Knowledge-Based Systems 227 (2021) 107198

Contents lists available at ScienceDirect

Knowledge-Based Systems

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

An improved gravity model to identify influential nodes in complex networks based on k-shell method

Xuan Yang, Fuyuan Xiao *

School of Computer and Information Science, Southwest University, Chongqing, 400715, China

ARTICLE INFO

Article history:
Received 5 November 2020
Received in revised form 22 April 2021
Accepted 2 June 2021
Available online 3 June 2021

Keywords:
Complex networks
Influential nodes
Gravity model
K-shell

ABSTRACT

To find the important nodes in complex networks is a fundamental issue. A number of methods have been recently proposed to address this problem but most previous studies have the limitations, and few of them considering both local and global information of the network. The location of node, which is a significant property of a node in the network, is seldom considered in identifying the importance of nodes before. To address this issue, we propose an improved gravity centrality measure on the basis of the k-shell algorithm named KSGC to identify influential nodes in the complex networks. Our method takes the location of nodes into consideration, which is more reasonable compared to original gravity centrality measure. Several experiments on real-world networks are conducted to show that our method can effectively evaluate the importance of nodes in complex networks.

1. Introduction

With the rapid development of network science, complex network has appealed to much attention in different fields [1], such as society [2,3], biology, physics [4–6], time series [7,8], transportation [9,10], immunization strategy [11,12] and so on [13–17]. The world is full of complexity and uncertainty, while the real-world systems can be modeled by complex networks [18,19]. Some special nodes in a complex system may decide many structural properties of the whole network [20–22]. As a result, researchers paid more attention to identify the important nodes, which can accurately reveal the hidden properties in complex networks.

Up to now, there exist many classical methods to identify the important nodes in complex networks including Degree centrality (DC) [23], Betweenness centrality (BC) [24], Closeness centrality (CC) [25] and Eigenvector centrality (EC) [26]. However, these classical methods have their own shortcomings. Degree centrality only focuses on the number of surrounding connected nodes and does not consider global structure of the whole network [27]. The Betweenness centrality and Closeness centrality rank the importance of nodes from global perspective, but they cannot perform well in large-scale complex networks due to their high computational complexity [28]. As for Eigenvector centrality, its property that cannot be used in weighted networks reduces its application. Recently some new measures have been proposed to improve the performance in identifying nodes [29–35]. For instance, Wen and Jiang proposed a method to identify influential nodes based on fuzzy local dimension [36]. Zhao et al. considered both the self-importance and global importance of a node simultaneously [37]. Zareie et al. has used gray-wolf optimization (GWO) technique which is a bio-inspired optimization algorithm to find the seed nodes [38]. Due to the good ability of data fusion, evidence theory has been used to deal with this problem [39].

Gravity law is a basic physical rule to calculate the interaction between two objects. The interaction between two objects is proportional to the masses of two objects and inversely proportional to the square of the distance between two objects. Inspired by gravity formula, Li et al. [40] proposed a gravity centrality (GC) model which utilizes the degree of a node as its mass and the shortest path distance between two nodes as the their distance. While in the gravity centrality model, the interaction between two nodes is only related with their degree values and distance, which indicates that the attractions between two nodes are the same. In reality, the attraction ability of each node may be different. Liu et al. [41] improved this model by considering the weight of a node in the network and the improved centrality measure named WGC can be more applicable in real-world networks.

Given the above discussion, in this paper, we take the location of node into consideration. That is to say, a node in the center of the network has more ability to attract other nodes than nodes at the edge of the network [42]. An improved gravity model on the basis of k-shell algorithm [43] is proposed to identify influential nodes in the networks. The location difference between nodes represented by the k-shell values difference is used as the attraction coefficient which adjusts the attractiveness of the central nodes in the network. The approach proposed by this

* Corresponding author.
E-mail address: doctorxiaofy@hotmail.com (F. Xiao).

https://doi.org/10.1016/j.knosys.2021.107198
0950-7051/

paper combines both local information and global information. Several experiments are conducted here and the final results show that our method performs well.

The rest of this paper is organized as follows. Some preliminary knowledges are briefly introduced in Section 2. In Section 3, the method which is named as KSGC to identify the influential nodes is proposed. Several experiments are applied in Section 4 to illustrate the reasonability of the proposed method. The conclusion is given in Section 5.

2. Preliminaries

2.1. Centrality measures

Considering a given complex network *G=(N,E)*, where N is the set of nodes and E is the set of edges, the topological structure of network is usually represented by a adjacency matrix A. The element a_{ij} in a adjacency matrix can describe the connective relationship between nodes. $a_{ij} = 1$ means that there is a connection edge between node i and node j, while $a_{ij} = 0$ otherwise. Some existing centrality measures are briefly introduced as follows.

Definition 2.1. The Degree centrality (DC) [23] is a simple method to measure the centrality of nodes. Its value is measured by the number of surrounding connected nodes. The DC of node i is defined as follows:

$$DC(i) = \sum_{j}^{|N|} a_{ij} \tag{1}$$

where $|N|$ is the number of nodes in the network. The value of DC(i) is also called the degree of node i which can be represented by k_i.

Definition 2.2. The Betweenness centrality (BC) [24] focus on the degree of concentration of the path. The BC of node i is defined as follows:

$$BC(i) = \frac{\sum_{j \neq k \neq i} L_{jk}(i)}{\sum_{j \neq k} L_{jk}} \tag{2}$$

where L_{jk} is the number of shortest paths from node j to node k, and $L_{jk}(i)$ represents the number of shortest paths from node j to node k through node i.

Definition 2.3. The Closeness centrality (CC) [25] measures the centrality score of a node by calculating the sum of shortest path distance from one node to the other nodes in the network. The CC of node i is defined as follows:

$$CC(i) = \frac{1}{\sum_{j \neq i} d_{ij}} \tag{3}$$

where d_{ij} is the shortest path distance between node i and node j.

Definition 2.4. The Eigenvector centrality (EC) [26] calculates the relative scores of all nodes in the network. Given a matrix A which represents a network, x_i is the value of the ith entry in the normalized eigenvector belonging to the largest eigenvalue of A. The EC of node i is defined as follows:

$$Ax = \lambda x, EC(i) = x_i = \frac{1}{\lambda} \sum_{j=1}^{|N|} a_{ij} x_j \tag{4}$$

where λ is the largest eigenvalue of A.

Definition 2.5. The Gravity centrality (GC) [40] measures the importance of a node by calculating the sum of interactions between a node and other nodes in the network. A truncation radius is introduced here to solve the problem of time-consuming for large networks, which is set as half of the average path length of the network. The GC of node i is defined as follows:

$$GC(i) = \sum_{j \neq i, d_{ij} \leq 0.5\langle d \rangle} \frac{k_i k_j}{d_{ij}^2} \tag{5}$$

where $\langle d \rangle$ is the average path length of the network.

Definition 2.6. The Weighted Gravity centrality (WGC) [41] improves the GC by taking the weight of each node into consideration. The WGC of node i is defined as follows:

$$WGC(i) = \sum_{j \neq i, d_{ij} \leq 0.5\langle d \rangle} e_i * \frac{k_i k_j}{d_{ij}^2} \tag{6}$$

where e_i is the ith value of normalized eigenvector belonging to the largest eigenvalue of A.

2.2. K-shell algorithm

The k-shell decomposition method is proposed by Kitsak [43] to indicate node's importance in the network. The first step of the k-shell algorithm is to remove all the nodes in the network whose degree $k = 1$. We shall continue to remove nodes whose degree $k <= 1$ after one round removal because this step may lead to the reduction of the degree values during the process of removal. Until there are no nodes in the network with degree $k <= 1$, all the nodes which have been removed in this step create 1-shell and their k-shell values are equal to one. Then we can repeat this procedure to obtain 2-shell, 3-shell…and so on. Finally all nodes are divided into different shells and we can get the k-shell value of each node. Obviously, a node with a higher k-shell value means that the node is located in a more central position in the network [28].

