

Vital nodes identification method integrating degree centrality and cycle ratio

Yu Zhao(赵玉) and Bo Yang(杨波)

Citation: Chin. Phys. B, 2025, 34 (**3**): 038901. DOI: 10.1088/1674-1056/ada42d
Journal homepage: http://cpb.iphy.ac.cn; http://iopscience.iop.org/cpb

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

Impact of environmental factors on the coevolution of information-emotions-epidemic dynamics in activity-driven multiplex networks

Liang'an Huo(霍良安), Bingjie Liu(刘炳杰), and Xiaomin Zhao(赵晓敏)
Chin. Phys. B, 2024, 33 (**12**): 128903. DOI: 10.1088/1674-1056/ad7df5

Percolation transitions in edge-coupled interdependent networks with directed dependency links

Yan-Li Gao(高彦丽), Hai-Bo Yu(于海波), Jie Zhou(周杰), Yin-Zuo Zhou(周银座), and Shi-Ming Chen(陈世明)
Chin. Phys. B, 2023, 32 (**9**): 098902. DOI: 10.1088/1674-1056/acd685

LCH: A local clustering H-index centrality measure for identifying and ranking influential nodes in complex networks

Gui-Qiong Xu(徐桂琼), Lei Meng(孟蕾), Deng-Qin Tu(涂登琴), and Ping-Le Yang(杨平乐)
Chin. Phys. B, 2021, 30 (**8**): 088901. DOI: 10.1088/1674-1056/abea86

Contagion dynamics on adaptive multiplex networks with awareness-dependent rewiring

Xiao-Long Peng(彭小龙) and Yi-Dan Zhang(张译丹)
Chin. Phys. B, 2021, 30 (**5**): 058901. DOI: 10.1088/1674-1056/abe1ab

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

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

---

SPECIAL TOPIC — Computational programs in complex systems

Vital nodes identification method integrating degree centrality and cycle ratio

Yu Zhao(赵玉)[1,2] and Bo Yang(杨波)[1,2,†]

[1] *Data Science Research Center, Kunming University of Science and Technology, Kunming* 650500, *China*
[2] *Faculty of Science, Kunming University of Science and Technology, Kunming* 650500, *China*

(Received 9 October 2024; revised manuscript received 7 December 2024; accepted manuscript online 31 December 2024)

Identifying vital nodes is one of the core issues of network science, and is crucial for epidemic prevention and control, network security maintenance, and biomedical research and development. In this paper, a new vital nodes identification method, named degree and cycle ratio (DC), is proposed by integrating degree centrality (weight α) and cycle ratio (weight $1-\alpha$). The results show that the dynamic observations and weight α are nonlinear and non-monotonicity (i.e., there exists an optimal value α^* for α), and that DC performs better than a single index in most networks. According to the value of α^*, networks are classified into degree-dominant networks ($\alpha^* > 0.5$) and cycle-dominant networks ($\alpha^* < 0.5$). Specifically, in most degree-dominant networks (such as Chengdu-BUS, Chongqing-BUS and Beijing-BUS), degree is dominant in the identification of vital nodes, but the identification effect can be improved by adding cycle structure information to the nodes. In most cycle-dominant networks (such as Email, Wiki and Hamsterster), the cycle ratio is dominant in the identification of vital nodes, but the effect can be notably enhanced by additional node degree information. Finally, interestingly, in Lancichinetti–Fortunato–Radicchi (LFR) synthesis networks, the cycle-dominant network is observed.

Keywords: cycle ratio, percolation, epidemic spreading, targeted immunization

PACS: 89.75.Fb, 64.60.aq

DOI: 10.1088/1674-1056/ada42d **CSTR:** 32038.14.CPB.ada42d

1. Introduction

The identification of vital nodes in complex networks is a long-standing problem in network science.[1–5] It is critical for controlling outbreaks, mitigating damage caused by attacks on transportation or telecommunications networks, and finding key genes for targeted receptor proteins and molecular biologics.[6,7] Although many simple and effective methods have been applied to these problems, they have their limitations due to the diversity of evaluation criteria. For example, degree centrality[8] is simple and intuitive but does not consider the global network structure, while betweenness centrality[9] and closeness centrality[10] consider the global structure but have high computational cost. Therefore, vital nodes identification is not a simple task, and it is particularly important to constantly explore new identification strategies. The LocalRank algorithm[11] identified vital nodes by considering the information contained in the fourth-order neighbors of each node. Kitsak *et al.*[12] obtained the coreness of node through k-core[13] decomposition, and believed that a larger coreness value indicates a more central location in network. Lv *et al.*[14,15] discussed an extension of the H-index concept to quantify the importance of nodes in its network. The H-index of a node is defined to be the maximum value H such that there exists at least H neighbors of degree no less than H. Node centrality ranking based on random walk is also a feasible method. For example, the PageRank algorithm[16] constructs a network based on the relationship between web pages, and distinguishes the importance of different websites through random walks. Lin and Zhang[17] proposed a random walk method based on non-retrospective centrality, NBCRW, in which walkers are more likely to visit neighbors with high non-retrospective centrality. TOPSIS[18] regarded centrality measures as the multi-attribute of a complex network and connected the multi-attribute to evaluate node importance.

With the rapid increases of network scale and complexity, mining vital nodes faces new challenges, and the cycle structure within networks has also received more and more attention. The cycle structure is an important part of a network and is defined as a closed path with the same start and end nodes. It plays an important role in structure organization and function realization. Fan *et al.*[19] proposed the cycle ratio, which quantifies the importance of an individual node by simply measuring to which extent it is involved in the shortest cycles associated with other nodes, and it performs well in scaling and control processes. Zhao *et al.*[20] proposed a centrality measurement method based on high-order network characteristics, including higher-order cycle ratio (HOC), higher-order degree (HOD), higher-order H-index (HOH), and higher-order PageRank (HOP), to quantify and rank the importance of

†Corresponding author. E-mail: yangbo@kust.edu.cn

© 2025 Chinese Physical Society and IOP Publishing Ltd. All rights, including for text and data mining, AI training, and similar technologies, are reserved.

cliques in a network. The results indicate that clique ranking is superior to other metrics in maintaining network connectivity. Jiang *et al.*[21] ranked the importance of different cycles in a network by comparing the sensitivity of the Fiedler value (the second smallest Laplacian eigenvalue), and identified the key cycles that contributed the most to the dynamic behavior of the network. Shi *et al.*[22] believed that the nodes participating in the cycle basis might be crucial to propagation, and proposed the basic cycle number (NC) to quantify the importance of a node.

Although the cycle ratio is effective in identifying vital nodes in a network, the obvious shortcoming of a cycle is that it cannot be applied to trees or tree-like networks. Even for ordinary networks, there may be some nodes that are not associated with any cycle, which have different effects but are given the same importance. The degree mainly evaluates the importance of a node based on the number of its direct neighbors, but ignores the vital bridging nodes with a small degree. Therefore, this paper proposes the index DC (degree and cycle ratio) to quantify the importance of single node by combining degree (weight α) and cycle ratio (weight $1-\alpha$), so that all nodes with equal cycle importance will be updated by considering the degree of node at the same time. In order to evaluate the performance of DC, three dynamic processes, which include connectivity test,[23–25] epidemic spreading simulations[26,27] and targeted immunization control,[28–30] are used to test the identification effect on degree-dominant network and cycle-dominant network respectively, and the results are compared with other node-ranking methods. It is found that the dynamic observations and weight α exhibit non-monotonicity characteristics, meaning that there exists an optimal parameter α that enables DC to perform better than a single index in most networks. Finally, in the Lancichinetti–Fortunato–Radicchi (LFR) synthesis network model, the cycle-dominant network was observed.

This paper is organized as follows. In Section 2, we introduce the vital node identification method DC and the model motivation. In Section 3, we use three dynamic processes, including node percolation, epidemic spreading and targeted immunization, to verify the rationality of the method, and carry out analysis and discussion. Then the robustness is simulated and analyzed in the LFR synthesis network. In Section 4, we draw some conclusions.

