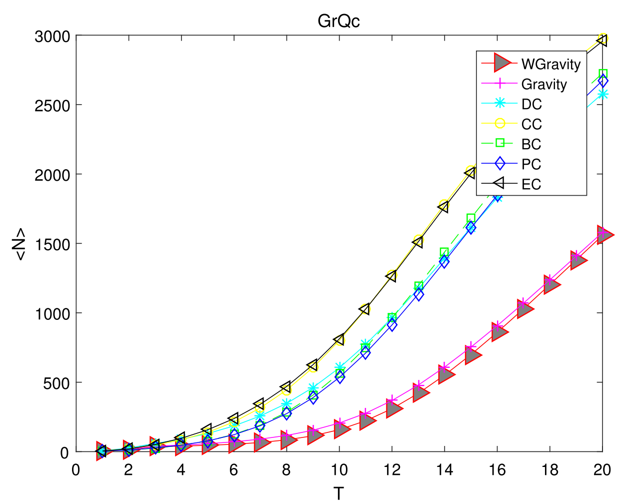

Fig. 19. This figure compares the number of average infected nodes in the top 100 nodes selected by GrQc in different methods. The trends of the number of influences of the top100 nodes of CC and EC are consistent, and the curve rises faster. WGravity's top100 nodes have the least number of impacts.

Where n_k is the maximum number of nodes within any node radius R.

4. Experiments

To test the effectiveness of the proposed method, five different experiments were performed in ten real networks and compared with six different methods.

4.1. Data

The experiments were conducted in a total of ten data sets in real world. Jazz [63] is a network of Jazz musicians. NS [64] is a network of scientists working together. GrQc [65] is a network published on preprints. Email [66] describes the network in which members send emails to each other. The EEC [67] describes a network in which European research members exchange mail. The PG [68] describes a network of file sharing. PB [69] is a blog network. USAir [70] is the US transportation aviation network. Power [71] is a power network. The specific information about the data is as shown in Table 1.

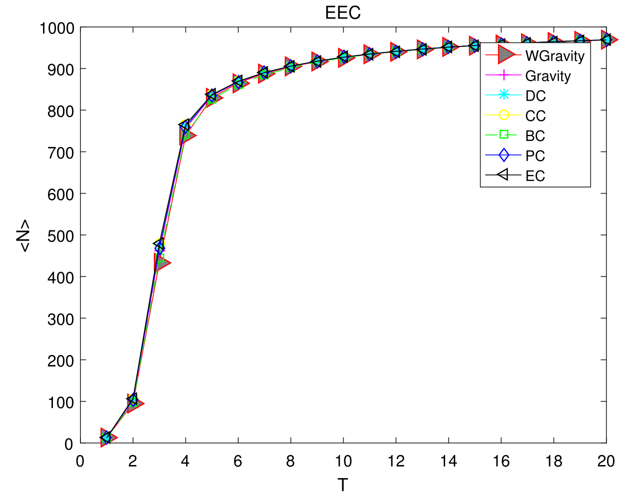

Fig. 20. This figure compares the number of average infected nodes in the top 100 nodes selected by EEC in different methods. The differences in EEC among all the methods are not obvious, and the upward trend is consistent.

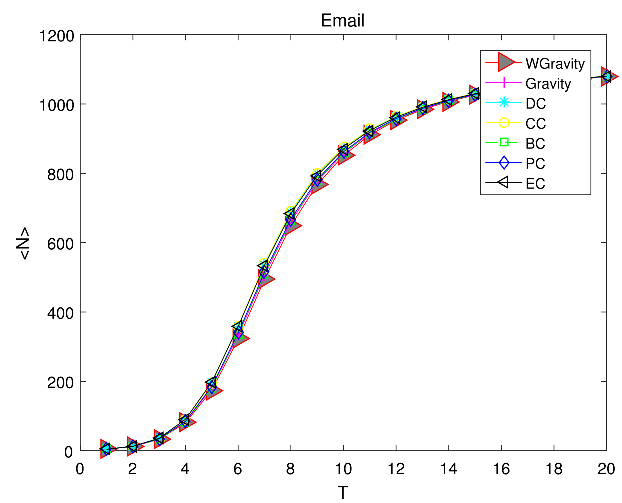

Fig. 21. This figure compares the number of average infected nodes in the top 100 nodes selected by Email in different methods. The differences in EEC among all the methods are not obvious, and the upward trend is consistent. However, it can be observed that the number of nodes affected by WGravity is slightly smaller.

Table 1
The basic information of ten networks. n and m are the number of nodes and edges, $\langle k \rangle$ and $\langle d \rangle$ are average degree and the shortest distance.

