

Important edge identification in complex networks based on local and global features

Jia-Hui Song(宋家辉)
Citation: Chin. Phys. B, 2023, 32 (**9**): 098901. DOI: 10.1088/1674-1056/aca6d8
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 multiple influential spreaders in complex networks based on spectral graph theory

Dong-Xu Cui(崔东旭), Jia-Lin He(何嘉林), Zi-Fei Xiao(肖子飞), and Wei-Ping Ren(任卫平)
Chin. Phys. B, 2023, 32 (**9**): 098904. DOI: 10.1088/1674-1056/acac0b

Self-similarity of complex networks under centrality-based node removal strategy

Dan Chen(陈单), Defu Cai(蔡德福), and Housheng Su(苏厚胜)
Chin. Phys. B, 2023, 32 (**9**): 098903. DOI: 10.1088/1674-1056/acd3e2

Identification of key recovering node for spatial networks

Zijian Yan(严子健), Yongxiang Xia(夏永祥), Lijun Guo(郭丽君), Lingzhe Zhu(祝令哲), Yuanyuan Liang(梁圆圆), and Haicheng Tu(涂海程)
Chin. Phys. B, 2023, 32 (**6**): 068901. DOI: 10.1088/1674-1056/acb75f

AG-GATCN: A novel method for predicting essential proteins

Peishi Yang(杨培实), Pengli Lu(卢鹏丽), and Teng Zhang(张腾)
Chin. Phys. B, 2023, 32 (**5**): 058902. DOI: 10.1088/1674-1056/acb9f9

AIGCrank: A new adaptive algorithm for identifying a set of influential spreaders in complex networks based on gravity centrality

Ping-Le Yang(杨平乐), Lai-Jun Zhao(赵来军), Chen Dong(董晨),Gui-Qiong Xu(徐桂琼), and Li-Xin Zhou(周立欣)
Chin. Phys. B, 2023, 32 (**5**): 058901. DOI: 10.1088/1674-1056/ac8e56

Important edge identification in complex networks based on local and global features

Jia-Hui Song(宋家辉)†

College of Science, China Three Gorges University, Yichang 443002, *China*

(Received 1 October 2022; revised manuscript received 16 November 2022; accepted manuscript online 29 November 2022)

Identifying important nodes and edges in complex networks has always been a popular research topic in network science and also has important implications for the protection of real-world complex systems. Finding the critical structures in a system allows us to protect the system from attacks or failures with minimal cost. To date, the problem of identifying critical nodes in networks has been widely studied by many scholars, and the theory is becoming increasingly mature. However, there is relatively little research related to edges. In fact, critical edges play an important role in maintaining the basic functions of the network and keeping the integrity of the structure. Sometimes protecting critical edges is less costly and more flexible in operation than just focusing on nodes. Considering the integrity of the network topology and the propagation dynamics on it, this paper proposes a centrality measure based on the number of high-order structural overlaps in the first and second-order neighborhoods of edges. The effectiveness of the metric is verified by the infection–susceptibility (SI) model, the robustness index R, and the number of connected branches θ. A comparison is made with three currently popular edge importance metrics from two synthetic and four real networks. The simulation results show that the method outperforms existing methods in identifying critical edges that have a significant impact on both network connectivity and propagation dynamics. At the same time, the near-linear time complexity can be applied to large-scale networks.

Keywords: complex networks, high-order structure, edge importance, connectivity, propagation dynamics

PACS: 89.75.Fb **DOI:** 10.1088/1674-1056/aca6d8

1. Introduction

Some real-world infrastructure systems are subject to targeted attacks and unpredictable failures, which can result in significant financial losses. Understanding the "roles" played by the different components of a system allows us to find the critical components so that we can take advanced measures to protect or increase the resilience of the system. We model realistic complex systems as complex networks whose nodes represent the constituent entities of the system. The edges between them represent the interrelationships between these entities. In this representation, the problem of critical component identification can be easily formulated as quantifying the importance of nodes or edges in the network. There are already many methods to identify important nodes in a network. For example, degree,[1] H-index[2] and k-shell[3] are based on nodes' neighbors; closeness centrality[4] and betweenness centrality[5] are based on the path; and quasi-Laplacian centrality,[6] the local gravity model[7] and AWLM[8] are based on semi-local structural information. In contrast, the identification and importance ranking of key edges has received less attention. This is because the density of edges in a complex network often determines the complexity of the network, and the sizes of the edges are much larger than those of nodes. Therefore, the contribution of individual edges to the functional properties of the whole network is usually very small, and some subtle changes are difficult to capture. In addition, the contributions from different types of edges to the network are usually interrelated, and this mechanism of "entanglement" in each other makes it difficult to describe their importance individually. These limitations increase the difficulty of critical edge identification to some extent.[9,10] In fact, the identification of critical edges plays an important role in the study of complex networks, as it helps to guide or control the network from a global perspective. First, a network can be made safe from target attacks by protecting critical edges. For example, we can enhance the robustness of a power grid by protecting critical transmission lines.[11] Second, removing critical nodes can be "costly" when we want to suppress the spread of diseases or information.[12] In contrast, strategies that block the path of virus or rumor transmission may be more suitable for certain situations. For example, isolating an airport is an impractical and costly measure, but cutting the route between two locations can be achieved more flexibly and at a lower cost. Current approaches to identifying critical edges focus on the topology of the network. Ball *et al.*[13] show that the importance of an edge can be measured by the relative amount of change in the average path length of the network after removing that edge. Newman *et al.*[14] used edge medial centrality (EB) to quantify the importance of edges. Yu *et al.*[15] proposed a method named $\mathrm{BCC_{MOD}}$ which calculates the number of clusters on an edge considering that the greater the number of clusters an edge belongs

†Corresponding author. E-mail: songjiahuizz123456@163.com

© 2023 Chinese Physical Society and IOP Publishing Ltd

to, the less important it is. However, it is not applicable to large-scale networks because it is time-consuming. Subsequently, many researchers started to use local information to describe the importance of edges in order to reduce the time complexity. For example, Kanwar *et al.*[16] proposed a metric named EC_{DCN} using the betweenness of edges, the degree of nodes at both ends of the edge and the number of common neighbors. Simulation results show that the method has a low time complexity to quickly identify the key edges in large networks. However, the accuracy of this metric is low because it only considers the situation within the first-order neighborhood. Holme *et al.*[17] assumed that edges between important nodes in the network are more important than other edges and proposed the degree product (DP) index to evaluate the importance of edges. Onnela *et al.*[18] proposed the topological overlap (TO) to quantify the effect of common neighbors of nodes on the importance of edges. Cheng *et al.*[19] found that the nodes at both ends of an edge in a small cluster that are also in a large cluster are important for the connectivity of the network and proposed the bridge degree (BN) index. Liu *et al.*[20] proposed diffusion intensity (DI) to identify critical edges from the perspective of diffusion dynamics. Matamalas *et al.*[21] considered the characteristics or parameters of the epidemic transmission process and proposed a set of discrete-time control equations (ELE) that can be closed and analyzed to evaluate the contribution of links to the propagation process in complex networks. This also provides a new way of thinking to analyze dynamic diffusion models at the link rather than node level. Xia *et al.*[22] proposed an importance metric that can be used to quantify a single edge or a set of edges based on the connectivity of nearest neighbors of nodes. Ren *et al.*[23] proposed an edge ranking algorithm that considers both local and global information by combining the node influence distribution law and the irreplaceability of edges. Chen *et al.*[24] proposed a novel and effective sub-graph overlap (*SO*) metric based on the degree of overlap of communities within an edge neighborhood. This method can identify the key edges that maintain the "communication" between different communities (tightly connected sub-networks) in the network. However, this metric is more suitable for networks with a high degree of modularity and is less accurate in the face of homogeneous networks with a high link density of edges. Zhao *et al.*[25] considered the second-order neighborhoods of the nodes at both ends of an edge and proposed a second-order neighborhood (*SN*) index to quantify the role played by an edge in maintaining network connectivity. The *SN* faces a similar problem to the *SO* in not being able to identify the important edges of a high-density homogeneous network. It can be seen that the current edge importance metrics focus on identifying "bridge edges" between communities in heterogeneous networks with high modularity, but not for homogeneous networks with high link density. In this paper, we summarize the shortcomings of the above studies and propose an edge importance metric based on the number of higher-order structures within the first and second-order neighborhoods of an edge to identify the important edges in a homogeneous network. In addition to the triangles in the first-order neighborhood of the edge, the fully connected quadrilateral and diagonal quadrilateral in the first-order neighborhood of the edge and the empty quadrilateral and diagonal quadrilateral in the second-order neighborhood are also taken into account, so that the higher-order structures considered are more diverse and richer in variety. The importance of edges in the global position of the network is also considered. The main idea of this metric is that the removal of edges from the network causes a decrease in the connectivity and propagation capacity of the network structure, and has an impact on the integrity of the network structure. We use this metric as a measure to compare with the current representative measures of edge importance EC_{DCN}, *SO* and *SN* metrics. The effectiveness of the metric is verified in six networks by five aspects: the SI model, robustness metric R, number of connected branches θ, Kendall's tau correlation coefficient τ, and monotonicity $M(R)$. The numerical simulation results show that the proposed edge importance metric can identify the edges that have the greatest influence on the network structure and function more quickly and accurately.