2. Motivation and model

Real-world networks exhibit a high degree of complexity, which is manifested not only in the number of nodes and edges but also in the ways in which nodes are connected. The connections between nodes in real-world networks are often not simply linearly arranged but instead form a complex topology. Cycles, as a fundamental topological structural element abundant in real-world networks, can well reflect the complexity and redundancy of network structures. Recent research has unveiled the topological characteristics of cycles, including the distribution of cycles of different sizes in both real-world and synthetic networks.[31] Cycles play significant roles in network functions related to storage,[32] synchronizability,[33] and controllability.[34] For instance, cycles can introduce redundant paths in network connectivity and feedback effects in network dynamics. Additionally, cycles are also used as a tool to measure the proximity of networks to tree networks, thereby revealing significant differences between synthetic networks and real-world networks.[35]

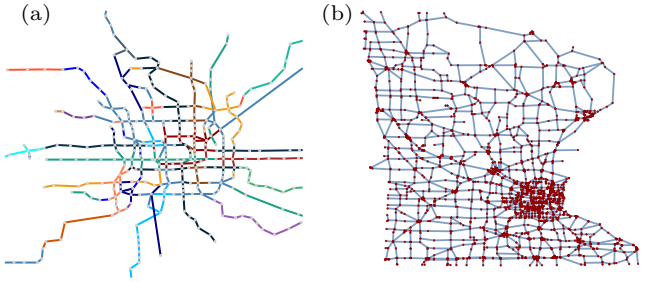


Fig. 1. Cycle structures are widespread in networks: (a) visualization of part of the Beijing subway network; (b) visualization of part of the Minnesota road network.

Different cycles are likely to play different roles in network structures and dynamics, among which the smallest cycles play an important role in analysis of graphs or networks. For example, the cohesion of a network is measured by the clustering coefficient, which is essentially the fraction of all possible triangles (the smallest cycle) in a graph. In graph theory, the minimum length bases of the cycle space of a graph (MCBs) of cycle spaces have broad applications in various scientific and engineering fields, such as structural flexibility analysis, power grids, and chemical structure storage and retrieval systems.[36] In social networks, the community structure, as a fundamental characteristic, reflects the high connectivity and closeness among members within a community, naturally leading to the emergence of numerous cycle structures. In public transportation networks, such as bus and subway networks, loop lines are widely available, which increases the possibility of transfer, thus significantly improving the efficiency and flexibility of the entire transportation system. As shown in Fig. 1(a), which depicts a portion of the Beijing subway network, loop lines are ubiquitous and constitute an important part of the network layout, not only enhancing internal connectivity but also greatly facilitating convenient transfers between different lines for passengers. It is generally believed that a road network is composed of many polygons. Figure 1(b) shows a visualization of the road network of Minnesota.[37] It can be seen that a large number of cycles exist in road networks. The cycle structures exhibit good

reachability. When both the initial node and the final node are specified, there are two distinct and well-connected paths, indicating good reachability. Furthermore, a cycle graph is characterized as a two-connected graph, implying that the removal of at least two vertices is necessary to disconnect it. In the field of network security, cycle structures play a crucial role in enhancing the robustness of networks.

Cycle structure is a unique characteristic of node importance that distinguishes it from other indicators. Degree centrality focuses on considering the number of connections of nodes. Betweenness and closeness centrality essentially involve path analysis, where the former calculates the number of times a node is passed by the shortest path, while the latter evaluates the centrality by the reciprocal length of the shortest path between a node and other nodes. H-index and eigenvector centrality primarily focus on the importance of the node itself and its neighbors, without directly considering the cycle structure. The local clustering coefficient only reflects the clustering property of a node by considering the number of triangles formed by three nodes, neglecting the significance of larger cycles that still exist in large numbers in the network. Coreness determines the core status of a node by analyzing the number of neighbors of a node and the number of neighbors of its neighbors, but it fails to account for the impact of cycle structures on network structure and function. Therefore, we propose a new method for identifying vital nodes by combining cycle with degree centrality. Certainly, it is also possible to combine cycles with the other indicators mentioned above, which will be discussed in Subsection 3.1, but the results show that the combination of degree and cycle ratio has a relatively good effect.

Based on the above considerations, this paper proposes a new vital node identification method. Considering a simple network $G = (V, E)$, where V is the set of N nodes and E is the set of edges. The size of a cycle is equal to the number of edges it contains. The cycle containing the smallest size of node i is defined as the shortest cycle associated with node i, and the corresponding size is called the girth of node i. S_i represents the set of all smallest cycles associated with i. $S = \cup_{i\in V} S_i$ is the minimum set of all cycles in the graph G. The cycle number matrix $C = [c_{ij}]_{N\times N}$, where c_{ij} is the number of cycles in S that pass through both i and j when $i \neq j$. If $i = j$, c_{ii} is the number of cycles in S that contain i. Based on the number of cycles matrix C, the formula of cycle ratio is as follows:[19]

$$r_i = \begin{cases} 0, & \text{if } c_{ii} = 0, \\ \sum_j \frac{c_{ij}}{c_{jj}}, & \text{if } c_{ii} > 0. \end{cases} \tag{1}$$

The identification of vital nodes under different indicators in Fig. 2 is evaluated according to the attack efficiency of complex networks. As is shown in Table 1, the metric values of each node in Fig. 2 under degree, betweenness, cycle ratio, coreness, eigenvector centrality, local clustering coefficient, H-index and DC are shown. By observing the ranking of the network in Fig. 2 and the corresponding node centrality index in Table 1, it is found that node 3 is regarded as the key node when measured according to degree, betweenness centrality and eigenvector centrality. When node 3 is removed, the size of the maximal connected subgraph of network is 6. In Fig. 2(b), nodes 4, 5, 6, 7, 8, and 9 are simultaneously considered as central nodes when measured according to cycle ratio, coreness, and H-index, so the optimal vital node cannot be determined. Local clustering coefficients are not effective in identifying important nodes in such networks. Obviously, it is inefficient to use only one metric for many nodes with the same metric. Therefore, in order to improve the attack efficiency, we put forward a new index DC to measure the importance of nodes. Each node is assigned a DC value that indicates how important the node is in the network. The coefficient α is used to adjust the importance of different metrics in DC, the DC formula is as follows:

$$\mathrm{DC}_i = \alpha \cdot k_i + (1-\alpha)\cdot r_i, \tag{2}$$

where k_i and r_i indicate the degree and cycle ratio of node i, respectively, $0 \leq \alpha \leq 1$.

Table 1. The metric values of each node in Fig. 2 under degree, betweenness, cycle ratio, coreness, eigenvector centrality, local clustering coefficient, H-index and DC.

Node ID	Degree	Betweenness	Cycle ratio	Coreness	Eigenvector	Clustering	H-index	DC ($\alpha = 0.8$)
0	1	0	0	1	0.23	0	1	0.8
1	2	0.2	0	1	0.28	0	1	1.6
2	1	0	0	1	0.23	0	1	0.8
3	4	0.64	0	1	0.53	0	2	3.2
4	3	0.60	6	2	0.48	0	2	3.6
5	2	0.21	6	2	0.30	0	2	2.8
6	2	0.10	6	2	0.21	0	2	2.8
7	2	0.04	6	2	0.18	0	2	2.8
8	2	0.10	6	2	0.21	0	2	2.8
9	2	0.21	6	2	0.30	0	2	2.8
10	1	0	0	1	0.12	0	1	0.8

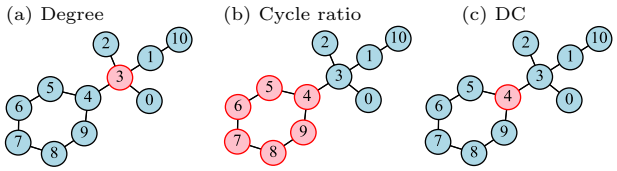


Fig. 2. Illustration of an example network. The nodes in the red circle are "vital nodes" under different vital node ranking metrics. (a) Degree ($\alpha = 1$); (b) cycle ratio ($\alpha = 0$); (c) DC ($0 < \alpha < 6/7$).