Networks	n	m	$\langle k \rangle$	$\langle d \rangle$
Jazz [63]	198	2472	27.6970	2.2350
NS [64]	379	914	4.8232	6.0419
GrQc [65]	4158	13 422	6.4560	6.0494
EEC [67]	986	16 064	32.5842	2.5869
Email [66]	1133	5451	9.6222	3.6060
PG [68]	6299	20 776	6.5966	4.6430
USAir [70]	332	2126	12.8072	2.7381
PB [69]	1222	16 714	27.3552	2.7375
Facebook [72]	4039	88 234	43.6910	3.6925
Power [71]	4941	6594	2.6691	18.9892

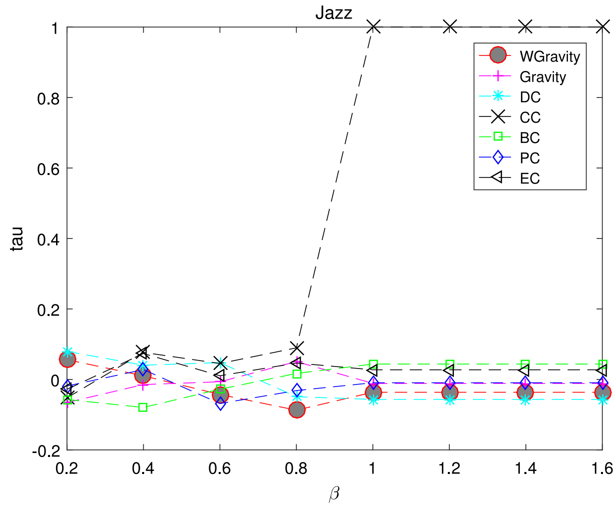

Fig. 22. When β changes, this figure describes the difference between the standard sequence generated by SI and the other methods generated in Jazz. When β increases, CC and tau are gradually close to the same as SI, and the difference between DC and SI is the largest.

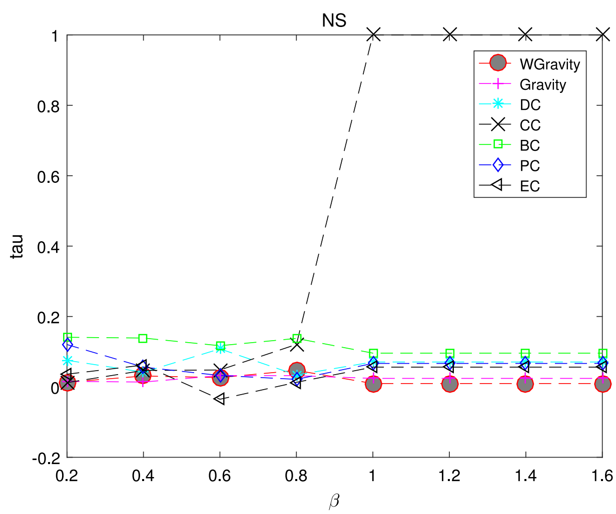

Fig. 23. When β changes, this figure describes the difference between the standard sequence generated by SI and the other methods generated in NS. When β increases, CC and tau are gradually close to the same as SI. EC fluctuated, when $\beta = 0.6$, and the tau of wGravity and SI sequences was the lowest.

4.2. Centrality scores of nodes

In this experiment, the proposed method was used to calculate the centrality scores of the nodes. For comparison, Gravity model, DC, CC, BC, PC, and EC were also applied to the same data set. The experimental results are shown in Figs. 2–11. In the heat map, the darker the color of the node, the greater the importance of the node.

First, calculate each method to derive the central scores of each node. Then, normalize all the scores. Finally, the comparisons of the experimental results are performed by heat maps. As can be seen Figs. 2–11, the distribution of the relative importance of the nodes obtained by all methods is consistent. The value of CC calculation node importance is too large. The distribution of WGravity model and Gravity model nodes is very close. Gravity model and WGravity performance are relatively consistent in judging the importance of each node, but WGravity is significantly different globally. The WGravity model believes that the nodes on the right network in the Jazz network are more important than the nodes on the left.

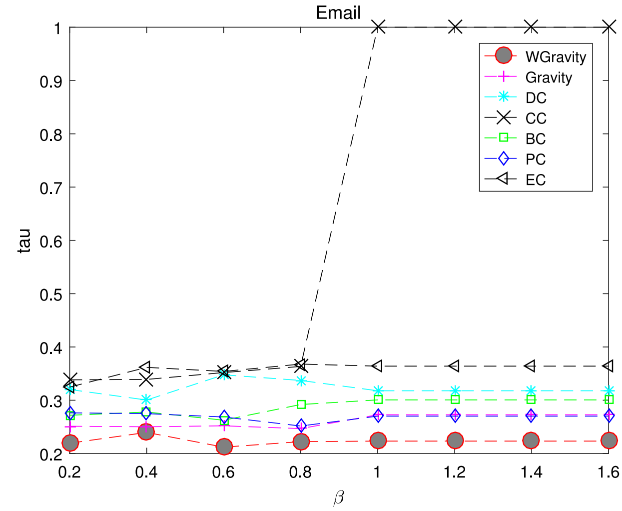

Fig. 24. When β changes, this figure describes the difference between the standard sequence generated by SI and the other methods generated in Email. The tau of wGravity and SI sequences is the lowest. The tau of EC and SI sequences is the second highest.

