

Influential nodes identification in complex networks based on global and local information

Yuan-Zhi Yang(杨远志), Min Hu(胡敏), Tai-Yu Huang(黄泰愚)
Citation: Chin. Phys. B . 2020, 29(8): 088903 . **doi:** 10.1088/1674-1056/ab969f
Journal homepage: http://cpb.iphy.ac.cn; http://iopscience.iop.org/cpb

What follows is a list of articles you may be interested in

Identifying influential spreaders in complex networks based on entropy weight method and gravity law

Xiao-Li Yan(闫小丽), Ya-Peng Cui(崔亚鹏), Shun-Jiang Ni(倪顺江)
Chin. Phys. B . 2020, 29(4): 048902 . **doi:** 10.1088/1674-1056/ab77fe

Theoretical analyses of stock correlations affected by subprime crisis and total assets: Network properties and corresponding physical mechanisms

Shi-Zhao Zhu(朱世钊), Yu-Qing Wang(王玉青), Bing-Hong Wang(汪秉宏)
Chin. Phys. B . 2019, 28(10): 108901 . **doi:** 10.1088/1674-1056/ab3f22

Pyramid scheme model for consumption rebate frauds

Yong Shi(石勇), Bo Li(李博), Wen Long(龙文)
Chin. Phys. B . 2019, 28(7): 078901 . **doi:** 10.1088/1674-1056/28/7/078901

Coordinated chaos control of urban expressway based on synchronization of complex networks

Ming-bao Pang(庞明宝), Yu-man Huang(黄玉满)
Chin. Phys. B . 2018, 27(11): 118902 . **doi:** 10.1088/1674-1056/27/11/118902

Dominant phase-advanced driving analysis of self-sustained oscillations in biological networks

Zhi-gang Zheng(郑志刚), Yu Qian(钱郁)
Chin. Phys. B . 2018, 27(1): 018901 . **doi:** 10.1088/1674-1056/27/1/018901

Influential nodes identification in complex networks based on global and local information

Yuan-Zhi Yang(杨远志)[1], Min Hu(胡敏)[2,†], and Tai-Yu Huang(黄泰愚)[3,‡]

[1] *Air Force Engineering University, Xi'an* 710038, *China*
[2] *China Petroleum Planning and Engineering Institute, Beijing* 100083, *China*
[3] *Sichuan University of Arts and Science, Dazhou* 635000, *China*

(Received 23 April 2020; revised manuscript received 18 May 2020; accepted manuscript online 27 May 2020)

Identifying influential nodes in complex networks is essential for network robust and stability, such as viral marketing and information control. Various methods have been proposed to define the influence of nodes. In this paper, we comprehensively consider the global position and local structure to identify influential nodes. The number of iterations in the process of k-shell decomposition is taken into consideration, and the improved k-shell decomposition is then put forward. The improved k-shell decomposition and degree of target node are taken as the benchmark centrality, in addition, as is well known, the effect between node pairs is inversely proportional to the shortest path length between two nodes, and then we also consider the effect of neighbors on target node. To evaluate the performance of the proposed method, susceptible-infected (SI) model is adopted to simulate the spreading process in four real networks, and the experimental results show that the proposed method has obvious advantages over classical centrality measures in identifying influential nodes.

Keywords: complex networks, influential nodes, global position, local structure, susceptible-infected (SI) model

PACS: 89.75.Fb **DOI:** 10.1088/1674-1056/ab969f

1. Introduction

Complex networks have received increasing attention in recent years. A number of systems in real life can be represented abstractly by complex networks, such as social network,[1,2] traffic system,[3,4] power grid,[5,6] and biological network.[7,8] A small number of special nodes, called influential nodes, will have a significant influence on the structure and function of the network. Structural damage or functional failure of influential nodes will cause the entire network to crash. Therefore, research on identifying influential nodes is of great significance for improving the robustness of systems and designing an efficient system structure. On the one hand, we can control and protect influential nodes to stabilize and secure the complex systems. For example, we can control some famous users in social network to prevent rumors from spreading, finding and protecting important airports in aviation network can avoid the large area delay of routes, protecting key substations in power grid can prevent large-scale power outages. On the other hand, we can also attack influential nodes to destroy the structure and function of a network. For instance, the Stuxnet virus, developed by the United States and Israel, damaged critical facilities of Iran's nuclear power network and successfully attacked and controlled Iran's nuclear power network.

