



OPEN

Identifying important nodes in complex networks based on extended degree and E-shell hierarchy decomposition

Jun Liu¹ & Jiming Zheng²✉

The identification of important nodes is a hot topic in complex networks. Many methods have been proposed in different fields for solving this problem. Most previous work emphasized the role of a single feature and, as a result, rarely made full use of multiple items. This paper proposes a new method that utilizes multiple characteristics of nodes for the evaluation of their importance. First, an extended degree is defined to improve the classical degree. And E-shell hierarchy decomposition is put forward for determining nodes' position through the network's hierarchical structure. Then, based on the combination of these two components, a hybrid characteristic centrality and its extended version are proposed for evaluating the importance of nodes. Extensive experiments are conducted in six real networks, and the susceptible–infected–recovered model and monotonicity criterion are introduced to test the performance of the new approach. The comparison results demonstrate that the proposed new approach exposes more competitive advantages in both accuracy and resolution compared to the other five approaches.

With the rapidly growing scale of networks, topics concerning complex networks are emerging in network science^{1,2}. Generally, complex networks are modeled by various systems in the real world, which are disorganized, self-similar, and small-world³. Massive systems can be remodeled by complex networks, for instance, protein networks in biology⁴, social networks in sociology⁵, financial networks in economics⁶, power networks in engineering⁷, and so on. Complex networks have received extensive attention in theory and practice in recent years^{8–10}. The identification of important nodes has become a fundamental problem in complex networks^{11–13}, which has not only theoretical value^{14,15}, but likewise practical applications. Some examples can be easily found. In social networks, the positive contribution of “celebrities” can effectively reduce the spread of negative social opinion. Ad-technology vendors often seek out the most influential users for advertising to maximize revenue in marketing networks. More applications can be given in other cases, such as disease control¹⁶ and sociology¹⁷.

The identification of important nodes in complex networks is an NP-hard problem¹⁸. Up to now, many schemes with polynomial complexity have been presented. For example, degree centrality¹⁹ counts the number of node's nearest neighbors and considers important nodes to have more neighbors. Closeness centrality¹⁹ considers the average length of the shortest path from the target node to other nodes in the network. Betweenness centrality¹⁹ calculates the fraction of the shortest paths that cover the target node. A number of variants and approximations have been proposed for speed and accuracy^{20–22}. These schemes, which are called “classical”, exploit the topological characteristics of nodes while ignoring community properties. However, it is well known that community organization is a main feature of complex networks²³.

Consequently, when one considers community organizations, nodes that are not of interest in classical centrality may have potential influence. Community centrality considers the heterogeneity between intra-community links and inter-community links^{24–26}. The intra-community links quantify the local influence of a node inside a community. Conversely, inter-community links account for the global influence of nodes on various communities. Many ideas of community centrality have been put out in light of the various ways in which the two type of links might be coupled^{27–29}. Zhao et al.²⁷ distinguish the weights of inter-community links by community size. The community-based mediator targets influential nodes through entropy and normalized degree²⁸. This

¹School of Science, Chongqing University of Posts and Telecommunications, Chongqing 400065, China. ²Key Lab of Intelligent Analysis and Decision on Complex System, Chongqing University of Posts and Telecommunications, Chongqing 400065, China. ✉email: zhengjm0903@qq.com

entropy is generated by the heterogeneity of the links between the inside and outside communities. Modular Vitality assesses the centrality of a node by investigating the modular variations caused when one is removed²⁹.

Besides the community, the hierarchical structure of the network has also attracted attention³⁰. Some previous outcomes have used the hierarchical structure to drive the K-shell decomposition algorithm and its improvements^{31–34}. Although the K-shell algorithm has promising applications in various fields^{35–38}, there has been a drive to improve its resolution. This is because it ties many nodes to the same shell, even though these nodes differ in influence^{39,40}. Zeng et al.⁴¹ introduced the exhausted degree and proposed a mixed degree decomposition (MDD) method to identify important nodes. Bae et al.⁴² presented an extended version, called neighborhood coreness, using the k-shell index of 1-order neighborhood nodes. Feng et al.⁴³ consider opinion leaders and structural hole nodes to have more ability to influence others. Wang et al.⁴⁴ suggested a integral k-shell algorithm (IKS) by accumulating historical k-shell index and second-order degree of neighborhood. Liu et al.⁴⁵ refined the K-shell algorithm utilizing TOPSIS technique.

Many facts indicate that the importance of the node depends on multiple characteristics, such as degree, neighbors, position, and so on. Sheikahmadi et al.⁴⁶ presented a multi-criteria approach (MCDE) which utilizes a combination of node's degree, k-shell index and information entropy. There is still relatively little work on multi-characteristic methods. As a result, a new approach based on multiple characteristics is proposed in this paper. The interests of our work are as follows,

- (1) An improved version of the classical degree, extended degree, is introduced.
- (2) A E-shell hierarchy decomposition is put forward for determining nodes' position information through the network's hierarchical structure.
- (3) A hybrid characteristic centrality (HCC) is presented, which combines the extended degree and E-shell hierarchy decomposition.

Extensive experiments were performed in six real networks and the performance of the proposed approach was examined using the monotonicity function⁴² and the susceptible–infected–recovered (SIR) model^{47,48}. The results indicate that the new approach is more competitive than the classical and community centrality in terms of accuracy and resolution.