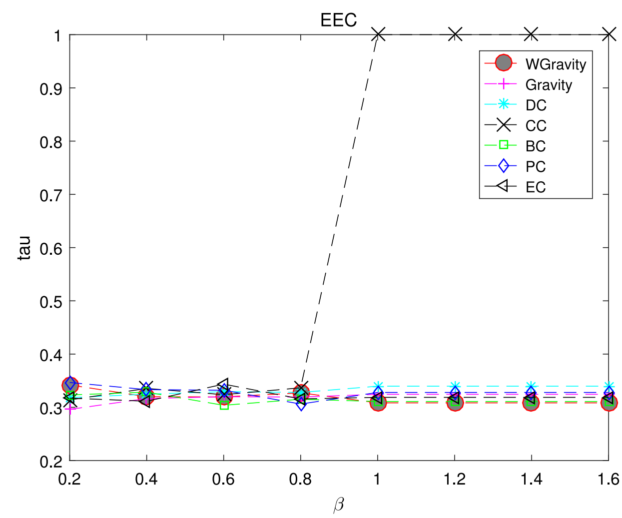

Fig. 25. When β changes, this figure describes the difference between the standard sequence generated by SI and the other methods generated in EEC. Besides CC, the tau of other methods and SI sequences is the lowest.

4.3. Compare the proposed method with other methods and the susceptible infected (SI) model

In this section, we will first introduce the SI model [73] and Kendall's tau [74]. Then, the relationship between the proposed method and other methods will be compared.

4.3.1. SI model

The SI model can be used to assess the transmission capacity of each node in the network and to reflect the impact of each node [73]. In the SI model, the nodes of the network have two states: susceptible state and infected state. Infected state nodes infect surrounding susceptible nodes with a certain probability. Table 2 describes the experimental parameters that will be used in the SI model.

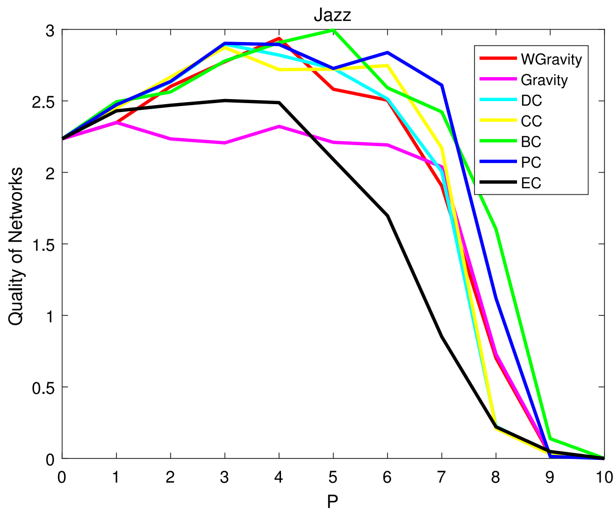

Fig. 26. The figure depicts the changes in the quality of the Jazz network when the node sequences generated by the different methods are attacked. If the network quality drops faster, it indicates that nodes selected by the method is more important in the network path. At first, the quality of the network in gravity declined the fastest, and then EC declined rapidly. BC falls relatively slowly in all methods.

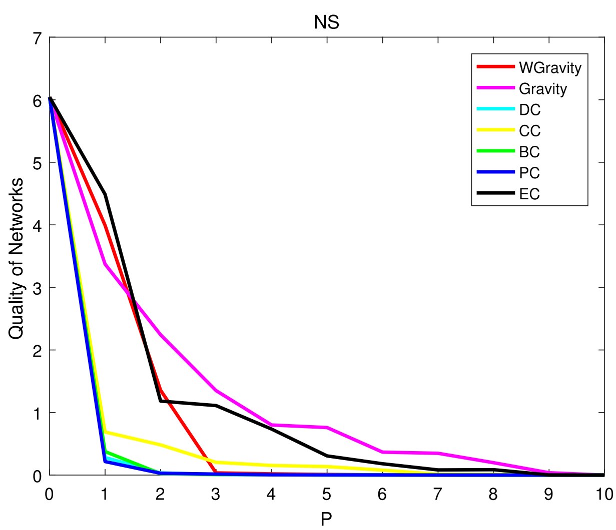

Fig. 27. The figure depicts the changes in the quality of the NS network when the node sequences generated by the different methods are attacked. Among all the methods, PC dropped the fastest, followed by CC, and gravity dropped the slowest.

