



OPEN ACCESS

EDITED BY

Xiaoming Zheng,
Central Michigan University, United States

REVIEWED BY

Yilun Shang,
Northumbria University, United Kingdom
Mohyeddin Sweidan,
Concord University, United States

*CORRESPONDENCE

Xiaodong Cui,
✉ xiaodong.cui@nwpu.edu.cn
Peican Zhu,
✉ ericcan@nwpu.edu.cn

SPECIALTY SECTION

This article was submitted to Social Physics, a section of the journal Frontiers in Physics

RECEIVED 13 February 2023

ACCEPTED 06 March 2023

PUBLISHED 17 March 2023

CITATION

Liang G, Cui X and Zhu P (2023), An effective method for epidemic suppression by edge removing in complex network.
Front. Phys. 11:1164847.
doi: 10.3389/fphy.2023.1164847

COPYRIGHT

© 2023 Liang, Cui and Zhu. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

An effective method for epidemic suppression by edge removing in complex network

Guangbo Liang¹, Xiaodong Cui^{2*} and Peican Zhu^{1*}

¹School of Computer Science, Northwestern Polytechnical University (NWPU), Xi'an, Shaanxi Province, China, ²School of Marine Science and Technology, NWPU, Xi'an, Shaanxi Province, China

Since the birth of human beings, the spreading of epidemics such as COVID-19 affects our lives heavily and the related studies have become hot topics. All the countries are trying to develop effective prevention and control measures. As a discipline that can simulate the transmission process, complex networks have been applied to epidemic suppression, in which the common approaches are designed to remove the important edges and nodes for controlling the spread of infection. However, the naive removal of nodes and edges in the complex network of the epidemic would be practically infeasible or incur huge costs. With the focus on the effect of epidemic suppression, the existing methods ignore the network connectivity, leading to two serious problems. On the one hand, when we remove nodes, the edges connected to the nodes are also removed, which makes the node isolated and the connectivity is quickly reduced. On the other hand, although removing edges is less detrimental to network connectivity than removing nodes, existing methods still cause great damage to the network performance in reality. Here, we propose a method to measure edge importance that can protect network connectivity while suppressing epidemic. In the real-world, our method can not only lower the government's spending on epidemic suppression but also persist the economic growth and protect the livelihood of the people to some extent. The proposed method promises to be an effective tool to maintain the functionality of networks while controlling the spread of diseases, for example, diseases spread through contact networks.

KEYWORDS

complex networks, epidemic suppression, network connectivity, edge importance, network science

1 Introduction

For the convenience of investigating practical systems, they are usually modeled as complex networks, such as biology, ecology, physics and medicine [1–8]. The abbreviated complex network usually consists of nodes and edges which represent individuals and connections among individuals, respectively. With these complex network models, various phenomena can be revealed through studying the properties of nodes, edges and communities. Among existing studies, the determination of edge/node importance is of great significance which can provide valuable guidance in addressing practical problems. For instance, in the aviation network, the nodes correspond to the important airports and the related studies are critical.

Since the birth of network science, scholars devote to studying the heterogeneity of nodes. Various approaches are presented to measure the importance of nodes, such as degree centrality [9,10], betweenness centrality [11,12], eigenvector centrality [13], k -shell [14], etc.

According to node degree centrality [9], we can obtain the number of routes for each airport in an aviation network; thus, airports with more routes are of higher importance. Furthermore, scholars also considered the problem of edge importance and different approaches are presented, including WERA-KPath (WKP) [15], edge betweenness centrality (EBC) [16,17], Link Importance (LI) [18], Nearest-neighbor edge centrality [19], etc. The above methods can be effectively applied to address practical problems, such as community detection [20–22], virus/information diffusion [23–27].

In recent years, many infectious diseases spread globally, countries around the world are actively looking for effective countermeasures to suppress epidemic. Studying heterogeneity of nodes and edges may play an important role in reducing virus impact. For example, nodes and edges to be removed in aviation networks corresponds to airports to be isolated and routes to be canceled, respectively. With these measures, some undesired effects might be limited. However, removing nodes has a more serious impact on the connectivity of the network than removing edges; furthermore, greater cost and impact might also be incurred in practice. To achieve the purpose of effectively suppressing epidemic and protecting network connectivity simultaneously, we propose a method E_{BD} to measure edge importance through combining edge degree centrality and edge betweenness centrality. After that, extensive experiments are conducted to verify its effectiveness.

The reminder of our work is organized as follows: In Section 2, some existing edge importance methods are reviewed. Then, in Section 3, our proposed method is explained explicitly while the difference between our method and existing ones are presented. Later, extensive experiments are conducted with the obtained results being provided in Section 4. The performance of our method is compared to that of existing ones with sufficient discussions. Eventually, Section 5 concludes this work with some future research directions being provided.

2 Related work

Firstly, we suppose a complex network is formulated as $G(V, E)$, where V and E stand for the node set and edge set, respectively, $n = |V|$ denotes the number of nodes, $m = |E|$ represents the number of edges. For brevity, we provide some edge importance derivation methods as follows.

2.1 EBC

In [28], Freeman presents the node betweenness centrality, and then Girvan and Newman extend the definition of betweenness centrality from nodes to edges [16]. Later, Newman implements it in [17]. The concept of edge betweenness centrality indicates that the edges which are more important than the others are likely to be central. We use C_B to denote quantitative value of edge importance which is defined as:

$$C_B(e) = \frac{2}{n \times (n-1)} \sum_{i \neq j} \frac{s_{ij}(e)}{S_{ij}}, \quad (1)$$

where $e \in E$, S_{ij} represents the number of shortest paths from node i to node j and $s_{ij}(e)$ indicates the number of the shortest paths from

node i to node j which pass through edge e . If there exists no path linking i and j , we set $\frac{s_{ij}(e)}{S_{ij}} = 0$ for convenience. According to Eq. 1, a larger C_B usually indicates better control ability over the network. For instance, in the aviation network, if the key routes obtained by C_B are canceled, the aviation network is likely to be decomposed into a number of small ones. This is important for epidemic suppression which will be discussed explicitly in Section 3.

2.2 WKP

De Meo *et al.* apply Random-Walk to measure edge importance [15]. A total of $m - 1$ rounds of loops are carried out with performing 10 Random-Walk in each loop from a random node (m indicates the number of edges in the network). Then, the number of times for each edge being traversed will be recorded. After that, the edge importance sequence is sorted in descending order according to the recorded traversal times (i.e., C_W).

