Identifying influential nodes in complex networks: A node information dimension approach

Tian Bian, and Yong Deng

Citation: Chaos **28**, 043109 (2018); doi: 10.1063/1.5030894
View online: https://doi.org/10.1063/1.5030894
View Table of Contents: http://aip.scitation.org/toc/cha/28/4
Published by the American Institute of Physics

Identifying influential nodes in complex networks: A node information dimension approach

Tian Bian[1,2] and Yong Deng[1,2,a)]
[1]Institute of Fundamental and Frontier Science, University of Electronic Science and Technology of China, Chengdu 610054, China
[2]School of Computer and Information Science, Southwest University, Chongqing 400715, China

(Received 22 March 2017; accepted 26 March 2018; published online 10 April 2018)

In the field of complex networks, how to identify influential nodes is a significant issue in analyzing the structure of a network. In the existing method proposed to identify influential nodes based on the local dimension, the global structure information in complex networks is not taken into consideration. In this paper, a node information dimension is proposed by synthesizing the local dimensions at different topological distance scales. A case study of the Netscience network is used to illustrate the efficiency and practicability of the proposed method. *Published by AIP Publishing.* https://doi.org/10.1063/1.5030894

How to identify the influential nodes in complex networks is one of the most important research directions in complex networks. Many methods have been proposed to identify the influential nodes in complex networks. Recently, a methodology has been proposed by obtaining the local dimension (LD) at different scales for each node. Based on the method, a measurement to identify influential nodes in the networks by comparing each node's local dimension (LD) at the farthest distance is proposed. However, it just considers the local information of the node. In this paper, a node information dimension (NID) is proposed to identify influential nodes by synthesizing each segmental local dimension. The results show that the identification of the influential nodes is of theoretical and practical significance.

I. INTRODUCTION

In recent years, because of its great theoretical significance and wide application, the evaluation of node importance in complex networks has attracted much attention, such as in the control of the disease spreading,[1–3] creating new marketing tools,[4–6] and research on public opinion and rumor dynamics.[7–10] As the requisite tool to identify the influence of nodes in complex networks, many centrality measures have been used commonly, such as the Degree Centrality (DC),[11] Betweenness Centrality (BC),[11] Closeness Centrality (CC),[11] Eigenvector Centrality (EC),[12] PageRank (PR),[13] LeaderRank (LR),[14] and many other methods.[15–20] The DC method is concise but does not take the global characteristic of the network into consideration. Although BC and CC are global metrics, they are difficult to be applied in large-scale networks because of their computational complexity.

Fractal is a fundamental property in physics and it can characterize complex graphics and the complex process of dynamical systems. The dimensionality in complex networks was first depicted by Csányi[21] and then developed by many researchers.[22–25] Song *et al.* found that a variety of real complex networks exhibits a self-similarity property[26] and proposed a box-covering algorithm to evaluate the fractal dimension of the complex networks.[27] However, the homogeneous distribution of the nodes cannot be expressed when calculating the fractal dimension by the box-covering algorithm, because the number of nodes in different boxes is different. Thus, a dimension called information dimension[28] is proposed to reveal the self-similarity properties of complex networks. From then on, many researchers have been studying the fractal and the self-similarity properties of complex networks extensively,[29–33] nevertheless not much attention has been paid to the local fractal property of the complex networks. Recently, Silva and Costa[34] proposed a methodology consisting of obtaining the local dimension at different scales for each node. It can describe the local dispersion of the network well. Based on the method, Pu *et al.*[35] proposed a measurement to identify influential nodes in the networks by comparing each node's local dimension (LD) at the farthest distance. However, it just considers the local information of the node. In this paper, a node information dimension is proposed to identify influential nodes by synthesizing each segmental local dimension. This methodology first calculates the local dimension at different distance scales for each node and then converts these local dimensions into the probabilities to calculate the information entropy of different topological distance scale for each node. Finally, the node information dimension can be calculated by the linear regression of information entropy and distance at different scales for each node. The method proposed in this paper not only considers the distance between the node and other nodes, but also considers the distribution of the surrounding nodes in the network. If the nodes' distribution in the network is satisfactory, the closer are other nodes to the node that is being measured, the larger the information dimension of the node will be, and then the node will be more important.

a)Author to whom correspondence should be addressed: prof.deng@hotmail.com and dengentropy@uestc.edu.cn. Tel.: +86-28-61830858. Fax: +86-28-61830858.

The proposed method fuses local and global characteristics of networks which can better identify influential nodes when the node's importance in the network is not entirely determined by the node's degree or robustness, such as social networks, infrastructure networks, and co-authorship networks, because the nodes that have influence in these networks need not only a have high degree, but also need to have high robustness in the center of the network structure and play a key role in communication. But the proposed method is inappropriate to be applied in large-scale networks due to its computational complexity, such as the Wikipedia networks[36] (nodes are pages in Wikipedia and edges correspond to hyperlinks) which have tens of millions of nodes. Although the proposed method has a complicated calculation the same as the BC and CC, its synthesis of the distribution information of nodes at different distances makes it a better method to analyze a network that is not large-scale. For example, compared to the CC, the proposed method not only considers the distance from the node to other nodes, but also considers the local dimension, making the result closely associated with the distribution of the surrounding nodes. Thus, the proposed method is more superior to the CC.