Table 2
Parameter description.

Parameter	Parameter description
t	SI model experimental simulation time
$F(t)$	The average number of infected nodes at time t
β	Probability of node infection
N	Number of experiments

4.3.2. Kendall's tau

Kendall's tau reflects the correlation of the two sequences. The larger the value of Kendall's tau, the greater the similarity between the two sequences.

Given two sets of sequences x and y, x_i and y_i are the values of the *ith* position of x and y, respectively. Let (x_i, y_i) be a set of sequence pairs. If $x_i > x_j$ and $y_i > y_j$ or $x_i < y_i$ and $y_i < y_j$, then (x_i, y_i) is recorded as a positive sequence pair. Otherwise, it is recorded as a negative sequence pair. The Kendall's tau is defined as follows [74].

$$tau = \frac{n_+ - n_-}{n(n-1)} \tag{10}$$

where n_+ and n_- are the number of positive sequence pairs and negatives sequence pairs, and n is the total number of sequence pairs.

Table 3
The proposed method and the different methods of Kendall's tau.

Networks	Gravity	DC	CC	BC	PC	EC
Jazz	−0.0334	0.0907	−0.0365	−0.0374	0.0359	−0.0002
NS	0.0145	0.0276	0.0096	0.0466	−0.0021	0.0525
GrQc	−0.0072	0.0043	−0.0165	0.0085	0.0156	−0.0019
EEC	0.3579	0.3496	0.3084	0.3218	0.3317	0.3005
Email	−0.0072	0.0043	−0.0165	0.0085	0.0156	−0.0019

4.3.3. Correction with other method

In this experiment, the effectiveness of the method is demonstrated by comparing the relationship between the proposed method and other methods. All methods are performed on the data sets Jazz, NS, GrQc, EEC, Email. In the experimental SI model, β was set to 0.1, time t was set to 45, and all experiments were set to 100 times.

First, through the experimental results of experiment 4.3, each node is sorted in descending order. If the node's centrality scores is larger, the importance of the nodes is greater. Then, the importance of each node is evaluated by the SI model. All experimental results are shown in Figs. 12–16. Table 3 shows the proposed method and other methods of Kendall's tau. It can be seen that the proposed method has a positive correlation with the results obtained by Gravity model, PC, CC, DC. Among them, there is a high degree of correlation with the Gravity model.

4.4. The average number of infected nodes of the top nodes

In a network, if a node is more important, its ability to transmit infection is higher. In this experiment, we compared the average infectious power of the selected nodes of the proposed method and other methods. First, we select the top 100 nodes from each of the two methods in experiment Section 4.3.3. Then, this top 100 node is used as a source of infection in the SI model. Finally, calculate the number of average infected nodes per method.

The experimental results are shown in Figs. 17–21. If the node is more important, the more nodes it can propagate. Therefore, the faster the curve rises, the more important the set of sequences selected as seeds. It can be seen from the experimental results that the proposed method has the same effect in data, Email, EEC and Jazz and other methods. Figs. 20 and 21 show the performance of seven methods in EEC and Email. In the EEC, in the first two steps, all the curves rise rapidly and gradually converge in the eighth step. However, the trend of EC to rise significantly is slightly higher than other methods.

4.5. Compare other methods with SI models

In all data sets, a propagation probability β is given in an SI model to obtain a standard centrality sequence, and then we compare the other methods to this sequence. In all experiments, each SI model was independently performed 100 times, and β changed from 0.2 to 1.6. The results of the experiment are shown in Figs. 22–25.

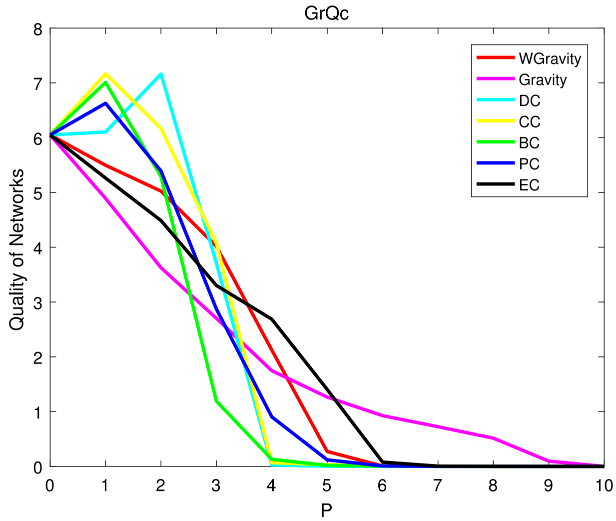

Fig. 28. The figure depicts the changes in the quality of the GrQc network when the node sequences generated by the different methods are attacked. At first, Gravity declined the fastest, then PC declined rapidly.