The paper is structured as follows. In the first part, a brief introduction is given to the basic concepts of complex networks and the three edge importance measures that are currently most popular for comparison, along with new definitions of higher-order structures on edges, first and second-order neighborhoods of edges, and the concepts of inner and outer edges. In the second part, the new measures for evaluating edge importance are introduced. In the third section, the data set used is described and simulations are performed in six networks to derive the superiority of the proposed metric. Finally, we conclude with a summary of the work done and some prospects for future research.

2. Centrality measure

Before we start, we introduce the basic concepts of complex networks. The complex network can be represented by $G = (V, E)$, where G is an undirected connected graph, with n nodes, m edges, V represents the set of nodes, E represents the set of edges, the adjacency matrix of G of $Q = [a_{ij}]$ has n rows and n columns, and the elements in Q are defined by a_{ij} as follows: (i) $a_{ij} = 1$, the ith node is connected to the jth node; (ii) $a_{ij} = 0$, the ith node is infinitely connected to the jth node.

2.1. EC_{DCN}

In a complex network, the following information is required in order to identify a critical edge and properly quantify its diffusion capacity.

① The global position of the edge in the network.

② Local topological features in the network that promote diffusion through that edge.

③ Local topological features that suppress diffusion through this edge in the network.

Based on this, Kanwar *et al*. proposed a hybrid edge centrality metric that combines edge betweenness, the product of node degrees at both ends of the edge and the number of common neighbors. It is defined as

$$EC_{\mathrm{DCN}}(u,v) = \frac{BC(u,v) \times DP(u,v)}{CN(u,v)}, \tag{1}$$

where u and v represent the two end nodes on the edge, and this edge is denoted by $e = (u,v)$. It locates an edge through a global perspective using the edge betweenness $BC(u,v)$. The local topological features that facilitate diffusion through edge e are represented by the degree product $DP(u,v)$ of nodes u and v. To quantify the local topological features that impede the diffusion of edge e, we use the common neighbor counts $CN(u,v)$ of nodes u and v. The time complexity of this metric is $O(|V||E|)$.[16]

2.2. *SO*

The literature[25,26] states that a common edge between two different communities in a network is often more important than an edge within a community in facilitating communication between communities in the network. Based on this, Chen *et al*. describe the importance of edges based on the degree of overlap of communities within an edge neighborhood. For a given static network G and an edge e_{ij}, there is a subgraph of G that contains nodes i, j and nodes whose distance from i or j is not greater than 2. The *SO* metric is defined as

$$SO(i,j) = \frac{\max\left(1, \left|\Gamma_i^{(2)} \bigcap \Gamma_j^{(2)}\right|\right)^2}{\left|G_{ij}\right|}, \tag{2}$$

where $\left|G_{ij}\right|$ denotes the number of nodes in G_{ij}. $\Gamma_i^{(2)}$ is the set of nodes in $G_{ij} \backslash e_{ij}$ (the subgraph from G_{ij} which edge e_{ij} is removed) whose distance to node i does not exceed 2. Lower *SO* values indicate that edges are more likely to connect large-scale communities with few other links between them. *SO* relies only on local topological information and has a time complexity of $O(M\langle k\rangle^2)$ (M is the number of edges and $\langle k\rangle$ is the average degree).[23]

2.3. *SN*

If node i and node j have few second-order common neighbors, then edge e_{ij} is more likely to be potential "bridges" between two different communities. Based on this, Zhao *et al*. proposed a second-order neighborhood (*SN*) index to quantify the importance of an edge. The *SN* index of an arbitrary edge e_{ij} is introduced. It is defined $n_{i\backslash j}^{(2)}$ as the number of nodes whose distance to node i is 2 in the subgraph $G\backslash e_{ij}$ where the edge e_{ij} is removed from G. The $n_{j\backslash i}^{(2)}$ is defined in the same way. the *SN* is defined as

$$SN(i,j) = \frac{\left|n_{i\backslash j}^{(2)} \bigcap n_{j\backslash i}^{(2)}\right|}{\left|n_{i\backslash j}^{(2)} \bigcup n_{j\backslash i}^{(2)}\right|}. \tag{3}$$

If node i and node j have few second-order common neighbors, edge e_{ij} is more likely to be a potential "bridge" between two different communities, which is essential to facilitate communication between them.[25,26] Therefore, the smaller the *SN* index, the more important the role played by edges. The computational complexity of *SN* is $O(M\langle k\rangle)$, where M is the number of edges and $\langle k\rangle$ is the average degree.[23]

2.4. Higher-order structures on the edge

Networks are inherently limited to describing pairwise interactions, whereas real-world systems are usually characterized by higher-order interactions involving three or more units. Therefore, higher-order structures, such as hypergraphs and simple complex shapes, are better tools for depicting the real structure of many social, biological, and man-made systems.[27] A triangle is the smallest higher-order structure, which is a third-order (node number 3) connected subgraph formed by three edges interconnected at the beginning and end. It serves as the basic component unit of the association and the basic topology of the network. The size of the higher-order structure is between a single node and a large dense association. It is a fully or partially connected subgraph consisting of a few nodes connected in a guaranteed closed loop, and it appears more frequently in real networks than in their corresponding random networks, and increases with the size of the network and the density of connections between nodes. There are five main cases of third and fourth-order structures of an edge e in its first and second-order neighborhoods (with the number of nodes 3 and 4). Here the first-order neighborhood of an edge e is the set of nodes whose maximum distance from the two end nodes of e is 1, and the second-order neighborhood of e is the set of nodes whose maximum distance from the two end nodes of e is 2 (for simplicity, we assume that the distance between two directly connected nodes is 1). If the maximum distance of all nodes except node i, j from i, j on the higher-order structure with e_{ij} as one side is 1, then the higher-order structure is said to be in the first-order neighborhood of e_{ij}. If the maximum distance of all the nodes except node i, j from i, j on the higher-order structure with e_{ij} as

one side is 2, then the higher-order structure is said to be located in the second-order neighborhood of e_{ij}. As shown in Fig. 1, ①–⑤ are the third-order and fourth-order structures within the first-order and second-order neighborhoods of edge e_{ij}, respectively. From Fig. 1, we can see that in ①, ②, ④ and ⑤, the break of edge e_{ij} causes the higher-order structure to change from a closed state to an open state. In ③, however, the disconnection of e_{ij} does not cause the higher-order structure to change from closed to open. In fact, in the neural network of the brain, even if e_{ij} is disconnected in ③, a basic functional module can still be maintained; for example, basic information transformation and storage can still be completed. However, the disconnection of edge e_{ij} in ①, ②, ④ and ⑤ results in the loss of this specific function.[28–30] However, we can be sure that in ①–⑤, the disconnection of e_{ij} makes the shortest path length from $i \to j$ larger in all cases. This predicts a decrease in communication or transmission efficiency in communication networks and grids.[31–34] We refer to e_{ij} in the first case as the outer edge and e_{ij} in the second case as the inner edge. The inner and outer edges play different "roles" in different complex environments. Furthermore, these five specific higher-order structures within the first- and second-order neighborhoods of the edges considered in this paper have different roles in different real-world complex systems. Some can reflect the resilience level of the system and serve as an early warning of system failure or risk, while others play an irreplaceable role in the functionality of the system. For example, in financial networks the decrease in the number of ①, ④ (fully connected) types predicts possible financial risks and in infrastructure networks the number of ①, ④ types can reflect the vulnerability of the network at a certain level,[35,36] while in biological networks the sudden change in the number of ②, ③ (partially connected) types is considered as a precursor of cellular carcinogenesis[37,38] and the structure of ⑤ (loop) types in neural networks of the brain assume important functions such as storage and transformation of information.[39,40] Therefore, we consider them together in combination.