Research of influential nodes identification has two lines: (i) influence ranking of individual node and (ii) identifying a group of nodes to maximize the influence. In the present paper, we focus on the former. Up to now, various centrality measures have been proposed to address this issue, such as degree centrality (DC),[9] closeness centrality (CC),[10] betweenness centrality (BC),[11] eigenvector centrality,[12] local centrality (LC),[13] k-shell decomposition (Ks),[14] *etc*. Degree centrality is a well-known local metric that determines the influence of node by the number of its nearest neighbors, but it only considers limited local information and ignores the effect of global information. Closeness centrality and betweenness centrality are classical measures belonging to global metric. They have high time complexity because they need to obtain the topological characteristics of the entire network, and thus they cannot be applied to large scale networks. To make a tradeoff between local metric and global metric, Chen *et al.*[13] proposed a local centrality to identify influential nodes in which the target node's neighbors within 4-hops are taken into consideration. Kitsak *et al.*[14] found that the influence of node is determined by node's location in the network and proposed k-shell decomposition centrality. The k-shell decomposition strips the nodes in the outer layer, and the nodes in inner layer are influential nodes in the network. However, this method is a coarse-grained sorting algorithm and cannot further distinguish the influences of nodes. Then many methods were put forward to modify the k-shell decomposition. Zeng and Zhang[15] considered the degree information of removing nodes and put forward the mixed degree decomposition method. Lin *et al.*[16] considered neighbors' k-core value and proposed an improved

†Corresponding author. E-mail: hu_min_min@163.com
‡Corresponding author. E-mail: 1097762865@qq.com

© 2020 Chinese Physical Society and IOP Publishing Ltd

k-shell decomposition method. In the field of search engine, a number of approaches have been put forward to identify influential nodes such as PageRank,[17] LeaderRank,[18] and Hits,[19] in these iterative refinement methods the influence of neighbors is taken into account. Note that centrality measure cannot work well from a single view, then some researchers indicated that combining several centrality measures to identify influential nodes can obtain more accuracy ranking results. The technique for order preference by similarity to an ideal object (TOPSIS) method was used to identify influential nodes in two actual networks,[20] Yang *et al.*[21] used Vlsekriterijumska Optimizacija I Kompromisno Resenje (VIKOR) method to combine several classical centrality measures, Dempster–Shafer evidence theory was used to comprehensively combine degree centrality, betweenness centrality, and closeness centrality.[22] The global position and local degree variation were both considered to identify influential nodes, and entropy method was used to assign weights to the two aspects.[23] Wang *et al.*[24] combined the improved k-shell decomposition and the neighborhood's effect to give an influential nodes identifying method with low time complexity and high accuracy. Zhao *et al.*[25] considered both local centrality and clustering coefficient, and provided a new method to identify influential nodes. In a word, it is still an open issue to establish a model for identifying the influential nodes.

In this paper, a novel method of identifying influential nodes, called GLI, is proposed based on global and local information. The greatest issue with k-shell decomposition is that it produces many nodes with the same ranking, we consider the number of iterations in the process of k-shell decomposition and put forward the improved k-shell decomposition. The improved k-shell decomposition and degree of target node are chosen as the benchmark centrality, and it is considered in the GLI that the effect between nodes will decrease with the increase of the shortest path length between nodes, in addition, neighbors within 3-hops are considered. To evaluate the performance of GLI, the susceptible-infected (SI) model[26] is utilized to simulate the spreading process in four real networks, and other classical centrality measures (DC, CC, BC, and Ks) are employed to compare the GLI. The superiority of GLI is demonstrated according to four experiments.

The rest of this paper is organized as follows. The related work is introduced in Section 2 and GLI is proposed in Section 3. In Section 4, four experiments are conducted based on four real networks and the superiority of GLI is proved. In Section 5 some conclusions are drawn from the present study.

2. Related work

An undirected and unweighted graph is represented as $G=(V,E,A)$, $V=\{v_1,v_2,\ldots,v_n\}$ denotes the set of nodes and $E=\{e_1,e_2,\ldots,e_m\}$ refers to the set of edges, $A=\{a_{ij}\}$ is the adjacency matrix of G: if node i and node j are connected, $a_{ij}=1$; otherwise, $a_{ij}=0$.

2.1. Degree centrality

Degree centrality[9] is the simplest and most famous centrality for identifying influential nodes, but it only considers the number of its nearest neighbors, which is defined as

$$\mathrm{DC}(i)=\sum_{j}^{n}a_{ij}, \tag{1}$$

where n is the total number of nodes, the greater the degree of a node, the more influential the node is.