The rest of this paper is organized as follows: Section II begins with a brief overview of the existing centrality measures and an introduction to the local dimension. In Sec. III, a node information dimension for identifying the influential nodes is proposed and illustrated by an example network. Then, the application of the proposed method is illustrated. A case study of Netscience is adopted to evaluate the performance of the proposed method in Sec. IV. Finally, some conclusions are presented in Sec. V.

II. PRELIMINARIES

A. Centrality measures

Considering a graph G = (V, E) with n = $|V|$ nodes and m = $|E|$ links. And the node centrality measurement of DC, CC, and BC is well defined as follows.

1. Degree centrality

The DC of node i, denoted as $CD(i)$, is defined as[37]

$$C_D(i) = \sum_{j}^{N} x_{ij}, \tag{1}$$

where i is the focal node, j represents all other nodes, N is the total number of nodes, and x_{ij} represents the connection between node i and node j. The value of x_{ij} is defined as 1 if node i is connected to node j, and 0 otherwise.

2. Betweenness centrality

The BC of node i, denoted as $CB(i)$, is defined as[37]

$$C_B(i) = \sum_{j,k\neq i} \frac{g_{jk}(i)}{g_{jk}}, \tag{2}$$

where g_{jk} denotes the number of the shortest paths between nodes j and k, and $g_{jk}(i)$ means the number of the shortest paths between nodes j and k that go through node i.

3. Closeness centrality

The CC of node i, denoted as $CC(i)$, is defined as[37]

$$C_C(i) = \left[\sum_{j}^{N} d_{ij}\right]^{-1}, \tag{3}$$

where d_{ij} denotes the distance between node i and node j.

4. Eigenvector centrality

Let $\mathbf{A}$ be an $n \times n$ similarity matrix.[12] The eigenvector centrality x_i of node i is defined as the ith entry in the normalized eigenvector belonging to the largest eigenvalue of $\mathbf{A}$. λ is the largest eigenvalue of $\mathbf{A}$ and n is the number of vertices

$$Ax = \lambda x, \quad x_i = u\sum_{j=1}^{n} a_{ij}x_j, \quad i = 1, 2, \ldots, n, \tag{4}$$

with proportionality factor $u = \frac{1}{\lambda}$ so that x_i is proportional to the sum of similarity scores of all nodes connected to it.

5. PageRank

The PageRank algorithm[39] is a famous variant of eigenvector centrality and is used to rank websites in the Google search engine and other commercial scenarios.[40] Similar to eigenvector centrality, PageRank supposes that the importance of a webpage is determined by both the quantity and the quality of the pages linked to it.[41] Initially, each node gets one unit PR value. Then, every node evenly distributes the PR value to its neighbors along its outgoing links. Mathematically, the PR value of node v_i at t step is[38]

$$PR_i(t) = \sum_{j=1}^{n} a_{ji}\frac{PR_j(t-1)}{k_j^{out}}, \tag{5}$$

where n is the total number of nodes in the network, and k_j^{out} is the out-degree of node v_j. The above iteration will stop if the PR values of all nodes reach the steady state.

6. LeadRank

The Leader Rank was proposed by Lü *et al.*[14] which is a variant of the Page Rank algorithm. In general, they introduced a ground node which bidirectionally connects to all other nodes. Then, the random walk process is used to find influential nodes and this process continues until the steady state is attained. The process can be described by the stochastic matrix $\mathbf{P}$ with a probability $p_{ij} = \frac{a_{ij}}{k_i^{out}}$ that a random walker at node i goes to j the next time, where $a_{ij} = 1$ means node i pointing to node j, and k_i^{out} denotes the out-degree. So the score of node i at time t can be defined as

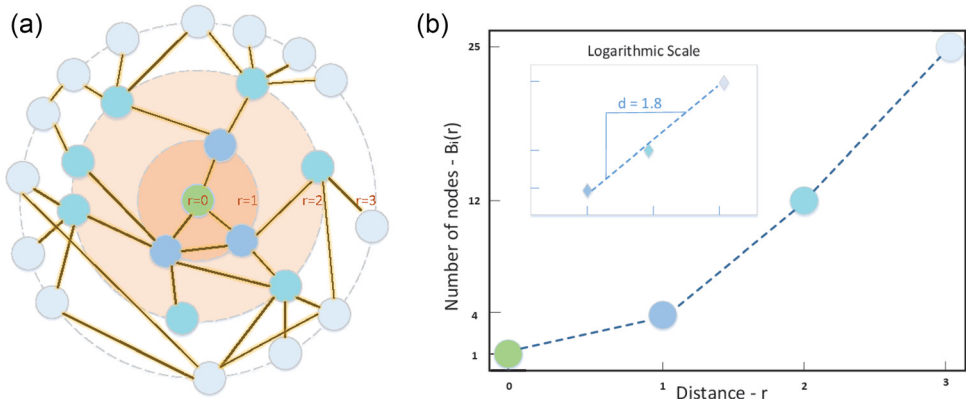

FIG. 1. Process of obtaining the dimension of a network. Starting from a node, along with the expansion of distance r = 0 up to r = 3, we can get four different $B_i(r_n)$. Thus, on a double-logarithmic scale the dimension can be obtained by the slope of the $B_i(r)$ curve, while $r = 0$ shall not be included. (Modified from. Ref. 34.).

$$s_i(t+1) = \sum_{j=1}^{N+1} \frac{a_{ji}}{k_j^{out}} s_j(t). \tag{6}$$

The initial scores for all node i is 1, while 0 for the ground node. When the score $s_i(t)$ for all i converges to a unique steady state which can be denoted as $s_i(t_c)$, where t_c is the convergence time. Thus, the final score of a node is defined as

$$S_i = s_i(t_c) + \frac{s_g(t_c)}{N}, \tag{7}$$

where $s_g(t_c)$ is the score of the ground node at steady state. Hence, we can rank each node with the S_i value down. (For more detailed information please refer to Ref. 14).