Although the k-shell algorithm may used to evaluate the position of nodes in the network, one obvious drawback of this method is that there are many of nodes sharing with the same k-shell values, which makes it fail to rank the nodes in the same shell. Therefore, many researchers have modified the original k-shell method to tackle the problem of influential nodes identification [44]. The mix degree decomposition [45] takes into account the contribution of the removed nodes and their links after each iteration which are ignored by k-shell algorithm. The hierarchical k-shell [46] is a hierarchical approach to improve the precision of k-shell algorithm. Maji et al. [47] improve the k-shell hybrid method [48] which combines the local information (degree) and global information (k-shell). Wang et al. [49] consider the iterations where the nodes are removed in the k-shell algorithm and propose a k-shell iteration factor. And the potential edge weight based k-shell degree neighborhood [50] uses the node degree and k-shell index to assign weight to links between two nodes.

3. Proposed method

In a system of real-world, the elements in the system kernel are more stable than those at the edge. A main reason is that the connection density between internal elements is higher than the edge. We can use the k-shell value to denote the position of a node in the network. Actually, a node, whose k-shell value is n, would form a $n + 1$-clique with other n nodes in the network. n-clique denotes a complete subgraph composed by n nodes. It is a stable local structure because its inner connection density is the highest. Obviously, a node with higher k-shell value means

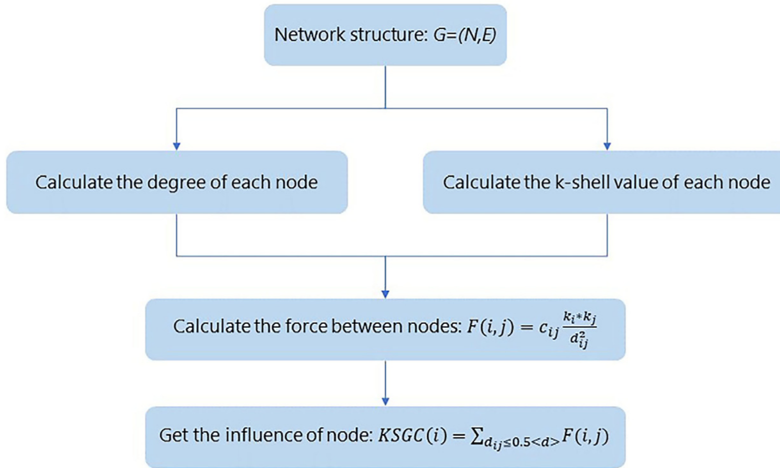


Fig. 1. The flow chart of this proposed method.

Table 1
Some basic topological features of real-world networks.

Network	$\|N\|$	$\|E\|$	$< d >$	d_{max}	$< k >$	k_{max}	β_{th}
Jazz [51]	198	2742	2.2350	6	27.6970	100	0.0266
USAir [52]	332	2126	2.7381	6	12.8072	139	0.0231
Netscience [53]	379	914	6.0419	17	4.8232	34	0.1424
Political blogs [54]	1222	16714	2.7375	8	27.3552	351	0.0125
PGP [55]	10680	24316	7.4855	24	4.5536	205	0.0559
LFR6000	6000	203665	2.3874	3	67.8883	598	0.0060
UCLA [56]	20453	747604	3.0194	8	73.1046	1180	0.0070
Astroph [57]	17903	196972	4.1940	14	22.0044	504	0.0155

it belongs to a more stable local structure, which denotes that it would get more help from other nodes in this clique when suffering from the external influence. In other words, a node with higher k-shell value may offset the external influence to some extent and may have more ability to influence nodes. In the previous studies about the k-shell based methods, the k-shell value is only used as the characteristic of the node itself. While in this paper, the above analysis motivates us to consider that it is necessary to take into account the location difference which is represented by k-shell difference when we calculate the mutual interaction between two nodes.

The gravity model calculates the mutual interactions between pairs of nodes to measure the influence of a node. And the influence of a node in this model consists of two aspects. The first is the degree of a node. In other words, if a node has more surrounding neighbors, the more nodes can be affected directly, as a result, the node has higher influence. The second is the distance between the node and the other nodes. The influence that a node imposes to another node is inversely proportional to the square of their distance. The smaller their distance, the greater their mutual interactions. In this paper, we assume that the ability of a node to influence another node is related with their location. That is to say, a node in the central part of the network may have more attraction than the nodes located in the periphery. And we use the k-shell value to represent the position of a node. The attraction coefficient that node i imposes to node j is c_{ij}:

$$c_{ij} = e^{\frac{ks(i) - ks(j)}{ks_{max} - ks_{min}}} \tag{7}$$

where $ks(i)$ and $ks(j)$ are the k-shell values of node i and node j respectively. And the ks_{max} and ks_{min} are the largest and smallest k-shell value in the network. So the force node i imposes to node j can be defined as $F(i,j)$:

$$F(i,j) = c_{ij}\frac{k_i * k_j}{d_{ij}^2} \tag{8}$$

obviously, if node i and node j hold different k-shell value, $F(i,j)$ is not equal to $F(j,i)$. And when $ks(i) > ks(j)$, the force node i imposes to node j is higher than gravity model; when $ks(i) < ks(j)$, the situation is opposite, and when $ks(i) = ks(j)$, our method degenerates into gravity model. And the importance of a node is the sum of the forces that the node imposes to other nodes in the network.

A flowchart of the proposed method is shown in Fig. 1. The specific steps of the proposed method is conducted as the following steps:

STEP 1: Constructing the network

Given a connected undirected network, we can use *G=(N,E)* to represent it. In order to indicate the connective relationship between nodes, the adjacency matrix A is introduced here. $a_{ij} = 1$ means there is a link between node i and node j, and $a_{ij} = 0$ is opposite.

STEP 2: Calculating the degree of each node and the shortest distance between different node pairs

As mentioned above, the force between two nodes is proportional to their masses and inversely proportional to the square of their distance. In our model, the shortest path (d_{ij}) is considered as the shortest distance between node i and node j. The mass of a node is represented by its degree.

$$k_i = \sum_j a_{ij} \tag{9}$$

where k_i is the degree of node i, and a_{ij} is the element in the adjacency matrix to show the connective relationship between node i and node j.

STEP 3: Calculating the k-shell value of each node

Due to the consideration of node's location, the k-shell value of each node needs to be obtained here. The specific calculating method has been mentioned in Section 2.2.

STEP 4: Calculating the force that a node imposes to another node

The force between two nodes is connected with three parts. There are the degrees of two nodes, distance between two nodes and their k-shell difference. Eq. (8) is the formulation to calculate the force that node i imposes to node j.

STEP 5: Getting the influence of nodes

Finally, the influence of a node is denoted by the sum of forces that the node imposes to other nodes. In order to reduce the computational complexity, we also inherit the truncation radius of the network proposed by Li. So the k-shell based on gravity centrality (KSGC) measure proposed in this paper is defined as follows:

$$KSGC(i) = \sum_{d_{ij} \leq 0.5\langle d \rangle} F(i, j) \tag{10}$$

This formulation has considered both the local and global information. If a node has a larger degree and shorter distances with other nodes, in addition, its k-shell value is also relatively higher than its surrounding nodes, then the importance of this node will be greater. And with the setting of the truncation radius, the influence of a node does not exceed half of the network diameter.