2.2. Closeness centrality

Closeness centrality[10] of node i is defined as the reciprocal of the sum of the shortest distances from node i to others in the network, it is expressed as

$$\mathrm{CC}(i)=\frac{1}{\sum_{j=1}^{n}d_{ij}}, \tag{2}$$

where d_{ij} represents the shortest path between node i and node j. Greater closeness value of a node implies that the node is more influential.

2.3. Betweenness centrality

Betweenness centrality[11] is a classical global metric, which calculates the number of the shortest paths passing through the node. It is defined as

$$\mathrm{BC}(i)=\sum_{s,t\neq i}\frac{g_{st}(i)}{g_{st}}, \tag{3}$$

where $g_{st}(i)$ is the number of shortest paths between node s and node t passing through node i, and g_{st} represents the number of all possible shortest paths. A node with greater betweenness value is more influential.

2.4. The k-shell decomposition

The k-shell decomposition[14] assigns an index (ks) to each node. Assume that there is no isolated node in the network, and take Fig. 1 for example. Firstly, all the nodes with degree DC = 1 (node 1) and their connected edges are removed, at this time, some new nodes with degree DC = 1 (node 2) may appear in the network. We continue to remove these nodes and their connected edges, and repeat this operation until no node with degree DC = 1 exists in the network, all the removed nodes are assigned with an index $ks=1$. Next, the nodes with degree DC = 2 (node 3, node 4, node 10, node 11, and node 9) and their connected edges are removed in a similar way, and the removed nodes are assigned with an index $ks=2$. The decomposition process will stop until all the nodes are removed, and all the nodes are assigned with a ks index finally.

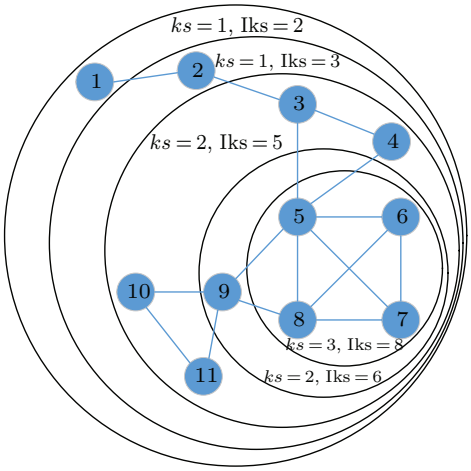


Fig. 1. A simple example graph to explain the identifying process of k-shell decomposition and improved k-shell decomposition.

2.5. Improved k-shell decomposition

In Fig. 1, the influence of node 1 and node 2 are obviously different, but k-shell decomposition assigns them with the same ks value. We can find that node 2 is removed after removing node 1, that is, the number of iterations in the process of k-shell decomposition of node 1 and node 2 are different. The same situation exists for node 9 and node 10. Therefore, we come up with the improved k-shell decomposition, which considers the number of iterations and ks value. We define the improved k-shell decomposition (Iks) as

$$\mathrm{Iks}(i) = ks(i) + \mathrm{nit}(i), \tag{4}$$

where $\mathrm{nit}(i)$ represents the number of iterations of node i.

The improved k-shell decomposition in Fig. 1 is shown in Table 1. We can find that the improved k-shell decomposition can identify the difference among nodes more effectively.

Table 1. The improved k-shell decomposition in Fig. 1.

The number of iterations	Removed nodes	ks	Iks
1	1	1	2
2	2	1	3
3	3, 4, 10, 11	2	5
4	9	2	6
5	5, 6, 7, 8	3	8

3. Proposed method

Various centrality measures have been proposed to determine an actor to be 'influential' from their perspectives, which will lead to some limitations for each measure. However, the accuracy and effectiveness of influential nodes identification can be improved by considering the influence of nodes from multiple perspectives. The global position and degree are classical centrality measures which consider global information and local information respectively, and we propose to integrate these two centrality measures. Besides, the main limitation of k-shell decomposition is that it leads to a few nodes with same ranking, then we propose the improved k-shell decomposition based on the number of iterations. What is more, we also consider the effect of neighbors, in fact, the effect between node pairs is inversely proportional to the shortest path length between two nodes. Therefore, we propose a novel measure based on global and local information (denoted as GLI). The GLI considers that the effect between nodes will decrease with the increase of the shortest path length between nodes, the improved k-shell decomposition and degree are chosen as the benchmark centrality, and we consider the neighbors within 3-hops. GLI is defined as

$$I_i = \exp\left[(\mathrm{Iks}(i) + \mathrm{DC}(i)) \Big/ \sum_{i=1}^{n} (\mathrm{Iks}(i) + \mathrm{DC}(i))\right] \times \sum_{j \in \Gamma_i^r} \frac{\mathrm{Iks}(j) + \mathrm{DC}(j)}{d_{ij}}, \tag{5}$$

where $\mathrm{Iks}(i)$ and $\mathrm{DC}(i)$ denote the improved k-shell value and degree of node i, respectively; d_{ij} is the shortest path length between node i and node j; Γ_i^r denotes the neighborhood set whose distance to node i is less than or equal to a given value r; to reduce the time complexity, we set $r = 3$, *i.e.*, we only consider the neighbors within 3-hops.