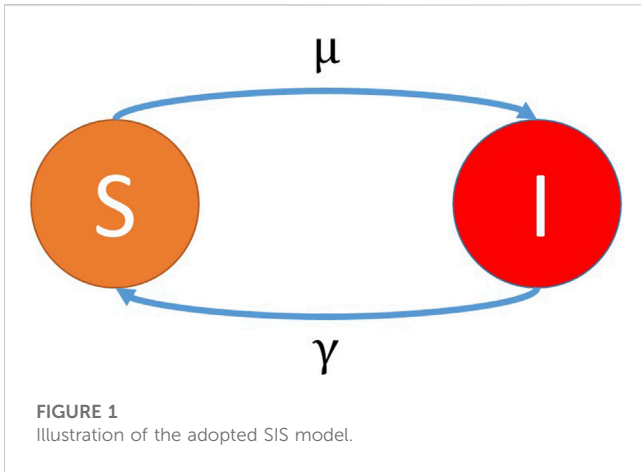
2.3 LI

Aiming to suppress epidemic effectively under SIS model, Joan T. Matamalas *et al.* propose a method for assessing the importance of edges during the spreading process, called Link Importance (LI) [18]. Specifically, they first define a model named ELE which is similar as the SIS model. Through the ELE model, they conduct the disease spreading process until the system converges to a steady state. Then, if nodes i and j are directly connected, they can obtain the probability $P(\bar{\omega}_i = S, \bar{\omega}_j = I)$ (i.e., the probability of node i is susceptible and node j is infectious) which is represented by Φ_{ij} , and the probability $P(\bar{\omega}_i = \bar{\omega}_j = I)$ (given by Θ_{ij}^I). Here, $\bar{\omega}_i$ represents the state of node i , S and I represent the susceptible and the infectious states, respectively. The edge importance is calculated by:

$$I_{ij} = \mu^2 \left(\Phi_{ji} \sum_{r=1}^n A_{jr} \frac{\Phi_{rj}}{\Phi_{rj} + \Theta_{rj}^I} + \Phi_{ji} \sum_{r=1}^n A_{ir} \frac{\Phi_{ri}}{\Phi_{ri} + \Theta_{ri}^I} \right), \quad (2)$$

where I_{ij} stands for the importance of edge e_{ij} and A represents the adjacency matrix of network G . According to Eq. 2, after a node being infected by its neighbors, we can effectively evaluate the extent of the impact being applied on neighboring nodes by the newly infected node. For instance, we assume that two nodes being connected by an edge e_{ij} are in different states, i.e., susceptible and infectious, respectively. When the susceptible node is infected by an infectious one, if the newly affected node has a larger number of neighbors, the authors suppose it will infect more nodes.

According to the analysis of existing studies, EBC tends to achieve better effect of spreading control compared with the others; however, greater damage will be incurred to network connectivity. WKP has high randomness; hence, it is not effective in spreading control and network connectivity protection. LI can protect network connectivity well, but the effect of spreading control is worse than that of EBC. In summary, existing methods cannot achieve excellent performance in both epidemic suppression and network connectivity protection. However, both aspects are of great significance. Taking the aviation network as an example, on the one hand, it is necessary to suppress epidemic to reduce the spreading



range; thus, it is very important to reduce the necessary overhead as much as possible. On the other hand, to meet the requirements of world-wide transportation, it is also of great significance to ensure the accessibility between countries. Therefore, we propose a method (being denoted as E_{BD}) which is effective in spreading control and protecting network connectivity simultaneously.

3 Methods

In this work, we propose a new method to measure the importance of edges through combining degree centrality and EBC. With our proposed approach, we can control the spreading phenomena and protect the network connectivity effectively after removing some of the most important edges.

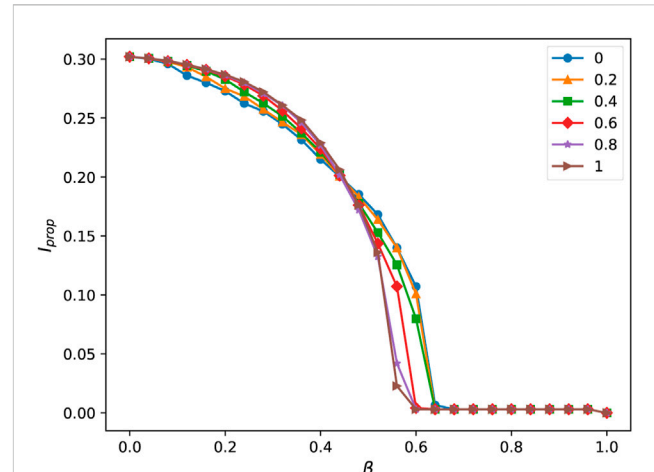
In this paper, we adopt the widely utilized SIS model to simulate the spreading process. Accordingly, each node can be either in a state of S (susceptible) or I (infectious). The node in the I state can infect the neighboring nodes in the S state with probability μ , while the node in the I state can recover with probability γ , the overall state transition process can be depicted as in Figure 1. According to [29], we can derive the epidemic threshold λ_c (i.e., if $\mu/\gamma > \lambda_c$, the virus will spread forever; otherwise, the virus will disappear quickly) in degree uncorrelated networks which is provided as:

$$\lambda_c = \frac{\langle k \rangle}{\langle k^2 \rangle}, \tag{3}$$

where $\langle k \rangle$ stands for the average degree of all nodes and $\langle k^2 \rangle$ represents the mean of squared degree of all nodes.

According to the definition of λ_c , it is obvious that we can control the virus spreading process through increasing the epidemic threshold. After a theoretical analysis, we find that when we remove an edge ($e_{ij} \in E, i, j \in V$), $\langle k \rangle$ is reduced by $2/n$; thus, we can only increase λ_c by minimizing $\langle k^2 \rangle$ as much as possible. Before removing e_{ij} , we assume that the degrees of nodes i and j are denoted as k_i and k_j , respectively. Then, after removing e_{ij} , $\langle k^2 \rangle$ is reduced by $[k_i^2 - (k_i - 1)^2 + k_j^2 - (k_j - 1)^2]/n$. Hence, λ_c becomes:

$$\lambda_c' = \frac{\langle k \rangle - 2/n}{\langle k^2 \rangle - 2(k_i + k_j - 1)/n}. \tag{4}$$



The effect of different values of η on epidemic suppression. We show the fraction of infectious nodes in steady state (I_{prop}), as a function of the fraction of removed edges (β) in usAir97 network ($\mu = 0.06, \gamma = 0.4$), where the numbers in the legend represent the values of η . Note that, the value of each point is the result of averaging 100 simulated values of I_{prop} .