Figure 2(c) shows the identification of vital nodes under the DC indicator when $0 < \alpha < 6/7$. The DC identifies node 4 as the key node. When node 4 is removed, the size of the maximal connected subgraph of network is 5, and the network disassembly efficiency reaches the maximum.

3. Results and discussion

We used Email (email communication network from the Rovira I Virgili University in Wisconsin from the email communication network of Rovira I Virgili University in Spain),[38] Wiki (Wikipedia network),[39] Hamsterster (Social network),[40] the bus networks Chengdu-BUS, Chongqing-BUS and Beijing-BUS to study the importance of nodes. Their basic topological features are shown in Table 2. Other datasets used are in Appendixes A–C.

Table 2. The topological properties of networks. Average degree $\langle k \rangle$; average path length $\langle L \rangle$; average clustering coefficient $\langle C \rangle$.

Network	N	E	$\langle k \rangle$	$\langle L \rangle$	$\langle C \rangle$
Email	1133	5451	9.62	3.61	0.22
Wiki	2405	11596	9.64	3.65	0.38
Hamsterster	2426	16630	13.71	3.59	0.54
Chongqing-BUS	5476	8767	3.2	13.44	0.15
Beijing-BUS	9416	14248	3.03	28.55	0.13
Chengdu-BUS	10460	16755	3.20	25.50	0.13

Three dynamic processes, including node percolation, epidemic spreading, and targeted immunization, are used to test the rationality of DC in identifying vital nodes. The first dynamic process considers the ability of nodes to maintain network connectivity, the second focuses on the accessibility of infected nodes during epidemic outbreaks, and the third concentrates on prevention and control in the early stages of disease outbreaks. We analyze the network characteristics by dividing the degree-dominant network (Email, Wiki, Hamsterster) and the cycle-dominant network (Chongqing-BUS, Beijing-BUS, Chengdu-BUS), and verify the effectiveness of DC in identification of vital nodes. The DC index provides a new perspective and tool for vulnerability analysis of networks, protection of critical infrastructure, and prevention and control of outbreaks.

3.1. Percolation

In order to evaluate the importance of nodes in maintaining network connectivity, this paper studies the percolation dynamics of nodes[41] and uses a static approach based on the initial state of the network to evaluate its robustness. Specifically, the nodes are ranked by their DC value in the initial network from smallest to largest, and then the nodes with the highest DC value are removed in turn. This method assumes that the network structure remains unchanged during the evaluation process, and can quickly identify nodes that have an initial impact on network connectivity. Removing nodes may change the network connectivity. The relative size of the maximal connected subgraph of a network is often used to indicate the robustness of the network to random or deliberate attack strategies. The relative size of the maximal connected subgraph of the network G can be expressed as

$$g(t) = \frac{n_t}{N}, \tag{3}$$

where N represents the total number of nodes in the initial network, and n_t represents the number of nodes in the maximal connected subgraph of network after t nodes are removed at step t, with one node removed in each step.

Figure 3 shows the ranking of vital nodes by six real-world networks under different weights of α for the DC index. The relative size of the maximal connected subgraph g of the network changes with the proportion of removed nodes P. Obviously, it can be observed that the relative size of the maximal connected subgraph of the network varies greatly with different α, especially in Figs. 3(d)–3(f). To quantify and compare the differences between the curves in Fig. 3, we calculate the area between each curve and the horizontal and vertical axes as an indicator of the differences between curves. We use robustness to measure the efficiency of attacks under different attack strategies. For a given network G, one node is removed at each time step and the size of the maximal connected subgraph in the remaining network is calculated until the remaining network is empty. The observation R is defined as[2,19]

$$R = \frac{1}{N}\sum_t g(t), \tag{4}$$

which describes the area enclosed by the g vs. P curve with the x-axis and y-axis. Obviously, a smaller R means a faster breakdown of the network, resulting in a more efficient sorting of vital nodes.

Based on the premise that the relative size of the maximal connected subgraph reaches 0.01 for the first time, this paper explores the changes of R under different coefficients α. As shown in Fig. 4, when removing nodes based on the DC attack strategy, the observation R is a function of coefficient α. Interestingly, as α increases, the observation R exhibits a nonlinear relationship with α and shows a trend of first decreasing and then increasing. This non-monotonic property reveals that there is an optimal configuration of α, which ensures that the extent of network dismantling under the ranking of vital nodes obtained according to DC index reaches the maximum. At the

same time, it also proves that there are deficiencies in identifying vital nodes under a single index, and blindly enlarging the degree advantage or cycle structure advantage in the DC value of nodes by changing α cannot guarantee the best results in network dismantling.

In order to show the effect, we take α corresponding to the lowest point of the curve in Fig. 4 as the coefficient of DC for vital node identification. Figure 5 shows the breakdown process of the real-world networks under different attack strategies, including removing nodes according to degree,

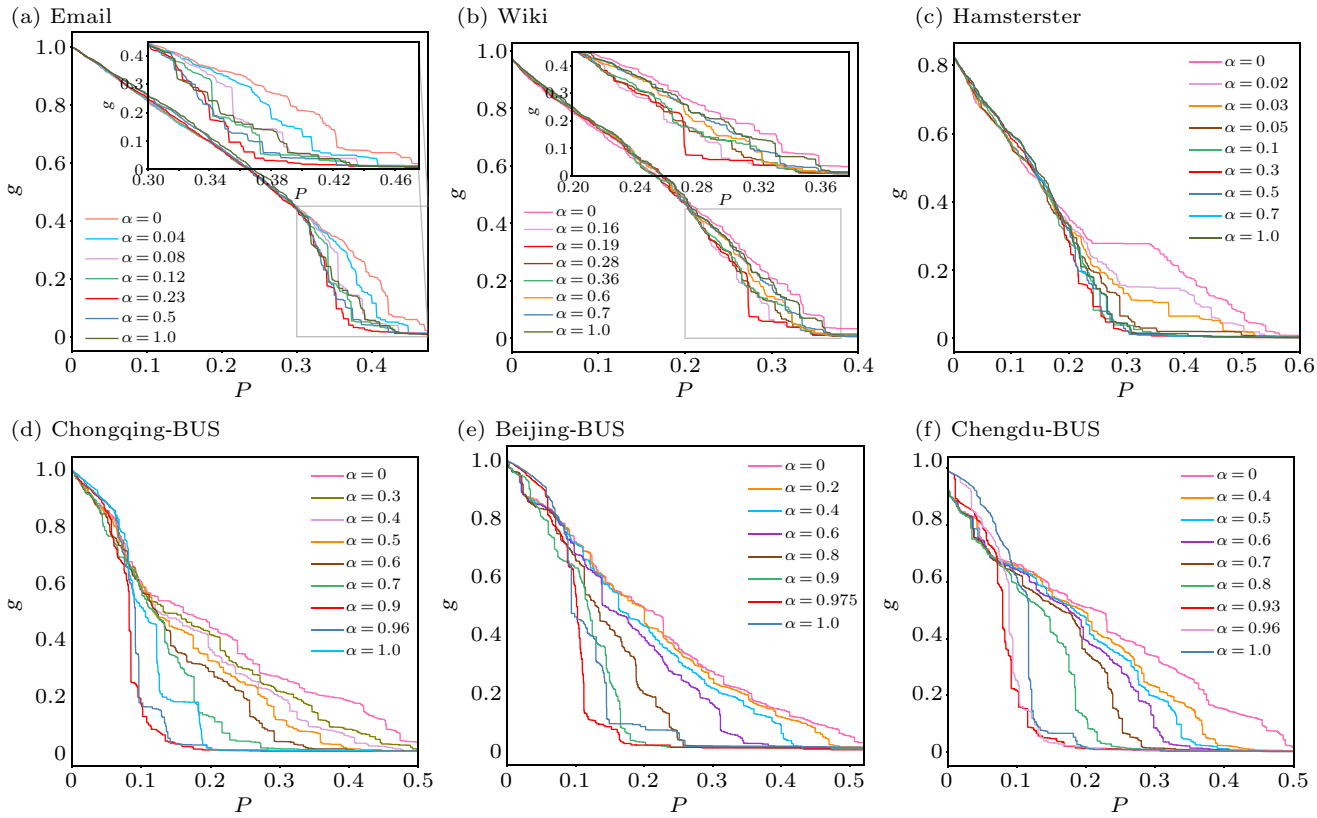


Fig. 3. The relative size of the maximal connected subgraph of the network g as a function of the proportion of nodes removed P, where $P = t/N$, with one node removed at each step.