The GLI considers the global position and local structure, and the neighbors within 3-hops are also taken into consideration. We provide a new method to identify influential nodes from our perspective. Compared with classical centrality measures, GLI can give very accuracy and effective ranking results.

4. Experimental results

4.1. Datasets

To verify the efficiency of GLI, we conduct the experiments with the following four real networks. They are (i) Karate club network,[27] which is a social network describing the relationship among 34 members of a Karate Club in the 1970s; (ii) Jazz musicians network,[28] which is a collaboration network among 198 Jazz musicians, with each edge representing two musicians play together in a band; (iii) USAir97 network, which is an aviation network with each node denoting an airport, and each edge representing a route between two airports; (iv) Email network,[29] which is a social network with each node denoting a user, and each edge representing e-mail interchange between members.

4.2. SI model

We employ susceptible-infected (SI) model[26] to simulate the spreading process and evaluate the performance of GLI. Each node in the SI model is either in susceptible state or in infected state. Nodes in susceptible state will be infected by nodes in infected state with a certain probability, and once susceptible nodes are infected, they cannot be recovered. Suppose that only one node (node i) in the network is infected at

$t = 0$, namely, node i is the source node of spreading process, and all the other nodes are in susceptible state. Then the source node infects its neighbors with a certain probability, and at the same time, the disease or information begins to spread through the network. The number of infected nodes will be n_{it} after t $(t = 1, 2, \ldots)$ time step, and we define the spreading ability of node i as $I_i(t) = n_{it}/n$, besides, we set the maximum time step to be $t = 50$. To eliminate random error, we run 1000 independent simulations.

4.3. Kendall's tau coefficient

The Kendall's tau coefficient[30] is adopted to measure the correlation coefficient between centrality measures and the actual ranking simulated by SI model. Suppose that there are two ranking lists X and Y, $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$ is a set of joint ranks from the two ranking lists. A pair of distinct nodes i and j is concordant 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 $(y_i < y_j))$ or $((x_i < x_j)$ and $(x_i > x_j))$, then the node pair is said to be discordant. And the Kendall's tau coefficient τ is defined as

$$\tau = \frac{n_c - n_d}{0.5 \times n \times (n-1)}, \tag{6}$$

where n_c and n_d are the number of concordant pairs and discordant pairs, respectively. Generally, $\tau \in [-1, 1]$, $\tau < 0$ indicates negative correlation and $\tau > 0$ represents that two lists are of positive correlation. The greater the τ value, the better the performance of a measure is.

4.4. Effectiveness

4.4.1. Experiment 1 for comparing top-10 lists generated by different methods

We rank the influences of nodes in four networks with GLI, and other classical centrality measures including DC, CC, BC, and Ks are applied to the same networks for comparison. We choose the top-10 lists generated by each method. The results to be ranked are presented in Table 2. In addition, we run SI model to obtain the actual spreading ability ranking $I = \text{sort}(I_1, I_2, \ldots, I_n)$, which is used to compare the top-10 lists from different methods.

Table 2. Comparisons among top-10 lists generated by different methods in four networks. Actual spreading ability ranking $I = \text{sort}(I_1, I_2, \ldots, I_n)$ is obtained by SI model with 1000 independent simulations.