B. Local dimension

One of the characteristics of complex networks is that the degree of nodes follows a power law distribution,[42,43] which means within the topological distance r (r included), the number of nodes $B_i(r)$ for each node i with regard to the topological distance r follows a power law as follows:

$$B_i(r) \sim r^d, \tag{8}$$

where the constant d denotes the dimension of a network. However, due to the concept of locality, Eq. (8) can be further developed by considering the dimension d which may vary both for each node i and the topological distance r.[34] Thus, the dimension d in Eq. (8) is redefined as the local dimension $D_i(r)$. Each local dimension coefficient $D_i(r)$ can be obtained by the slope of the $B_i(r)$ curve on a double logarithmic scale[44] (shown in Fig. 1) as follows:

$$B_i(r) = ar^{D_i(r)}, \tag{9}$$

$$D_i(r) = \frac{d}{d\log r} \log B_i(r) \tag{10}$$

for the discrete nature of complex network,[45] the formula above can be expressed as follows:[34]

$$D_i(r) = \frac{d}{d\log r} \log B_i(r), \tag{11}$$

$$D_i(r) = \frac{r}{B_i(r)} \frac{d}{dr} B_i(r), \tag{12}$$

$$D_i(r) \simeq r \frac{n_i(r)}{B_i(r)}, \tag{13}$$

where $n_i(r)$ denotes the number of nodes that the topological distance from the central node i is exactly equal to the topological distance r. In brief, the local dimension $D_i(r)$ of the node i is different with different starting node i and the topological distance r. For example, to obtain the local dimension of the node a in which the arrow points in Fig. 2, the number of nodes within distance $r = 3$ should be calculated as $B_a(3)$ and at $r = 3$ should be counted as $n_a(3)$. According to Eq. (13), the local dimension $D_a(3)$ of the node a is approximately 1.

C. Information entropy

In the field of information science, when we want to analyze the uncertainty of any information, we need to make a quantitative analysis of the information. Shannon proposed

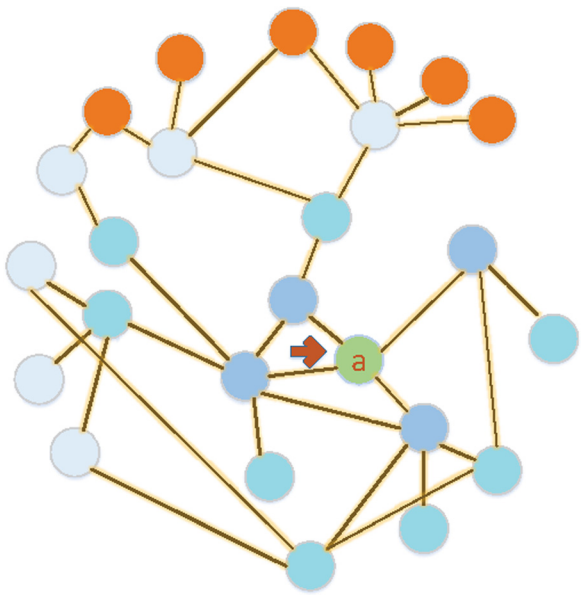

FIG. 2. The process of obtaining local dimension $D_a(3)$ for node a.

information entropy[46] to represent the uncertain degree of information. Moreover, the larger the entropy, the more uncertain the information. Here, the information captures the global structure in the network, and the larger the entropy, the more influential the node is. The formula can be expressed as follows:[46]

$$H = -\sum_{i=1}^{N} \theta_i \log_b \theta_i, \tag{14}$$

where N is the distance from the farthest node in the network to the node which we are calculating, θ_i is the ratio of the node's local dimension of distance i to the sum of the node's local dimension at each distance, and b is the base of the logarithm. In the paper, we let $b = e$.

III. NODE INFORMATION DIMENSION

In this section, a node information dimension (NID) is proposed by synthesizing each segmental local dimension. In other words, the local dimension may vary both for each node i and the topological distance r, but the node information dimension is the aggregation of the local dimension at different topological distance r for each node i, it only varies for each node i. Thus, it is more comprehensive than the local dimension method proposed by Pu *et al.*[35] to identifying the influential nodes in complex networks. The process to obtain the node information dimension is developed as follows.