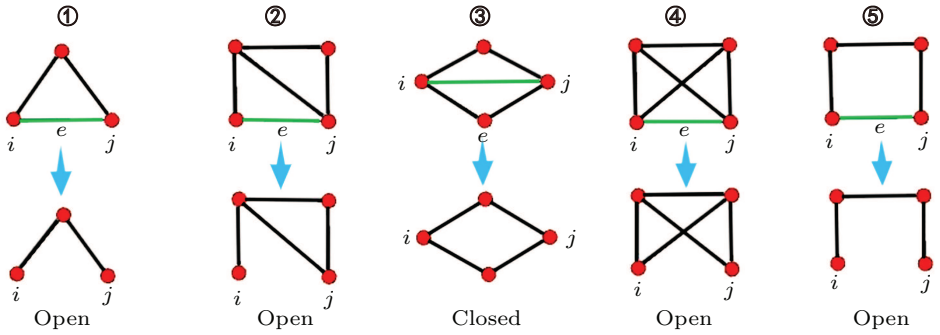


Fig. 1. ①–⑤ denote the higher-order structures in the first-order and second-order neighborhoods of edge e_{ij}, where ① denotes a triangle in the e_{ij} first-order neighborhood, ② denotes a diagonal quadrilateral in the e_{ij} second-order neighborhood, ③, ④ denotes a diagonal quadrilateral and a fully connected quadrilateral in the e_{ij} first-order neighborhood, and ⑤ denotes an empty quadrilateral in the e_{ij} second-order neighborhood.

3. Proposed method

This paper proposes an edge importance measure, i.e., edge clustering coefficients, and extends it to the second-order neighborhood. The scope is extended from the first-order neighborhood of edges to the second-order neighborhood, and the types of higher-order structures are extended from a single triangle to a diagonal quadrilateral, fully connected quadrilateral, and empty quadrilateral. The effects of triangles, diagonal quadrilaterals, fully connected quadrilaterals and empty quadrilaterals on the network topology and propagation dynamics are considered comprehensively. The effect of edges on the global characteristics of the network (average path length) is also considered. The equation for this metric is expressed as follows:

$$EI(e_{ij}) = \frac{N'_{\Delta} + N'_{\alpha} + N'_{\eta} + N'_{\beta} + N'_{\blacksquare}}{N_{\Delta} + N_{\alpha} + N_{\eta} + N_{\beta} + N_{\blacksquare}} \times \left| L - L' \right|, \tag{4}$$

where k_i, k_j are the degrees of nodes i, j, respectively. $N_{\Delta} = (k_i + k_j - 2)$ denotes the maximum possible number of triangles in a first-order neighborhood with e_{ij} as the outer edge. $N_{\alpha} = (k_i - 1)(k_j - 1)$ denotes the maximum possible number of empty quadrilaterals in a second-order neighborhood thought to be the outer edge of e_{ij}. $N_{\beta} = (N_{\Delta}/2)$ denotes the maximum possible number of diagonal quadrilaterals in the first-order neighborhood with e_{ij} as the inner edge. $N_{\eta} = (k_i + k_j - 2)(k_i + k_j - 3)$ denotes the maximum possible number of diagonal quadrilaterals in the second-order neighborhood with e_{ij} as the outer edge. $N_{\blacksquare} = (N_{\Delta}/2)$ denotes the maximum possible number of fully connected quadrilaterals in the first-order neighborhood with e_{ij} as the outer edge. In summary, the denominator represents the sum of the maximum number of triangles, diagonal quadrilaterals, fully connected quadrilaterals and empty quadrilaterals in the first and second-order neighborhoods with e_{ij} as the inner and outer edges in the ideal case. Let $\Gamma_{i\setminus j}$ be the set of nodes in the first-order

neighborhood of node i excluding node j. $|\Gamma_{i\setminus j}|$ represents the number of nodes in the set. $\Gamma^{(2)}_{i\setminus j}$ is the set of nodes in the second-order neighborhood of node i excluding node j. $|\Gamma^{(2)}_{i\setminus j}|$ represents the number of nodes in the set. $\Gamma_{j\setminus i}$ and $|\Gamma^{(2)}_{j\setminus i}|$ are defined in the same way. Let $\Phi = |\Gamma_{i\setminus j} \bigcap \Gamma^{(2)}_{j\setminus i}| + |\Gamma_{j\setminus i} \bigcap \Gamma^{(2)}_{i\setminus j}|$; then, the actual number of empty quadrilaterals and diagonal quadrilaterals in the second-order neighborhood with e as the outer edge is given as

$$N'_{\alpha}+N'_{\eta} = \begin{cases} \frac{\Phi}{2} + \left|\Gamma_{i\setminus j} \bigcap \Gamma_{(\Gamma_{i\setminus j}\cap\Gamma_{j\setminus i})\setminus i,j}\right| \\ \quad + \left|\Gamma_{j\setminus i} \bigcap \Gamma_{(\Gamma_{i\setminus j}\cap\Gamma_{j\setminus i})\setminus i,j}\right| \Phi \in Q, \\ \frac{\Phi+1}{2} \Phi \in w, \end{cases}$$

where $\Gamma_{(\Gamma_{i\setminus j}\cap\Gamma_{j\setminus i})\setminus i,j}$ denotes the set of nodes that do not include$i$, j in the first-order neighborhood of the common neighbor nodes of i and j. Q and w represent the even set and the odd set, respectively. $N'_{\beta} = \left(N'_{\Delta}/2\right)$ denotes the actual number of diagonal quadrilaterals in the first-order neighborhood with e_{ij} as the inner edge. $N'_{\blacksquare} = \left|e_{ab(a\neq b.a,b\in\Gamma_{i\setminus j}\cap\Gamma_{j\setminus i}}\right|$ denotes the actual number of fully connected quadrilaterals in the first-order neighborhood with e_{ij} as the outer edge. In summary, the numerator denotes the sum of the number of triangles, diagonal quadrilaterals, fully connected quadrilaterals and empty quadrilaterals in the first- and second-order neighborhoods that are thought to be the inner and outer edges in the real case. $L = \frac{1}{n(n-1)}\sum d$ denotes the average path length of the original network. $L' = \frac{1}{(n-1)(n-2)}\sum d$ denotes the average path length of the network after nodes i and j at both ends of edge e_{ij} shrink to one node. d is the shortest distance between any two nodes in the network. If the average path length of the network changes significantly before and after the node contraction, this edge is considered important for the transmission efficiency of the network. In the face of large-scale networks, for molecules this information is available in the adjacency and association matrices generated by the statistical software Gephi. For L and L' it can also be obtained by Gephi.

In the following, we illustrate briefly by giving an example. As shown in Figure 2, the EI value of the edge e_{12} is to be calculated. $k_1 = 7$, $k_2 = 6$, $\Gamma_{1\setminus 2} = \{5,3,9,8,7,6\}$, $\Gamma_{2\setminus 1} = \{3,4,7,8,9\}$, $\Gamma^{(2)}_{1\setminus 2} = \{4\}$, and $\Gamma^{(2)}_{2\setminus 1} = \{5,6\}$. First, we start by calculating the terms in the denominator. $N_{\Delta} = 11$, $N_{\alpha} = 30$, $N_{\eta} = 110$, $N_{\beta} = 55$, and $N_{\blacksquare} = 55$. It should be emphasized here that the maximum number of diagonal quadrilaterals and fully connected quadrilaterals in the e_{12} first-order neighborhood is equal. For example, connecting vertices 3 and 8 of diagonal quadrilateral 1328 yields a fully connected quadrilateral. Connecting the diagonal vertices 6 and 2 of the diagonal quadrilateral 1286 in the second-order neighborhood of e_{12} also yields a fully connected quadrilateral but does not count this case in $N_{\blacksquare}$. This is mainly because node 6 is a second-order neighbor of 2. Instead, our initial range is locked in the first-order neighborhood of nodes 1, 2 to find the fully connected quadrilateral, i.e., the degree of nodes 1, 2 is assumed to be constant in advance. Connecting nodes 6, 2 increase the degree of 2, which is obviously inconsistent. Second, we first calculate the terms in the numerator. $N'_{\Delta} = 4$, $\Phi = 3 \in w$, $N'_{\alpha} + N'_{\eta} = 2$, $N'_{\beta} = 6$, and $N'_{\blacksquare} = 1$. Hence, $EI = 4+2+6+1/11+30+110+55+55\times(1.64-1.61) \approx 0.001$. Next, let us consider the computational complexity of the proposed importance metric. Given a network with n nodes and m edges, the average degree of the nodes is $\langle k\rangle$ and the computational complexity of computing the number of first-order neighbor nodes for each edge is $\mathrm{O}(\langle k\rangle)$. The computational complexity of computing the number of second-order neighbor nodes for each edge is $O(\langle k\rangle^2)$. Then the total computational complexity is $\mathrm{O}(m(\langle k\rangle + \langle k\rangle^2))$. Thus, the near-linear computational complexity makes the proposed metric scalable to large-scale networks.

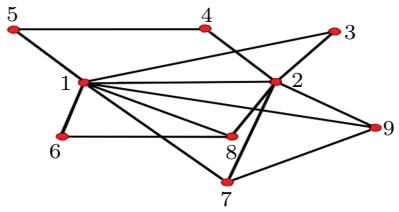


Fig. 2. Toy network.