In conclusion, this work combines the gravity model and k-shell algorithm to identify the importance of nodes in the complex networks. The highlight of this paper is that we propose an attraction coefficient to modify the mutual interactions between pairs of nodes. By introducing this attraction coefficient, when there are two nodes with different k-shell values in the network, the attraction between them would be different. A node with higher k-shell value would have more ability to attract other nodes than lower k-shell value. Hence, it is a more comprehensive method which captures the difference between pair of nodes from a global perspective.

4. Experimental analysis

In this Section, six real-world networks and a synthetic network generated from the LFR benchmark model [58] are used here to test the performance of the proposed model. And the LFR6000 network is generated from the LFR benchmark model with the parameter setting that the nodes number is 6000, the average degree is 60, the maximum degree is 500 and the mixing parameter of the community structure is 0.8. The detailed information about the data can be seen in Table 1. $|N|$ and $|E|$ are the numbers of nodes and edges respectively; $\langle d \rangle$ and d_{max} denote the average and maximum value of the shortest distance in the network; $\langle k \rangle$ and k_{max} represent the average and maximum degree of the network; β_{th} is the SI epidemic threshold probability of the network which would be described in detail later.

At the experimental stage, to verify the correctness of our proposed method, the DC, BC, CC and EC are used to explore their relationship with KSGC. And for testing the performance of our method, some state-of-the-art k-shell based methods like mixed degree decomposition (MDD), hierarchical k-shell (HKS), improved k-shell hybrid (KSHI), k-shell iteration factor (KS-IF) and potential edge weight based k-shell index (WKS), and the gravity model based methods including GC and WGC are used as comparisons.

4.1. Models and measures used in this paper

4.1.1. SI model

The ability to spread information is a key factor to evaluate the importance of node [59]. There are many information spreading models like independent cascade model [60], linear threshold model [61] and disease spreading models [62]. Among them, the SI model is the most popular and simplest. Therefore, in this paper, we use the SI model to access the nodes' ability to spread information in the network.

In this model, each node in the network has two states: infected states and susceptible states. A susceptible node can turn into infected state, while the process is irreversible. A infected node can only affect its uninfected neighbor nodes with a probability β, which is also regarded as a spreading rate or infection rate. When the infection process begins, the infected nodes have a spreading rate β to infect their susceptible neighbor nodes at each time t. We use the $F(t)$ to denote the number of infected nodes at specific time t in the network. After the spreading process, the initial infected node which has higher ability of spreading information will lead to a greater numbers of nodes be infected in the network. And one important point to note is that the value of β plays an important role in controlling the spreading speed. In order to accurately evaluate the spreading ability of nodes, Castellano and Pastor-Satorras [63] hold that the β should be kept small and just little more than the epidemic threshold probability $\beta_{th} = \frac{\langle k \rangle}{\langle k^2 \rangle - \langle k \rangle}$.

4.1.2. Kendall's tau coefficient

Kendall's tau coefficient [64] can measure the correlation between two sequences. A larger Kendall's tau coefficient indicates a greater similarity between two sequences. The definition of Kendall's tau coefficient is presented as follows. Given two sequences X and Y which have the same length, their ith values are denoted by x_i and y_i, respectively. Let each pair of elements x_i and y_i forms a set, which contains (x_i, y_i). If $x_i > x_j$ and $y_i > y_j$ or $x_i < x_j$ and $y_i < y_j$, (x_i, y_i) and (x_j, y_j) are considered concordant. They are considered discordant if $x_i > x_j$ and $y_i < y_j$ or $x_i < x_j$ and $y_i > y_j$. If $x_i = x_j$ and $y_i = y_j$, the pair is neither concordant nor discordant. Therefore, the kendall's tau coefficient τ is defined as

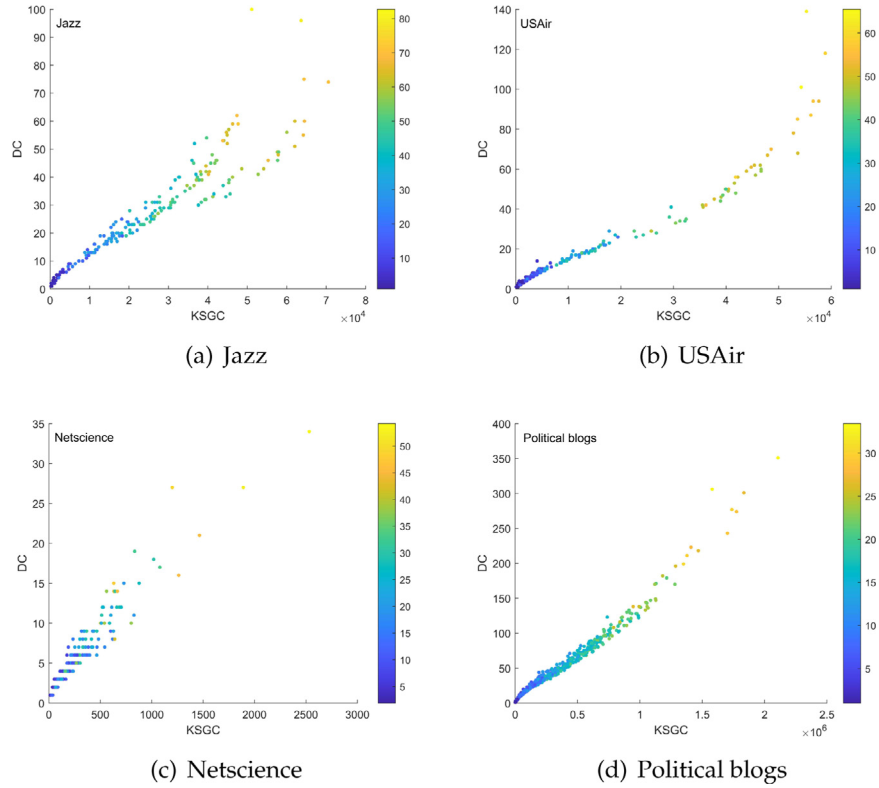
$$\tau = \frac{n_c - n_d}{0.5n(n-1)} \tag{11}$$

where n_c and n_d represent the numbers of concordant and discordant pairs, respectively, and n is the length of sequence. The value of τ lies between -1 and 1. Obviously, a higher τ value means a greater similarity between two ranking lists whereas a low value indicates dissimilarity.

4.2. The relationship of KSGC with other methods

In order to find the relationship between KSGC and other methods, the relationship graphs between different methods in several real-world networks are presented here. The DC, BC, CC and EC are chosen for comparison. The experimental results are given in Figs. 2–5. Each point in the graph represents a node in the network. The coordinates of the points in the graph mean the values obtained by different methods, and the color of points shows the infection ability of nodes which obtained by $F(10)$ when $\beta = \beta_{th}$ in different networks using SI model.