A. Proposed method

Step 1: Calculate the maximum shortest path R_i for each node i.
Step 2: Calculate the information entropy $I_i(r)$ of different topological distance scale r for each node i.
First, we should obtain the step $S_i(r)$ we need to calculate by each topological distance scale r as follows:

$$S_i(r) = ceil\left(\frac{R_i}{r}\right), \quad (r = 1, 2, \ldots, R_i(r)). \tag{15}$$

Second, the local dimensions at different steps are calculated based on Eq. (13) as follows:

$$d_j(r) = j\frac{n_r(j)}{B_r(j)}, \quad (j = 1, 2, \ldots, S_i(r)), \tag{16}$$

where $d_j(r)$ denotes the jth step probability with respect to topological distance scale r, $n_r(j)$ denotes the number of nodes that are involved in the jth step, and $B_r(j)$ denotes the number of nodes within j step (j included) with respect to topological distance scale r.
Third, the probability is denoted by

$$P_j(r) = \frac{d_j(r)}{\sum_{j=1}^{S_i(r)} d_j(r)}, \quad (j = 1, 2, \ldots, S_i(r)). \tag{17}$$

Finally, the information entropy $I_i(r)$ is obtained by the formulation as follows:

$$I_i(r) = -\sum_{j=1}^{S_i(r)} P_j(r) ln P_j(r). \tag{18}$$

Step 3: The node information dimension can be given by

$$D_i = -\lim_{r\to 0} \frac{I_i(r)}{lnr}. \tag{19}$$

B. Example explanation

Based on the method proposed above, a simple example is given to explain how NID performs in this part. From Fig. 2, we can obtain that the maximum shortest path of node a, $R_a = 4$.

For r is equal to 1, the step $S_a(1)$ is obviously equal to 4. Then, we can calculate the local dimensions $d_j(1)$.

$$d_j(1) = \left(\frac{4}{5}, \frac{16}{13}, \frac{18}{19}, \frac{24}{25}\right), \quad (j = 1, 2, 3, 4).$$

Next, the probability $P_j(1)$ is calculated by Eq. (17).

$$P_j(1) = \left(\frac{569}{2801}, \frac{761}{2435}, \frac{325}{1351}, \frac{988}{4053}\right), \quad (j = 1, 2, 3, 4).$$

The information entropy $I_a(1)$ is obtained by Eq. (18)

$$I_a(1) = -\left(\frac{569}{2801} ln\left(\frac{569}{2801}\right) + \frac{761}{2435} ln\left(\frac{761}{2435}\right) + \frac{325}{1351} ln\left(\frac{325}{1351}\right) + \frac{988}{4053} ln\left(\frac{988}{4053}\right)\right) = 1.3741.$$

In this way, the information entropy with respect to different topological distance scale r is obtained as follows:

$$I_a(r) = (1.3741, 0.6930, 0.6385, 0), \quad (r = 1, 2, 3, 4).$$

Finally, the NID of node a is calculated by Eq. (19), and the process is shown in Fig. 3.

$$D_a = -\lim_{r\to 0} \frac{I_a(r)}{lnr} = -0.8927.$$

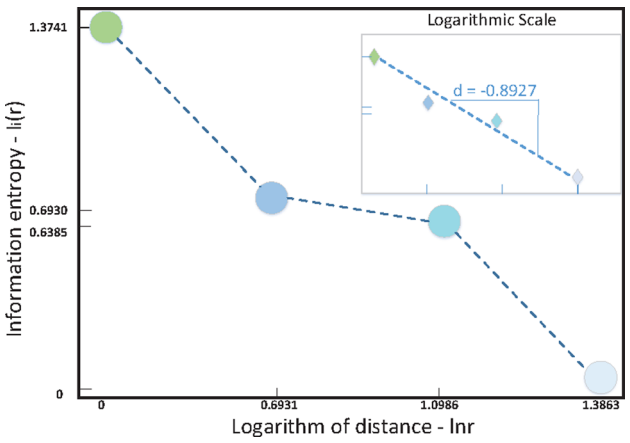

FIG. 3. The process of obtaining node information dimension for node a.

IV. CASE STUDY

In this section, we use a network of co-authorships between scientists interested in network science (namely, netscience network)[47] to demonstrate that the proposed method does a lot better than other centrality approaches when the influential nodes in the network are not entirely determined by a high degree or good robustness. The netscience network is a network of co-authorships between 379 scientists whose research centers on the properties of networks of one kind or another.[48]

In this case, we compare the proposed method with the results obtained in Newman's work.[48] In Newman's work, he proposed a method named community centrality to rank the nodes in the network of co-authorships between scientists. The results of this measure show that it has the ability to find the center in the local communities. Indeed, in the netscience network, all of the authors having the highest community centrality are group leaders or senior researchers of groups working in this area. At the same time, the DC, CC, BC, EC, PageRank (PR), and LeadRank (LR) are taken to compare with the Newman's results, too. The lists of top-10 influential nodes for different measures are shown in Table I. In the top-10 lists, the proposed method, DC, CC, BC, EC, PR, and LR have the same 7, 8, 3, 4, 3, 6, and 8 nodes with the community centrality, respectively, only the proposed method, DC, and LR have a good performance to identify the influential nodes in the netscience network. At the same time, from Table I, we can see that the results of local measure and global measures are very different; thus, a local and global comprehensive approach is very necessary.