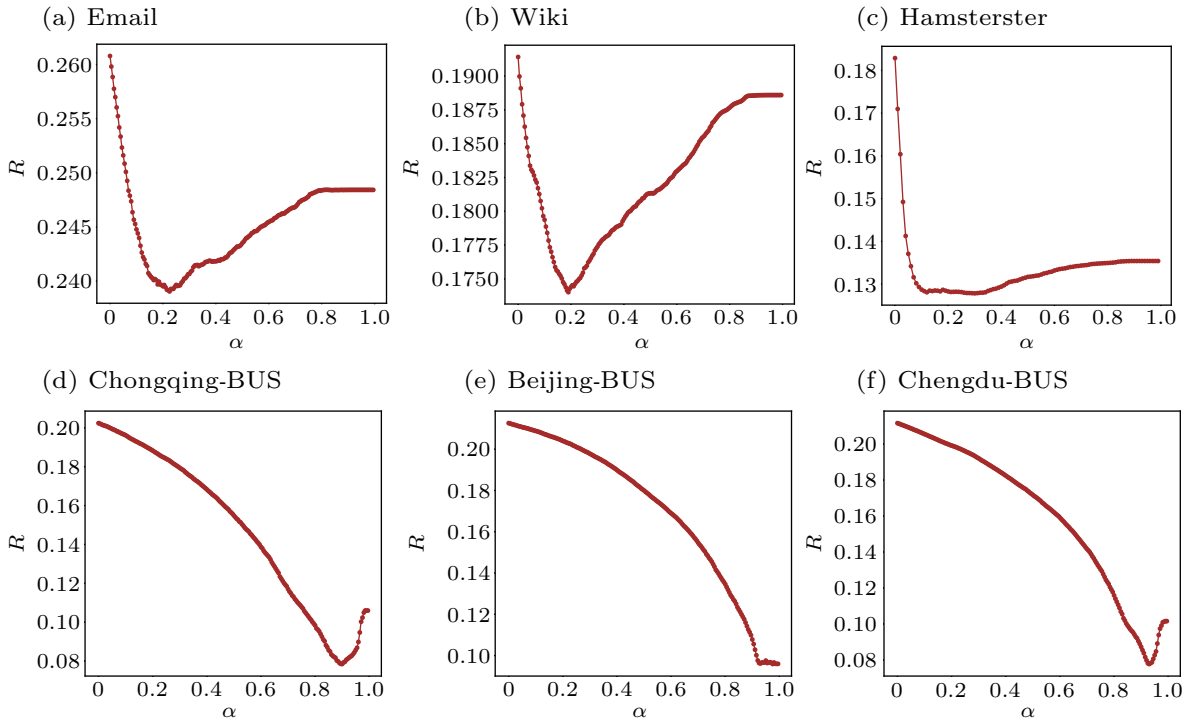


Fig. 4. The observation R as a function of the coefficient α for different networks.

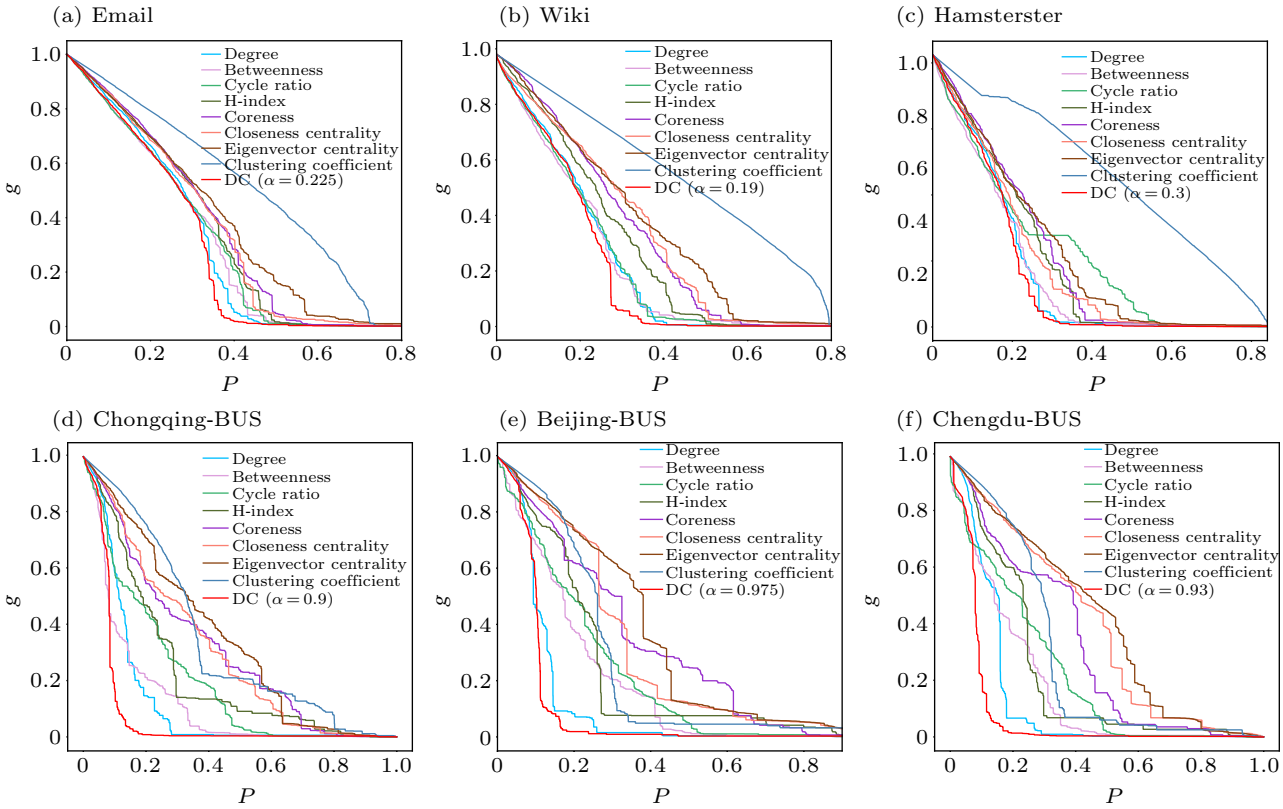


Fig. 5. The relative size of the maximal connected subgraph of the network g as a function of the proportion of nodes removed P, where $P = t/N$, with one node removed at each step.

betweenness, cycle ratio, H-index, coreness, closeness centrality, eigenvector centrality, local clustering coefficient and DC. Under removing nodes one by one according to different attack strategies, the relative size of the maximal connected subgraph of the network g as a function of the proportion of nodes removed P. Obviously, in most networks, DC causes the network to crash faster than other metrics. Specifically, when the node removal ratio P is small, the curves representing the relative size of the maximal connected subgraph of the nine attack strategies are constantly interleaving during the network dismantling process, which indicates that the nine attack strategies at this time have quite similar effects on network connectivity, showing equivalence. As P gradually increases, DC attack strategy gradually shows better performance than other metrics in network dismantling, and the network begins to crash faster. As can be seen from the above, in most networks, DC index shows effective performance in network disassembly and is the best index for identifying vital nodes that maintain network connectivity.