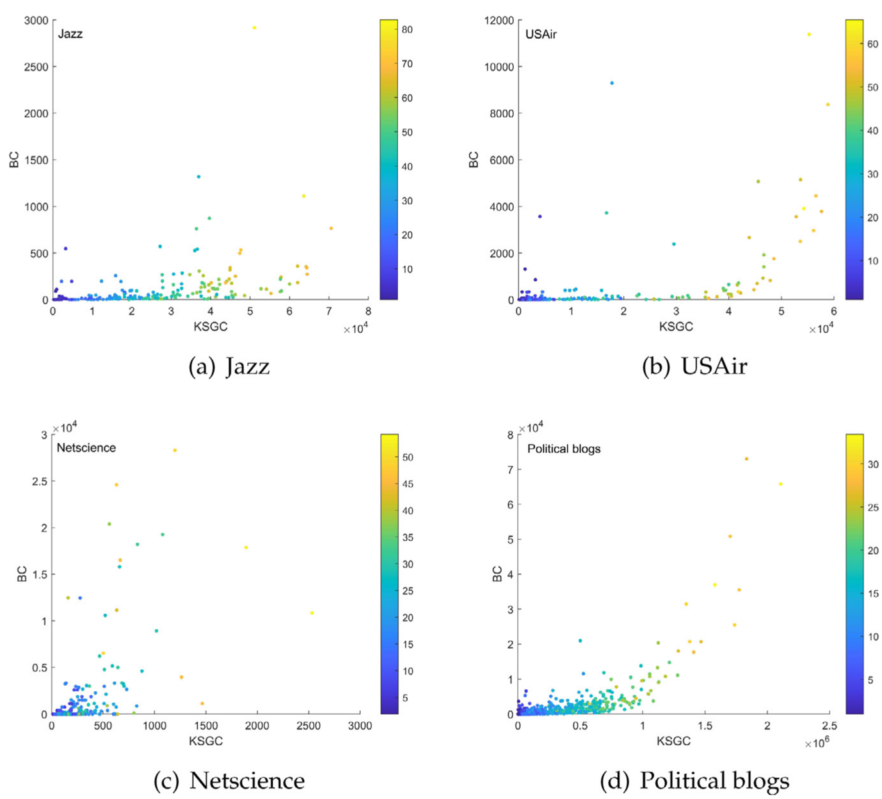
As can be seen from Fig. 2, DC and KSGC have a strong positive correlation, which means the KSGC value of node with large DC would also be great. However, just as the result presented in Fig. 3, KSGC is almost unrelated with BC. The correlation between KSGC and CC can be observed from Fig. 4. It can be found that the relationship between KSGC and CC is positive, except in the Netscience network. But the correlation degree is not as strong as DC. In Fig. 5, just like CC, there is little correlation between KSGC and EC in the Netscience network. While in another three networks, EC has shown more obvious positive correlation with KSGC than CC. In general, the proposed KSGC has different degrees of correlation with some classical centrality methods except



(a) Jazz (b) USAir

(c) Netscience (d) Political blogs

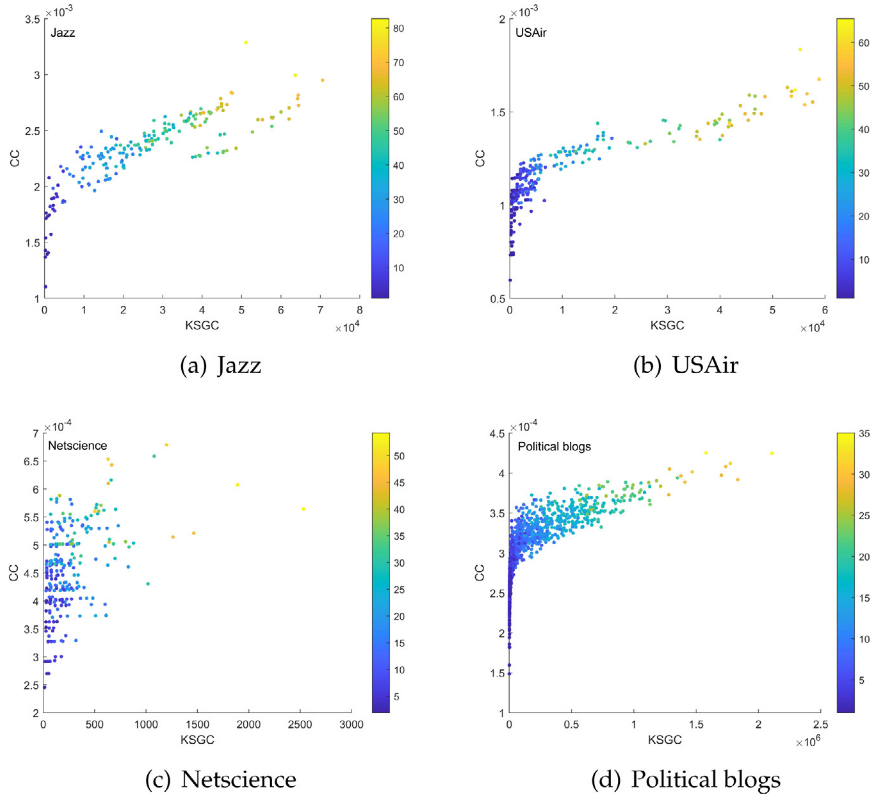
Fig. 2. This figure describes the relationship between KSGC and DC when $\beta = \beta_{th}$ in corresponding network. The values of the horizontal and vertical axes mean the values obtained by KSGC and DC respectively. The color of point means the infection ability obtained by SI model. According to the colorbar, the node with lighter color indicates its greater importance.



(a) Jazz (b) USAir

(c) Netscience (d) Political blogs

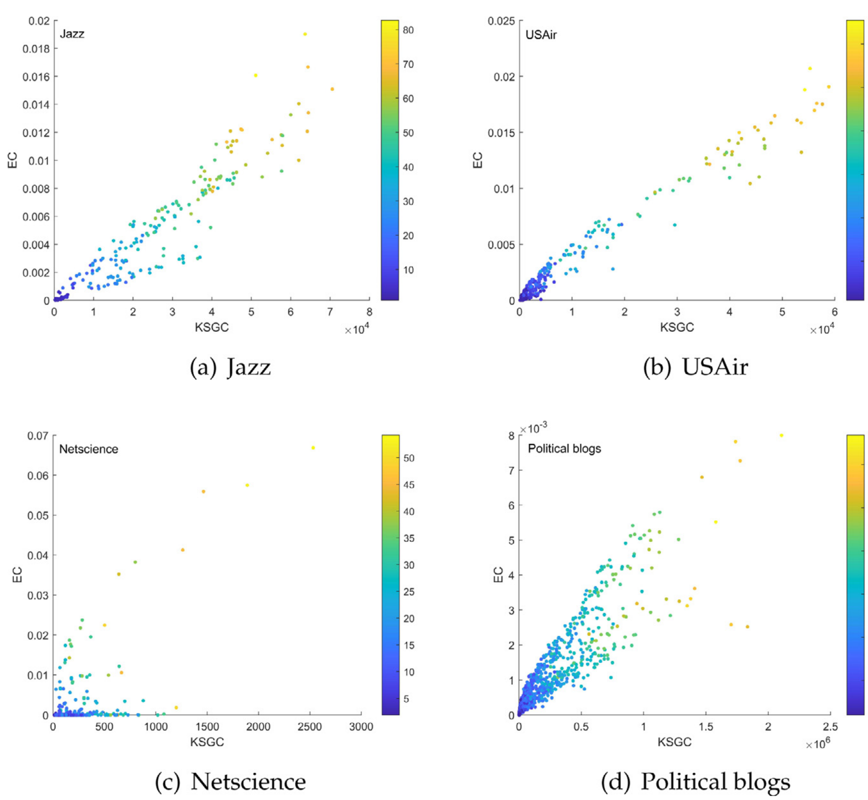
Fig. 3. This figure describes the relationship between KSGC and BC when $\beta = \beta_{th}$ in corresponding network. The values of the horizontal and vertical axes mean the values obtained by KSGC and BC respectively. The color of point means the infection ability obtained by SI model. According to the colorbar, the node with lighter color indicates its greater importance.