4. Experimental setup

4.1. Data sets

In order to test the effectiveness of the proposed method, we apply it to real-world networks and artificial networks. Real-world networks include (i) electricity: the network of power grids; (ii) network science: the coauthor network in network science consists of 1589 nodes and we choose the largest dense community of 400 nodes in this network; (iii) email: the email network of the University Rovirai Virgili; and (iv) yeast: the protein–protein interaction network. These data sets are available from: http://vlado.fmf.uni-lj.si/pub/networks/data/. Table 1 shows the statistical characteristics of four networks of different sizes. The two artificial networks are WS small world network and BA scale-free network. First, a small world network is constructed, in which nodes $n = 1000$, links $m = 1475$, and nodes are connected with a probability of $p = 0.01$. The average degree of the network is $\langle k\rangle \approx 3$. The scale-free network takes $N = 1000$, $m_0 = 5$, $m = 4$ (for the initial network with m_0 nodes, each time a new node is introduced to connect with m existing nodes, $m \leq m_0$), and generates a scale-free network with $N = 1000$ nodes and $mt = 3980$ edges (t steps are required to generate the required scale-free network). In this paper, all these real-world networks and analog networks are regarded as undirected and unweighted.

Table 1. Some statistical properties of four networks. Topological features include: number of nodes ($|V|$), number of edges ($|E|$), average degree ($\langle k\rangle$), maximum degree ($k_{\max}$), average clustering coefficient ($\langle k\rangle$), and propagation threshold ($\beta_{\rm th}=\langle k\rangle/\langle k\rangle$).

Network	$\lvert V\rvert$	$\lvert E\rvert$	$\langle k\rangle$	$k_{\max}$	$\langle k\rangle$	$\beta_{\rm th}$
Power grid	4941	6594	2.66	19	0.08	0.255
Network science	379	914	4.82	34	0.74	0.124
Email	1133	5451	9.62	71	0.22	0.053
Yeast	2361	7182	5.62	64	0.13	0.060
WS	1000	1475	3.78	25	0.84	0.154
BA	1000	3980	1.45	42	0.36	0.082

4.2. Evaluation methodology

4.2.1. SI model

To compare the accuracy of the proposed metric in identifying the propagation ability of edges, the SI model is used to simulate the propagation and evolution in real networks. The SI model contains only two types of nodes, susceptible nodes (S) and infected nodes (I). Once a node is infected, it is never able to recover. As the time increases, the number of infected nodes in the network increases until all nodes are infected and the process stops. In our experiments, we keep the propagation rate exactly above the propagation threshold of the epidemic in the network $\beta_{\rm th}$. The experiment starts with any one node in the network as an infected node and the number of infected nodes in the network after 100 time steps is $F(t_1)$, which can be used to reflect the propagation ability of the nodes in the network. The formula is as follows:

$$F(t_1)=\frac{1}{M}\sum_{m=1}^{M}f(m), \tag{5}$$

where $M=200$ is the total number of experimental repetitions. $f(m)=n/N$ is the proportion of infected nodes in the network after 100 time steps for each experiment, n is the number of infected nodes in the network after 100 time steps for each experiment, and N is the total number of nodes in the network. We rank the edges in the network from largest to smallest according to the four edge importance measures, EI, $EC_{\rm DCN}$, SO, and SN, and select the top 20 percent of edges for each. The edges are removed from the network sequentially starting from 100 time steps, and 2 percent of the top 20 percent of the edges are removed at each time step. After 50 time steps (when the top 20 percent of edges are removed), let the proportion of infected nodes in the network be $F'(t_2)$. The same experiment is conducted 200 times to take the average value. We then define ρ as

$$\rho=\frac{F'(t_2)-F(t_1)}{F'(t_2)}, \tag{6}$$

where $0\le\rho\le1$. From the perspective of propagation dynamics, the greater the influence of an edge, the greater the ability of that edge to be removed from the network to inhibit proliferation, and the fewer the number of nodes in the network that end up in the infected state, i.e., the smaller ρ.

4.2.2. Robustness R

The performance of EI is evaluated by an edge percolation process.[30,31] Edges are sequentially removed from the network according to the ranking results of the four edge importance metrics EI, $EC_{\rm DCN}$, SO, and SN and the greater the change in the topology of the network after removing the same proportion of edges, the more important the removed edges are. In this paper, the impact on network connectivity after edge removal is evaluated by the robustness metric R.[32] The formula is defined as

$$R=\frac{1}{M}\sum_{m=1}^{M}\delta(m_{\rm r}/m), \tag{7}$$

where $m_{\rm r}$ denotes the number of edges removed and m denotes the total number of edges in the network. $\delta(m_{\rm r}/m)$ denotes the proportion of nodes in the largest connected branch in the network after removing $m_{\rm r}/m$ proportion of edges. Obviously, a smaller R indicates that the metric can better destroy the connectivity of the network, which means it can better rank the importance of edges.

4.2.3. Number of connected components θ

The number of connected components is an indicator of the degree of "fragmentation" of the network, i.e., the extent to which the network structure is disrupted. The original network has a relatively complete structure where most nodes are located on a giant connected branch, and the number of connected branches in the network is small. As we perform edge removal, the number of connected branches in the network increases gradually as the giant connected branches are cut off and broken into many smaller connected branches. We also rank the edges in the network from largest to smallest according to the four edge importance measures EI, $EC_{\rm DCN}$, SO, and SN and each select the top 50 ($p=50\%$) percent of edges for removal. We observe the number of connected branches in the final network.

4.2.4. Kendall's tau correlation coefficient τ

Kendall's tau correlation coefficient is selected to determine the consistency between the ranking list obtained by the specific measurement method and the $F(t_{\rm c})$ ranking list obtained by the SI model. Given two ranking lists X and Y, (x_i,y_i) and (x_j,y_j) are the two edge pairs in these lists. The Kendall tau correlation coefficient is defined as

$$\tau(X,Y)=\frac{n_1}{\frac{1}{2}N(N-1)}-\frac{n_2}{\frac{1}{2}N(N-1)}. \tag{8}$$

In this paper, X represents the ranking result of edges obtained from a centrality measure, and Y represents the ranking result of the number of infected nodes obtained from the SI model. If $x_i>x_j$, $y_i>y_j$ or $x_i<x_j$, $y_i<y_j$, we say that the ordering of x_i in X and the ordering of y_i in Y are the same. Otherwise, it

is called inconsistent. n_1 and n_2 represent the number of consistent and inconsistent pairs, respectively, and N is the length of the sorting list X. Obviously, τ is between -1 and 1. Here, $\tau = -1$ means that the order of these two lists is completely opposite. At that time, $\tau = 1$, the order of the two lists is the same.

4.2.5. Monotonicity $M(R)$

In order to distinguish the propagation ability of all edges, each edge should be assigned a unique indicator through centrality measurement. The proportion of repeating elements in a sequence is called the monotonicity of the sequence. In order to quantify the monotonicity of different sorting methods, we use $M(R)$[8] and define it as

$$M(R) = \left[1 - \frac{\sum_{r\in R} N_r(N_r - 1)}{N(N-1)}\right], \tag{9}$$

where N is the length of the ranking list R, N_r indicating the number of edges with the same sorting value R. The range of M values is 0 to 1. The best value of M is 1, which means that each edge in the network has a unique and identifiable sorting value. In contrast, the worst value of M is 0, which means that all edges in the network have the same ranking.

4.3. Experimental results and analysis

In the experiments, the proposed edge importance metric is compared with three other metrics, including EC_{DCN}, *SO*, and *SN*. We use different methods to analyze the performance of the proposed centrality metric in terms of propagation ability, robustness, etc.

4.3.1. Spread capacity analysis

The simulation results of the four metrics on six networks are shown in Fig. 3. It can be seen that in (a)–(e), the *EI* metric always makes p the smallest, i.e., the smallest number of nodes eventually infected. In (f), EC_{DCN} outperforms *EI* by about 110 time steps; this is because nodes with larger betweenness or degree in BA networks are generally the "hubs" in the network, and EC_{DCN} has the advantage of identifying the edges connecting nodes with large degree or larger betweenness nodes in the network. For BA networks, once the "hub" node is infected, the spread of the disease is accelerated. The ability of EC_{DCN} to identify the links connected to the "hub" nodes and block them in time slows down the spread of the disease to some extent. However, as time increases, it can be seen that *EI* outperforms EC_{DCN} and gradually closes the gap. This is because as the degree of the "hub" node decreases the network becomes more homogeneous, and the effect of *EI* gradually emerges. At the same time, it can be seen that *EI* is more effective in the face of high-density homogeneous networks. In addition, we can also see that the suppression effect of *SO* and *SN* on propagation is approximately equal. This is because both are metrics used to identify links between different communities in the network. From (b) and (e), it can be seen that the suppression effect of *SO* and *SN* is second only to *EI*. This is because the modularity of network science and WS networks is higher as there are more communities in the network; *SO* and *SN* can identify the links between communities and cutting these links will make the virus only spread within the community. In Fig. 4, the same result can be seen after 150 time steps (d) i.e. the *EI* metric makes the network infected with the least number of nodes. In summary, we can conclude that the *EI* metric can be effective in identifying important edges in the propagation process.