Certainly, in addition to the aforementioned points, we also demonstrated the variation of R with respect to the weight α when integrating respectively betweenness, H-index, coreness, closeness centrality, eigenvector centrality, and local clustering coefficient metrics with the cycle ratio using the weight α. As shown in Fig. 6, it can be observed that the integration of degree centrality with the cycle ratio outperforms the integration of other metrics with the cycle ratio in terms of node percolation. Similarly, as illustrated in Fig. 7, we examined the variation of R with respect to the weight α when degree centrality was integrated with betweenness, H-index, coreness, closeness centrality, eigenvector centrality, and local clustering coefficient metrics using the weight α, respectively. Again, it is evident that the effectiveness of integrating other metrics with the cycle ratio is indeed less pronounced compared to the integration of degree centrality with the cycle ratio.

In summary, this paper is divided into a cycle-dominant network and a degree-dominant network to discuss the influence of weight α on the dynamic process. (i) There is a nonlinear relationship between the weight α of the DC index and the observation R, and the curve shows a non-monotonic characteristic of first decreasing and then increasing or first decreasing and then flattening, which reveals that the existence of an optimal α of the DC index can cause the network to crash quickly. (ii) Email, Wiki and Hamsterster datasets are similar to cycle-dominant networks, and the optimal α is small ($\alpha < 0.5$); that is to say, the DC index is inclined to identify nodes with high cycle ratio and attacks on these nodes usually destroy the cycle structure in network, so the robustness of the network is quickly weakened. (iii) Chongqing-BUS, Beijing-BUS and Chongqing-BUS are similar to degree-dominant networks, the optimal α is large ($\alpha > 0.5$), and the DC index focuses on the degree of nodes. In this type of network, robustness is very sensitive to high-degree nodes (such

as transportation hubs or large substations), which are often connected to multiple high-traffic sections or subnets, and congestion and power outages can quickly spread throughout the network, leading to congestion or widespread power outages.

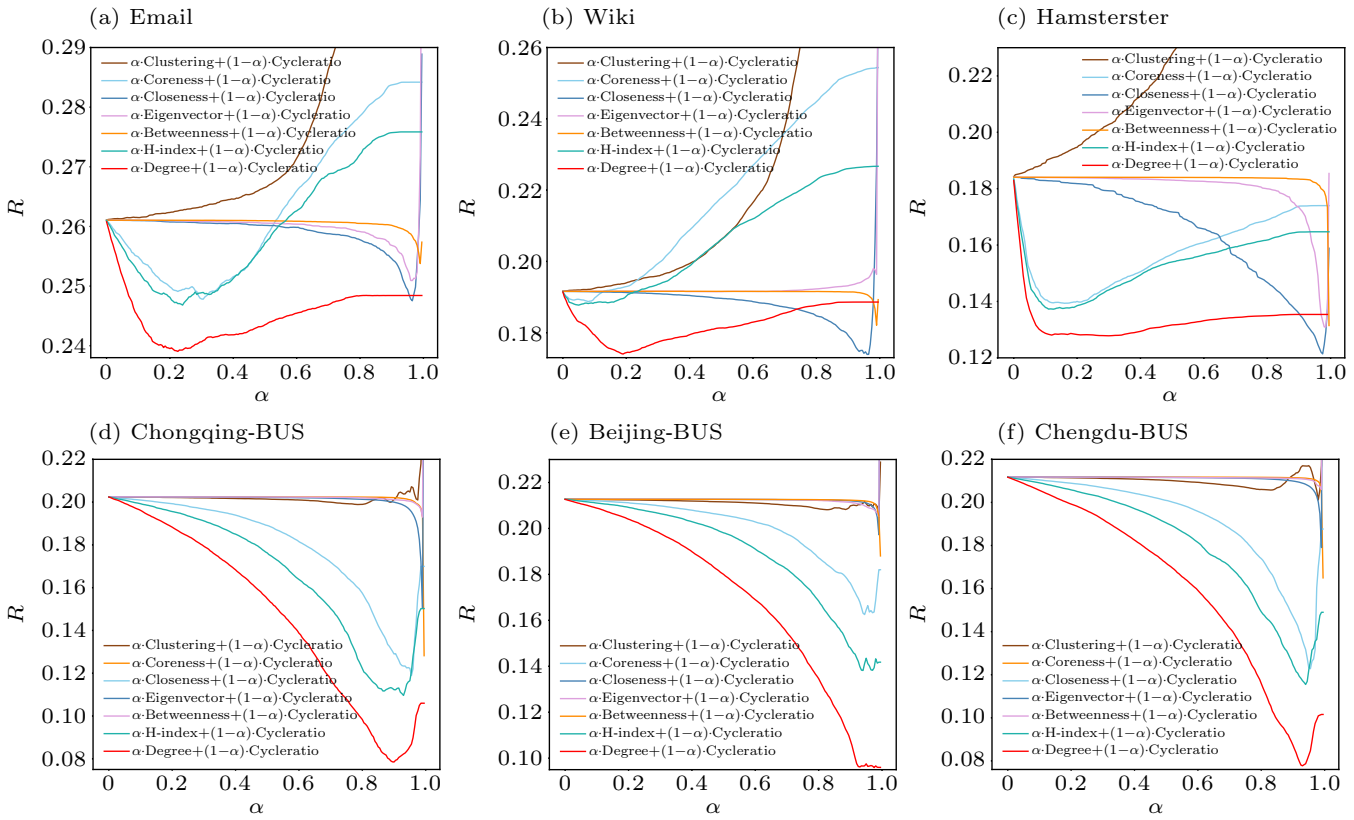


Fig. 6. Cycle ratio integrated with different indicators for robustness analysis of networks, the observation R as a function of the coefficient α.