In this part, we also compare the proposed method with other measures based on the knowledge about the context of the network. The netscience network is a co-authorship network of scientists working on network theory and experiment before 2006. According to Table I, we can see that the proposed method has the same 7, 7, 2, 4, 4, 5, and 6 nodes with the community centrality, DC, CC, BC, EC, PR, and LR, respectively. Because of the poor performance of global centrality such as CC, BC, and EC in this network, we only compare the proposed method with the community centrality, DC, PR, and LR in this part. To evaluate the performance of the proposed method, we use the number of citations and the H index to compare the different scientists with the higher ranking in various methods, the data are shown in Table II. By comparing the proposed method and community centrality, and the proposed method ranks "VICSEK," "ALMAAS," and "KOVACS" in the top-10, while the community centrality ranks "MORNEO," "SOLE," and "DIAZGVILERA" in the top-10. From Table II, we can see that "VICSEK" has much higher citations and H-index than "MORNEO," "SOLE," and "DIAZGVILERA." Even "ALMAAS" and "KOVACS" are worse than three other scientists on the contrast of two indicators, but we can see from their publications and they all have numerous influential papers published in well-known journals such as Nature, PRE, PRL, and EPL. At the same time, they also have a lot of cooperation with the famous scientist "BARABASI," this means that their academic

TABLE I. The top-10 ranked nodes by community centrality, degree centrality (DC), closeness centrality (CC), betweenness centrality (BC), eigenvector centrality (EC), PageRank (PR), LeadRank (LR), and the proposed method in Netscience.

	Netscience								
Rank	Community centrality	NID	DC	CC	BC	EC	PR	LR
1	NEWMAN	BARABASI	BARABASI	HOLME	BOCCALETTI	BARABASI	BARABASI	BARABASI
2	BARABASI	ALMAAS	JEONG	EDLING	VAZQUEZ	JEONG	NEWMAN	NEWMAN
3	JEONG	OLTVAI	NEWMAN	LILJEROS	BOGUNA	ALBERT	SOLE	JEONG
4	BOCCALETTI	JEONG	OLTVAI	JEONG	KURTHS	OLTVAI	JEONG	PASTORSATORRAS
5	SOLE	PASTORSATORRAS	BOCCALETTI	NEWMAN	ARENAS	VICSEK	PASTORSATORRAS	VESPIGNANI
6	OLTVAI	NEWMAN	KURTHS	CHUNG	SCHNITZLER	RAVASZ	BOCCALETTI	MORENO
7	PASTORSATORRAS	VICSEK	SOLE	STANLEY	PASTORSATORRAS	NEDA	VESPIGNANI	SOLE
8	KURTHS	KURTHS	VICSEK	YOON	HANMALAINEN	YOOK	BAVELAS	WATTS
9	DIAZGVILERA	KOVACS	DIAZGVILERA	HAN	CALDARELLI	BIANCONI	AMARAL	KURTHS
10	MORNEO	BOCCALETTI	YOUNG	CHOI	NEWMAN	FARKAS	STAUFFER	BOCCALETTI

TABLE II. The citations and H-index of different nodes with higher ranking in various methods. "-" means that we cannot find the exact data of the network scientist.

NID			Community centrality			NID			DC		
	Citations	H-index		Citations	H-index		Citations	H-index		Citations	H-index
VICSEK	43 891	77	MORNEO	24 322	59	PASTORSATORRAS	26 317	54	SOLE	19 426	70
ALMAAS	2977	23	SOLE	19 426	70	ALMAAS	2977	23	DIAZGVILERA	11 610	40
KOVACS	1251	16	DIAZGVILERA	11 610	40	KOVACS	1251	16	YOUNG	335	10
NID			PR			NID			LR		
KURTHS	62 775	107	VESPIGNANI	45 613	96	VICSEK	43 891	40	WATTS	79 873	49
VICSEK	43 891	77	AMARAL	38 684	78	OLTVAI	42 727	40	VESPIGNANI	45 613	96
OLTVAI	42 727	40	SOLE	19 426	70	ALMAAS	2977	23	MORENO	24 322	59
ALMAAS	2977	23	BAVELAS	…	…	KOVACS	1251	16	SOLE	19 426	70

achievements will attract a lot of attention and give them more potential influence. This is the advantage of the proposed method, which not only takes account of the distance between the node and other nodes, but also considers the distribution of the surrounding nodes. There are the same 7 nodes in the top-10 lists by comparing the proposed method and DC, and the proposed method ranks "PASTORSATORRAS," "ALMAAS," and "KOVACS" in the top-10, while the DC ranks "SOLE," "DIAZGVILERA," and "YOUNG" in the top-10. According to Table II, the citations of "PASTORSATORRAS" is much higher than "SOLE," "DIAZGVILERA," and "YOUNG," and both "KOVACS" and "ALMAAS" have much higher citations than "YOUNG." Then, comparing the proposed method and PR, it is obvious that "KURTHS" has much higher citations and H-index than other scientists. The citations of "VICSEK" and "OLTVAI" are higher than those of "AMARAL" and "SOLE," respectively. Finally, we compare the proposed method with LR, one can observe that although the LR found two famous network scientists "VESPIGNANI" and "WATTS" both have high citations, the method proposed in this paper also identified that two network scientists, "VICSEK" and "OLTVAI," have higher citations than "MORENO" and "SOLE," respectively. From the above comparison, we found that the proposed method can always find some influential network scientists not found in other methods, so our method performed better than other methods in this application.