As in Eq. 4, we find that a larger $(k_i + k_j)$ usually indicates a larger λ_c' . Then, we should find the edge e_{ij} with the largest $k_i + k_j$. This means that the edge with more neighbors is more important than the others and the important edge plays a greater role in the epidemic diffusion process. Similar as $k_i + k_j$, $k_i \times k_j$ can also indicate the number of neighbors being connected by e_{ij} . Van Mieghem *et al.* verify that the performance of utilizing $k_i \times k_j$ is much better than the adoption of $k_i + k_j$ during epidemic suppression process [30]. To verify this, we conduct some experiments. Firstly, we combine $k_i + k_j$ with $k_i \times k_j$ by an adjustable parameter $\eta \in [0, 1]$ to obtain a new criteria to measure edge importance which is provided as

$$MA(e_{ij}) = \frac{(k_i \times k_j - v_M) \times \eta}{\sigma_M} + \frac{(k_i + k_j - v_A) \times (1 - \eta)}{\sigma_A}, \tag{5}$$

where v_M and σ_M represent the average value and standard deviation of $k_i \times k_j$ in all edges, respectively; v_A and σ_A stand for the average value and standard deviation of $k_i + k_j$ in all edges, respectively; v and σ are used to normalize $k_i + k_j$ and $k_i \times k_j$. As in Eq. 5, we find that the influence of the two aspects in the MA can be adjusted by changing η and we set $\eta = \{0, 0.2, 0.4, 0.6, 0.8, 1\}$. Finally, we repeatedly remove the edge with the largest MA for different η . Then, we obtain the fraction of infectious nodes I_{prop} in steady state by simulating the SIS model after every $0.04m$ edges are removed in usAir97 network. The experimental results are shown in Figure 2. According to Figure 2, we find that when the fraction of removed edges is less than 0.45, the effect of spreading control becomes better when a smaller η is adopted. However, there are many nodes that are infected in all η . Then, when $0.65 > \beta > 0.45$, large η leads to better effect of spreading control and small spread range of virus. In addition, the effect of spreading control for the scenario of $\eta = 0$ is much worse than that for $\eta = 1$ when $0.65 > \beta > 0.45$. Note that, since the lines in Figure 2 are relatively dense, the error bars are not provided. Moreover, the maximum standard deviation of the data points

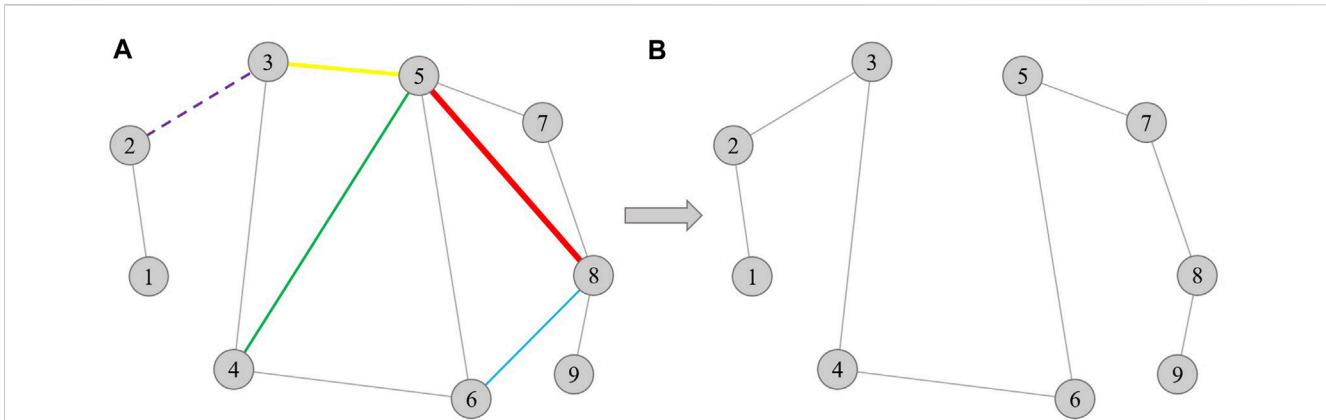


FIGURE 3 An example showing the effect of removing edges on network connectivity. (A) shows the initial network which consists of 9 nodes and 12 edges. The thickness of these colored solid lines corresponds to the C_D of the edge, the derived set according to C_D in descending order is $\{e_{58}, e_{35}, e_{45}, e_{68}\}$. The purple dot line indicates the edge with the largest EBC. (B) shows the network after removing $\{e_{58}, e_{35}, e_{45}, e_{68}\}$.

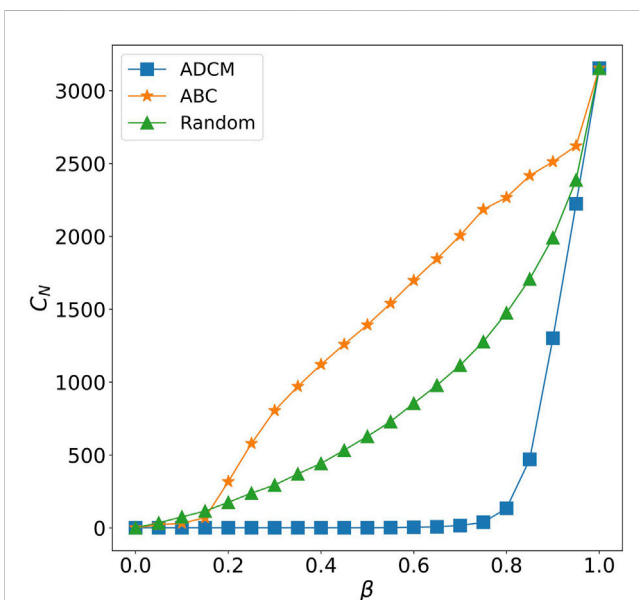


FIGURE 4 The effect of edge removing on network connectivity based on different methods. We show the number of connected components (C_N), as a function of the fraction of removed edges (β) in the Global airline network. We compare three different approaches: ADCM, ABC and Random.

on each line in Figure 2 is in $[0.03, 0.04]$. From above analysis, compared with $k_i + k_j$, the adoption of $k_i \times k_j$ can achieve the purpose of reducing the epidemic on a large scale while removing fewer edges. Therefore, we adopt

$$C_D(e_{st}) = k_s \times k_t \tag{6}$$

to represent the importance of edge, where $C_D(e_{st})$ represents degree centrality of edge e_{st} .

In addition to the spreading ability of edges, we also consider the effect of edge removal on network connectivity when studying the edge sorting algorithm. Generally speaking, $C_D(e_{st})$ is large given a large number of neighbors of s and t . Thus, when we remove the edge

e_{st} with large C_D , their high-order neighbors can likely be connected because nodes s and t are of large degrees. As in Figure 3A, the edge e_{58} has the largest C_D . When we remove it from the network, nodes 5 and 8 are still connected via other paths, such as $\{5, 6, 8\}$. Then, we continuously remove the edge that has the largest C_D in the remaining network. After removing four edges, the network is still connected as in Figure 3B. However, if the degrees of s and t are relatively small, when edge e_{st} is removed, the probability of s and t being in two connected components respectively increases greatly, such as edge e_{89} in Figure 3A. If we continuously remove the edge with the largest EBC, when we remove the most important edge (e_{23}), the network becomes an unconnected one. Aiming to further verify the performance of network connectivity under different edge removing strategies, we conduct experiments on the Global airline network with the obtained results being provided in Figure 4. As in Figure 4, we utilize the number of connected components (C_N) to reflect the network connectivity (i.e., the better methods have fewer connected components when the same number of edges are removed) and compare the effects of adopting different methods (i.e., ADCM, ABC and Random). Here, ADCM and ABC repeatedly remove the edge with the highest C_D and C_B , respectively. Random randomly selects an edge to remove at each time. As in Figure 4, we find that when the proportion of removed edges is less than 0.7, ADCM is of limited effect on the network connectivity; however, ABC and Random have already decomposed the network into many connected components. Among these methods, ABC has the greatest impact on network connectivity. These results can further validate the effectiveness of C_D in protecting network connectivity.