Karate club							Jazz musicians						
Rank	DC	CC	BC	Ks	GLI	I	Rank	DC	CC	BC	Ks	GLI	I
1	34	1	1	1	1	1	1	136	136	136	60	60	136
2	1	3	34	34	3	34	2	60	60	153	168	136	149
3	33	34	33	33	34	3	3	132	168	60	108	132	96
4	3	32	3	31	33	33	4	168	70	149	122	168	60
5	2	9	32	3	14	2	5	70	83	168	33	108	194
6	4	14	9	14	2	9	6	99	132	167	58	99	70
7	32	33	2	2	9	14	7	108	122	189	66	70	168
8	9	20	14	4	32	32	8	83	194	115	100	131	132
9	14	2	20	8	4	4	9	158	174	96	132	194	83
10	24	4	6	9	20	20	10	7	158	83	179	83	99
USAir97							Email						
Rank	DC	CC	BC	Ks	GLI	I	Rank	DC	CC	BC	Ks	GLI	I
1	118	118	118	67	118	118	1	105	333	333	299	105	333
2	261	261	8	94	261	261	2	333	23	105	389	333	105
3	255	67	261	109	255	147	3	16	105	23	434	42	42
4	152	255	201	112	182	166	4	23	42	578	552	23	41
5	182	201	47	118	152	67	5	42	41	76	571	76	54
6	230	182	182	131	230	47	6	41	76	233	726	41	52
7	166	47	255	146	166	182	7	196	233	135	756	233	233
8	67	166	152	147	67	152	8	233	52	41	788	3	3
9	112	248	313	150	112	255	9	21	135	355	885	52	76
10	201	112	13	152	147	112	10	76	378	42	886	378	23

According to Table 2, the top-10 lists ranked by different methods are different. Here we focus on the differences between the results from each method and the actual spreading ability ranking. In Karate Club network, the top-10 nodes of GLI and CC are all the same with I, DC and BC have 9 identical nodes in the top-10 lists with I, and Ks only has 8 identical nodes. In Jazz musicians network, GLI owning 8 identical nodes with I performs best, DC and CC have 7 identical nodes in the top-10 lists with I, there are 6 identical nodes between BC and I, while Ks only has 3 identical nodes. In USAir97 network, there are 8, 8, 6, 5, 9 identical nodes in the top-10 lists among DC, CC, BC, Ks, GLI, and I, respectively. As for Email network, GLI has 9 identical nodes with I, but there is no identical node between Ks and I, and the DC, CC, and BC

have 7, 8, and 7 identical nodes with I, respectively. From the above analysis, we can find that GLI has the most identical nodes in the top-10 lists with the actual spreading ability ranking, GLI can identify more influential nodes in the top-10 lists than other methods.

4.4.2. Experiment 2 for comparing frequencies of nodes with same ranking

When ranking the influence of nodes with different methods, we find that some nodes are given the same ranking, which makes it impossible to distinguish them. In addition, the indistinguishable degrees of nodes generated by different methods are not the same. Therefore, we focus on the frequencies of nodes with the same ranking of each method in four networks, the higher the frequency, the more the nodes in the same ranking are and the worse the distinguishing ability. Figure 2 shows the frequencies of nodes with the same ranking of each method in four networks. Of the frequencies of the four networks, the frequency of GLI is the lowest, but DC and Ks have the highest frequency, which demonstrates that GLI has the best distinguishing ability.

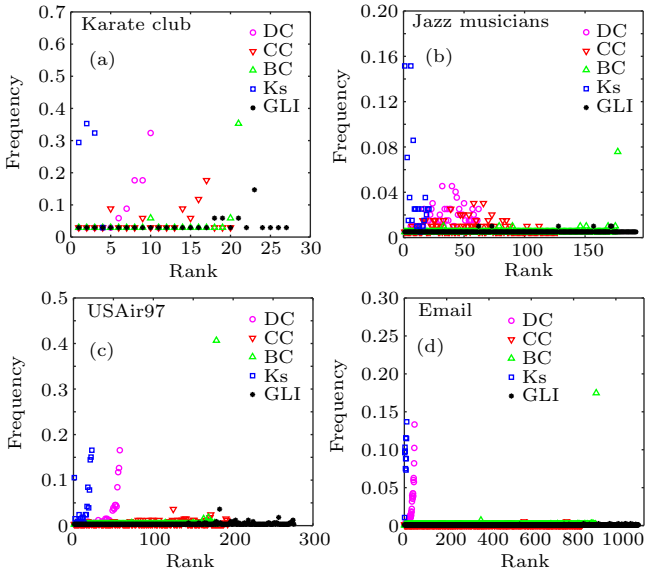


Fig. 2. Frequencies of nodes with the same ranking in four networks.

To further analyze the distinguishing ability of different methods, we introduce the monotonicity index $M(R)$ and it is defined as[31]

$$M(R) = \left[1 - \frac{\sum_{r\in R} n_r(n_r-1)}{n(n-1)}\right]^2, \tag{7}$$

where n is the number of nodes in a network, n_r represents the number of nodes with the same ranking value r, and R denotes a ranking list. The greater the $M(R)$ value, the better the monotonicity is. The $M(R) = 1$ indicates that each node is assigned with a different ranking. The monotonicity indices of different methods in four networks are shown in Table 3.