The framework of this paper is organized as follows. The “Preliminaries” section briefly introduces some basic preliminaries. The “Methods” section presents the new method and provides a simple example. The “Experiments” section examines the performance of the new method and compares the existing algorithms. The “Discussion” section summarizes the work.

Preliminaries

Let $G = (V, E)$ be an unweighted undirected network, where V and E are the set of nodes and edges, respectively. Denote $n = |V|$ and $m = |E|$. The adjacency of network G can be represented by $A = (a_{uv})_{n \times n}$, where a_{uv} indicates the connection between nodes u and v . $a_{uv} = 1$, if nodes u and v are directly connected; $a_{uv} = 0$, other cases.

K-order neighborhood. For any two nodes u and v , v is said to be a k -order neighbor of u if there exists the smallest positive integer such that Eq. (1) holds,

$$\prod_{i=0}^{k-1} a_{v_i v_{i+1}} = 1 \quad (1)$$

where $v_i \in G$ ($i = 0, 1, \dots, k$), $v_0 = u$ and $v_k = v$. Denote the k -order neighborhood of node u as the set $\phi^{(k)}(u)$, which consists of all k -order neighbors of u . If not otherwise specified, note that $\phi^{(1)}(u) = \phi(u)$. The number of 1-order neighbors of node u is called its degree¹⁹ and is denoted by $k(u)$, i.e.,

$$k(u) = |\phi(u)| \quad (2)$$

K-shell hierarchy decomposition. The main idea of the K-shell hierarchy decomposition: for a given k , a subgroup of the network called k -shell is obtained by iteratively deleting nodes with degree less than or equal to k . All node within the k -shell have the same index k . If there are no isolated nodes (degree equal to 0) in the network, then the nodes with degree equal to 1 have the lowest importance. Therefore, these nodes and their connected edges are deleted from the current network. Again, the new nodes with degrees less than or equal to 1 and their connected edges need to be deleted. Continue the above process until the degree of each node is greater than 1 in the current network. The nodes deleted in this round form 1-shell. Based on the above description, the procedure of the K-shell algorithm can be summarized as follows:

Step 1 delete nodes with degree k .

Step 2 repeatedly remove the new nodes whose degree is less than or equal to k until the degree of each new node is greater than k .

Step 3 all nodes deleted in steps 1 and 2 form the k -shell.

Step 4 let $k = K = 1$, and repeat the above steps. In the above program, if the initial and maximum values of k are 1 and K respectively, then we can get K shells, i.e., 1-shell, 2-shell, ..., K -shell.

Susceptible–infected–recovered (SIR) model. The SIR model is a classical model of disease transmission in which individuals are classified into 3 states: susceptible (S), infected (I), and recovered (R). Initially, a

single (or multiple) target individual is selected and its state is specified as I. At each step, the infected individuals spread the disease to each of its susceptible 1-order neighbors and enters state recovered. Herein, the infection rate and recovery rate are α and β , respectively. Then, the spread procedure is continued until there are no individuals with infected state in the network. The final cumulative count of infected individuals is considered as the real spread ability of the initial target individuals.

Methods

Main idea. Degree centrality and K-shell decomposition are two traditional single-characteristic schemes. However, numerous experiments have demonstrated that using a single characteristic to measure the importance of nodes is unreliable. Accordingly, a new idea is to combine multiple characteristics. Previous works have simply combined degree centrality and K-shell decomposition, but this fails to address the underlying problem: their low resolution limits the performance of the combination. Based on the above consideration, the interest of our work lies in the improvement of the classic degree and the K-shell decomposition, as well as their combination. As a result, we propose an extended degree and an E-shell hierarchy decomposition, and a combination of them called hybrid characteristic centrality.

Extended degree and E-shell hierarchy decomposition. The classical degree only counts the number of neighbors of the node itself. Thus, we define the extended degree to overcome the limitation of classical degree. Let G be an unweighted undirected network. The degree and 1-order neighbors of node $u \in G$ are denoted as $k(u)$ and $\phi(u)$, respectively. Then, the extended degree of node u , denoted by $k^{ex}(u)$, is defined by

$$k^{ex}(u) = \delta * k(u) + (1 - \delta) * \sum_{v \in \phi(u)} k(v) \quad (3)$$

where $\delta \in [0, 1]$ is a weight which reflects the dependence of $k^{ex}(u)$ on $k(u)$. If $\delta = 1$, then Eq. (3) degenerates to the classical degree (Eq. 2).

Next, we propose a E-shell hierarchy decomposition to determine the position of each node by decomposing the network's hierarchical structure. In this method, the nodes with minimum extended degree are found and deleted from the current network in each iteration. Then, these nodes are tagged with a position index, which is represented here by the iterations number. The procedures of E-shell hierarchy decomposition are described below.

Step 1: Input a network G .

Step 2: Initialize the iteration number $p = 1$ and $G_1 = G$.

Step 3: Find the set of minimum nodes $S_p = \arg \max_{u \in G_p} \{k^{ex}(u)\}$.

Step 4: Tag a position index $index_p$ for each node within S_p , where $index_p = p$.

Step 5: Remove S_p from G_p and then get a new network G_{p+1} . If G_{p+1} is non-empty, proceed to step 6; otherwise, the procedure terminates.

Step 6: Update the extended degree of each node within G_{p+1} .

Step 7: Update the iteration number $p = p + 1$.

Step 8: Return to step 3.