From another perspective, since infectious diseases always originate from one or a few people, the spreading range is usually limited in the connected components where the seed nodes are located in. Therefore, when some edges are removed, the method which dismantles the network into many connected components is effective in epidemic suppression, while C_D is less effective in epidemic suppression (the corresponding experimental results will be provided in Section 4). To solve this problem, it is necessary to find a method that incorporates C_D to increase the effect of epidemic suppression when fewer edges are removed. Furthermore, individual aggregation is a common

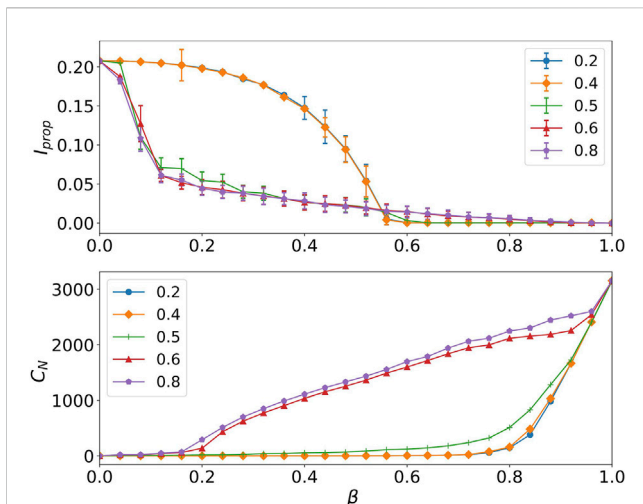


FIGURE 5
The effect of different values of α on network connectivity and epidemic suppression. We show the spread range (I_{prop}) and the number of connected components (C_N), as a function of the fraction of removed edge (β) in GA network, where the numbers in the legend represent the values of α . The error bars of all points are shown in the top plot.

phenomenon in practical networks. For example, in the aviation network, there are close connections between airports belonging to the same country, but fewer routes between airports in different countries; this results in the fact that modules are usually connected by few edges. According to [31], edges with higher C_B are usually the links connecting different modules. Therefore, removing edges with high C_B first can quickly decompose the network into small connected components. Figure 4 also shows that EBC-based ABC can quickly decompose the network into many connected components.

Based on the above discussions, we propose our method E_{BD} through combining EBC (C_B) with C_D and a new edge importance evaluation index C_{BD} is defined as:

$$C_{BD}(e_{st}) = (1 - \alpha) \frac{C_D(e_{st})}{\max(C_D(e))} + \alpha \frac{C_B(e_{st})}{\max(C_B(e))}, \quad (7)$$

where e is an element from the edge set E and α stands for an adjustable parameter ($\alpha \in [0, 1]$). Since C_D and C_B belong to different orders of magnitude (e.g., $C_D(e_{58}) = 16$ and $C_B(e_{58}) = 0.25$ in Figure 3A), these two indicators should be normalized before fusion. In this paper, we adopt the max normalization (i.e., $\max(C_D(e))$ and $\max(C_B(e))$). Here, $\max(C_D(e))$ represents the largest C_D among all the edges in the current network, $\max(C_B(e))$ is similar as $\max(C_D(e))$. Obviously, the influence of the two aspects in C_{BD} can be changed by adjusting the parameter α .

In this article, we dynamically update C_{BD} to derive the edge importance ranking sequence (i.e., S_{BD}). The updating processes are provided as:

- **Step 1:** Initialization, we select a complex network $G(V, E)$ and set an empty edge ranking sequence $S_{BD} = \{\}$;
- **Step 2:** Calculate C_{BD} of each edge in the current network G according to Eq. 7;

TABLE 1 Seven practical networks used for experimental analysis.

Networks	n	m	$\langle k \rangle$	c	r
UA	332	2126	12.8072	0.6252	-0.2079
CN	379	914	4.8232	0.7412	-0.0817
SH	2426	16630	13.7098	0.5375	0.0474
GA	3154	18592	11.7895	0.4914	-0.0170
FC	4039	88234	43.6910	0.6055	0.0636
CC	21363	91286	8.5461	0.6334	0.1340
EE	32430	54397	3.3547	0.1127	-0.3816

- **Step 3:** Remove the edge with the largest C_{BD} from G (if there exist multiple edges with same C_{BD} , we will randomly select one and then remove it); then, the removed edge is added to the end of sequence S_{BD} ;
- **Step 4:** If there still exist edges in G , we need to go back to **Step 2**;
- **Step 5:** Determine the edge importance ranking sequence S_{BD} .

Through the above steps, an edge ranking sequence sorted in declining order according to the edge importance index C_{BD} will be obtained ($S_{BD} = \{e_1, e_2, \dots, e_m\}$). Removing edges according to this sequence can protect the network connectivity efficiently while suppressing epidemic simultaneously.

4 Results and analysis

4.1 Baseline models

To evaluate the performance of our proposed method, some baselines are adopted, including ADCM, Random, WKP, ABC and LI. For ADCM, we need to repeatedly remove the edge e_{ij} with the largest $k_i \times k_j$ from the remaining network (here, k_i indicates the degree of node i). As to Random, we randomly remove edges. For WKP, we need to repeatedly remove the edge with the largest C_W from the remaining network. For ABC [32], we repeatedly remove the edge with the largest C_B from the remaining network. As to LI, the edge with the largest I_{ij} is repeatedly removed from the remaining network. Similarly, in our approach, we continuously remove the edge with the largest E_{BD} from the remaining network.

4.2 Data description

In the following experiments, we mainly consider seven practical networks including usAir97 (UA) [33], Global airline (GA), Facebook combined (FC) [34], Ca netscience (CN) [33], Soc hamsterster (SH) [33], ca CondMat (CC) [33] and email EU (EE) [33]. The UA network is a small aviation network which is constructed based on the US aviation network in 1997 (nodes and edges stand for airports and routes, respectively). The GA network is obtained from OpenFlights (<https://openflights.org>), in which airports are denoted by nodes and airlines between airports are captured by the corresponding edges. As to