Table 3. Monotonicity indices of different methods in four networks.

Network	M(DC)	M(CC)	M(BC)	M(Ks)	M(GLI)
Karate club	0.7079	0.8993	0.7723	0.4958	0.9542
Jazz musicians	0.9659	0.9878	0.9885	0.7944	0.9994
USAir97	0.8586	0.9892	0.6968	0.8114	0.9951
Email	0.8874	0.9988	0.9399	0.8088	0.9999

From Table 3, of the monotonicity values from the above-mentioned methods, the monotonicity value of GLI is the greatest, and almost in the four networks, the monotonicity of Ks is the worst. In the four networks, the fluctuation of monotonicity for Ks is the most apparent. The difference between monotonicity values of Ks in Karate club network and Jazz musicians network is approximately 0.3, and the GLI achieves the greatest and most stable value in the four networks. Then we can conclude that the GLI has better distinguishing ability than other methods.

4.4.3. Experiment 3 for comparing average spreading ability of top-10 nodes

In this part, we compare the average spreading ability of top-10 nodes ranked by each method in four networks. The average spreading ability of top-10 nodes is defined as $I(t) = \sum_{i\in \text{top}-10} I_i(t)$ and we set the maximum time step to be $t = 50$. The SI model is used to simulate the spreading process in a network and 1000 independent simulations are conducted to eliminate randomness. Figure 3 shows the spreading ability top-10 nodes of each method in four networks. The greater the infected nodes, the stronger the spreading ability is.

According to Fig. 3, the infected nodes increase with time step increasing and finally reaches a stable value. In general, the GLI achieves better spreading performance than other methods almost in the four networks. Specifically, in the Karate club network, the curves of GLI and CC overlap because their top-10 nodes are the same, and they have more infected nodes than DC, Ks, and BC at each time step. In the Jazz musicians network, the average spreading ability of GLI is slightly stronger than that of CC, and significantly stronger than those of DC, Ks, and BC. In USAir97 network, GLI and DC have the similar average spreading ability because their curves almost overlap, and they slightly outperform CC and Ks, but BC has the worst performance in average spreading ability. As for the Email network, the average spreading ability of GLI is marginally stronger than those of CC, DC, and BC, and the infected nodes of Ks are less than those of other methods at each time step. Overall speaking, the average spreading ability of GLI is stronger than those of other methods, which indicates that the GLI can identify influential nodes more effectively.