(a) Jazz (b) USAir

(c) Netscience (d) Political blogs

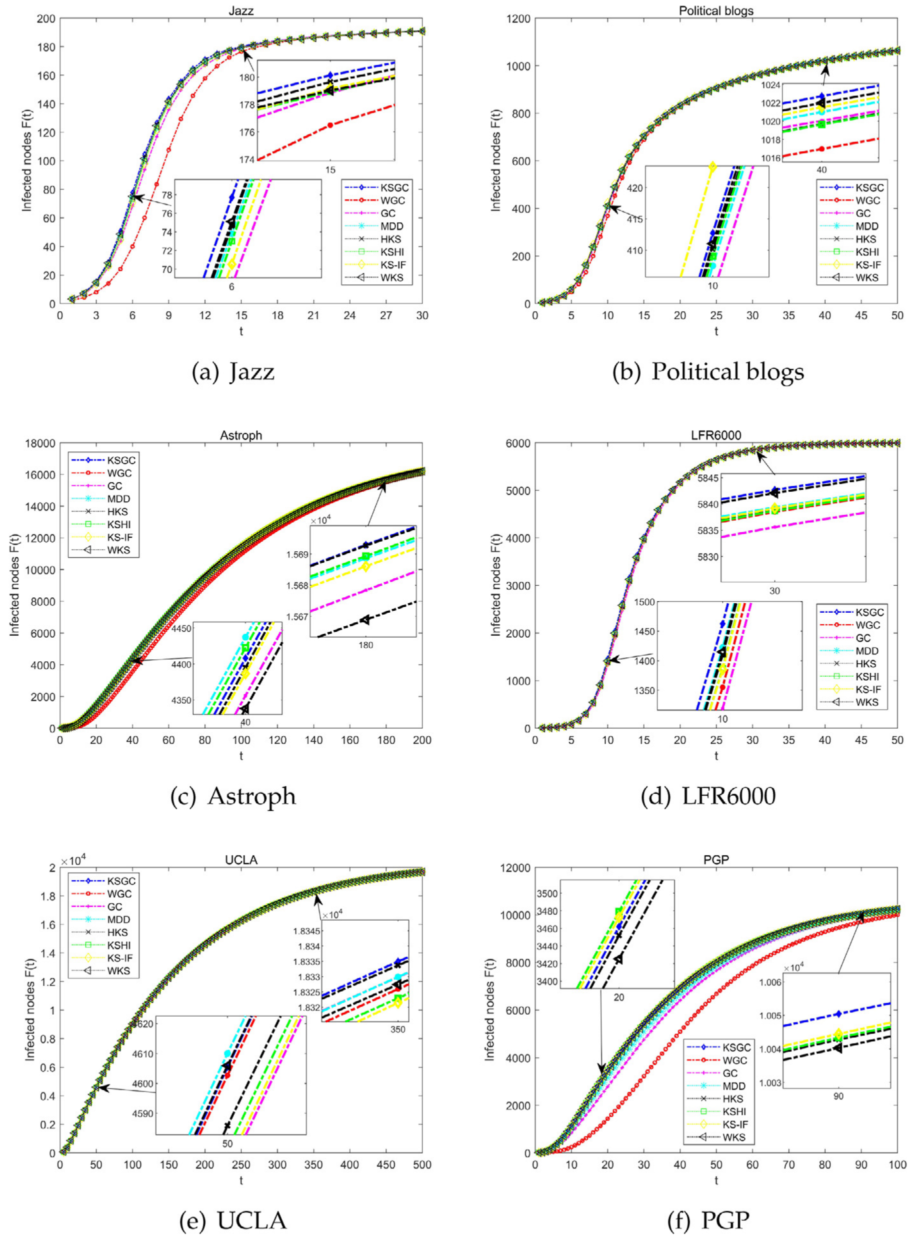
Fig. 4. This figure describes the relationship between KSGC and CC when $\beta = \beta_{th}$ in corresponding network. The values of the horizontal and vertical axes mean the values obtained by KSGC and CC respectively. The color of point means the infection ability obtained by SI model. According to the colorbar, the node with lighter color indicates its greater importance.



(a) Jazz (b) USAir

(c) Netscience (d) Political blogs

Fig. 5. This figure describes the relationship between KSGC and EC when $\beta = \beta_{th}$ in corresponding network. The values of the horizontal and vertical axes mean the values obtained by KSGC and EC respectively. The color of point means the infection ability obtained by SI model. According to the colorbar, the node with lighter color indicates its greater importance.



(a) Jazz (b) Political blogs

(c) Astroph (d) LFR6000

(e) UCLA (f) PGP

Fig. 6. This figure shows the comparison of the number of average infected nodes with the top 5 nodes as the initial nodes obtained by different methods. And the spreading rate β is set as the epidemic threshold β_{th}. The faster growing curve indicates the selected high ranking nodes have a higher spreading ability in the network.

BC. Another key point to notice is that a node with high KSGC value has a higher infection ability which would demonstrate the reasonableness of our proposed KSGC.

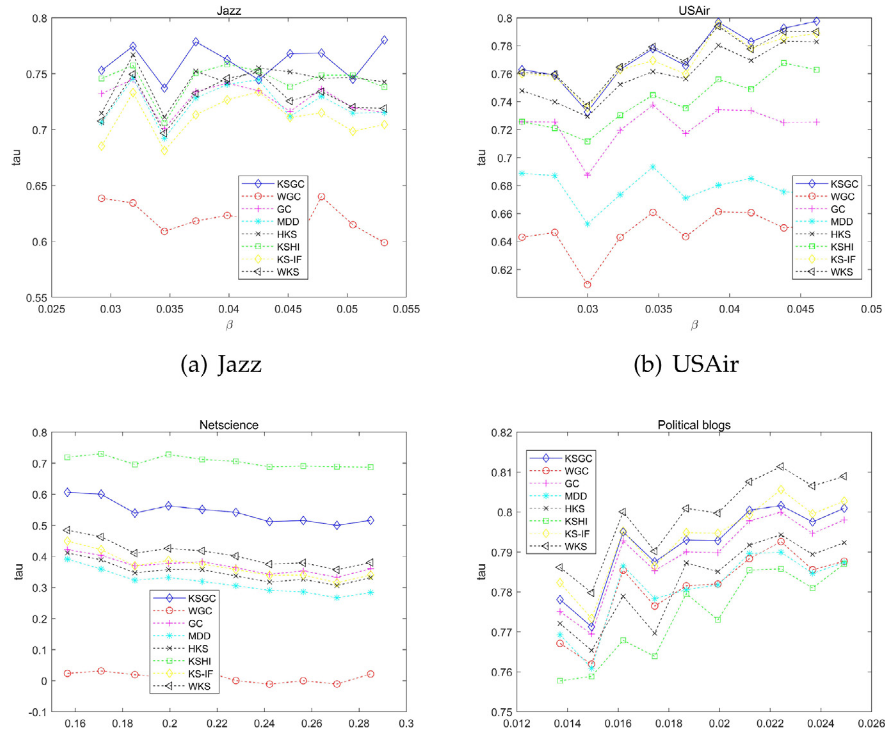
4.3. The comparison of spreading ability

As mentioned above, a node with a greater significance would have a stronger spreading ability in the network. Therefore, in order to compare the spreading ability of high ranking nodes obtained by different methods, we adopt the SI model and the high ranking nodes in different methods are used as the initial infected nodes. We use the $F(t)$ to denote the number of infected nodes at specific time t in the network. $F(t)$ increases with the infection process t and eventually reaches a stable value when most of nodes in the network are infected. We select the top 5% nodes as initial infected nodes in each method. Further, we assign different weights to nodes with different rankings. The weight of node whose ranking index is i can be defined as,

$$w_i = \frac{L+1-i}{\sum_{i=1}^{L} i} \tag{12}$$

where L is the number of initial infected nodes we select. It can be easily found that a higher ranking node (with small ranking index) would have a higher weight. And the $F(t)$ could be obtained as,

$$F(t) = \sum_{i=1}^{L} w_i * F_i(t) \tag{13}$$



(a) Jazz (b) USAir

(c) Netscience (d) Political blogs

Fig. 7. The Kendall's tau coefficient between the ranking list generated by the SI model and ranking lists of different algorithms with the different spreading probability β in the Jazz, USAir, Netscience and Political blogs network. And the value of β varies from epidemic threshold β_{th} to $2 * \beta_{th}$ with each step of 10% increase.

where $F_i(t)$ is the number of infected nodes at time t caused by the initial infected node whose ranking index is i. Each experiment has been carried for 100 times with $\beta = \beta_{th}$ in corresponding network. The final $F(t)$ is the average result to indicate the infection ability.