Hybrid characteristic centrality and its extension. The hybrid characteristic centrality combines the extended degree and the E-shell decomposition. Because of this, it gives a better understanding of how important a node is. On the one hand, the extended degree measures the local influence using degree about the node itself and its neighbors. On the other hand, E-shell decomposition reflects the global influence of a node by its position in the network. In brief, a node receives greater importance by having more neighbors and a higher position. Let $k^{ex}(u)$ be the extended degree of node u . Denote the position index of node u given by E-shell decomposition as $pos(u)$. The HCC of node u , written by $HCC(u)$, can be calculated by

$$HCC(u) = \frac{k^{ex}(u)}{k_{max}^{ex}} + \frac{pos(u)}{pos_{max}} \quad (4)$$

where k_{max}^{ex} and pos_{max} signify the maximum extended degree and position index of the nodes, respectively. Further, EHCC, an extensive version of HCC, is introduced as follows.

$$EHCC(u) = HCC(u) + \sum_{v \in \phi(u)} HCC(v) \quad (5)$$

EHCC draws on the idea of the extended degree, but it utilizes more information about a node, which includes the degree and position of its neighbors as well as itself. Nodes with higher EHCC values are commonly rewarded with a greater importance.

Computational complexity. The complexity of the EHHC calculation is as follows.

- (1) The complexity of calculating the extended degree is $O(n)$.
- (2) The complexity of the E-shell procedure for determining the position of the node is $O(m)$ (similar to K-shell).
- (3) Therefore, the total complexity is $O(n + m)$.

Computational process. To interpret the calculation procedure of the new method, a simple example is given. Figure 1 shows a simple network with 10 nodes and 14 edges.

First, the extended degree of each node is calculate by Eqs. (2) and (3). Take $\delta = 0.5$ in Eq. (3). The computational results are provided in Table 1. In the case of node a , the extended degree is $k^{ex}(a) = 0.5 * k(a) + 0.5 * (k(b) + k(e)) = 4.5$.

Second, the position index of each node is obtained according to the E-shell decomposition procedure. The complete process listed in Table 2. We observe that the position index of node a is 4.

Third, the HCC of each node is computed by Eq. (4). From Tables 1 and 2, the maximum extended degree and the maximum position index are 11 and 6, respectively. As a result, HCC of nodes a is $HCC(a) = \frac{4.5}{11} + \frac{4}{6} = 1.08$.

At last, the EHCC value of each node is computed using Eq. (5). The computation of node a is $EHCC(a) = HCC(a) + HCC(b) + HCC(e) = 4.15$. The results of HCC and EHCC are given by Table 3.

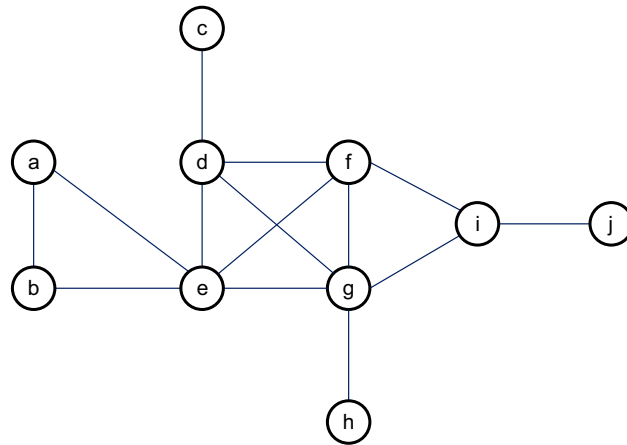


Figure 1. A simple network.

Node	a	b	c	d	e	f	g	h	i	j
Classical egree	2	2	1	4	5	4	5	1	3	1
Extended degree	4.5	4.5	2.5	9.5	11	10.5	11	3	6.5	2

Table 1. The class degree and extended degree of each node in the simple network.

Iteration number	Minimum extended degree	Delete node set	Position index
1	2	{ j }	1
2	2.5	{ c }	2
3	3	{ h }	3
4	4.5	{ a, b }	4
5	5	{ i }	5
6	6	{ d, e, f, g }	6

Table 2. The E-shell decomposition procedure of the simple network.

Node	a	b	c	d	e	f	g	h	i	j
HCC	1.08	1.08	0.56	1.86	2	1.95	2	0.77	1.42	0.35
EHCC	4.15	4.15	2.42	8.38	9.97	9.24	10.01	2.77	5.73	1.77

Table 3. The results of EHCC and HCC for the simple network.

Experiments

Experimental setup. Experimental environment: Windows 10 system with Intel(R) Core i7-12700H (2.1 GHz), 16 GB RAM and 512 GB Hard Disk. The methods and experiments were implemented in Python 3.11.0.

Data sets. Six types of networks were selected for the experiment. A social network of dolphin populations (Dolphins)⁴⁹. A metabolic network of *C. elegans* (Celegans)⁵⁰. A social network for exchanging e-mails (Email)⁵¹, a power network in the western United States (Power)⁵². A collaborative network of Arxiv paper authors (GrQc)⁵³. A relational network of PGP users (PGP)⁵⁴. These data are available from the <https://networkrepository.com/>. The topological parameters of these networks are listed in Table 4.

Comparison algorithm. The proposed method is compared with five well-known methods, as shown in Table 5.