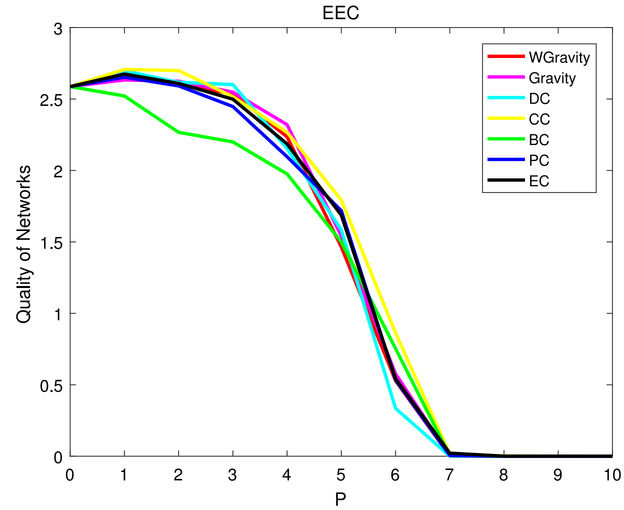

Fig. 29. The figure depicts the changes in the quality of the EEC network when the node sequences generated by the different methods are attacked. Among all the methods, BC decreased the fastest, and other methods had little difference.

4.6. Compare the quality of the network by attacking important nodes

The connectivity of the network can be used as an indicator of network quality. If the connectivity of the network is better, the quality of the network is higher. In this experiment, we test the quality of the network by attacking each method to give a sequence of top nodes. The quality of the network is defined as follows.

$$Q = \sum_{i,j} \frac{d_{ij}}{n(n-1)} \tag{11}$$

where $d_{i,j}$ is the shortest distance between node i and node j, and n is the number of nodes.

In this experiment, we attacked $P \times 10\%$ of the nodes in order of node importance. First, according to each method, the order of node importance is derived. Then, attack each node and invalidate the node. Finally, calculate the quality of the network. The results of the experiment are shown in Figs. 26–30.

It can be seen that the proposed method shows good performance in the data set Jazz, GrQc. As the percentage of important nodes of the attack continues to increase, the quality of the network drops rapidly.

It can be seen that as β continues to increase, the sequence of CC and SI is constantly approaching. The proposed method is close to the sequence derived by other methods on the network EEC.

5. Conclusion

In this paper, a generalized mechanical model is proposed to assess the importance of nodes. In order to verify the effectiveness of the proposed method, we compared the proposed method with five different methods. The experimental results show the effectiveness of the proposed method. In addition, an innovative experiment was proposed to validate the importance of network nodes. First, a network quality is defined. Then, by attacking important nodes, the quality of the network is calculated.

In conclusion, the GMM considers both the local and global information. If the degree of one node is larger , the shorter distance and the greater weight with other nodes, the importance of the node is greater. In addition, the GMM inherits the truncation radius of the gravity model to deal with the time-consuming problem in large-scale networks. Gravity mode and GMM performance are relatively consistent in judging the importance of each node, but GMM is significantly different globally. We think that when global information (the relative weight of nodes) is introduced, the difference in the importance of nodes can be judged from a global perspective, which is more refined.

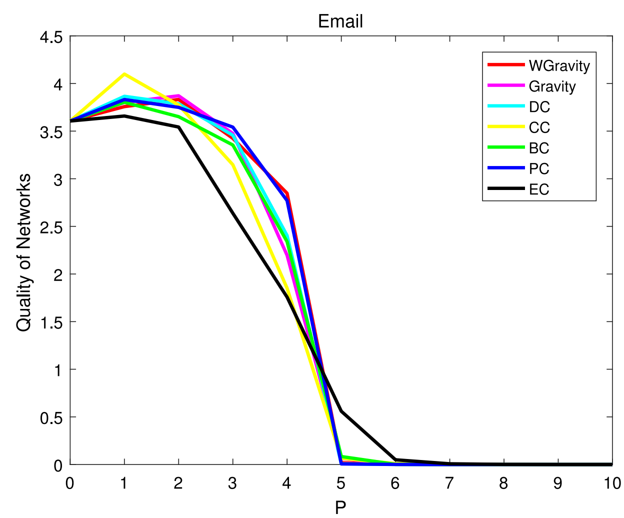

Fig. 30. The figure depicts the changes in the quality of the Email network when the node sequences generated by the different methods are attacked. EC dropped the fastest, followed by CC. There is little difference in other methods.