Fig. 6 describes the comparison of the spreading ability of different methods. The rising trends of $F(t)$ obtained by KSGC and other methods are basically the same in all networks. In the case of Jazz and LFR6000 network, KSGC performs better than other methods during almost the whole process. In some networks such as political blogs, UCLA and PGP, KSGC is not the best at the early infection process, but its growth rate is faster than other methods. And in the late period of infection process, KSGC has the highest $F(t)$. In Astroph network, the growth of KSGC and WKS reach the stable state almost at the same time and perform better than other methods. In conclusion, the most results presented in Fig. 6 show that this proposed method has a relative superiority performance, which can demonstrate that the top ranking nodes selected by the proposed KSGC have a stronger spreading ability in the network.

4.4. The ranking accuracy of the proposed method

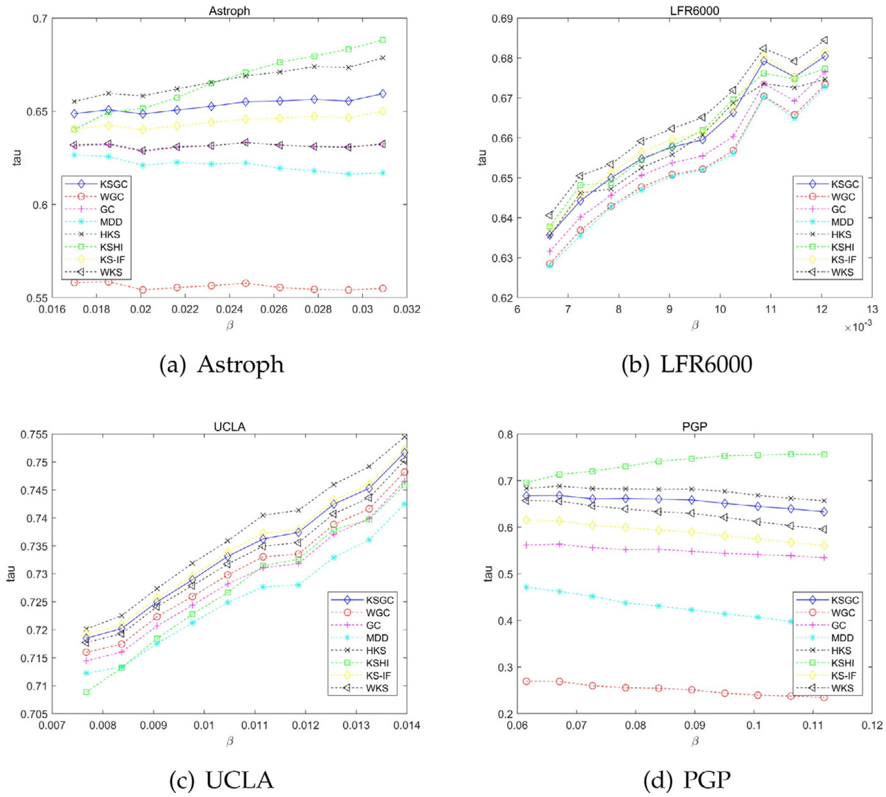
In this section, we have investigated the ranking accuracy of different algorithms. We use the SI simulation to generate the ranking according to the numbers of infected nodes in 10 steps. And the value of spreading rate β varies from β_{th} to $2 * \beta_{th}$ with each step of 10% increase. The final ranking result would be obtained by the average of 100 independent experiments. Then, we use the Kendall's tau coefficient to measure the correlation between the ranking lists of different algorithms and the ranking list generated by the SI simulation.

Figs. 7 and 8 show the result of change in β on the Kendall's tau coefficient (τ) for different ranking algorithms in the eight networks. It can be seen that, compared with other methods, this proposed method performs well in general. For example, in the case of Jazz and USAir network, KSGC has the highest Kendall's tau coefficient. And in other networks, even if KSGC is not the best, but its values are still in relatively higher positions. We can find that the KSHI has the highest Kendall's tau coefficient in Netscience and PGP network, but it performs badly in Political blogs and UCLA network. Another point to note is that the proposed KSGC performs better than WGC and GC in all networks which can demonstrates that our improvement is effective. In order to present the results more intuitively, the average Kendall's tau coefficient value are listed in Table 2. And we could obtain the same conclusion that mentioned above. The proposed KSGC and KSHI perform better than other methods.

5. Conclusion

In this paper, a new method named KSGC to identify influential nodes in the network is presented. The new approach improves the existing gravity centrality by taking the location of nodes into consideration. It combines both the global information, such as location, distance with other nodes and local information, for example, the degree information. A node located in the center of network which has a higher degree, shorter distance with other nodes has more attraction to other nodes. Finally, the importance of nodes can be measured by the sum of attractions that one node imposes to the other in the network. Several experiments on both real-world and synthetic networks are conducted here to demonstrate the reasonableness and effectiveness of the proposed method.

The location is a fundamental property of a node in identifying its importance. Due to the extension of classical k-shell decomposition method to weighted network [65,66], how to apply this method to weighted networks is the future work.



(a) Astroph (b) LFR6000

(c) UCLA (d) PGP

Fig. 8. The Kendall's tau coefficient between the ranking list generated by the SI model and ranking lists of different algorithms with the different spreading probability β in the Astroph, LFR6000, UCLA and PGP network. And the value of β varies from epidemic threshold β_{th} to $2 * \beta_{th}$ with each step of 10% increase.

Table 2
This table shows the average Kendall's tau coefficient ($\tau_A(\cdot)$) between the ranking lists obtained by different nodes ranking algorithms and the ranking list generated by the SI model. The maximum value in each row is marked with red, and the second and third are marked with blue and green, respectively.

Network	KSGC	GC	WGC	MDD	HKS	KSHI	KS-IF	WKS
Jazz	0.7612	0.7276	0.6184	0.7230	0.7429	0.7445	0.7102	0.7282
USAir	0.7731	0.7231	0.6469	0.6781	0.7603	0.7404	0.7694	0.7722
Netscience	0.5663	0.3708	0.0116	0.3158	0.5446	0.7047	0.3697	0.4095
Political blogs	0.7918	0.7893	0.7808	0.7809	0.7826	0.7740	0.7934	0.7791
PGP	0.6546	0.5494	0.2516	0.4285	0.6764	0.7367	0.5900	0.6294
LFR6000	0.6603	0.6557	0.6525	0.6519	0.6588	0.6607	0.6617	0.6649
UCLA	0.7339	0.7290	0.7307	0.7256	0.7326	0.7277	0.7346	0.7369
Astroph	0.6533	0.6313	0.5560	0.6211	0.6667	0.6662	0.6445	0.6315

CRediT authorship contribution statement

Xuan Yang: Writing - original draft, Investigation, Conceptualization, Methodology, Software, Validation. **Fuyuan Xiao:** Writing - review & editing, Supervision, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors greatly appreciate the reviews' suggestions and the editor's encouragement. This research is supported by the National Natural Science Foundation of China (No. 62003280).

References

[1] R. Albert, A. Barabasi, Statistical mechanics of complex networks, Rev. Modern Phys. 74 (1) (2001) 47–97, http://dx.doi.org/10.1103/RevModPhys.74.47.