Evaluation indicators. *Accuracy.* Accuracy is a criterion to evaluate the algorithm's performance. Inspired by previous works^{41,42,44,46}, the accuracy of the algorithm was investigated by the SIR model in this experiment. Here, we conducted 500 simulations for each node and used its average value as the final results. Let σ_u and x_u be the ranking of node u provided by the SIR model and algorithm X , respectively. The ranking sequences σ and x are sorted in descending order. In order to quantify the accuracy of algorithm X , we take the sequence σ as the benchmark and calculates the consistency coefficient between the ranking sequences σ and x . For this purpose, a pair of nodes u and v is considered as follows:

- if $(x_u - x_v) \times (\sigma_u - \sigma_v) > 0$, then nodes u and v are said to have a positive relationship in x .
- if $(x_u - x_v) \times (\sigma_u - \sigma_v) < 0$, then nodes u and v are said to have a negative relationship in x .
- if $(x_u - x_v) \times (\sigma_u - \sigma_v) = 0$, then nodes u and v are said to be independent in x . Based on the above, the consistency coefficient of ranking sequence X , denoted by $Coe(X)$, is defined as

$$Coe(x) = \frac{2 \times (n_a - n_b)}{n \times (n - 1)} \quad (6)$$

where n is the number of nodes, n_a and n_b represent the number of node pairs with positive and negative relationships in the ranking sequence x , respectively. Obviously, $Coe(x) \in [0, 1]$. $Coe(X) = 1$ means that each node pair has a positive relationship and thus the algorithm X has the best accuracy. The worst case is $Coe(X) = 0$.

Resolution. There may be multiple nodes with the same ranking in the ranking sequence. For this reason, we introduce the cumulative distribution function (CDF) to describe the distribution of nodes. Let x be the ranking sequence generated by algorithm X . Let $\omega \in [0, 1]$ be the identity of the nodes, then the mathematical equation of $CDF(\omega)$ can be written by

Network	Type	n	m	$\langle k \rangle$	k_{max}	α_{th}
Dolphins	Community	62	159	5.12	12	0.1470
Celegans	Metabolic	297	2148	14.46	134	0.0384
Email	Social	1133	5451	9.62	71	0.0535
Power	Power	4941	6594	2.67	19	0.2583
GrQc	Collaboration	5241	14,484	5.52	81	0.0593
PGP	Relationship	10,680	24,316	4.55	205	0.0530

Table 4. Topological parameters of six real networks: node number n , edge number m , maximum degree k_{max} , and spread threshold α_{th} ⁵⁵.

Method	Type	Complexity
Degree ¹⁹	Local	$O(n)$
K-shell ²²	Global	$O(n)$
WKSND ²⁸	Combined	$O(n + m)$
CHB ²⁴	Community	$O(n^2)$
MCDE ⁴⁶	Multiple	$O(n + m)$
EHCC (our method)	Multiple	$O(n + m)$

Table 5. The proposed method and five well-known methods.

$$CDF(\omega) = \frac{n_{x,\omega}}{n} \quad (7)$$

where $n_{x,\omega}$ signifies the number of nodes whose identity is greater than or equal to ω in sequence x . Here, the identity a node is the reciprocal of the number of nodes with the same ranking. Obviously, the faster the CDF increases, the greater the number of nodes with high identity, and the higher the resolution.

The monotonicity function measures the resolution by calculating the fraction of the number of nodes that rank differently⁴². Let d_x be the count of nodes in x that are distinguishable (nodes with different rankings are distinguishable). The monotonicity of sequence x is calculated as follows

$$Monotonicity(x) = \left(1 - \frac{\sum_{i=1}^{d_x} n_i \times (n_i - 1)}{n \times (n - 1)} \right)^2 \quad (8)$$

where n_i denotes the number of nodes ranked i in sequence x . $Monotonicity(x) = 1$ indicates the best monotonicity of the ranking sequence x . $Monotonicity(x) = 0$ means that the ranking of all nodes is the same and monotonicity is the worst in this case.

Experimental results. *Results of accuracy.* Table 6 reports the consistency coefficients (Eq. 6) between the different algorithms and the SIR model in the six real networks. Here, the infection rate $\alpha = 1.05 \times \alpha_{th}$ and the recovery rate $\beta = 1$ are chosen in the SIR model. The numerical results show that, except for the network GrQc, EHCC provides the most accurate results compared to the other five methods. In the network GrQc, MCDE performs slightly better than EHCC. It is also noted that EHCC still outperforms the other four methods. In the networks Email, Power, and PGP, EHCC has the highest consistency coefficients, followed by MCDE. In networks Dolphins and Celegans, EHCC and WKSDN have the best and second best performance respectively. Moreover, we observe that the results of the K-shell method are the worst in all networks.

Results of resolution. Figure 2 presents the CDF curves (Eq. 7) of different algorithms in six real networks. It can be observed that the CCDF of EHCC increases the fastest, which illustrates that the node distribution given by EHCC is the best. In network Dolphins, EHCC performs best and WKSDN is second. In networks Celegans and Email, the CDF curves of EHCC and WKSDN almost overlap and are followed by that of MCDE. We also observe performance of EHCC is the best in network Power. Moreover, EHCC is far superior to other methods in networks Power and PGP. Table 7 shows the monotonicity (Eq. 8) of the different methods in the six real networks. The monotonicity scores of EHCC in all the networks except network GRQC were higher than 0.99. Obviously, the monotonicity of EHCC is the highest and K-shell is the worst, which supports the results in Fig. 2. In networks Dolphins, Email, GrQc, and PGP, the best performer is EHCC, followed by WKSDN. In the network Celegans, EHCC follows CHB, but still beats the other methods.