Acknowledgments

The authors greatly appreciate the reviews' suggestions and the editor's encouragement. The work is partially supported by National Natural Science Foundation of China (Grant No. 61973332), Key Area R & D Program of Guangdong Province (No. 2019B010137004), National Natural Science Foundation of China (Grant No. U1803263), the Fundamental Research Funds for the Central Universities (No. 3102019PJ006), and Shanxi Natural Science Foundation-Key Program (Grant No. 2019ZDLGY17-07).

References

[1] M. Zhao, T. Zhou, B. Wang, W. Wang, Enhanced synchronizability by structural perturbations, Phys. Rev. E 72 (5, 2) (2005).

[2] Y. Yu, G. Xiao, J. Zhou, Y. Wang, Z. Wang, J. Kurths, H.J. Schellnhuber, System crash as dynamics of complex networks, Proc. Natl. Acad. Sci. USA 113 (42) (2016) 11726–11731.

[3] Q. Jin, L. Wang, C.-Y. Xia, Z. Wang, Spontaneous symmetry breaking in interdependent networked game, Sci. Rep. 4 (2014).

[4] Z. Wang, C.T. Bauch, S. Bhattacharyya, A. d'Onofrio, P. Manfredi, M. Perc, N. Perra, M. Salathe, D. Zhao, Statistical physics of vaccination, Phys. Rep. 664 (2016) 1–113.

[5] L. Lue, Y.-C. Zhang, C.H. Yeung, T. Zhou, Leaders in social networks, the delicious case, PLoS One 6 (6) (2011).

[6] X. Li, M. Jusup, Z. Wang, H. Li, L. Shi, B. Podobnik, H.E. Stanley, S. Havlin, S. Boccaletti, Punishment diminishes the benefits of network reciprocity in social dilemma experiments, Proc. Natl. Acad. Sci. USA 115 (1) (2018) 30–35.

[7] Z. Wang, M.A. Andrews, Z.-X. Wu, L. Wang, C.T. Bauch, Coupled disease-behavior dynamics on complex networks: A review, Phys. Life Rev. 15 (2015) 1–29.

[8] Z. Wang, M. Jusup, R.-W. Wang, L. Shi, Y. Iwasa, Y. Moreno, J. Kurths, Onymity promotes cooperation in social dilemma experiments, Sci. Adv. 3 (3) (2017).

[9] D. Zhao, L. Wang, Z. Wang, G. Xiao, Virus propagation and patch distribution in multiplex networks: Modeling, analysis and optimal allocation, IEEE Trans. Inf. Forensics Secur. 14 (2019) 1755.

[10] F. Liu, Y. Deng, A fast algorithm for network forecasting time series, IEEE Access 7 (1) (2019) 102554–102560.

[11] P. Xu, R. Zhang, Y. Deng, A novel visibility graph transformation of time series into weighted networks, Chaos Solitons Fractals 117 (2018) 201–208.

[12] H. Yang, Y. Deng, R. Mayne, A bio-inspired network design method for intelligent transportation, Int. J. Unconv. Comput. 14 (3,4) (2019) 199–215.

[13] B. Wei, F. Xiao, Y. Shi, Synchronization in kuramoto oscillator networks with sampled-data updating law, IEEE Trans. Cybern. (2019) 1–9.

[14] B. Wei, F. Xiao, Y. Shi, Fully distributed synchronization of dynamic networked systems with adaptive nonlinear couplings, IEEE Trans. Cybern. (2019) 1–9.

[15] W. Deng, Y. Deng, Entropic methodology for entanglement measures, Physica A 512 (2018) 693–697.

[16] H. Yang, Y. Deng, J. Jones, Network division method based on cellular growth and physarum-inspired network adaptation, Int. J. Unconv. Comput. 13 (6) (2018) 477–491.

[17] Q. Cai, Y. Deng, A fast Bayesian iterative rule in amoeba algorithm, Int. J. Unconv. Comput. (20) (2020).

[18] A.-L. Barabasi, N. Gulbahce, J. Loscalzo, Network medicine: a network-based approach to human disease, Nat. Rev. Genet. 12 (1) (2011) 56–68.

[19] Z. Li, D. Jia, H. Guo, Y. Geng, Y. Shen, Z. Wang, X. Li, The effect of multigame on cooperation in spatial network, Appl. Math. Comput. 351 (2019) 162.

[20] C. Liu, C. Shen, Y. Geng, S. Li, C. Xia, Z. Tian, L. Shi, R. Wang, S. Boccaletti, Z. Wang, Popularity enhances the interdependent network reciprocity, New J. Phys. 20 (2019) 123012.

[21] C. Chu, X. Hu, C. Shen, T. Li, S. Boccaletti, L. Shi, Z. Wang, Self-organized interdependence among populations promotes cooperation by means of coevolution, Chaos 29 (1) (2019) 013139.

[22] D. Li, X. Wang, P. Huang, A fractal growth model: Exploring the connection pattern of hubs in complex networks, Physica A 471 (C) (2017) 200–211.