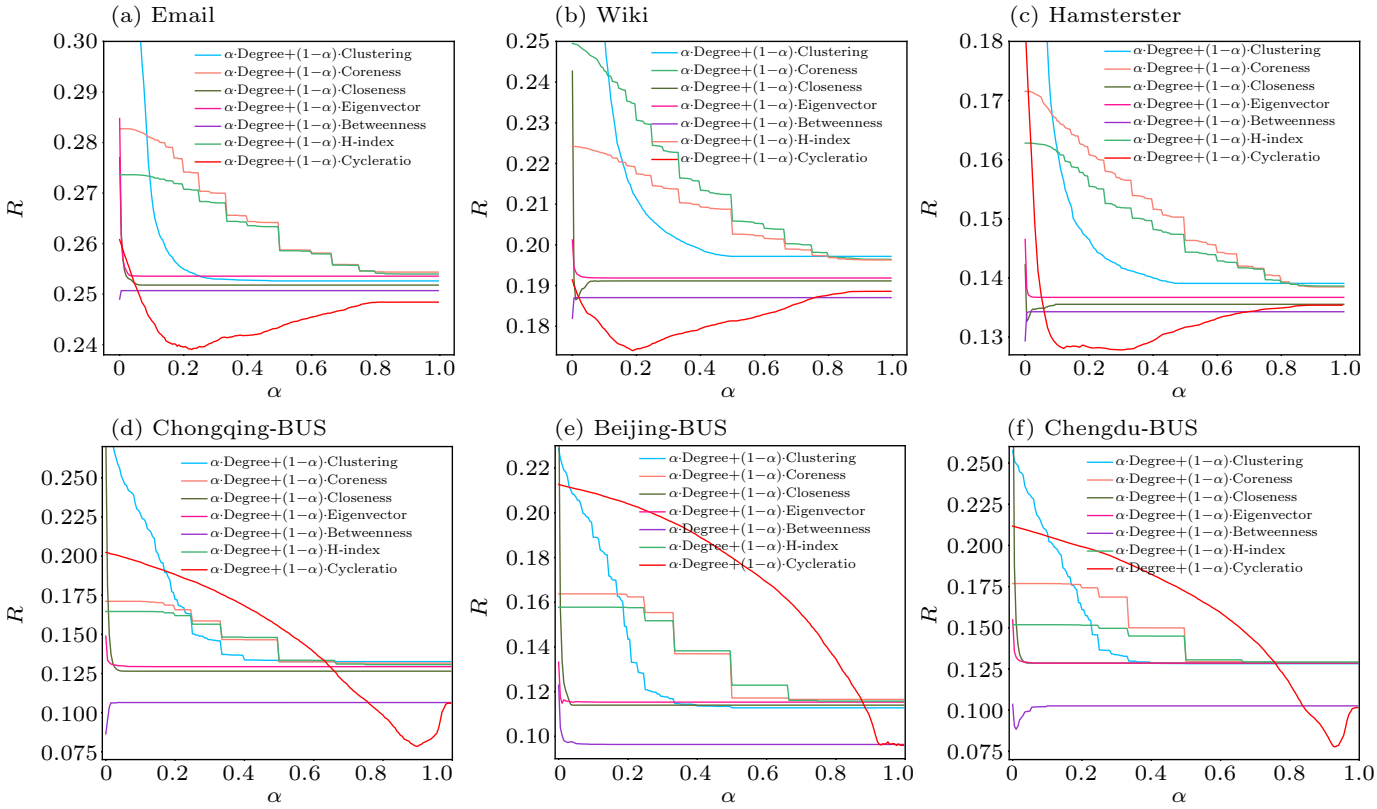


Fig. 7. Degree integrated with different indicators for robustness analysis of networks, the observation R as a function of the coefficient α.

3.2. Epidemic spreading

In the field of epidemic spreading, identifying vital nodes in network is crucial for understanding the spreading mechanism of diseases, predicting epidemic trends, and formulating effective prevention and control strategies.[42,43] In this paper, DC is used to define the propagation centrality of nodes, which is used to rank the propagation influence of nodes. The importance of nodes can be evaluated by their role in the process of disease spreading. To quantify the impact of a selected set of nodes, we simulate the standard susceptibility–infection–recovery (SIR) propagation dynamics, where, at each time step, each susceptible node will be infected by an infected neighbor node with a probability of β, and each infected node will recover with a probability of γ. Initially, the first $0.1N$ nodes selected by each metric are set as infected nodes, and the other nodes are set as susceptible nodes. We consider the following infection rate β [26,27] and recovery rate $\gamma = 1$, where

$$\beta = \frac{\langle k \rangle}{\langle k^2 \rangle - \langle k \rangle}, \tag{5}$$

here $\langle k \rangle$ and $\langle k^2 \rangle$ are the average degree and average squared degree of network, respectively.

In this paper, the cumulative number of infected populations is used to measure the spread influence of nodes. Figure 8 shows how important nodes of six real-world networks are sorted, respectively, under different weights of α for DC indicators, and how the cumulative number of infected populations I_t changes with the time step. Obviously, it can be observed that with different α, the cumulative number of infected groups in the network I_t varies greatly, especially in Figs. 8(b), 8(c), 8(e) and 8(f). To quantify and compare the differences between the curves in Fig. 8, we propose observation I to measure the performance of different metrics affected by infection, which is defined as the cumulative number of infected individuals at a steady state. The formula is as follows:

$$I = \sum_{t=1}^{T_{\max}} I_t, \tag{6}$$

where I_t is the number of infections at time t. I represents the cumulative number of infected nodes at $T_{\max}$. The larger I is, the better the infection effect is.

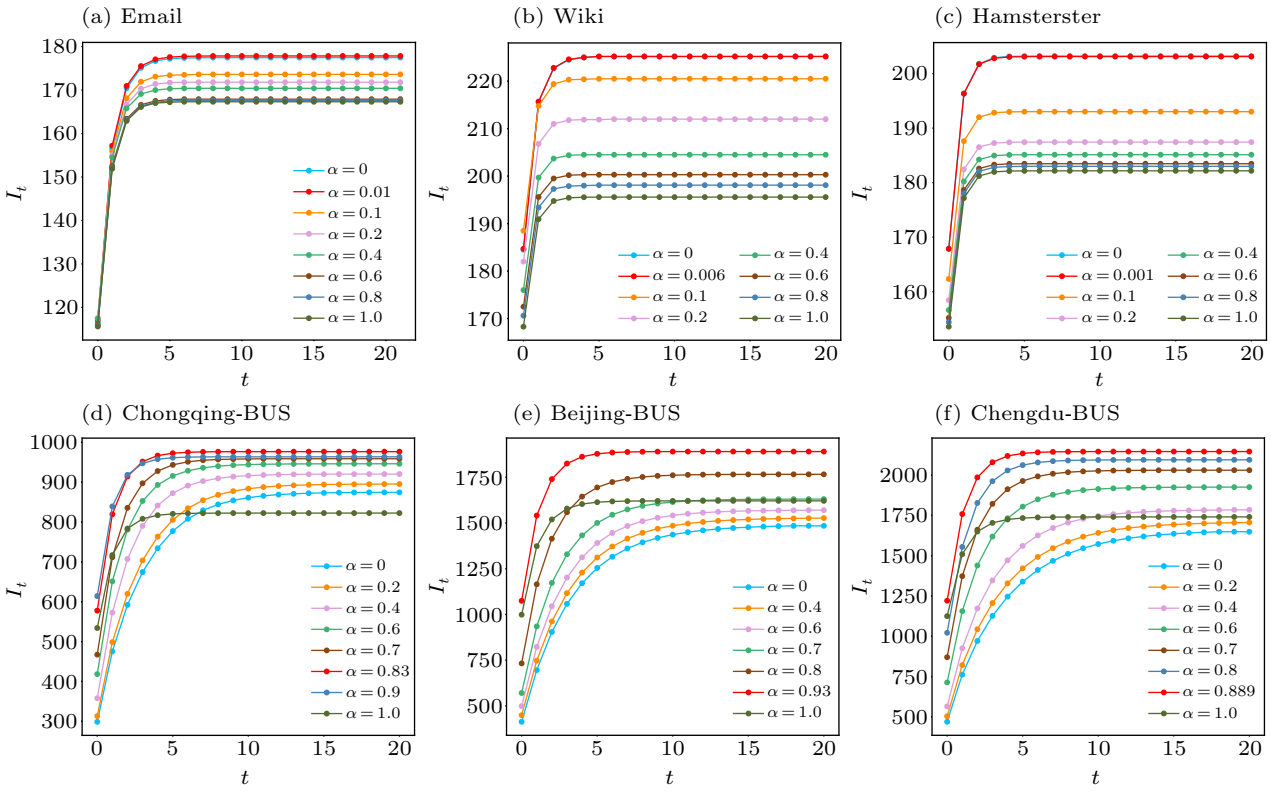


Fig. 8. The SIR propagation dynamics model is simulated in six real-world networks, and based on different α for the DC index, the number of infections I_t as the function of the running times t. The first $0.1N$ nodes under each metric are set as infected nodes, and the other nodes are set as susceptible nodes. The value of infection rate β is shown in formula (5), and the recovery rate $\gamma = 1$.

Firstly, based on the premise that the number of infections and the number of recoveries in network reach a balanced state, this paper explores the change of the cumulative number of infected individuals I under stable state with different coefficients α based on the networks in Table 2. As shown in Figs. 9(a)–9(c), with the increase of α, the cumulative number of infected groups I in the network gradually decreases to a flat level. The more cumulative infections, the more important the

node is. We find that the optimal value of α is generally less than 0.1, which indicates that degree plays a very weak role in the ranking of important nodes based on the DC method, and the cycle ratio predominates the ranking of important nodes to a large extent.