Discussion

This paper investigates the identification of important node in complex networks. In previous studies, degree centrality and K-shell decomposition were introduced as two benchmark methods for identifying important nodes. Degree centrality estimates the influence of a node based on its local characteristics (the number of neighbors). Whereas, K-shell evaluates the influence of a node according to its global characteristics (hierarchy position). However, a single local or global characteristic fails to effectively estimate the importance of the node. As a result, this paper proposes a multi-characteristic approach based on the extended degree and E-shell hierarchy decomposition. The extended degree improves the classical degree by introducing the neighbors' degree. E-shell hierarchy decomposition is used to determine nodes' position through the network's hierarchical structure. Combining these two components (extended degree and position), we define a hybrid characteristic centrality (HCC) that can be a comprehensive indicator of the node's importance. Furthermore, we propose an extended version of HCC called EHCC. As the numerical results listed in Tables 6, 7, and Fig. 2, the accuracy and resolution of the new approach are superior to those of the five well known approaches within the SIR model and the monotonicity function. In addition, it is observed that multi-characteristic methods (EHCC and MCDE)

Network	Degree	K-shell	WKSDN	CHB	MCDE	EHCC
Dolphins	0.8003	0.7409	<u>0.8145</u>	0.7730	0.7657	0.9037
Celegans	0.8030	0.7621	<u>0.8170</u>	0.7886	0.7664	0.9060
Email	0.7894	0.8131	0.8031	0.7856	<u>0.8246</u>	0.9227
Power	0.5979	0.5575	0.6570	0.7309	<u>0.7485</u>	0.8325
GrQc	0.6835	0.6967	0.7955	0.7590	0.8253	<u>0.8205</u>
PGP	0.6178	0.6545	0.7033	0.7425	<u>0.7657</u>	0.7816

Table 6. The consistency coefficients between the node ranking sequences of the different algorithms and the ranking sequences of the SIR model in the six real networks, where the infection rate $\alpha = 1.05 \times \alpha_{th}$ in the SIR model. The highest and second-highest values in each network are highlighted in bold and underlined, respectively.