[23] T. Wen, W. Jiang, Identifying influential nodes based on fuzzy local dimension in complex networks, Chaos Solitons Fractals 119 (2019) 332–342.

[24] L. Lu, D. Chen, X.-L. Ren, Q.-M. Zhang, Y.-C. Zhang, T. Zhou, Vital nodes identification in complex networks, Phys. Rep. 650 (2016) 1–63.

[25] T. Zhou, Reply to: "Comment on: 'Epidemic spreading on heterogeneous networks with identical infectivity' [Phys. Lett. A 364 (2007) 189]" - [Phys. Lett. 372 (2008)1722], Phys. Lett. A 372 (10) (2008) 1725–1726.

[26] X. Huang, I. Vodenska, F. Wang, S. Havlin, H.E. Stanley, Identifying influential directors in the United States corporate governance network, Phys. Rev. E 84 (4, 2) (2011).

[27] X. Zhang, J. Zhu, Q. Wang, H. Zhao, Identifying influential nodes in complex networks with community structure, Knowl.-Based Syst. 42 (2013) 74–84.

[28] M. Kitsak, L.K. Gallos, S. Havlin, F. Liljeros, L. Muchnik, H.E. Stanley, H.A. Makse, Identification of influential spreaders in complex networks, Nat. Phys. 6 (11) (2010) 888–893.

[29] J. Tang, R. Zhang, P. Wang, Z. Zhao, L. Fan, X. Liu, A discrete shuffled frog-leaping algorithm to identify influential nodes for influence maximization in social networks, Knowl.-Based Syst. (2019).

[30] Z.Z. Alp, Ö. Şule Gündüz, Identifying topical influencers on twitter based on user behavior and network topology, Knowl.-Based Syst. 141 (2018) 211–221.

[31] Z.Z. Alp, Ö. Şule Gündüz, Influence factorization for identifying authorities in twitter, Knowl.-Based Syst. 163 (2019) 944–954.

[32] X. Lei, X. Yang, H. Fujita, Random walk based method to identify essential proteins by integrating network topology and biological characteristics, Knowl.-Based Syst. 167 (2019) 53–67.

[33] T. Zhou, B. Wang, Catastrophes in scale-free networks, Chin. Phys. Lett. 22 (5) (2005) 1072–1075.

[34] X. Wang, J. Li, Detecting communities by the core-vertex and intimate degree in complex networks, Physica A 392 (10) (2013) 2555–2563.

[35] Y. Cui, X. Wang, J. Li, Detecting overlapping communities in networks using the maximal sub-graph and the clustering coefficient, Physica A 405 (2014) 85–91.

[36] J. Li, X. Wang, J. Eustace, Detecting overlapping communities by seed community in weighted complex networks, Physica A 392 (23) (2013) 6125–6134.

[37] X. Wang, W. Zhou, R. Li, J. Cao, X. Lin, Improving robustness of interdependent networks by a new coupling strategy, Physica A 492 (C) (2018) 1075–1080.

[38] F. Nian, C. Hu, S. Yao, L. Wang, X. Wang, An immunization based on node activity, Chaos Solitons Fractals 107 (2018) 228–233.

[39] X. Wang, T. Zhao, Model for multi-messages spreading over complex networks considering the relationship between messages, Commun. Nonlinear Sci. Numer. Simul. 48 (2017) 63–69.

[40] P. Bonacich, Factoring and weighting approaches to status scores and clique identification, J. Math. Sociol. 2 (1) (1972) 113–120.

[41] L.C. Freeman, Centrality in social networks conceptual clarification, Soc. Netw. 1 (3) (1978) 215–239.

[42] M.E. Newman, A measure of betweenness centrality based on random walks, Soc. Netw. 27 (1) (2005) 39–54.

[43] L. Lue, T. Zhou, Q.-M. Zhang, H.E. Stanley, The H-index of a network node and its relation to degree and coreness, Nat. Commun. 7 (2016).

[44] H. Li-Ya, T. Ping-Chuan, H. You-Liang, Z. Yi, C. Xie-Feng, Node importance based on the weighted K-order propagation number algorithm, Acta Phys. Sin. 68 (12) (2019).

[45] J. Zhou, X. Yu, J.-A. Lu, Node importance in controlled complex networks, IEEE Trans. Circuits Syst. II-Express Briefs 66 (3) (2019) 437–441.

[46] S. Brin, L. Page, The anatomy of a large-scale hypertextual web search engine, Comput. Netw. ISDN Syst. 30 (1–7) (1998) 107–117.

[47] M. Li, H. Xu, Y. Deng, Evidential decision tree based on belief entropy, Entropy 21 (9) (2019) 897.