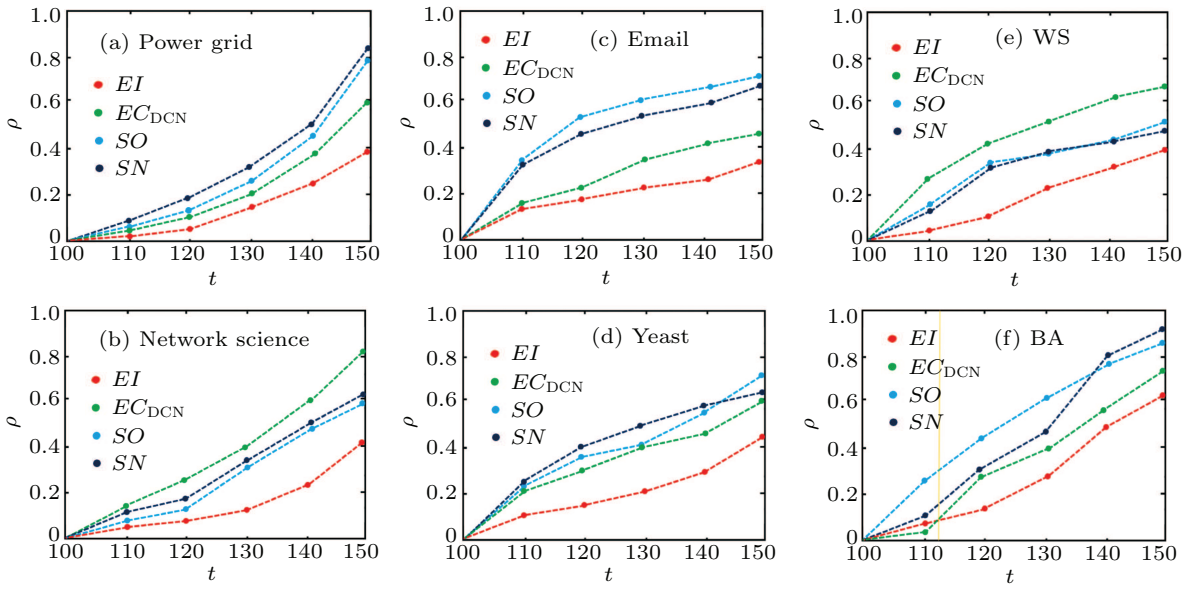


Fig. 3. The ρ as a function of t. The changes in infected nodes in the six networks with the continuous removal of edges are obtained by averaging the results over 200 independent simulations of the *SI* model for each network, with β slightly larger than β_{th}.

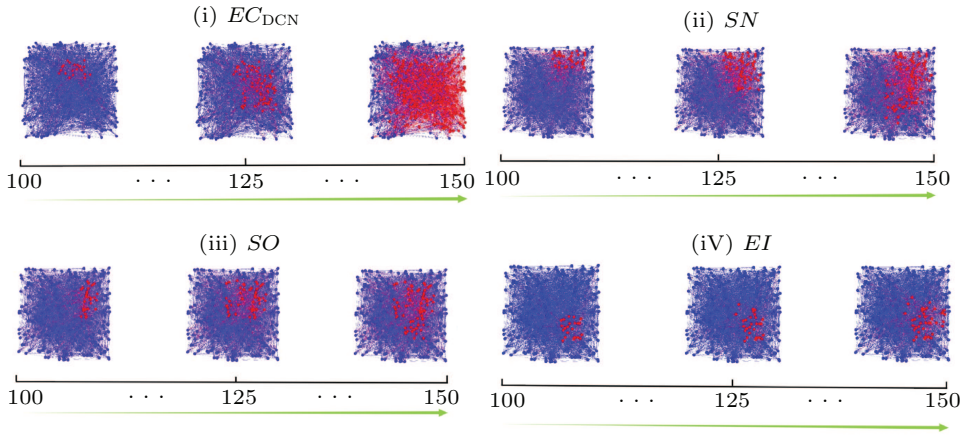


Fig. 4. The visualization result of SI model on network science. The blue nodes are susceptible states and the red nodes are infected states. Each of them corresponds to an edge importance metric.

4.3.2. Robustness

The simulation results of the four metrics on the six networks are shown in Fig. 5. It can be seen that in (a)–(f) *EI* is always the fastest decreasing as the proportion of edge removal increases. In (b), *SO* outperforms *EI* until δ drops to 0.55, and *SN* outperforms *EI* until δ drops to 0.64. This is because network science contains more communities, so *SO* and *SN* can identify the links connecting different communities and disconnect them in the early stage, which can cause great damage to the connectivity of the network. However, with δ at [0, 0.55] and [0, 0.64], it can be seen that *EI* is always better than *SO* and *SN*. *SO* and *SN* cannot accurately identify edges in these sub-networks, so the damage to the connectivity of the network is gradually diminished. Similarly, we can see the same phenomenon in (e). For δ in [0, 0.62] and [0, 0.67], *SN* and *SO* outperform *EI*, but in [0.62, 1.0] and [0.67, 1.0] *EI* is always better than *SO* and *SN*. In (f), EC_{DCN} outperforms *EI* at [0, 0.7] because EC_{DCN} is able to identify the links attached to the "hub" nodes and remove them in time. As the number of links connected to the "hub" nodes decreases, the network is gradually "fragmented" into many small clusters, which can cause some damage to the connectivity of the network in a short period of time. Since the degree of heterogeneity of nodes in these small clusters is low and there are no nodes with large degrees or meshes, EC_{DCN} cannot accurately identify the edges connected to nodes with large degrees or betweenness. Thus we see that *EI* is always better than EC_{DCN} when δ is between [0.7,1.0]. To more accurately illustrate the computational merits of the proposed metric, we remove the top ten percent of edges from the network at each time step Δt according to the ranking results of the four edge importance metrics and observe the time taken to finally minimize the maximum connected branches of the network.

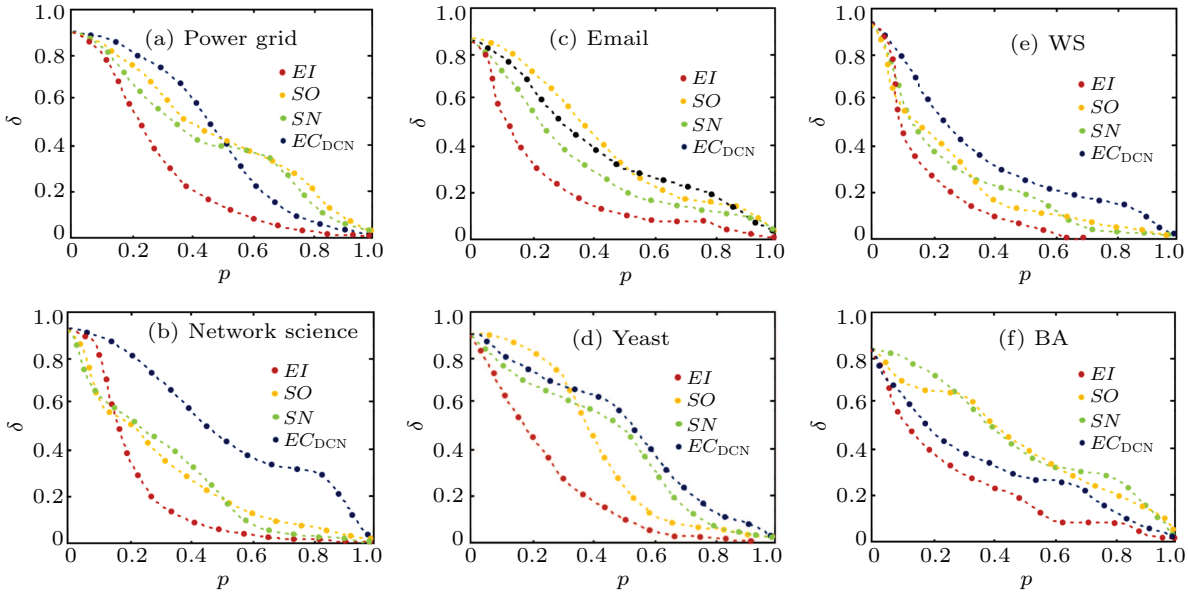


Fig. 5. Ratio of nodes on the largest connected branch as δ function of the ratio of edges removed p. Ratio of nodes of the largest connected component after removing p ratio of edges in six networks. Results obtained by averaging 200 independent simulations for each network.