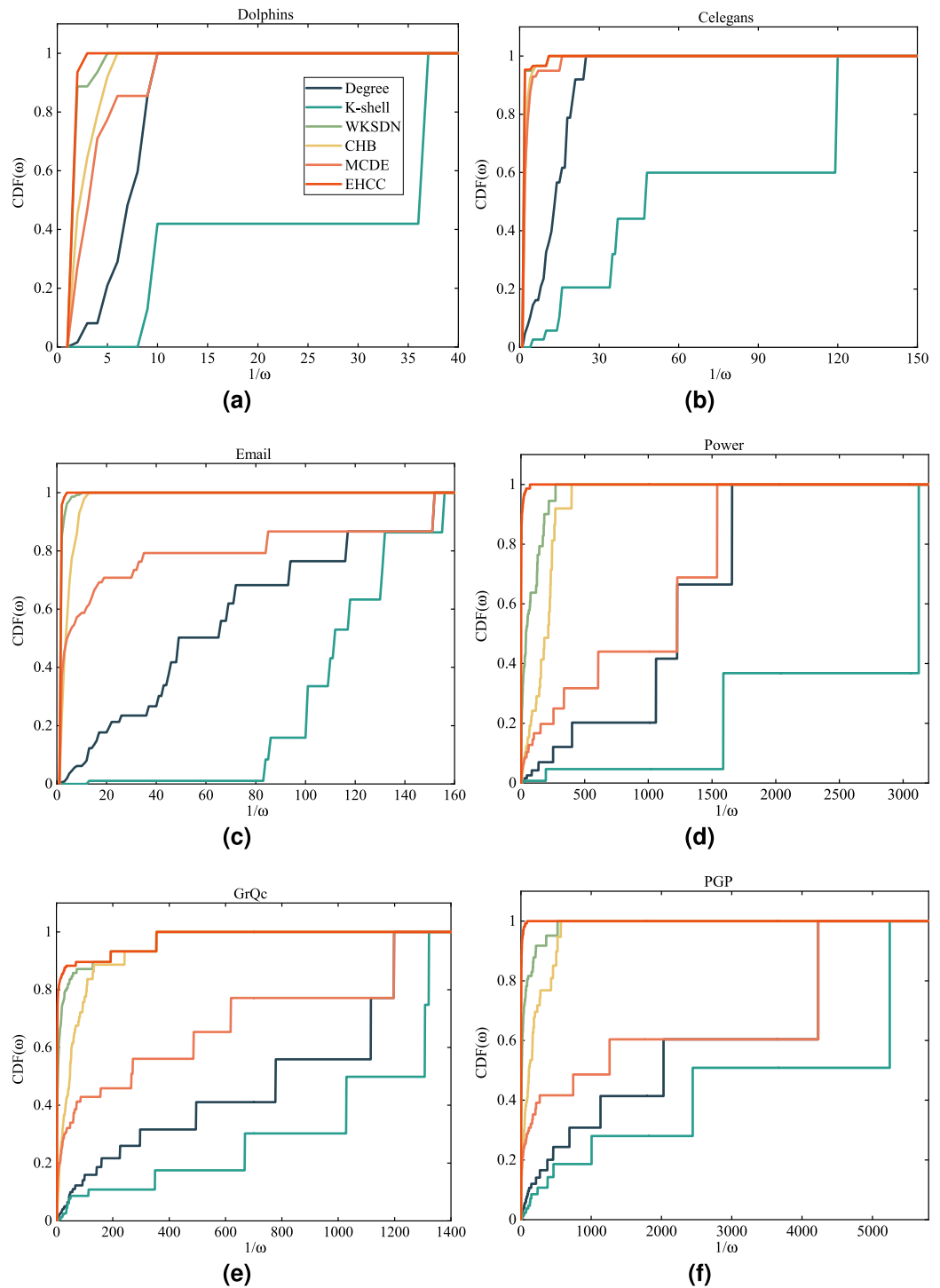


Figure 2. CDF (The cumulative distribution function) of different algorithms in six real networks. ω is the identity of the nodes. (a)–(f) are the networks Dolphins, Celegans, Email, Power, GrQc and PGP respectively.

outperform single-characteristic methods (Degree and K-shell). The work in this paper provides new ideas for learning about and analyzing important nodes. However, some challenges still need to be further investigated. For example, the relationship between local and global features of nodes has not been well-documented. As a result, we will continue to explore new methods in future.

Network	Degree	K-shell	WKSDN	CHB	MCDE	EHCC
Dolphins	0.8312	0.3769	<u>0.9905</u>	0.9612	0.9243	0.9979
Celegans	0.9217	0.6094	0.9925	<u>0.9939</u>	0.9909	0.9977
Email	0.8874	0.8088	<u>0.9994</u>	0.9951	0.9460	0.9999
Power	0.5927	0.2459	<u>0.9688</u>	0.9273	0.6671	0.9987
GrQc	0.7459	0.6630	<u>0.9841</u>	0.9717	0.8417	0.9872
PGP	0.6193	0.4806	<u>0.9872</u>	0.9654	0.6753	0.9992

Table 7. Monotonicity scores of different algorithms in six real networks. The highest and second highest values in each network are highlighted in bold and underlined, respectively.

Data availability

All relevant data are available at <https://networkrepository.com/>.

Code availability

The code covering this article can be accessed via [Github](#).