[2] L. Lee, P. Hu, Vulnerability analysis of cascading dynamics in smart grids under load redistribution attacks, Int. J. Electr. Power Energy Syst. 111 (2019) 182–190, http://dx.doi.org/10.1016/j.ijepes.2019.03.062.

[3] D. Zhao, Z. Wang, G. Xiao, B. Gao, L. Wang, The robustness of interdependent networks under the interplay between cascading failures and virus propagation, Europhys. Lett. 115 (5) (2016) 58004, http://dx.doi.org/10.1209/0295-5075/115/58004.

[4] Q. Wang, J.S. Hesthaven, D. Ray, Non-intrusive reduced order modeling of unsteady flows using artificial neural networks with application to a combustion problem, J. Comput. Phys. 384 (2019) 289–307, http://dx.doi.org/10.1016/j.jcp.2019.01.031.

[5] T.W. Grimm, C. Li, E. Palti, Infinite distance networks in field space and charge orbits, J. High Energy Phys. 2019 (3) (2019) 16, http://dx.doi.org/10.1007/JHEP03(2019)016.

[6] A. Andreassen, I. Feige, C. Frye, M.D. Schwartz, JUNIPR: a framework for unsupervised machine learning in particle physics, Eur. Phys. J. C 79 (2) (2019) 102, http://dx.doi.org/10.1140/epjc/s10052-019-6607-9.

[7] S. Mao, F. Xiao, Time series forecasting based on complex network analysis, IEEE Access 7 (2019) 40220–40229, http://dx.doi.org/10.1109/ACCESS.2019.2906268.

[8] Z. Tian, Chaotic characteristic analysis of network traffic time series at different time scales, Chaos Solitons Fractals 130 (2020) 109412, http://dx.doi.org/10.1016/j.chaos.2019.109412.

[9] M. Zanin, X. Sun, S. Wandelt, Studying the topology of transportation systems through complex networks: Handle with care, J. Adv. Transp. 2018 (2018) 1–17, http://dx.doi.org/10.1155/2018/3156137.

[10] D. Tsiotas, S. Polyzos, Analyzing the maritime transportation system in Greece: a complex network approach, Netw. Spat. Econ. 15 (4) (2015) 981–1010, http://dx.doi.org/10.1007/s11067-014-9278-y.

[11] Y. Liu, X. Wang, J. Kurths, Framework of evolutionary algorithm for investigation of influential nodes in complex networks, IEEE Trans. Evol. Comput. 23 (6) (2019) 1049–1063, http://dx.doi.org/10.1109/TEVC.2019.2901012.

[12] F. Morone, H.A. Makse, Influence maximization in complex networks through optimal percolation, Nature 524 (7563) (2015) 65–68, http://dx.doi.org/10.1038/nature14604.

[13] L. Sui, J. Yu, D. Cang, W. Miao, H. Wang, J. Zhang, S. Yin, K. Chang, The fractal description model of rock fracture networks characterization, Chaos Solitons Fractals 129 (2019) 71–76, http://dx.doi.org/10.1016/j.chaos.2019.07.055.

[14] Y. Huang, H. Dong, W. Zhang, J. Lu, Stability analysis of nonlinear oscillator networks based on the mechanism of cascading failures, Chaos Solitons Fractals 128 (2019) 5–15, http://dx.doi.org/10.1016/j.chaos.2019.07.030.

[15] W. Li, A. Bashan, S.V. Buldyrev, H.E. Stanley, S. Havlin, Cascading failures in interdependent lattice networks: The critical role of the length of dependency links, Phys. Rev. Lett. 108 (22) (2012) 228702, http://dx.doi.org/10.1103/PhysRevLett.108.228702.

[16] M. Perc, J.J. Jordan, D.G. Rand, Z. Wang, S. Boccaletti, A. Szolnoki, Statistical physics of human cooperation, Phys. Rep. 687 (2017) 1–51, http://dx.doi.org/10.1016/j.physrep.2017.05.004.

[17] J. Zhao, Y. Deng, Complex network modeling of evidence theory, IEEE Trans. Fuzzy Syst. (2020) http://dx.doi.org/10.1109/TFUZZ.2020.3023760.

[18] L.K. Gallos, M. Sigman, H.A. Makse, The conundrum of functional brain networks: Small-world efficiency or fractal modularity, Front. Physiol. 3 (2012) 123, http://dx.doi.org/10.3389/fphys.2012.00123.

[19] B. Wei, F. Xiao, Y. Shi, Synchronization in kuramoto oscillator networks with sampled-data updating law, IEEE Trans. Cybern. 50 (6) (2020) 2380–2388, http://dx.doi.org/10.1109/TCYB.2019.2940987.

[20] A. Zareie, A. Sheikhahmadi, M. Jalili, M.S.K. Fasaei, Finding influential nodes in social networks based on neighborhood correlation coefficient, Knowl. Based Syst. (2020) 105580, http://dx.doi.org/10.1016/j.knosys.2020.105580.

[21] Q. Shang, B. Zhang, H. Li, Y. Deng, Identifying influential nodes: a new method based on network efficiency of edge weight updating, Chaos (2021) http://dx.doi.org/10.1063/5.0033197.

[22] J. Zhao, Y. Song, F. Liu, Y. Deng, The identification of influential nodes based on structure similarity, Connect. Sci. 32 (2020) Accept, http://dx.doi.org/10.1080/09540091.2020.1806203.

[23] L.C. Freeman, Centrality in social networks conceptual clarification, Social Networks 1 (3) (1978) 215–239, http://dx.doi.org/10.1016/0378-8733(78)90021-7.

[24] M.E.J. Newman, A measure of betweenness centrality based on random walks, Social Networks 27 (1) (2003) 39–54, http://dx.doi.org/10.1016/j.socnet.2004.11.009.

[25] G. Sabidussi, The centrality index of a graph, Psychometrika 31 (4) (1966) 581–603, http://dx.doi.org/10.1007/BF02289527.

[26] P. Bonacich, P. Lloyd, Eigenvector-like measures of centrality for asymmetric relations, Social Networks 23 (3) (2001) 191–201, http://dx.doi.org/10.1016/S0378-8733(01)00038-7.

[27] D. Chen, L. Lü, M.-S. Shang, Y.-C. Zhang, T. Zhou, Identifying influential nodes in complex networks, Physica A 391 (4) (2012) 1777–1787, http://dx.doi.org/10.1016/j.physa.2011.09.017.

[28] L. Lü, 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, http://dx.doi.org/10.1016/j.physrep.2016.06.007.

[29] S. Kumar, B.S. Panda, Identifying influential nodes in social networks: Neighborhood coreness based voting approach, Physica A 553 (2020) 124215, http://dx.doi.org/10.1016/j.physa.2020.124215.

[30] M. Alshahrani, Z. Fuxi, A. Sameh, S. Mekouar, S. Huang, Efficient algorithms based on centrality measures for identification of top-k influential users in social networks, Inform. Sci. 527 (2020) 88–107, http://dx.doi.org/10.1016/j.ins.2020.03.060.

[31] T. Wen, D. Pelusi, Y. Deng, Vital spreaders identification in complex networks with multi-local dimension, Knowl.-Based Syst. 195 (2019) 105717, http://dx.doi.org/10.1016/j.knosys.2020.105717.

[32] L. Xiao, S. Wang, G. Mei, Efficient parallel algorithm for detecting influential nodes in large biological networks on the graphics processing unit, Future Gener. Comput. Syst. 106 (2020) 1–13, http://dx.doi.org/10.1016/j.future.2019.12.038.

[33] A. Zareie, A. Sheikhahmadi, M. Jalili, Influential node ranking in social networks based on neighborhood diversity, Future Gener. Comput. Syst. 94 (2019) 120–129, http://dx.doi.org/10.1016/j.future.2018.11.023.