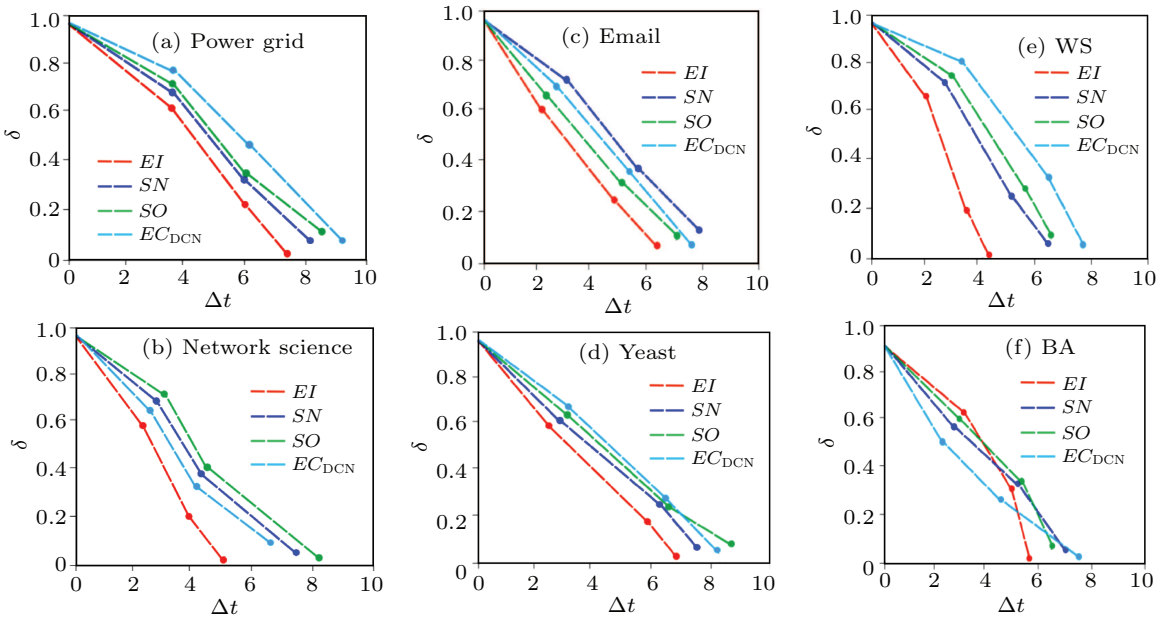


Fig. 6. Ratio of nodes on the maximum connected branch δ as a function of time step Δt on six networks. Each network is the result obtained by averaging 100 independent simulations on MATLAB.

As can be seen in Fig. 6, the edge metric *EI* always takes the least time to maximize the decomposition of the connected branches in these six networks, with the best performance of *EI* in network science and WS taking the least time. This is mainly due to the larger link density and average clustering coefficient of the two networks, which also confirms our previous statement that *EI* performs better in the face of high-density homogeneous networks. In addition, the computational efficiency of *SO*, *SN*, and EC_{DCN} is different in each network and the computational superiority in each network is also different, which is mainly due to the different topological characteristics of different networks, so we do not elaborate too much here. All the results show that decomposing the network according to the ranking results of the edge importance measure *EI* can destroy the robustness of the network the fastest.

4.3.3. Number of connected branches

The simulation results of the four metrics on the six networks are shown in Fig. 7. It can be seen in (a), (c), (d) that as the proportion of edge removal keeps increasing, *EI* always makes the number of connected branches in the network increase fastest and eventually the most. In (b), *SN* and *SO* are superior to *EI* for θ at [0, 0.07], [0, 0.13]. This is because network science contains more associations, and *SN* and *SO* preferentially identify links connecting different communities to each other. The removal of these links leads to a rapid increase in the number of connected branches (small associations) in the network in a short period of time. After $p = 0.07$, 0.13 the removal of edges is concentrated in each community and the disruptive effect of *SN* and *SO* on network connectivity is reduced, so *EI* is always superior to *SN* and *SO*. Similarly, in (f), EC_{DCN} outperforms *EI* when p is within [0, 0.15], thanks to its ability to identify links connected to "hub" nodes and remove them in a timely manner. The gradual "fragmentation" of the network into smaller clusters leads to an increase in the number of connected branches. After $p = 0.15$, the removal of edges is concentrated in each small dense homogeneous cluster, and it can be seen that *EI* gradually outperforms EC_{DCN}. The computational efficiency of the proposed metric in terms of decomposing the network is similarly observed by removing the top ten percent of the ranked edges sequentially from the network at each time step Δt according to the ranking results of the four edge importance metrics, and separately observing the time used to reach a certain number (in the vicinity of this number) of connected branches in each network. Here we set the fixed number of connected branches to be reached in each network as 2000, 200, 500, 1000, 600 and 600, respectively. The results are shown in Fig. 8, which shows that the use of the metric *EI* to remove edges results in the shortest time to reach a fixed value for the number of connected branches in the six networks, with the best performance in network science and WS. This confirms the conclusion of Subsection 4.3.2 and reflects that the proposed metric *EI* can quickly and accurately identify the important edges in the "skeleton" of the network, which are the basis for the derivation and development of the network. In addition, the three metrics *SO*, *SN* and EC_{DCN} have their own advantages and disadvantages in each network. In the BA network, it can be seen that the decomposition efficiency of EC_{DCN} before $\Delta t = 4$ is still very high, which is mainly due to the special topological characteristics of the BA network as mentioned before. In summary, the edge importance metric *EI* can effectively and efficiently disrupt the structure of the network.

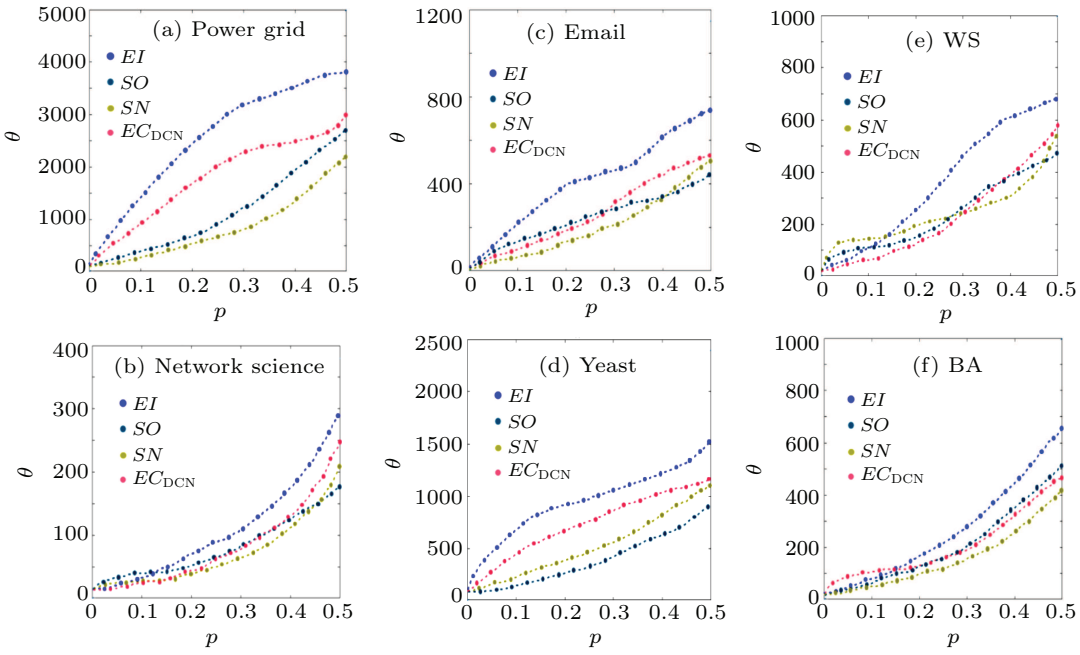


Fig. 7. Number of connected branches θ as a function of the edge removal ratio p. Each network is the result of averaging 200 independent simulations.