Received: 12 November 2022; Accepted: 21 February 2023

Published online: 23 February 2023

References

- Duan, Z. S., Chen, G. R. & Huang, L. Complex network synchronizability: Analysis and control. *Phys. Rev. E* **76**, 056103. <https://doi.org/10.1103/PhysRevE.76.056103> (2007).
- Cornelius, S. P., Kath, W. L. & Motter, A. E. Realistic control of network dynamics. *Nat. Commun.* **4**, 1942. <https://doi.org/10.1038/ncomms2939> (2013).
- Hofmann, S. G., Curtiss, J. & McNally, R. J. A complex network perspective on clinical science. *Perspect. Psychol. Sci.* **11**, 597–605. <https://doi.org/10.1177/1745691616639283> (2016).
- Lei, X. J., Yang, X. Q. & Fujita, H. Random walk based method to identify essential proteins by integrating network topology and biological characteristics. *Knowl. Based Syst.* **167**, 53–67. <https://doi.org/10.1016/j.knsys.2019.01.012> (2019).
- Zhang, Y. J. *et al.* Social brain network predicts real-world social network in individuals with social anhedonia. *Psychiatry Res. Neuroimaging* **317**, 111390. <https://doi.org/10.1016/j.pscychresns.2021.111390> (2021).
- Bardoscia, M. *et al.* The physics of financial networks. *Nat. Rev. Phys.* **3**, 490–507. <https://doi.org/10.1038/s42254-021-00322-5> (2021).
- Pagani, G. A. & Aiello, M. The power grid as a complex network: A survey. *Physica A Stat. Mech. Appl.* **392**, 2688–2700. <https://doi.org/10.1016/j.physa.2013.01.023> (2013).
- da Silva, D. C., Bianconi, G., da Costa, R. A., Dorogovtsev, S. N. & Mendes, J. F. F. Complex network view of evolving manifolds. *Phys. Rev. E* **97**, 032316. <https://doi.org/10.1103/PhysRevE.97.032316> (2018).
- He, K., Li, Y. R., Soundarajan, S. & Hoperoft, J. E. Hidden community detection in social networks. *Inf. Sci.* **425**, 92–106. <https://doi.org/10.1016/j.ins.2017.10.019> (2018).
- Zou, Y., Donner, R. V., Marwan, N., Donges, J. F. & Kurths, J. Complex network approaches to nonlinear time series analysis. *Phys. Rep.* **787**, 1–97. <https://doi.org/10.1016/j.physrep.2018.10.005> (2019).
- Zhao, J., Song, Y. T., Liu, F. & Deng, Y. The identification of influential nodes based on structure similarity. *Connect. Sci.* **33**, 201–218. <https://doi.org/10.1080/09540091.2020.1806203> (2021).
- Zhou, Y. M., Wang, Z., Jin, Y. & Fu, Z. H. Late acceptance-based heuristic algorithms for identifying critical nodes of weighted graphs. *Knowl. Based Syst.* **211**, 106562. <https://doi.org/10.1016/j.knsys.2020.106562> (2021).
- Wang, B., Zhang, J. K., Dai, J. Y. & Sheng, J. F. Influential nodes identification using network local structural properties. *Sci. Rep.* **12**, 1833. <https://doi.org/10.1038/s41598-022-05564-6> (2022).
- Srinivas, S. & Rajendran, C. Community detection and influential node identification in complex networks using mathematical programming. *Expert Syst. Appl.* **135**, 296–312. <https://doi.org/10.1016/j.eswa.2019.05.059> (2019).
- Zhang, B. *et al.* A most influential node group discovery method for influence maximization in social networks: A trust-based perspective. *Data Knowl. Eng.* **121**, 71–87. <https://doi.org/10.1016/j.datak.2019.05.001> (2019).
- Barabasi, A. L., Gulbahce, N. & Loscalzo, J. Network medicine: A network-based approach to human disease. *Nat. Rev. Genet.* **12**, 56–68. <https://doi.org/10.1038/nrg2918> (2011).
- Cho, Y., Hwang, J. & Lee, D. Identification of effective opinion leaders in the diffusion of technological innovation: A social network approach. *Techol. Forecast. Soc. Change* **79**, 97–106. <https://doi.org/10.1016/j.techfore.2011.06.003> (2012).
- Zhu, T., Wang, B., Wu, B. & Zhu, C. X. Maximizing the spread of influence ranking in social networks. *Inf. Sci.* **278**, 535–544. <https://doi.org/10.1016/j.ins.2014.03.070> (2014).
- Freeman, L. C. Centrality in social networks conceptual clarification. *Soc. Netw.* **1**, 215–239. [https://doi.org/10.1016/0378-8733\(78\)90021-7](https://doi.org/10.1016/0378-8733(78)90021-7) (1979).
- Buechel, B. & Buskens, V. The dynamics of closeness and betweenness. *J. Math. Sociol.* **37**, 159–191. <https://doi.org/10.1080/0022250X.2011.597011> (2013).
- Du, Y. X. *et al.* A new closeness centrality measure via effective distance in complex networks. *Chaos* **25**, 033112. <https://doi.org/10.1063/1.4916215> (2015).
- Lv, L. Y., Zhou, T., Zhang, Q. M. & Stanley, H. E. The h-index of a network node and its relation to degree and coreness. *Nat. Commun.* **7**, 10168. <https://doi.org/10.1038/ncomms10168> (2016).
- Su, G., Kuchinsky, A., Morris, J. H., States, D. J. & Meng, F. Glay: Community structure analysis of biological networks. *Bioinformatics* **26**, 3135–3137. <https://doi.org/10.1093/bioinformatics/btq596> (2010).
- Ghalmene, Z., Hassouni, M. E. & Cherifi, H. Immunization of networks with non-overlapping community structure. *Soc. Netw. Anal. Min.* **9**, 45. <https://doi.org/10.1007/s13278-019-0591-9> (2019).
- Rajeh, S., Savonnet, M., Leclercq, E. & Cherifi, H. Characterizing the interactions between classical and community-aware centrality measures in complex networks. *Sci. Rep.* **11**, 10088. <https://doi.org/10.1038/s41598-021-89549-x> (2021).