[34] A. Zareie, A. Sheikhahmadi, M. Jalili, Identification of influential users in social networks based on users' interest, Inform. Sci. 493 (2019) 217–231, http://dx.doi.org/10.1016/j.ins.2019.04.033.

[35] A. Sheikhahmadi, M.A. Nematbakhsh, A. Zareie, Identification of influential users by neighbors in online social networks, Physica A 486 (2017) 517–534, http://dx.doi.org/10.1016/j.physa.2017.05.098.

[36] T. Wen, W. Jiang, Identifying influential nodes based on fuzzy local dimension in complex networks, Chaos Solitons Fractals 119 (2018) 332–342, http://dx.doi.org/10.1016/j.chaos.2019.01.011.

[37] J. Zhao, Y. Wang, Y. Deng, Identifying influential nodes in complex networks from global perspective, Chaos Solitons Fractals 133 (2020) 109637, http://dx.doi.org/10.1016/j.chaos.2020.109637.

[38] A. Zareie, A. Sheikhahmadi, M. Jalili, Identification of influential users in social network using grey wolf optimization algorithm, Expert Syst. Appl. 142 (2019) 112971, http://dx.doi.org/10.1016/j.eswa.2019.112971.

[39] J. Zhao, Y. Song, Y. Deng, A novel model to identify the influential nodes: Evidence theory centrality, IEEE Access 8 (1) (2020) 46773–46780, http://dx.doi.org/10.1109/ACCESS.2020.2978142.

[40] Z. Li, T. Ren, X. Ma, S. Liu, T. Zhou, Identifying influential spreaders by gravity model, Entific Rep. 9 (1) (2019) 8387, http://dx.doi.org/10.1038/s41598-019-44930-9.

[41] F. Liu, Z. Wang, Y. Deng, GMM: A generalized mechanics model for identifying the importance of nodes in complex networks, Knowl.-Based Syst. 193 (2020) 105464, http://dx.doi.org/10.1016/j.knosys.2019.105464.

[42] J. Zhang, B. Wang, J. Sheng, J. Dai, J. Hu, L. Chen, Identifying influential nodes in complex networks based on local effective distance, Inf. Int. Interdiscip. J. 10 (10) (2019) 311, http://dx.doi.org/10.3390/info10100311.

[43] 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, http://dx.doi.org/10.1038/nphys1746.

[44] G. Maji, S. Mandal, S. Sen, A systematic survey on influential spreaders identification in complex networks with a focus on K-shell based techniques, Expert Syst. Appl. 161 (2020) 113681, http://dx.doi.org/10.1016/j.eswa.2020.113681.

[45] A. Zeng, C.-J. Zhang, Ranking spreaders by decomposing complex networks, Phys. Lett. A 377 (14) (2013) 1031–1035, http://dx.doi.org/10.1016/j.physleta.2013.02.039.

[46] A. Zareie, A. Sheikhahmadi, A hierarchical approach for influential node ranking in complex social networks, Expert Syst. Appl. 93 (2018) 200–211, http://dx.doi.org/10.1016/j.eswa.2017.10.018.

[47] G. Maji, A. Namtirtha, A. Dutta, M.C. Malta, Influential spreaders identification in complex networks with improved k-shell hybrid method, Expert Syst. Appl. 144 (2020) http://dx.doi.org/10.1016/j.eswa.2019.113092.

[48] A. Namtirtha, A. Dutta, B. Dutta, Identifying influential spreaders in complex networks based on kshell hybrid method, Phys. A 499 (2018) 310–324, http://dx.doi.org/10.1016/j.physa.2018.02.016.

[49] Z. Wang, Y. Zhao, J. Xi, C. Du, Fast ranking influential nodes in complex networks using a k-shell iteration factor, Phys. A 461 (2016) 171–181, http://dx.doi.org/10.1016/j.physa.2016.05.048.

[50] G. Maji, Influential spreaders identification in complex networks with potential edge weight based k-shell degree neighborhood method, J. Comput. Sci. 39 (2020) 101055, http://dx.doi.org/10.1016/j.jocs.2019.101055.

[51] P.M. Gleiser, L. Danon, Community Structure in Jazz, Adv. Complex Syst. 06 (04) (2003) 565–573, http://dx.doi.org/10.1142/S0219525903001067.

[52] V. Colizza, R. Pastor-Satorras, A. Vespignani, Reaction–diffusion processes and metapopulation models in heterogeneous networks, Nat. Phys. 3 (4) (2007) 276–282, http://dx.doi.org/10.1038/nphys560.

[53] M.E.J. Newman, Finding community structure in networks using the eigenvectors of matrices, PhRvE 74 (2006) http://dx.doi.org/10.1103/PhysRevE.74.036104.

[54] L.A. Adamic, N. Glance, The political blogosphere and the 2004 U.S. election: Divided they blog, in: Proceedings of the 3rd International Workshop on Link Discovery, 2005, pp. 36–43, http://dx.doi.org/10.1145/1134271.1134277.

[55] M. Boguñá, R. Pastor-Satorras, A. Díaz-Guilera, A. Arenas, Models of social networks based on social distance attachment, Phys. Rev. E 70 (2004) 056122, http://dx.doi.org/10.1103/PhysRevE.70.056122.

[56] A.L. Traud, E.D. Kelsic, P.J. Mucha, M.A. Porter, Comparing community structure to characteristics in online collegiate social networks, SIAM Rev. 53 (3) (2011) 526–543, http://dx.doi.org/10.1137/080734315.

[57] J. Leskovec, J. Kleinberg, C. Faloutsos, Graph evolution: Densification and shrinking diameters, ACM Trans. Knowl. Discov. Data 1 (1) (2007) 2–es, http://dx.doi.org/10.1145/1217299.1217301.

[58] A. Lancichinetti, S. Fortunato, F. Radicchi, Benchmark graphs for testing community detection algorithms, Phys. Rev. E 78 (2008) 046110, http://dx.doi.org/10.1103/PhysRevE.78.046110.

[59] M. Jalili, M. Perc, Information cascades in complex networks, J. Complex Netw. 5 (5) (2017) 665–693, http://dx.doi.org/10.1093/comnet/cnx019.

[60] D. Watts, A simple model of global cascades on random networks, Proc. Natl. Acad. Sci. U S A 99 (9) (2002) 5766–5771, http://dx.doi.org/10.1073/pnas.082090499.

[61] K. Saito, R. Nakano, M. Kimura, Prediction of information diffusion probabilities for independent cascade model., Lecture Notes in Comput. Sci. 5179 (2008) 67–75, http://dx.doi.org/10.1007/978-3-540-85567-5_9.

[62] I. Nsell, The quasi-stationary distribution of the closed endemic sis model, Adv. Appl. Probab. 28 (3) (1996) 895–932, http://dx.doi.org/10.1017/S0001867800046541.

[63] C. Castellano, R. Pastor-Satorras, Thresholds for epidemic spreading in networks, Phys. Rev. Lett. 105 (2010) 218701, http://dx.doi.org/10.1103/PhysRevLett.105.218701.
[64] M.G. Kendall, A new measure of rank correlation, Biometrika 30 (1–2) (1938) 81–93, http://dx.doi.org/10.2307/2332226.
[65] A. Garas, F. Schweitzer, S. Havlin, A k-shell decomposition method for weighted networks, New J. Phys. 14 (8) (2012) 083030, http://dx.doi.org/10.1088/1367-2630/14/8/083030.
[66] M. Eidsaa, E. Almaas, *s*-Core network decomposition: A generalization of *k*-core analysis to weighted networks, Phys. Rev. E 88 (2013) 062819, http://dx.doi.org/10.1103/PhysRevE.88.062819.