Figures 4(a)–4(g) show the correlation between the community centrality and the proposed method, DC, BC, CC, PR, and LR in the netscience network. From Figs. 4(a) to 4(g), we can see that the results of our approach and Newman's result have a similar tendency to show the influence of nodes, even community centrality has a bad performance to identify the nodes with lower influence, and most of them have the same community centrality values. This is because the community centricity method can only be good at distinguishing the communities in the network then consider the most important nodes in each community as the most important nodes in the network, however, it will not be a good way to distinguish the nodes in different communities, thus resulting in a large number of different nodes in the network that have the same evaluation. It is like we can find the leaders in every laboratory, but it is not easy to compare the other scientists in the various laboratories. It is a similar correlation between community centrality and DC, what is different is that the DC even has a worse ability to identify the nodes with lower influence. With the requirement that needs to count the shortest path via the node, BC always gives a lot of nodes' value equal to 0 as shown in Fig. 4(c). We can see that CC, PR, and LR really have a weak correlation with community centrality. It is obvious that the EC and community centrality are positively correlated; however, a lot of node's EC values are the same around 0 that cannot be effectively categorized. The Spearman's rank correlation coefficient[49] is also used to show the correlation between the community centrality and other measures. From Table III, we can come to the conclusion that the proposed method and the community centrality have the strongest correlation compared to other centrality measures. Therefore, the proposed method is an effective method to identify the influential nodes in the complex networks. As the community centrality gives a measure of the "strength" with which vertices belong to their assigned communities, it is not a good way to measure the importance of all nodes in a vast network consisting of different communities. Because of the differences between communities, we cannot measure the importance of different nodes across the network, simply based on how nodes behave in their communities, it does not tally with the reality. The proposed method considers not only the distance of other nodes to a node, but also the distribution of the surrounding nodes. It will no longer be limited to the information within the same community or in the vicinity of the node. Therefore, this method can make a good comparison of the nodes in the whole network.

V. CONCLUSION

In this paper, a new measure based on node information dimension is proposed to identify the influential nodes in complex networks. In the proposed method, the local dimensions of different topological distance scales for each node constitute the node information dimension. Thus, compared with the method that identifies the influential nodes based on the local dimension, the node information dimension which takes the global structure information of a complex network into consideration is more reasonable. To evaluate the performance, a case study of Netscience is used to show the effectiveness of the proposed method. The experimental

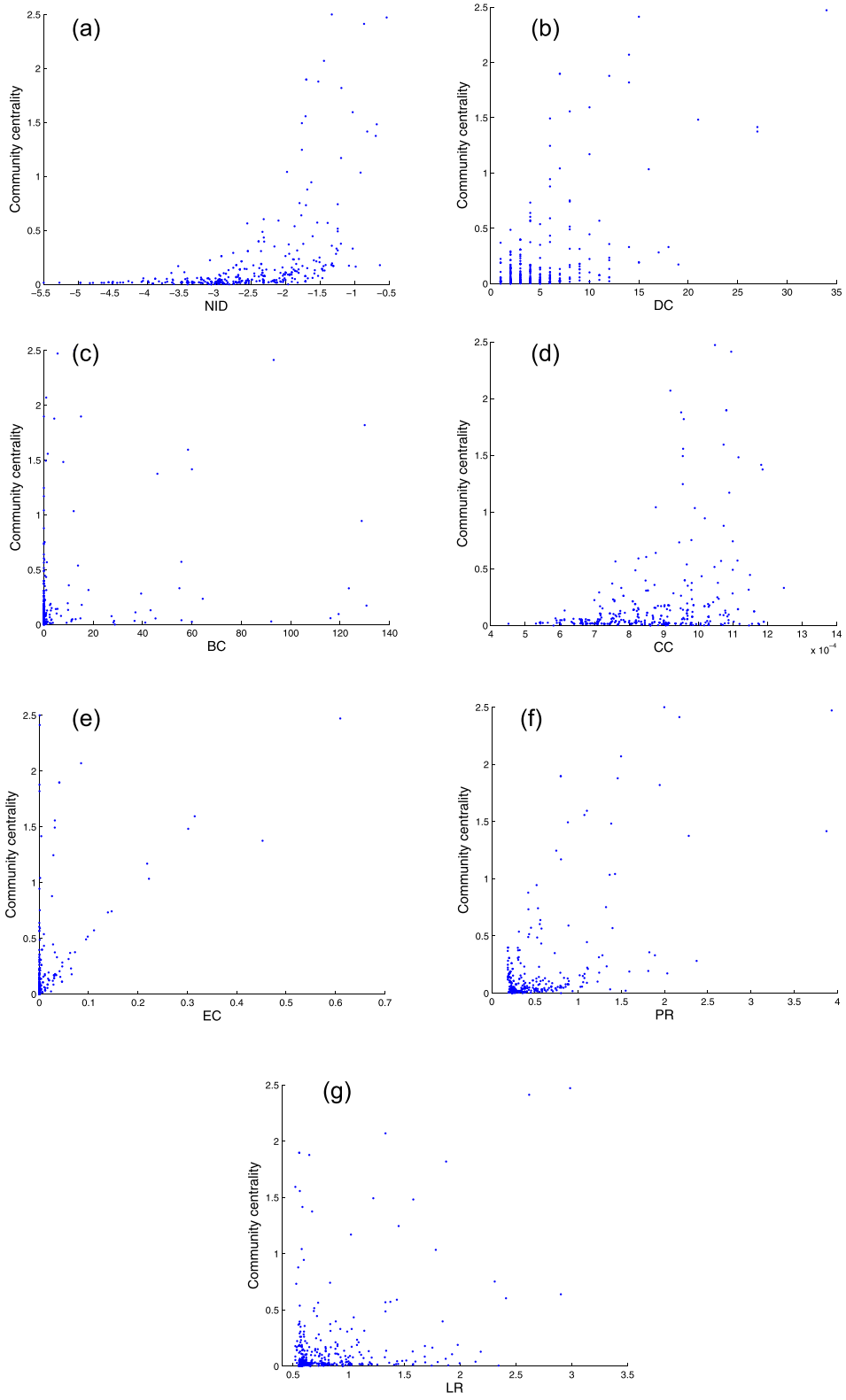

FIG. 4. The correlation between the community centrality and NID, DC, BC, CC, PR, and LR in the netscience network.