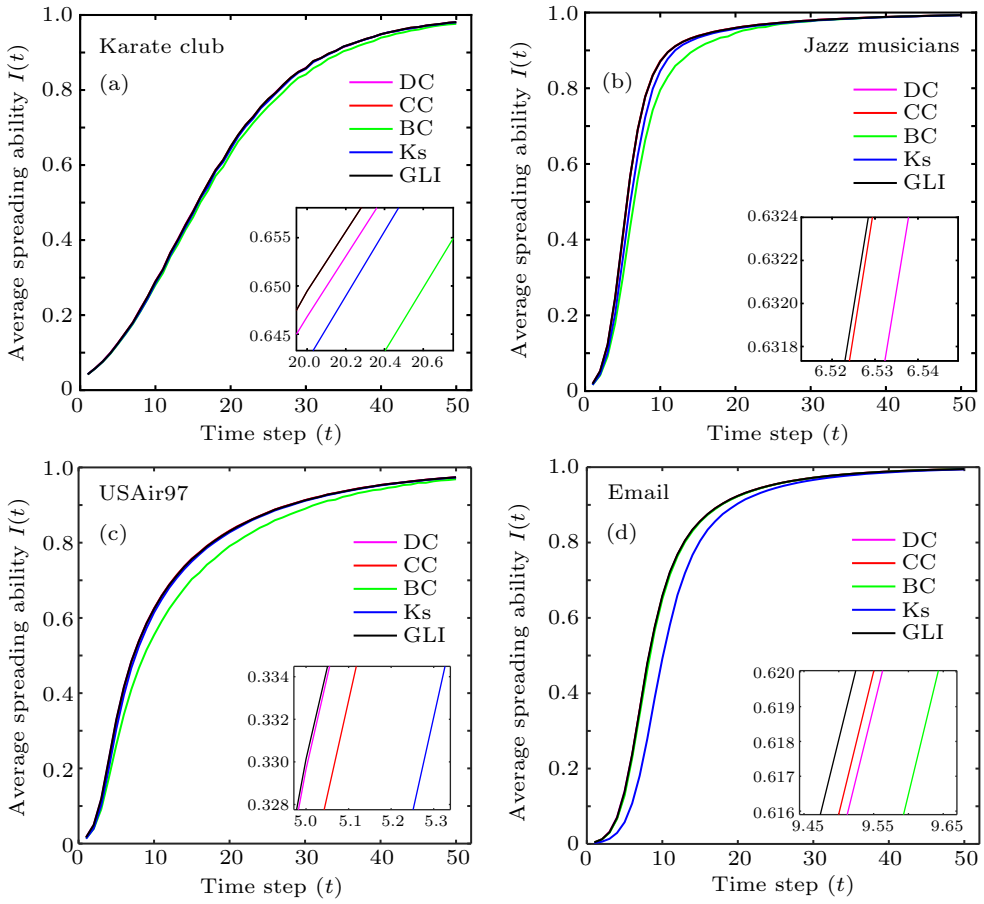


Fig. 3. Average spreading ability of top-10 nodes ranked by each method in four networks.

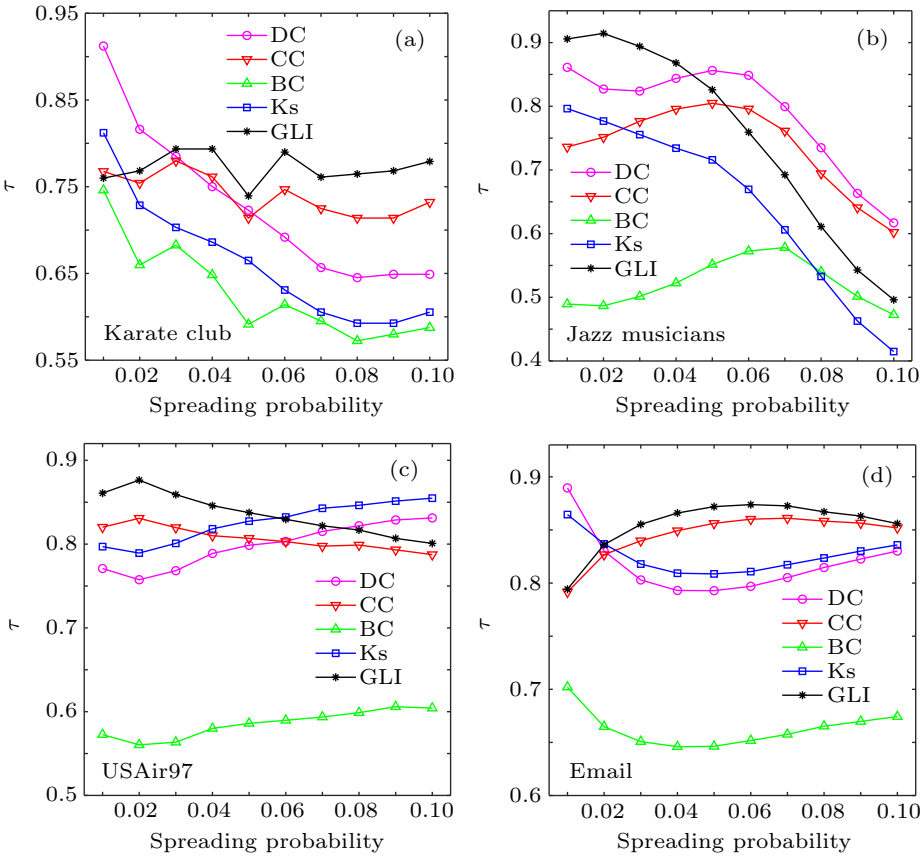


Fig. 4. Kendall's tau coefficients from different methods and actual ranking γ that is simulated by SI model at $t = 20$ with 1000 independent runs.

4.4.4. Experiment 4 for comparing Kendall's tau coefficients between different methods and the actual ranking list γ

In this part, we pay attention to the correlation between methods and the actual ranking list. The actual ranking list γ simulated by SI model is taken as an actual spreading situation, and the other ranking lists generated by each method are compared with γ. Figure 4 shows the correlation between each method and γ when the spreading probability varies from 0 to 0.1. In the Karate club network, the Kendall's tau coefficient τ of GLI is higher than those of other methods when the spreading probability is greater than 0.03. In the Jazz musicians network, when the spreading probability is less than 0.04, the GLI has a higher correlation with γ than others, but the τ values of DC and CC are greater than that of GLI with the increase of spreading probability. Under the same condition, for the USAir97 network when the spreading probability is less than 0.06, the GLI has a higher correlation with γ than others. As for Email network, GLI has higher τ value than other methods in almost the entire spreading probability. The BC has the weakest correlation with γ in the four networks. The above analyses indicate that GLI can accurately identify the influential nodes.