[48] X. Cao, Y. Deng, A new geometric mean FMEA method based on information quality, IEEE Access 7 (1) (2019) 95547–95554.

[49] Y. Song, Y. Deng, Divergence measure of belief function and its application in data fusion, IEEE Access 7 (1) (2019) 107465–107472.

[50] L. Pan, Y. Deng, An association coefficient of belief function and its application in target recognition system, Int. J. Intell. Syst. 35 (2020) 85–104.

[51] H. Fujita, A. Gaeta, V. Loia, F. Orciuoli, Hypotheses analysis and assessment in counter-terrorism activities: a method based on OWA and fuzzy probabilistic rough sets, IEEE Trans. Fuzzy Syst. (2019) http://dx.doi.org/10.1109/TFUZZ.2019.2955047.

[52] H. Mo, Y. Deng, Identifying node importance based on evidence theory in complex networks, Physica A (2019) http://dx.doi.org/10.1016/j.physa.2019.121538.

[53] L. Fei, Q. Zhang, Y. Deng, Identifying influential nodes in complex networks based on the inverse-square law, Physica A 512 (2018) 1044–1059.

[54] Z. Li, T. Ren, X. Ma, S. Liu, Y. Zhang, T. Zhou, Identifying influential spreaders by gravity model, Sci. Rep. 9 (1) (2019) 8387.

[55] L. Pan, W. Wang, S. Cai, T. Zhou, Optimal interlayer structure for promoting spreading of the susceptible-infected-susceptible model in two-layer networks, Phys. Rev. E 100 (2019) 022316.

[56] Y. Dai, J. Zhang, W. Wang, Interconnecting strategy of bridging multilayer networks to maximize synchronizability, Europhys. Lett. 125 (1) (2019) 18003.

[57] L. Gao, W. Wang, P. Shu, H. Gao, L.A. Braunstein, Promoting information spreading by using contact memory, Europhys. Lett. 118 (1) (2017) 18001.

[58] L. Pan, W. Wang, S. Cai, T. Zhou, Optimizing spreading dynamics in interconnected networks, Chaos 29 (10) (2019) 103106.

[59] T. Wen, Y. Deng, Identification of influencers in complex networks by local information dimensionality, Inform. Sci. 512 (2020) 549–562.

[60] Z. Li, T. Ren, X. Ma, S. Liu, Y. Zhang, T. Zhou, Identifying influential spreaders by gravity model, Sci. Rep. 9 (2019).

[61] P. Bonacich, P. Lloyd, Eigenvector-like measures of centrality for asymmetric relations, Soc. Netw. 23 (3) (2001) 191–201.

[62] D.L. Hansen, B. Shneiderman, M.A. Smith (Eds.), Preface, in: Analyzing Social Media Networks with NodeXL, Morgan Kaufmann, Boston, 2011, p. ix.

[63] P.M. Gleiser, L. Danon, Community structure in jazz, Adv. Complex Syst. 6 (04) (2003) 565–573.

[64] M.E. Newman, Finding community structure in networks using the eigenvectors of matrices, Phys. Rev. E 74 (3) (2006) 036104.

[65] J. Leskovec, J. Kleinberg, C. Faloutsos, Graph evolution: Densification and shrinking diameters, ACM Trans. Knowl. Discov. Data 1 (1) (2007) 2.

[66] R. Guimera, L. Danon, A. Diaz-Guilera, F. Giralt, A. Arenas, Self-similar community structure in a network of human interactions, Phys. Rev. E 68 (6) (2003) 065103.
[67] H. Yin, A.R. Benson, J. Leskovec, D.F. Gleich, Local higher-order graph clustering, in: Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, ACM, 2017, pp. 555–564.
[68] J. Leskovec, K.J. Lang, A. Dasgupta, M.W. Mahoney, Community structure in large networks: Natural cluster sizes and the absence of large well-defined clusters, Internet Math. 6 (1) (2009) 29–123.
[69] L.A. Adamic, N. Glance, The political blogosphere and the 2004 US election divided they blog, in: Proceedings of the 3rd International Workshop on Link Discovery, ACM, 2005, pp. 36–43.
[70] V. Batageli, A.P. Mrvar, Available at, http://vlado.fmf.uni-lj.si/pub/networks/data/, 2007.
[71] D.J. Watts, S.H. Strogatz, Collective dynamics of 'small-world' networks, Nature 393 (6684) (1998) 440.
[72] J. Leskovec, J.J. Mcauley, Learning to discover social circles in ego networks, in: Advances in Neural Information Processing Systems, 2012, pp. 539–547.
[73] M. Yang, G. Chen, X. Fu, A modified SIS model with an infective medium on complex networks and its global stability, Physica A 390 (12) (2011) 2408–2413.
[74] M.G. Kendall, A new measure of rank correlation, Biometrika 30 (1/2) (1938) 81–93.