TABLE III. The Spearman's rank correlation coefficient between the community centrality and the proposed method, degree centrality (DC), closeness centrality (CC), betweenness centrality (BC), eigenvector centrality (EC), PageRank (PR), and LeadRank (LR) in Netscience.

	NID	DC	CC	BC	EC	PR	LR
Community centrality	0.6662	0.2602	0.2791	0.3959	0.5483	0.2438	0.0739

results demonstrate that the proposed method can successfully identify the influential nodes in networks when the influential nodes in the network are not entirely determined by a high degree or good robustness.

ACKNOWLEDGMENTS

The authors greatly appreciate the reviews' suggestions and the editor's encouragement. This work was partially

supported by the National Natural Science Foundation of China (Grant Nos. 61573290 and 61503237).

[1]Z. Wang, M. A. Andrews, Z.-X. Wu, L. Wang, and C. T. Bauch, "Coupled disease–behavior dynamics on complex networks: A review," Phys. Life Rev. **15**, 1–29 (2015).
[2]A.-L. Barabási, N. Gulbahce, and J. Loscalzo, "Network medicine: A network-based approach to human disease," Nat. Rev. Genet. **12**, 56–68 (2011).
[3]J. Borge-Holthoefer and Y. Moreno, "Absence of influential spreaders in rumor dynamics," Phys. Rev. E **85**, 026116 (2012).
[4]F. Xiao, C. Zhan, H. Lai, L. Tao, and Z. Qu, "New parallel processing strategies in complex event processing systems with data streams," Int. J. Distrib. Sens. Networks **13**, 1–15 (2017).
[5]H. Zheng and Y. Deng, "Evaluation method based on fuzzy relations between Dempster-Shafer belief structure," Int. J. Intell. Syst. (2017).
[6]F. Xiao, M. Aritsugi, Q. Wang, and R. Zhang, "Efficient processing of multiple nested event pattern queries over multi-dimensional event streams based on a triaxial hierarchical model," Artif. Intell. Med. **72**, 56–71 (2016).
[7]X. Zhang, J. Zhu, Q. Wang, and H. Zhao, "Identifying influential nodes in complex networks with community structure," Knowl.-Based Syst. **42**, 74–84 (2013).
[8]T. Liu, Y. Deng, and F. Chan, "Evidential supplier selection based on DEMATEL and game theory," Int. J. Fuzzy Syst. **20**, 1321–1333 (2018).
[9]X. Zheng and Y. Deng, "Dependence assessment in human reliability analysis based on evidence credibility decay model and IOWA operator," Ann. Nucl. Energy **112**, 673–684 (2018).
[10]Z. Wang, C.-Y. Xia, S. Meloni, C.-S. Zhou, and Y. Moreno, "Impact of social punishment on cooperative behavior in complex networks," Sci. Rep. **3**, 3055 (2013).
[11]L. C. Freeman, "Centrality in social networks conceptual clarification," Soc. Networks **1**, 215–239 (1978).
[12]P. Bonacich and P. Lloyd, "Eigenvector-like measures of centrality for asymmetric relations," Soc. Networks **23**, 191–201 (2001).
[13]L. Page, S. Brin, R. Motwani, and T. Winograd, "The pagerank citation ranking: Bringing order to the web," (1999); available at http://ilpubs.stanford.edu:8090/422/.
[14]L. Lü, Y.-C. Zhang, C. H. Yeung, and T. Zhou, "Leaders in social networks, the delicious case," PLoS One **6**, e21202 (2011).
[15]T. Bian, H. Zheng, L. Yin, and Y. Deng, "Failure mode and effects analysis based on D numbers and TOPSIS," Qual. Reliab. Eng. Int. QRE2268 (2018).
[16]D.-B. Chen, R. Xiao, A. Zeng, and Y.-C. Zhang, "Path diversity improves the identification of influential spreaders," Europhys. Lett. **104**, 68006 (2014).
[17]F. Hu and Y. Liu, "Multi-index algorithm of identifying important nodes in complex networks based on linear discriminant analysis," Mod. Phys. Lett. B **29**, 1450268 (2015).
[18]Q. Zhang, M. Li, and Y. Deng, "Measure the structure similarity of nodes in complex networks based on relative entropy," Physica A **491**, 749–763 (2018).
[19]X. Zhou, Y. Hu, Y. Deng, F. T. S. Chan, and A. Ishizaka, "A DEMATEL-based completion method for incomplete pairwise comparison matrix in AHP," Ann. Oper. Res. (2018).
[20]H. Yu, L. Chen, X. Cao, Z. Liu, and Y. Li, "Identifying top-k important nodes based on probabilistic-jumping random walk in complex networks," in *International Workshop on Complex Networks and Their Applications* (Springer, 2017), pp. 326–338.
[21]G. Csányi and B. Szendrői, "Structure of a large social network," Phys. Rev. E **69**, 036131 (2004).