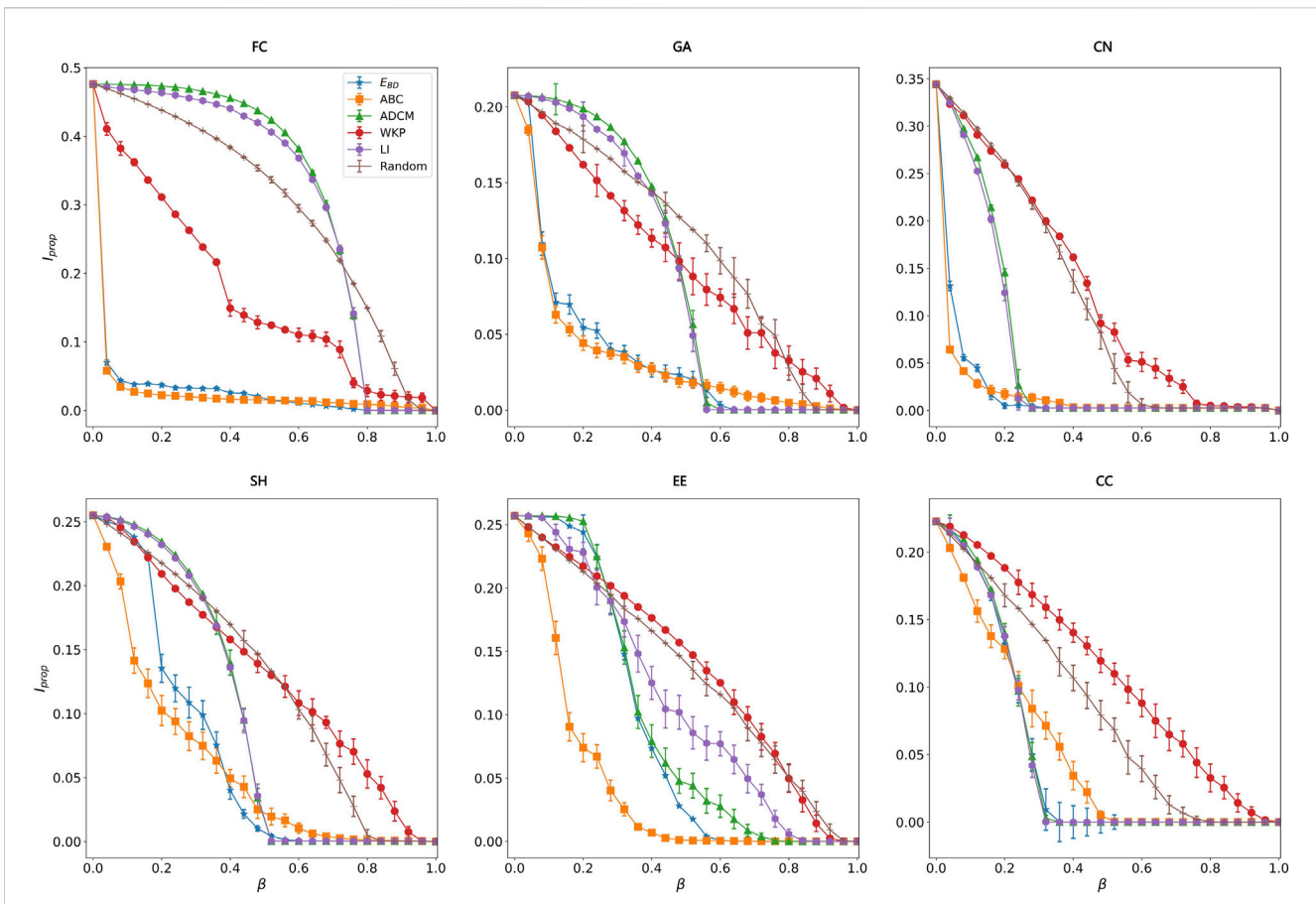


FIGURE 6 Performance of our method E_{BD} ($\alpha = 0.5$) and the others in epidemic suppression with experiments being conducted on six practical networks (i.e., FC, GA, CN, SH, EE and CC). The vertical axis represents the proportion of infectious nodes (i.e., I_{prop}) in the stationary state and the horizontal axis represents the proportion of edges being removed (i.e., β), the error bars are displayed at each point. Note that, $\mu = 0.06$ and $\gamma = 0.5$ in GA, FC and SH, $\mu = 0.2$ and $\gamma = 0.5$ in CN and EE, $\mu = 0.06$ and $\gamma = 0.4$ in CC.

FC, it consists of “circles” (or “friends lists”) from Facebook, the data is collected from survey participants using this Facebook app (nodes stand for Facebook users, if two users are friends then we need to add an edge between them). CN and CC denote collaborative networks of researchers (nodes stand for researchers, if two researchers have collaboration on articles, then an edge need to be added between them). As to SH, it indicates a social network (nodes stand for people, if two people are friends, an edge exists between them). EE is an email network (nodes stand for email users, if two users are contacted by an email, then an edge exists between them). Table 1 illustrates the characteristics of the adopted networks, where n , m , $\langle k \rangle$, c and r stand for the number of nodes, the number of edges, the average degree of all nodes, the average clustering coefficient, and the assortativity, respectively.

To evaluate the performance, two indicators are adopted (i.e., I_{prop} and C_N). Firstly, we utilize the proportion of infectious nodes in stable state (I_{prop}) to stand for the effect of epidemic suppression. Here, the SIS model is considered; with the provided network topology, we can simulate the spreading phenomena accordingly. Certain number of iterations of the SIS model will result in a steady number of infectious and susceptible nodes in the network. Smaller I_{prop} indicates the method is more

effective in epidemic suppression. Secondly, as in Section 3, we also consider network connectivity after removing edges. Same as Figure 4, we consider the number of connected components in the network (C_N) to evaluate the performance of different methods.

A small C_N indicates that the corresponding method is less harmful to network connectivity. Like epidemic suppression, network connectivity also possesses great practical significance. For example, the aviation network is decomposed into many connected components due to epidemic suppression, which leads to a large number of routes that have been eliminated must be added back to the aviation network if there is an emergency situation that necessitates traveling from one connected component to another. However, if there are few connected components in the network, then only a few routes need to be added. To sum up, when we remove the same number of edges, the smaller the values of I_{prop} and C_N , the better the performance of the method.

4.3 Performance Evaluation

To evaluate the performance of E_{BD} , the corresponding results are compared with those obtained by ADCM, random, WKP, ABC