26. Rajeh, S., Savonnet, M., Leclercq, E. & Cherifi, H. Comparative evaluation of community-aware centrality measures. *Qual. Quant.* <https://doi.org/10.1007/s11135-022-01416-7> (2022).
27. Zhao, Z. Y., Wang, X. F., Zhang, W. & Zhu, Z. L. A community-based approach to identifying influential spreaders. *Entropy* **17**, 2228–2252. <https://doi.org/10.3390/e17042228> (2015).
28. Tulu, M. M., Hou, R. & Younas, T. Identifying influential nodes based on community structure to speed up the dissemination of information in complex network. *IEEE Access* **6**, 7390–7401. <https://doi.org/10.1109/ACCESS.2018.2794324> (2018).
29. Magelinski, T., Bartulovic, M. & Carley, K. M. Measuring node contribution to community structure with modularity vitality. *IEEE Trans. Netw. Sci. Eng.* **8**, 707–723. <https://doi.org/10.1109/TNSE.2020.3049068> (2021).
30. Seidman, S. B. Network structure and minimum degree. *Soc. Netw.* **5**, 269–287. [https://doi.org/10.1016/0378-8733\(83\)90028-X](https://doi.org/10.1016/0378-8733(83)90028-X) (1983).
31. Kitsak, M. *et al.* Identification of influential spreaders in complex networks. *Nat. Phys.* **6**, 888–893. <https://doi.org/10.1038/nphys1746> (2010).
32. Liu, Y., Tang, M., Zhou, T. & Do, Y. Improving the accuracy of the k-shell method by removing redundant links: From a perspective of spreading dynamics. *Sci. Rep.* **5**, 13172. <https://doi.org/10.1038/srep13172> (2015).
33. Maji, G. Influential spreaders identification in complex networks with potential edge weight based k-shell degree neighborhood method. *J. Comput. Sci.* **39**, 101055. <https://doi.org/10.1016/j.jocs.2019.101055> (2020).
34. Li, Z. & Huang, X. Y. Identifying influential spreaders in complex networks by an improved gravity model. *Sci. Rep.* **11**, 22194. <https://doi.org/10.1038/s41598-021-01218-1> (2021).
35. Maji, G., Mandal, S. & Sen, S. A systematic survey on influential spreaders identification in complex networks with a focus on k-shell based techniques. *Expert Syst. Appl.* **161**, 113681. <https://doi.org/10.1016/j.eswa.2020.113681> (2020).
36. Sun, P. G., Miao, Q. G. & Staab, S. Community-based k-shell decomposition for identifying influential spreaders. *Pattern Recognit.* **120**, 108130. <https://doi.org/10.1016/j.patcog.2021.108130> (2021).
37. Lahav, N. *et al.* K-shell decomposition reveals hierarchical cortical organization of the human brain. *New J. Phys.* **18**, 083013. <https://doi.org/10.1088/1367-2630/18/8/083013> (2016).
38. Angelou, K., Maragakis, M. & Argyrakis, P. A structural analysis of the patent citation network by the k-shell decomposition method. *Physica A Stat. Mech. Appl.* **521**, 476–483. <https://doi.org/10.1016/j.physa.2019.01.063> (2019).
39. Ren, Z. M., Liu, J. G., Shao, F., Hu, Z. L. & Guo, Q. Analysis of the spreading influence of the nodes with minimum k-shell value in complex networks. *Acta Phys. Sin.* **62**, 108902. <https://doi.org/10.7498/aps.62.108902> (2013).
40. Liu, J. G., Ren, Z. M. & Guo, Q. Ranking the spreading influence in complex networks. *Physica A Stat. Mech. Appl.* **392**, 4154–4159. <https://doi.org/10.1016/j.physa.2013.04.037> (2013).
41. Zeng, A. & Zhang, C. J. Ranking spreaders by decomposing complex networks. *Phys. Lett. A* **377**, 1031–1035. <https://doi.org/10.1016/j.physleta.2013.02.039> (2013).
42. Bae, J. & Kim, S. Identifying and ranking influential spreaders in complex networks by neighborhood coreness. *Physica A Stat. Mech. Appl.* **395**, 549–559. <https://doi.org/10.1016/j.physa.2013.10.047> (2014).
43. Feng, J., Shi, D. D. & Luo, X. Y. An identification method for important nodes based on k-shell and structural hole. *J. Complex Netw.* **6**, 342–352. <https://doi.org/10.1093/comnet/cnx035> (2018).
44. Wang, Y. M., Chen, B., Li, W. D. & Zhang, D. P. Influential node identification in command and control networks based on integral k-shell. *Wirel. Commun. Mob. Comput.* **2019**, 6528431. <https://doi.org/10.1155/2019/6528431> (2019).
45. Liu, X. Y., Ye, S., Fiumara, G. & De Meo, P. Influential spreaders identification in complex networks with topsis and k-shell decomposition. *IEEE Trans. Comput. Soc. Syst.* <https://doi.org/10.1109/TCSS.2022.3148778> (2022).
46. Sheikhamadi, A. & Nematbakhsh, M. A. Identification of multi-spreader users in social networks for viral marketing. *J. Inf. Sci.* **43**, 412–423. <https://doi.org/10.1177/0165551516644171> (2017).
47. Satsuma, J., Willox, R., Ramani, A., Grammaticos, B. & Carstea, A. S. Extending the sir epidemic model. *Physica A Stat. Mech. Appl.* **336**, 369–375. <https://doi.org/10.1016/j.physa.2003.12.035> (2004).
48. Holme, P. Fast and principled simulations of the sir model on temporal networks. *PLoS ONE* **16**, 0246961. <https://doi.org/10.1371/journal.pone.0246961> (2021).
49. Lusseau, D. *et al.* The bottlenose dolphin community of doubtful sound features a large proportion of long-lasting associations. *Behav. Ecol. Sociobiol.* **54**, 396–405. <https://doi.org/10.1007/s00265-003-0651-y> (2003).
50. Duch, J. & Arenas, A. Community detection in complex networks using extremal optimization. *Phys. Rev. E* **72**, 027104. <https://doi.org/10.1103/PhysRevE.72.027104> (2005).
51. Guimerà, R., Danon, L., Diaz-Guilera, A., Giralt, F. & Arenas, A. Self-similar community structure in a network of human interactions. *Phys. Rev. E* **68**, 065103. <https://doi.org/10.1103/PhysRevE.68.065103> (2003).
52. Watts, D. J. & Strogatz, S. H. Collective dynamics of 'small-world' networks. *Nature* **393**, 440–442. <https://doi.org/10.1038/30918> (1998).
53. Leskovec, J., Kleinberg, J. & Faloutsos, C. Graph evolution: Densification and shrinking diameters. *ACM Trans. Knowl. Discov. Data* **1**, 2–43. <https://doi.org/10.1145/1217299.1217301> (2007).
54. Boguñá, M., Pastor-Satorras, R., Diaz-Guilera, A. & Arenas, A. Models of social networks based on social distance attachment. *Phys. Rev. E* **70**, 056122. <https://doi.org/10.1103/PhysRevE.70.056122> (2004).
55. Moreno, Y., Pastor-Satorras, R. & Vespignani, A. Epidemic outbreaks in complex heterogeneous networks. *Eur. Phys. J. B* **26**, 521–529. <https://doi.org/10.1140/epjb/e20020122> (2002).

Acknowledgements

This work is partially supported by the National Natural Science Foundation of China (Grant Nos. 11901073 and 62272077).

Author contributions

J.Z. designed the research project. J.L. performed numerical computations and analysis. J.Z. and J.L. wrote the paper.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to J.Z.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023