As shown in Figs. 9(d)–9(f), with the increase of α, the cumulative number of infected groups I of the network shows a trend of first increasing and then decreasing. The

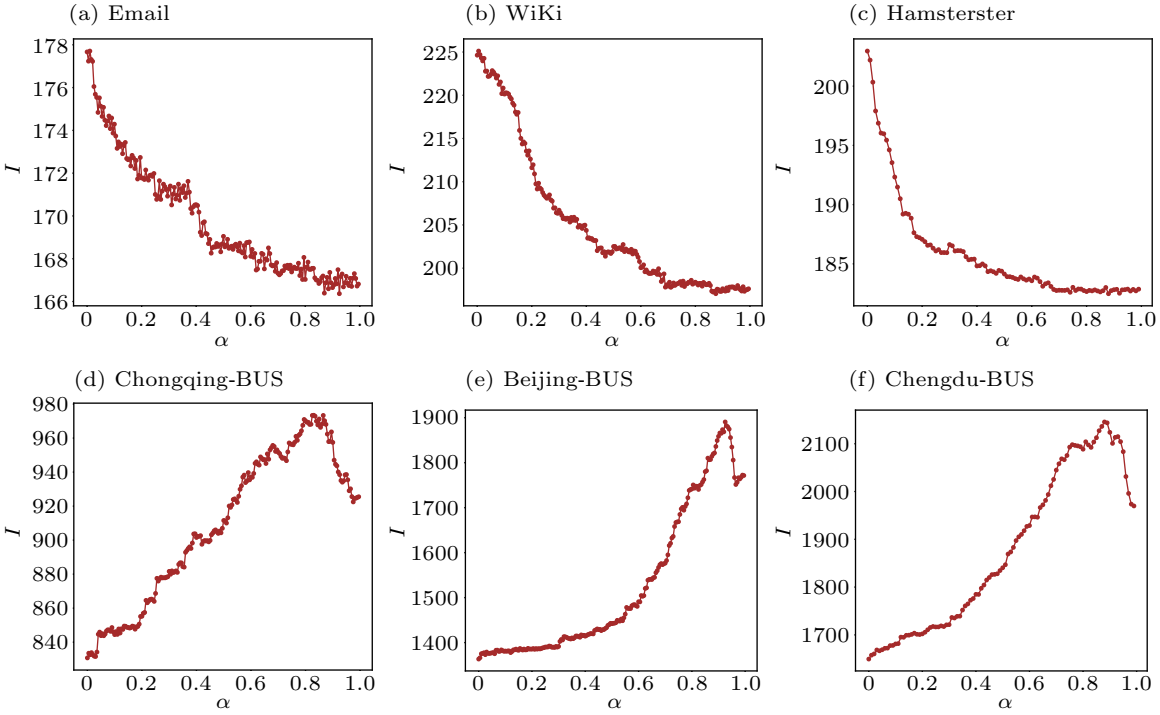


Fig. 9. The SIR propagation dynamics model is simulated in networks, and the observation I is a function of the coefficient α. The first $0.1N$ nodes under the DC index are set as infected nodes, and the other nodes are set as susceptible nodes. The value of infection rate β is shown in formula (5), and the recovery rate $\gamma = 1$.

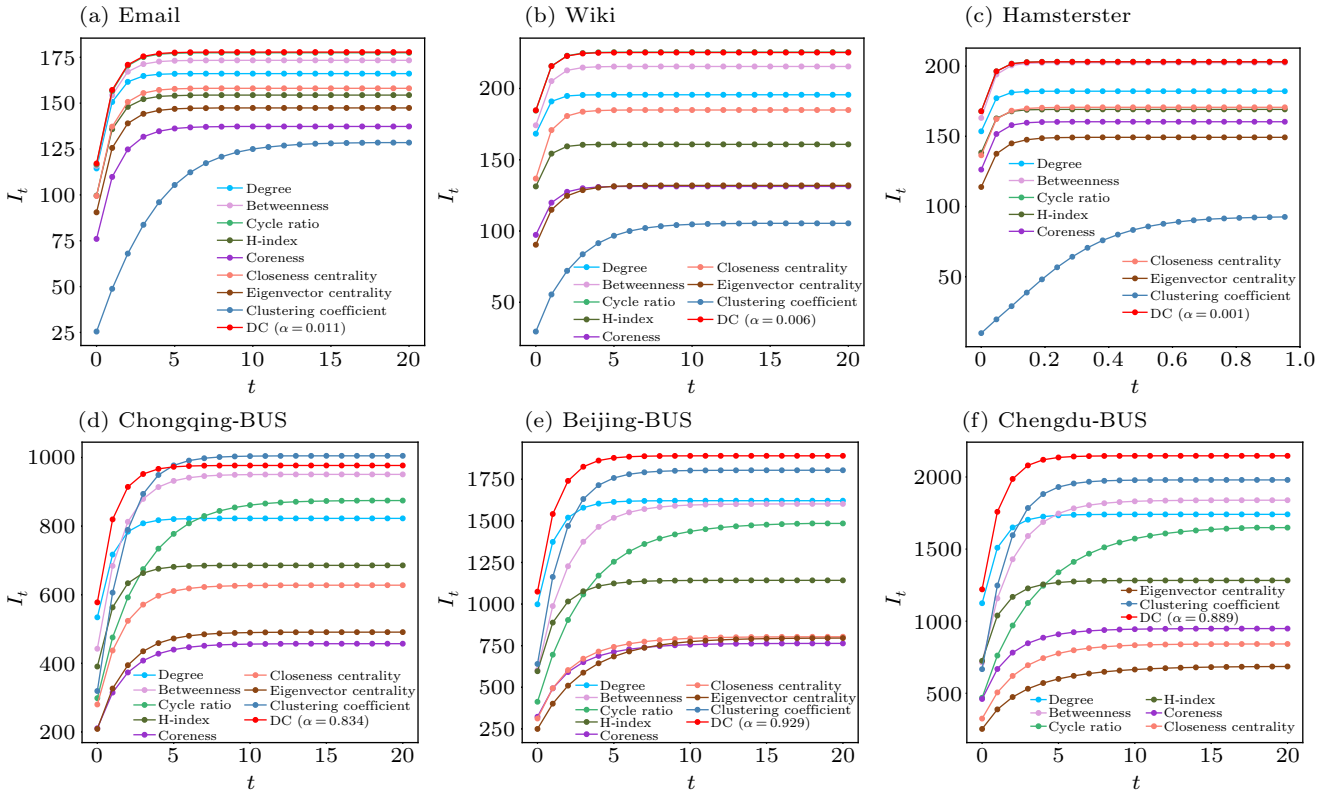


Fig. 10. The SIR propagation dynamics model is simulated in networks, and the number of infections I_t as the function of the running times t based on different metrics.

nonlinear characteristics of the curve make DC have the optimal weight α, and when α is larger, the propagation effect is maximized, and the node degree occupies the dominant advantage in the identification of vital nodes. Take α corresponding to the highest I as the weight of the DC method. Using degree, betweenness, cycle ratio, H-index, coreness, closeness centrality, eigenvector centrality, local clustering coefficient, and DC influence indicators reflects the importance and influence of nodes in the network, so as to provide a basis for identifying vital nodes.