[22]S. Havlin, D. ben Avraham *et al.*, "Fractal and transfractal recursive scale-free nets," New J. Phys. **9**, 175 (2007).
[23]B. Kang, G. Chhipi-Shrestha, Y. Deng, K. Hewage, and R. Sadiq, "Stable strategies analysis based on the utility of Z-number in the evolutionary games," Appl. Math. Comput. **324**, 202–217 (2018).
[24]Z. Zhang, S. Zhou, T. Zou, and G. Chen, "Fractal scale-free networks resistant to disease spread," J. Stat. Mech.: Theory Exp. **2008**, P09008.
[25]X. Deng and Y. Deng, "D-AHP method with different credibility of information," Soft Comput. 1–9 (2018).
[26]C. Song, S. Havlin, and H. A. Makse, "Self-similarity of complex networks," Nature **433**, 392–395 (2005).
[27]C. Song, L. K. Gallos, S. Havlin, and H. A. Makse, "How to calculate the fractal dimension of a complex network: The box covering algorithm," J. Stat. Mech.: Theory Exp. **2007**, P03006.
[28]S. N. Dorogovtsev and J. F. F. Mendes, "Evolution of networks with aging of sites," Phys. Rev. E **62**, 1842 (2000).
[29]Z.-Z. Zhang, S.-G. Zhou, and T. Zou, "Self-similarity, small-world, scale-free scaling, disassortativity, and robustness in hierarchical lattices," Eur. Phys. J. B **56**, 259–271 (2007).
[30]R. Zhang, B. Ashuri, and Y. Deng, "A novel method for forecasting time series based on fuzzy logic and visibility graph," Adv. Data Anal. Classif. **11**, 759–783 (2017).
[31]W. Deng, X. Lu, and Y. Deng, "Evidential model validation under epistemic uncertainty," Math. Probl. Eng. (2018).
[32]J. Kim, K.-I. Goh, B. Kahng, and D. Kim, "Fractality and self-similarity in scale-free networks," New J. Phys. **9**, 177 (2007).
[33]L. Yin and Y. Deng, "Measuring transferring similarity via local information," Physica A **498**, 102–115 (2018).
[34]F. N. Silva and L. d. F. Costa, "Local dimension of complex networks," preprint arXiv:1209.2476 (2012).
[35]J. Pu, X. Chen, D. Wei, Q. Liu, and Y. Deng, "Identifying influential nodes based on local dimension," Europhys. Lett. **107**, 10010 (2014).
[36]S. Auer, C. Bizer, G. Kobilarov, J. Lehmann, R. Cyganiak, and Z. Ives, "Dbpedia: A nucleus for a web of open data," in *The Semantic Web* (2007), pp. 722–735.
[37]Y. Du, C. Gao, Y. Hu, S. Mahadevan, and Y. Deng, "A new method of identifying influential nodes in complex networks based on TOPSIS," Physica A **399**, 57–69 (2014).
[38]U. Sehgal, K. Kaur, and P. Kumar, "The anatomy of a large-scale hyper textual web search engine," in *2009 Second International Conference on Computer and Electrical Engineering* (IEEE, 2009), pp. 491–495.
[39]S. Brin and L. Page, "Reprint of: The anatomy of a large-scale hypertextual web search engine," Comput. Networks **56**, 3825–3833 (2012).
[40]A. N. Langville and C. D. Meyer, *Google's PageRank and beyond: The Science of Search Engine Rankings* (Princeton University Press, 2011).
[41]L. Lü, D. Chen, X.-L. Ren, Q.-M. Zhang, Y.-C. Zhang, and T. Zhou, "Vital nodes identification in complex networks," Phys. Rep. **650**, 1–63 (2016).
[42]D. P. Dailey, "Uniqueness of colorability and colorability of planar 4-regular graphs are NP-complete," Discrete Math. **30**, 289–293 (1980).
[43]R. Lambiotte, V. D. Blondel, C. De Kerchove, E. Huens, C. Prieur, Z. Smoreda, and P. Van Dooren, "Geographical dispersal of mobile communication networks," Physica A **387**, 5317–5325 (2008).
[44]L. Daqing, K. Kosmidis, A. Bunde, and S. Havlin, "Dimension of spatially embedded networks," Nat. Phys. **7**, 481–484 (2011).
[45]E. Ben-Naim, H. Frauenfelder, and Z. Toroczkai, *Complex Networks* (Springer Science & Business Media, 2004), Vol. 650.
[46]C. E. Shannon, "A mathematical theory of communication," ACM SIGMOBILE Mobile Comput. Commun. Rev. **5**, 3–55 (2001).
[47]M. E. Newman and M. Girvan, "Finding and evaluating community structure in networks," Phys. Rev. E **69**, 026113 (2004).
[48]M. E. Newman, "Finding community structure in networks using the eigenvectors of matrices," Phys. Rev. E **74**, 036104 (2006).
[49]H. Xu and Y. Deng, "Dependent evidence combination based on Shearman coefficient and Pearson coefficient," IEEE Access **6**, 11634–11640 (2018).