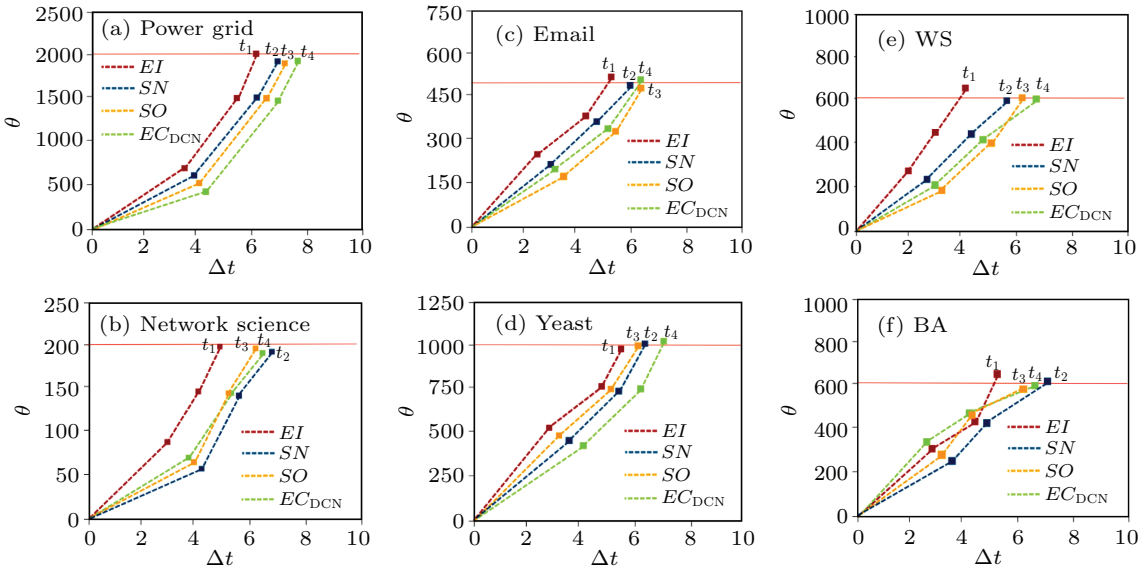


Fig. 8. Number of connected branches θ as a function of time step Δt on six networks. Each network is the result obtained by averaging 100 independent simulations on MATLAB.

4.3.4. Analysis of BA scale-free networks

The simulation results in Subsections 4.3.1, 4.3.2 and 4.3.3 show that the performance of the proposed metric *EI* for BA networks is less satisfactory in the early stage, while the performance of EC_{DCN} in the early stage is superior compared to the other three metrics. We have emphasized earlier that this is due to the special topological properties of BA networks, and here we analyze these two points in more detail. A random scale-free network with 600 nodes and 2760 edges is constructed using the Python networkx package as an example, where the network has three communities and two hub nodes. As shown in Fig. 9, *A* represents the largest community and *C* represents the smallest community. In the propagation experiment in Subsection 4.3.1, at the initial moment we randomly select a node from the network as the source of infection and start the edge removal operation after 100 time steps. In BA networks, nodes with a large betweenness or degree are generally hubs in the network, through which the virus can propagate to various communities and eventually cause widespread infection, while EC_{DCN} has the advantage of being able to identify the edges connecting nodes with a large degree or large betweenness in the network. As these edges are removed so that the virus cannot spread to other communities in the network but is restricted to the local area (in its own community), we can see that the inhibitory effect on propagation in the early stage is better than *EI*. With the removal

of edges, the link between the hub node and each community gradually decreases and the network is broken down into several high-density homogeneous small communities; at this time the superiority of *EI* becomes apparent, so the inhibitory effect on propagation is better than EC_{DCN} in the later stage. Also in Subsections 4.3.2 and 4.3.3, the initial network as a whole has only one maximum connected branch, and as we gradually remove the edges identified by EC_{DCN} this causes a sharp decrease in the number of nodes on the maximum connected branch and a rapid increase in the number of connected branches in a relatively short period of time, as the network is disintegrated into many smaller communities and a number of isolated nodes. However, the metric *EI* comes into play as the focus of work gradually shifts to each community. Although the proposed metric *EI* works better in the face of homogeneous networks, we can also adapt it to heterogeneous networks better by making appropriate modifications to the formula. The modified formula is defined as

$$EII(e_{ij}) = \frac{N'_{\Delta}+N'_{\alpha}+N'_{\eta}+N'_{\beta}+N'_{\blacksquare}}{N_{\Delta}+N_{\alpha}+N_{\eta}+N_{\beta}+N_{\blacksquare}} \times \left|L-L'\right| \times \sum_{s\neq t} \frac{m_{st}(e_{ij})}{m_{st}} \left|B_i - B_j\right|, \quad (10)$$

where m_{st} is the total number of shortest paths between nodes s and t, and $m_{st}(e_{ij})$ represents the number of these paths through edge e_{ij}. That is, $\sum_{s\neq t} m_{st}(e_{ij})/m_{st}$ represents the global betweenness of edge e_{ij}. B_i and B_j denote the betweenness of the nodes at each end of edge e_{ij}, respectively. In this formulation we consider some special properties of importance edges in the face of heterogeneous networks. First, when the betweenness of edge e_{ij} is larger it indicates that it may be a bridge edge connecting communities to each other. Secondly, when the absolute value of the betweenness of the nodes at both ends of edge e_{ij} is larger it indicates that one of the nodes may be a bridge node between communities. In fact, the introduction of these two points does not increase the complexity of the original computation because the edge and node betweenness can be obtained from the network statistics software Gephi. In the following, we compare the modified metric EII with *EI* and EC_{DCN} in three aspects (propagation capability, robustness, and number of connected branches) on the BA network of Table 1. The results are shown in Fig. 10, which shows that EII always outperforms EC_{DCN} in (a)–(c) and performs more consistently than *EI* throughout. Therefore, in practical applications, the appropriate metric should be selected according to the topological characteristics of the network under testing.

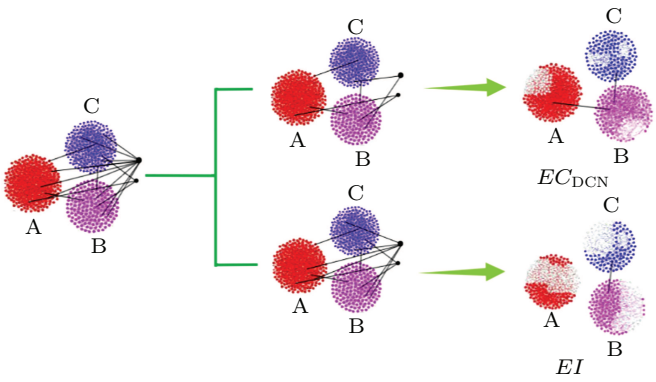


Fig. 9. The *EI*, EC_{DCN} visualization of the evolutionary process of two metrics applied on a random scale-free network, where three moment images in the time configuration window are intercepted. The size of the white area represents how the community has been broken down internally.

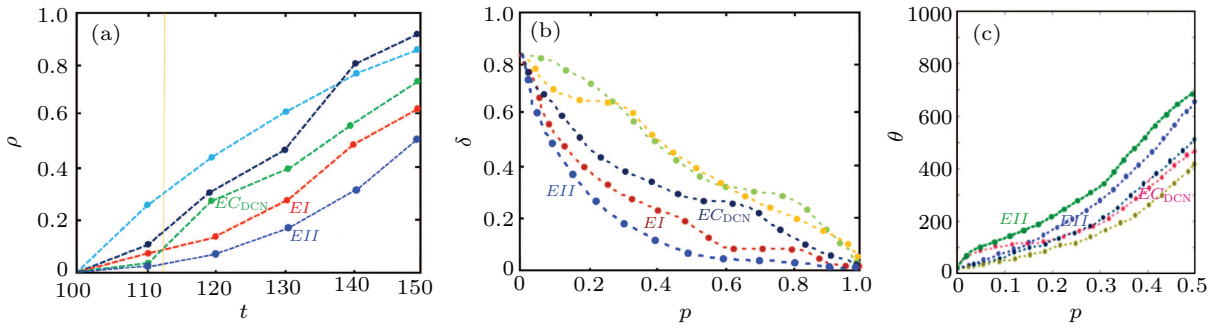


Fig. 10. Comparison of the three metrics *EII*, *EI* and EC_{DCN}, where each BA network is the result of averaging 200 independent simulations.

4.3.5. Kendall's tau correlation coefficient

In order to test the accuracy of the proposed centrality measure, it is compared with three different centrality measures on six networks under the SI model. Theoretically, the closer the value of τ is to 1, the better the performance of the sorting method. The selection value β_{th} changes near the epidemic threshold. The value β_{th} is related to the topological characteristics of a given network, so the abscissa of different networks in Fig. 11 is different. Figure 11 describes the Kendall's tau correlation coefficient of four central indicators varying with the probability of infection. As shown in (a)–(f), the value of τ for *EI* is always the largest. In (b) and (e), *SO* and *SN* have the second highest τ-value after *EI*. this is because there are more associations in network science and WS networks, and *SO* and *SN* can identify the links between communities. However, after the links between communities are

identified, *SO* and *SN* then start to identify the links in each association leading to a decrease in accuracy. Thus, it can be seen that the consistency between the ranked lists of both metrics is from strong to weak. In (f), the τ-value of EC_{DCN} is second only to *EI* and at [0.11, 0.16] they are approximately equal. This is because EC_{DCN} is able to identify edges in BA networks connecting nodes of large degree or larger betweenness. However, after the links with large degree are identified, the accuracy of identification decreases because the degrees of all nodes in the BA network except the "hub" node are not very different, so we can see that the consistency is gradually weakening.