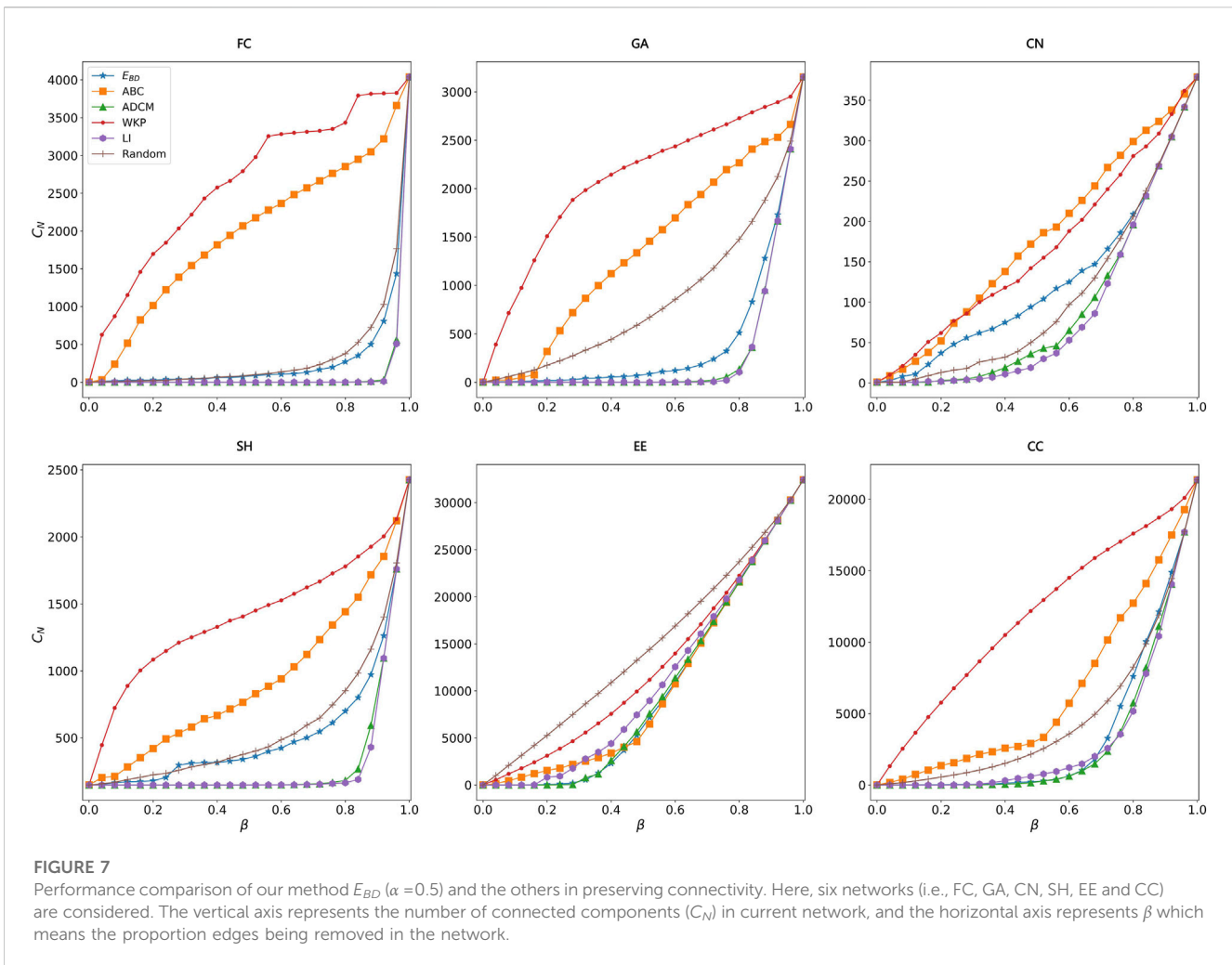


TABLE 2 Average I_{prop} in different networks. The obtained numbers represent the average value of ordinates of points on the same line in Figure 6.

Methods	GA	FC	CN	SH	EE	CC
E_{BD}	0.038	0.037	0.025	0.071	0.091	0.050
ADCM	0.092	0.319	0.064	0.980	0.097	0.050
ABC	0.037	0.034	0.023	0.060	0.046	0.054
WKP	0.097	0.173	0.129	0.132	0.138	0.112
LI	0.089	0.311	0.064	0.098	0.112	0.049
random	0.110	0.297	0.118	0.126	0.133	0.085

and LI. Firstly, we use the comparison methods and our method to obtain the edge importance sequence, respectively. Secondly, we remove edges according to these edge importance sequences respectively. We remove the edges at a certain portion (i.e., 4%) each time until all edges are removed and simulate 100 times using SIS model. Finally, we calculate the average I_{prop} and standard deviation of 100 simulation results after each removing. Note that when we simulate the propagation process through the SIS model, we randomly select only one node as the seed node initially, and the number of iterative times is large enough for the system to

converge. The means of calculating C_N is provided as follows: we remove the top $0.04m$ edges until all edges are removed according to edge importance sequence and calculate C_N per time.

Based on the above preparation, we first study the effect of different values of α on our method in the GA network, as the results are shown in Figure 5. According to Figure 5, given the same number of edges are removed, with the increase of α , I_{prop} decreases continuously, in contrast to the constant increase of C_N . In addition, the standard deviation remains relatively stable and does not exhibit significant change. Thus, these results confirm our assumption. With the increase α , the effect of betweenness centrality increases gradually leading to the network connectivity is more severely destroyed and enhancing epidemic suppression effectiveness when a few edges are removed. Controlling the spread of virus is relatively effective when $\alpha = 0.5$, and network connectivity is not significantly affected. When α is less than 0.5, even though network connectivity protection is improved, epidemic suppression is ineffective; when α is greater than 0.5, the effect of epidemic suppression becomes better, but network connectivity cannot be protected. As we normalize C_B and C_D before the fusion process, the above phenomena can be explained as follows. Firstly, when α is greater than 0.5, EBC dominates E_{BD} . Furthermore, since EBC selects the most central edge to remove, the network will be quickly decomposed into a number of small connected components which leads to better

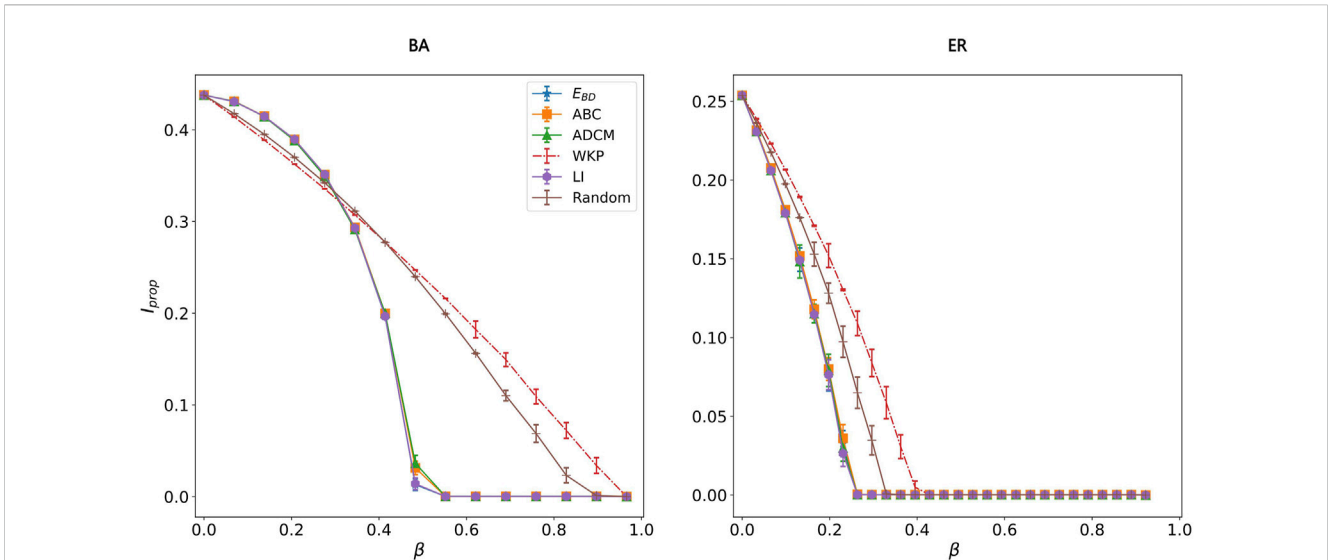


FIGURE 8 Performance comparison of our method E_{BD} ($\alpha = 0.5$) and the other methods in epidemic suppression. We show the spread range (I_{prop}), as a function of the fraction of removed edge (β) in two generated networks (BA network and ER network). For BA network, the number of nodes equals to 5,000 and the corresponding average degree is 6; as to ER network, the number of nodes is 5,000 and the average degree is 12. For both networks, the recovery probability γ is set to 0.5; the infectious probability μ is assigned to 0.2 and 0.06 for BA and ER networks, respectively.