5. Conclusions and perspectives

In this paper, various methods are proposed to identify influential nodes in complex networks from different perspectives, each method has its own viewpoint to define an actor to be 'influential'. We comprehensively consider global and local information to identify influential nodes, and the GLI method is then put forward. Firstly, we improve k-shell decomposition through considering the number of iterations in the process of k-shell decomposition, and the improved k-shell decomposition is considered as a global structure. Then, degree centrality is considered as a local structure, and the combination of global and local structure is taken as the benchmark centrality. Finally, we consider the effect of neighbors within 3-hops on the target node, and GLI method is established. The GLI comprehensively combines global position and local degree, thus providing a more accurate and effective approach to identifying the influential nodes. We utilize the SI model to simulate the spreading process in network, and four experiments are conducted to demonstrate the superiority of GLI. Experimental results show that the GLI method has obvious advantages over the classical centrality measures.

Acknowledgment

The authors are grateful to the reviewers for their valuable comments and suggestions.

References

[1] Shao P and Chen H 2019 *IEEE Access* **7** 118509
[2] Arularasan A N, Suresh A and Seerangan K 2019 *Cluster Computing* **22** 4035
[3] Shukla A, Bhattacharyya A, Kuppusamy L, Srivas M and Thattai M 2017 *Plos One* **12** e0180692
[4] Mou J, Liu C, Chen S, Huang G and Lu X 2017 *Sci. Rep.* **7** 1275
[5] Tan Z Z, Asad J H and Owaidat M Q 2017 *Int. J. Circuit Theor. Appl.* **45** 1942
[6] Owaidat M Q, Al-Badawi A A, Asad J H and Al-Twiessi Mohammed 2018 *Chin. Phys. Lett.* **35** 020502
[7] Chasman D, Siahpirani A F and Roy S 2016 *Current Opinion in Biotechnology* **39** 157
[8] Yuan M, Hong W and Li P 2017 *Biochemical Society Transactions* **45** 1015
[9] Bonacich P 1972 *Journal of Mathematical Sociology* **2** 113
[10] Freeman L C 1978 *Social Networks* **1** 215
[11] Newman M E J 2005 *Social Networks* **27** 39
[12] Bonacich P and Lloyd P 2001 *Social Networks* **23** 191
[13] Chen D, Lü L, Shang M S, Zhang Y C and Zhou T 2012 *Physica A* **391** 1777
[14] Kistak M, Gallos L K, Havlin S, Liljeros F, Muchnik L, Stanley H E and Makse H A 2010 *Nat. Phys.* **6** 888
[15] Zeng A and Zhang C J 2013 *Phys. Lett. A* **377** 1031
[16] Lin J H, Guo Q, Dong W Z, Tang L Y and Liu J G 2014 *Phys. Lett. A* **378** 3279
[17] Arasu A, Cho J, Garcia-Molina H, Paepcke A and Raghavan S 2001 *ACM Transactions on Internet Technology* **1** 2
[18] Lü L, Zhang Y C, Yeung C H and Zhou T 2011 *Plos One* **6** e21202
[19] Kleinberg J M 1999 *J. ACM* **46** 604
[20] Liu Z, Jiang C, Wang J and Yu H 2015 *Knowledge-Based Systems* **84** 56
[21] Yang Y, Yu L, Wang X, Zhou Z, Chen Y and Kou T 2019 *Physica A* **526** 121118
[22] Mo H, Gao C and Deng Y 2015 *Journal of Systems Engineering and Electronics* **26** 381
[23] Ibnoulouafi A, Haziti M E and Cherifi H 2018 *Journal of Statistical Mechanics Theory and Experiment* **7** 073407
[24] Wang Z, Du C, Fan J and Xing Y 2017 *Neurocomputing* **260** 466
[25] Zhao X, Liu F, Wang J and Li T 2017 *International Journal of Geo-Information* **6** 35
[26] Zhou T, Liu J G, Bai W J, Chen G and Wang B H 2006 *Phys. Rev. E* **74** 056109
[27] Zachary W W 1977 *Journal of Anthropological Research* **33** 452
[28] Gleiser P M and Danon L 2003 *Advances in Complex Systems* **6** 565
[29] Guimerà R, Danon L, Díaz-Guilera A, Giralt F and Arenas A 2003 *Phys. Rev. E* **68** 065103
[30] Kendall M G 1938 *Biometrika* **30** 81
[31] Bae J and Kim S 2014 *Physica A* **395** 549