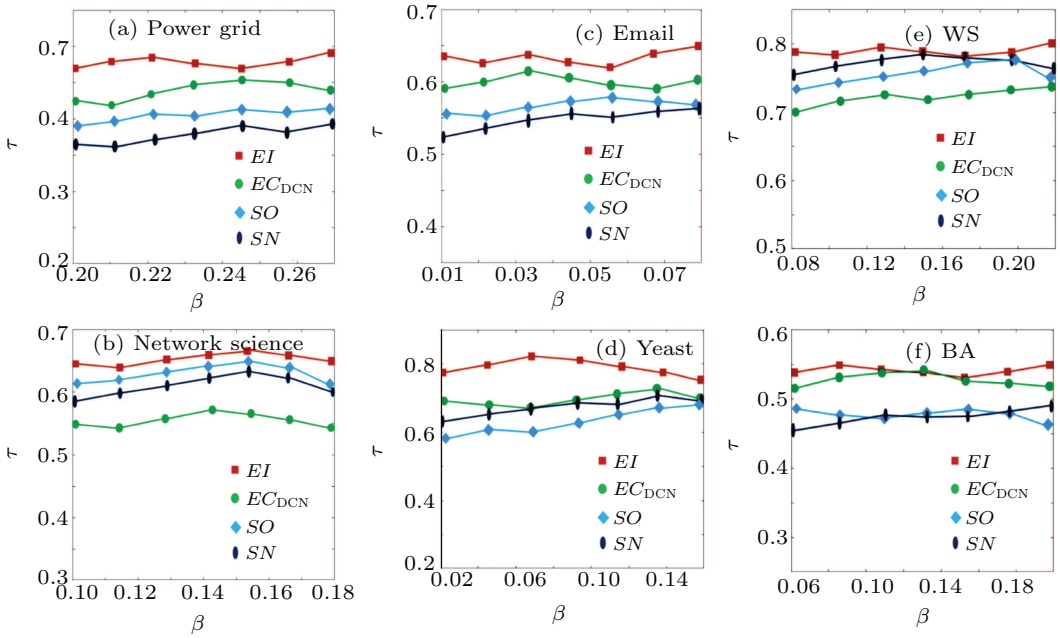


Fig. 11. Kendall's tau correlation coefficient (τ) between the ranking results obtained from the four different metrics and the $F(t)$ ranking list obtained from the SI model with different infection probabilities β. β changes near the neighborhood of β_{th}. Results obtained by averaging 100 independent simulations of the SI model.

4.3.6. Analysis of monotonicity

A good centrality measure should be able to distinguish between edges with different propagation capabilities. Table 2 shows the monotonicity of the four centrality measures in six different networks. The closer *M* is to 1, the better the performance of the centrality measure is. It can be seen that among the four centrality measures, the monotonicity value of *EI* is always close to 1 in all networks. In summary, *EI* can better distinguish edges with different propagation capabilities.

Table 2. Monotonicity of four centrality measures in six networks.

Complex type	*EI*	*SO*	*SN*	EC_{DCN}
Power grid	**0.9543**	0.8276	0.7424	0.6983
Network science	**0.9072**	0.8355	0.6784	0.7631
Email	**0.8549**	0.7244	0.6552	0.6954
Yeast	**0.8756**	0.7563	0.5944	0.7328
WS	**0.8032**	0.7126	0.6852	0.6349
BA	**0.7842**	0.7305	0.7024	0.6842

5. Conclusion

This paper proposes a new edge importance metric for identifying the most influential edges in a network. The metric innovatively considers loops and diagonal quadrilaterals in the second-order neighborhood of edges as well as fully connected quadrilaterals and diagonal quadrilaterals in the first-order neighborhood. Moreover, the higher-order clustering coefficients of edges are defined. We consider that the removal of edges affects the connectivity and completeness of the local structure of the network. The superiority of the proposed metric is obtained by simulating the actual network. This paper innovatively introduces the idea of probability for the first time to identify higher-order structures within edge neighborhoods, which provides a powerful analytical tool for future developments in the field of network science and hopefully will see more integration between the two fields. Furthermore, only higher-order forms within the first and second-order neighborhoods of edges are considered in this paper. In fact, as the scope of the analysis (the neighborhood of the edge) is extended outward, we see that the variety of higher-order structures grows exponentially in richness far beyond our imagination. How will the new structures that emerge in this section affect the structure and function of the network? This work will be explored in the next step. In the meantime, we will also propose some targeted protection strategies for the higher-order organization of the network.

References

[1] Bonacich P 1972 *J. Math. Sociol.* **2** 113
[2] Lü L Y, Zhou T and Zhang Q M 2016 *Nat. Comm.* **7** 10168
[3] Kitsak M, Gallos L K and Havlin S 2010 *Nat. Phys.* **6** 888

[4] Crescenzi P, D'angelo G and Severini L 2016 *ACM. T. Know. Discov. D* **11** 1
[5] Freeman L C 1977 *Sociometry* **40** 35
[6] Ma Y, Cao Z and Qi X 2019 *Physica A* **527** 121130
[7] Li Z, Ren T and Ma X 2019 *Sci. Rep.* **9** 8387
[8] Li Z, Ren T and Xu Y 2020 *IEEE Access* **8** 66068
[9] Holme P, Kim B and Yoon C 2002 *Stat. Non.* **65** 056109
[10] Xia Y and Hill D J 2008 *IEEE Trans. Biomed. Circ. S* **55** 65
[11] Cuadra L, Salcedo-Sanz S and Del Ser J 2015 *Energies* **8** 9211
[12] Goltsev A V, Dorogovtsev S N and Oliveira J G 2012 *Phys. Rev. Lett.* **109** 128702
[13] Ball M O, Golden B L and Vohra R V 1989 *Oper. Res. Lett.* **8** 73
[14] Girvan M, Newman M, Girvan M and Newman M E J 2002 *P. Natl. A. Sci.* **99** 7821
[15] Yu E Y, Chen D B and Zhao J Y 2018 *Sci. Rep.* **8** 14469
[16] Kanwar K, Kumar H and Kaushal S 2019 *Soc. Netw. Anal. Min.* **12** 49
[17] Holme P, Kim B and Yoon C 2002 *Phys. Rev. E* **65** 056109
[18] Onnela J P, Saramäki J and Hyvönen J 2007 *Nat. Aca. Sci.* **104** 7332
[19] Cheng X Q, Ren F X and Shen H W 2010 *J. Stat. Mech-theory. E* **2010** P10011
[20] Liu Y, Tang M and Zhou T 2015 *Sci. Rep.* **5** 131725
[21] Matamalas J T, Arenas A and Gómez S 2018 *Sci. Adv.* **4** eaau4212
[22] Ouyang B, Xia Y and Wang C 2018 *IEEE Trans. Biomed. Circ. S* **65** 1244
[23] Xu Y, Ren T and Sun S 2021 *Math.* **9** 2531
[24] Yu E Y, Chen D B and Zhao J Y 2018 *Sci. Rep.* **8** 14469
[25] Zhao N, Li J and Wang J 2018 *Physica A* **548** 123877
[26] Bröhl T and Lehnertz K 2019 *Chaos* **29** 1098
[27] Kossinets G and Watts D J 2006 *Science* **311** 88
[28] Battiston F 2021 *Nat. Phys.* **17** 1093
[29] Rubinov M and Sporns O 2010 *Neuroimage* **52** 1059
[30] Reijneveld J C, Ponten S C and Berendse H W 2007 *Clin. Neurophysiol.* **118** 2317
[31] Wu T, Zhang X and Liu Z 2022 *Front. Phys.* **17** 31504
[32] Wu J, Tse C K and Lau F C M 2013 *IEEE Trans. Biomed. Circ.S* **60** 3303
[33] Chen Z, Wu J and Xia Y 2018 *IEEE Trans. Circ. I* **65** 115
[34] De la Cruz Cabrera O, Jin J and Noschese S 2022 *Appl. Numer. Math.* **172** 186
[35] Cai S M, Hong L and Zhong Q 2011 *Complex Syst. Complexity Sci.* **10** 1099
[36] Milanović J V and Zhu W 2017 *IEEE T. Smart. Grid* **9** 4637
[37] Yang F, Zhu J and Sun J 2019 *J. Netw. Comput. Appl.* **39** 72
[38] Muldoon, Sarah and Feldt 2018 *Phys. Life. Rev.* **24** 143
[39] Rubinov M and Sporns O 2010 *Neuroimage* **52** 1059
[40] Papo D, Buldú J M and Boccaletti S 2014 *Biol. Sci.* **369** 20130520