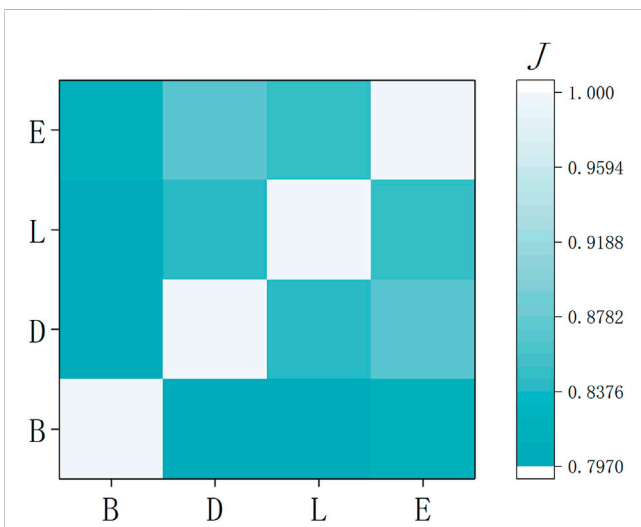


FIGURE 9 The correlation heat map for the four indices of edge importance over BA network. Here, B, D, L and E represent ABC, ADCM, LI and E_{BD} , respectively.

performance in epidemic suppression at a cost of sacrificing connectivity protection. Secondly, when α is less than 0.5, ADCM dominates E_{BD} . As ADCM selects the edge with the largest degree multiplication for removal, E_{BD} with small α has an advantage of connectivity protection at a cost of sacrificing epidemic suppression when some edges are removed. Note that since increasing the epidemic threshold is the starting point of ADCM, when α decreases, fewer edges need to be removed to total epidemic control which means $I_{prop} \rightarrow 0$. Thus, we set $\alpha = 0.5$ to achieve a trade-off between the two aspects.

In the following experiments, we set $\alpha = 0.5$ in Eq. 7. In Figure 6, Figure 7 and Table 2, E_{BD} achieves better performance at epidemic suppression and protection of network connectivity. According to Figure 6 and Table 2, we can carefully make the following conclusions. Firstly, the performance of E_{BD} with $\alpha = 0.5$ in epidemic suppression is better than that of LI, WKP, ADCM and Random, which proves the efficiency of our proposed method. Secondly, the error bar of E_{BD} falls within the acceptable range (the maximum standard deviations of E_{BD} in FC, GA, CN, SH, EE and CC are 0.012, 0.015, 0.011, 0.025, 0.022 and 0.027, respectively) which demonstrates the effectiveness of results. Finally, according to Figure 7, when $\alpha = 0.5$ in E_{BD} , it has a good effect in protecting network connectivity while the other methods with similar effects in protecting connectivity, are much worse than E_{BD} at epidemic suppression. The authors of the LI method emphasize the advantages of connectivity protection at the cost of losing epidemic suppression. Our method has both the above advantages.

In addition to the practical networks, we also conduct experiments on generated networks (i.e., BA and ER networks). As shown in Figure 8, we find that the performance of our proposed method in epidemic suppression seems to be similar as three methods (i.e., ABC, ADCM and LI) which is much better than the others. In the ER network, since the properties of all edges are relatively comparable, the performance of these methods in epidemic suppression are similar.

In the BA network, we design a sequence similarity verification experiment to explore the potential power of the proposed method in epidemic suppression. Furthermore, we adopt Jaccard index [35] to reflect the correlations between edge ranking sequences derived by different methods. The specific calculation steps of Jaccard index are as follows. Firstly, the correlation between two sequences is calculated as:

$$J(a) = \frac{S_1(a) \cap S_2(a)}{S_1(a) \cup S_2(a)}, \quad (8)$$

where a indicates a parameter ($a \in [0, 1]$), $S_1(a)$ and $S_2(a)$ represent the sets formed by the first $a \times m$ elements of sequence S_1 and sequence S_2 , respectively. Then, we calculate the correlation for the whole sequence:

$$J = \frac{1}{100} \sum_{a=0.01}^1 J(a), \quad (9)$$

where 100 indicates that we divide the sequence into 100 short sequences of the same length, incrementally calculate the correlation of the sequences, and eventually obtain the average. According to Eq. 8 and Eq. 9, we conduct some experiments. As shown in Figure 9, the Jaccard indexes between our method and ABC, ADCM, and LI in the BA network are all greater than 0.79, which reveals the edge sequence obtained by our method is similar to those obtained by the other methods. Therefore, the four methods can achieve good performance in epidemic suppression.

5 Conclusion

In summary, we propose a method to rank edges according to the degree of the connected nodes and betweenness centrality of the edge. From the perspectives of epidemic suppression and connectivity protection, we first consider using the product of the degrees of the two nodes to increase epidemic threshold and protect network connectivity. However, when some portion of edges are removed, it has bad effect on epidemic suppression. Here, we propose a new method (i.e., E_{BD}) combined with EBC to eliminate the bad effect. The experiments demonstrate that the proposed approach achieves better results compared with the other methods when some edges are removed. Furthermore, our method can not only protect the connectivity of networks, but also possess much better epidemic suppression effects. In the future, we will study and expand our method in other spread models (e.g., SIR and SEIR) or other types of networks (e.g., time-varying network and double-layer network) to increase the universality of our method.

References

- Boccaletti S, Latora V, Moreno Y, Chavez M, Hwang DU. Complex networks: Structure and dynamics. *Phys Rep* (2006) 424:175–308. doi:10.1016/j.physrep.2005.10.009
- Arenas A, Diaz-Guilera A, Kurths J, Moreno Y, Zhou C. Synchronization in complex networks. *Phys Rep* (2008) 469:93–153. doi:10.1016/j.physrep.2008.09.002
- Heckmann T, Schwanghart W, Phillips JD. Graph theory—Recent developments of its application in geomorphology. *Geomorphology* (2015) 243:130–46. doi:10.1016/j.geomorph.2014.12.024
- Baronchelli A, Ferrer-i Cancho R, Pastor-Satorras R, Chater N, Christiansen MH. Networks in cognitive science. *Trends Cogn Sci* (2013) 17:348–60. doi:10.1016/j.tics.2013.04.010
- Barabási A-L, Gulbahce N, Loscalzo J. Network medicine: A network-based approach to human disease. *Nat Rev Genet* (2011) 12:56–68. doi:10.1038/nrg2918
- Li J, Sun G, Jin Z. Interactions of time delay and spatial diffusion induce the periodic oscillation of the vegetation system. *Discrete and Continuous Dynamical Systems-B* (2022) 27:2147–72. doi:10.3934/dcdsb.2021127
- Sun G, Zhang H, Song Y, Li L, Jin Z. Dynamic analysis of a plant-water model with spatial diffusion. *J Differential Equations* (2022) 329:395–430. doi:10.1016/j.jde.2022.05.009
- Li J, Sun G, Guo Z. Bifurcation analysis of an extended klausmeier-gray-scott model with infiltration delay. *Stud Appl Maths* (2022) 148:1519–42. doi:10.1111/sapm.12482
- Christley RM, Pinchbeck G, Bowers RG, Clancy D, French NP, Bennett R, et al. Infection in social networks: Using network analysis to identify high-risk individuals. *Am J Epidemiol* (2005) 162:1024–31. doi:10.1093/aje/kwi308
- Pastor-Satorras R, Vespignani A. Immunization of complex networks. *Phys Rev E* (2002) 65:036104. doi:10.1103/PhysRevE.65.036104
- Zhang H, Li K, Fu X, Wang B. An efficient control strategy of epidemic spreading on scale-free networks. *Chin Phys Lett* (2009) 26:068901. doi:10.1088/0256-307X/26/6/068901
- Newman ME. A measure of betweenness centrality based on random walks. *Social Networks* (2005) 27:39–54. doi:10.1016/j.socnet.2004.11.009

Data availability statement

Publicly available datasets were analyzed in this study. This data can be found here: <https://networkrepository.com>, <https://openflights.org>, <http://snap.stanford.edu/data/>.

Author contributions

GL wrote the original manuscript. XC and PZ proposed the idea, supervised the research work, and revised the manuscript. GL and PZ discussed and analyzed the results.

Funding

This work was supported in part by the National Key R&D Program of China (No. 2020AAA0107704), National Natural Science Foundation of China (Grant Nos. 62073263, 62003273), Natural Science Basic Research Program of Shaanxi (Program No. 2020JQ-217), Key Research and Development Program of Shaanxi Province (Grant No. 2022KW-26), Technological Innovation Team of Shaanxi Province (Grant No. 2020TD-013).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

13. Bonacich P. Some unique properties of eigenvector centrality. *Soc Networks* (2007) 29:555–64. doi:10.1016/j.socnet.2007.04.002
14. Kitsak M, Gallos LK, Havlin S, Liljeros F, Muchnik L, Stanley HE, et al. Identification of influential spreaders in complex networks. *Nat Phys* (2010) 6: 888–93. doi:10.1038/nphys1746
15. De Meo P, Ferrara E, Fiumara G, Ricciardello A. A novel measure of edge centrality in social networks. *Knowledge-Based Syst* (2012) 30:136–50. doi:10.1016/j.knsys.2012.01.007
16. Girvan M, Newman ME. Community structure in social and biological networks. *Proc Natl Acad Sci* (2002) 99:7821–6. doi:10.1073/pnas.122653799
17. Newman ME, Girvan M. Finding and evaluating community structure in networks. *Phys Rev E* (2004) 69:026113. doi:10.1103/PhysRevE.69.026113
18. Matamalas JT, Arenas A, Gómez S. Effective approach to epidemic containment using link equations in complex networks. *Sci Adv* (2018) 4:eau4212. doi:10.1126/sciadv.aau4212
19. Bröhl T, Lehnertz K. A straightforward edge centrality concept derived from generalizing degree and strength. *Scientific Rep* (2022) 12:4407–12. doi:10.1038/s41598-022-08254-5
20. Aghaalizadeh S, Afshord ST, Bouyer A, Anari B. A three-stage algorithm for local community detection based on the high node importance ranking in social networks. *Physica A: Stat Mech its Appl* (2021) 563:125420. doi:10.1016/j.physa.2020.125420
21. El Kouni IB, Karoui W, Romdhane LB. Node importance based label propagation algorithm for overlapping community detection in networks. *Expert Syst Appl* (2020) 162:113020. doi:10.1016/j.eswa.2019.113020
22. Wang Z, Hou D, Gao C, Huang J, Xuan Q, Castellanos FX, et al. Cerebro-cerebellar dysconnectivity in children and adolescents with attention-deficit/hyperactivity disorder. *Proc ACM Web Conf* (2022) 61:1372–84. doi:10.1016/j.jaac.2022.03.035
23. Liu J, Ren Z, Guo Q. Ranking the spreading influence in complex networks. *Physica A: Stat Mech its Appl* (2013) 392:4154–9. doi:10.1016/j.physa.2013.04.037
24. Wang Z, Jusup M, Shi L, Lee J-H, Iwasa Y, Boccaletti S. Exploiting a cognitive bias promotes cooperation in social dilemma experiments. *Nat Commun* (2018) 9:2954. doi:10.1038/s41467-018-05259-5
25. Wang Z, Wang C, Gao C, Li X, Li X. An evolutionary autoencoder for dynamic community detection. *Sci China Inf Sci* (2020) 63:212205–16. doi:10.1007/s11432-020-2827-9
26. Zhu P, Cheng L, Gao C, Wang Z, Li X. Locating multi-sources in social networks with a low infection rate. *IEEE Trans Netw Sci Eng* (2022) 9:1853–65. doi:10.1109/TNSE.2022.3153968
27. Li Q, Cheng L, Wang W, Li X, Li S, Zhu P. Influence maximization through exploring structural information. *Appl Maths Comput* (2023) 442:127721. doi:10.1016/j.amc.2022.127721
28. Freeman LC. A set of measures of centrality based on betweenness. *Sociometry* (1977) 40:35–41. doi:10.2307/3033543
29. Pastor-Satorras R, Vespignani A. Epidemic dynamics in finite size scale-free networks. *Phys Rev E* (2002) 65:035108. doi:10.1103/PhysRevE.65.035108
30. Van Mieghem P, Stevanović D, Kuipers F, Li C, Van De Bovenkamp R, Liu D, et al. Decreasing the spectral radius of a graph by link removals. *Phys Rev E* (2011) 84:016101. doi:10.1103/PhysRevE.84.016101
31. Ferreira SC, Castellano C, Pastor-Satorras R. Epidemic thresholds of the susceptible-infected-susceptible model on networks: A comparison of numerical and theoretical results. *Phys Rev E* (2012) 86:041125. doi:10.1103/PhysRevE.86.041125
32. Wang S, Gong M, Liu W, Wu Y. Preventing epidemic spreading in networks by community detection and memetic algorithm. *Appl Soft Comput* (2020) 89:106118–8. doi:10.1016/j.asoc.2020.106118
33. [Dataset] Rossi RA, Ahmed NK. The network data repository with interactive graph analytics and visualization. *Proc AAAI Conf Artif Intelligence* (2015) 29. doi:10.1609/aaai.v29i1.9277
34. Leskovec J, McAuley J. Learning to discover social circles in ego networks. *Adv Neural Inf Process Syst* (2012) 25:1–9.
35. Hamers L, Hemeryck Y, Herweyers G, Janssen M, Keters H, Rousseau R, et al. Similarity measures in scientometric research: The jaccard index versus salton's cosine formula. *Inf Process Manag* (1989) 25:315–8. doi:10.1016/0306-4573(89)90048-4