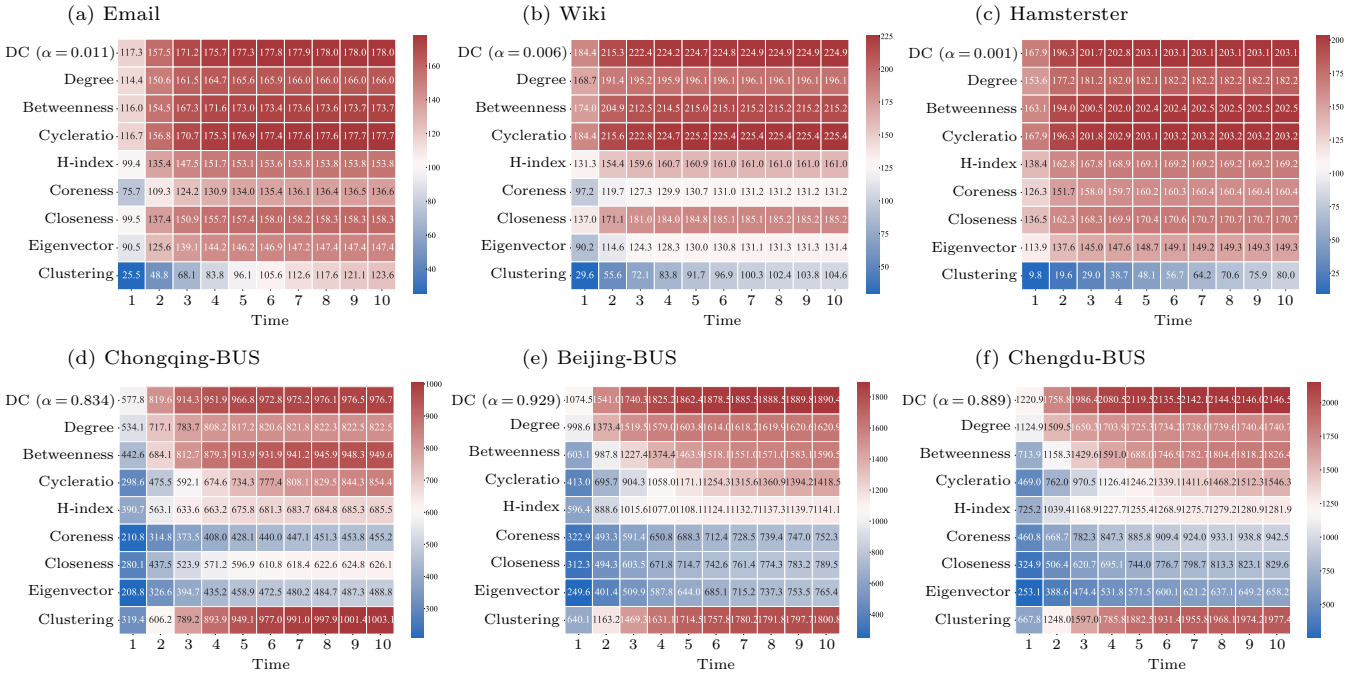


Fig. 11. The SIR propagation dynamics model is simulated in six real-world networks. The elements in each matrix are the cumulative number of infected nodes for nine indicators in the corresponding time step, and these values are represented by color: the deeper the red, the better.

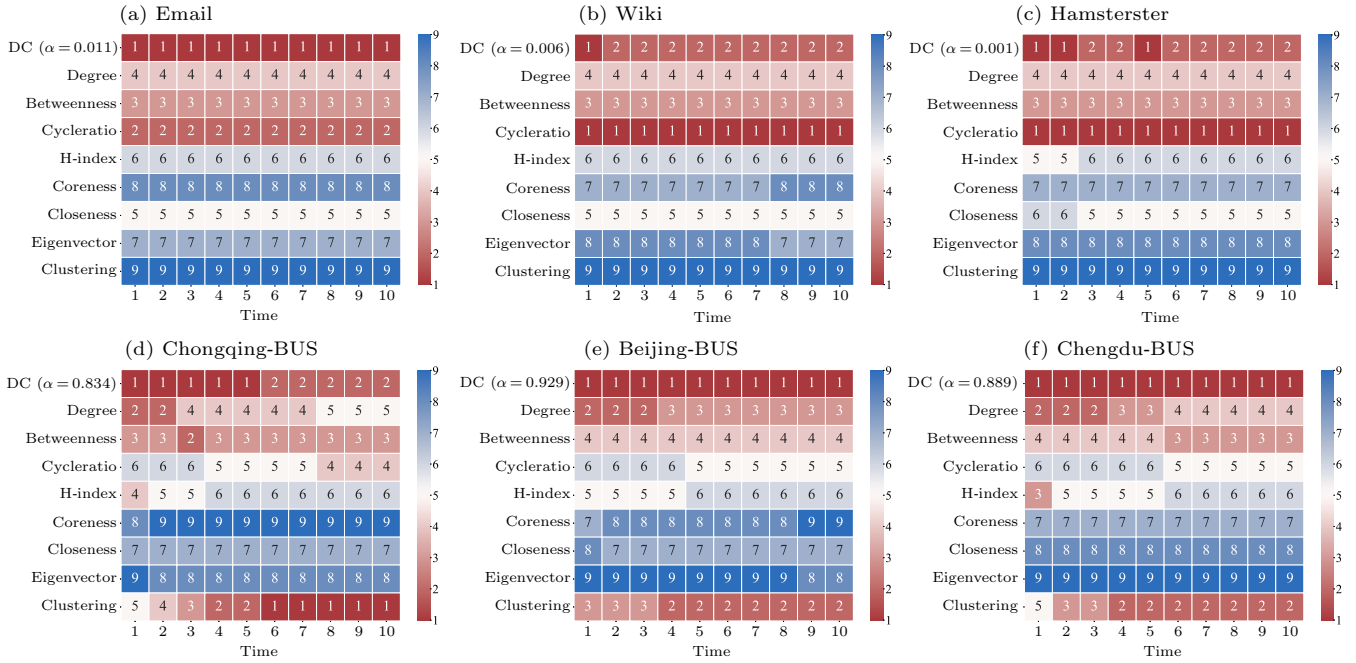


Fig. 12. Nine indicators characterize the performance of propagation dynamics on real-world networks. The elements in each matrix are the ranking of the nine indicators in the corresponding time step, determined by the cumulative infected nodes of each indicator in the SIR model simulation. The index with the largest number of infected nodes is number 1, and the other indexes are numbers 2, 3, 4, 5, 6, 7, 8 and 9. Each ranking is averaged over 8000 individual runs and is displayed by color: the deeper the red, the better.

Figure 10 shows the cumulative number of infected persons at different time steps t for nine indicators, which are averaged over 6000 independent runs. When the time step t gradually increases, the node propagation ability of each index gradually increases. In Figs. 10(a)–10(c), the effect of cycle ratio and DC index is the same, followed by betweenness,

and the local clustering coefficient has the worst propagation effect. This indicates that in the Email, Wiki and Hamsterster datasets, the vital nodes identified by the DC index have little improvement in disease transmission, while the vital nodes identified by the cycle ratio have reached the optimal effect in SIR transmission. It is probably because these networks have an obvious community structure and a large number of small cycles, and nodes with a high cycle ratio are usually located in multiple cycles, which are easier to form transmission links with other nodes, and can quickly spread diseases or information to multiple susceptible people, thus accelerating the spread of diseases or information.

In Figs. 10(d)–10(f), nodes have a good propagation ability under the DC index or local clustering coefficient. It may be because in Chongqing-BUS, Beijing-BUS and Chengdu-BUS, high nodes (such as a transportation hub or a large substation) are often connected to multiple high-traffic sections or subnet cores, and changes in their state (such as congestion, closure) can quickly propagate throughout the network, leading to congestion or widespread power outages. However, the nodes with a high cycle ratio are distributed in different layers or modules of the network. Because of their redundancy, they are connected to multiple cycles in the network, and their influence is relatively dispersed, and it is not easy to form a global fault effect.

Figure 11 shows the cumulative number of infected populations in the SIR model in the first 10 time steps under different node importance influence indicators and at different transmission times. The cumulative number of the first $0.1N$ nodes selected by each metric in terms of β infection probability. Figure 12 reports the ranking of the nine metrics at the time steps $t = 1$, $t = 2$, $t = 3$, $t = 4$, $t = 5$, $t = 6$, $t = 7$, $t = 8$, $t = 9$ and $t = 10$, averaged over 8000 independent runs. The index with the best performance is ranked first, followed by the second, …, the worst indicator was ninth. It can be seen from the results that DC ranking is always at the top, significantly better than other indicators.

In summary, in epidemic spreading, there is a non-monotonicity relationship between the cumulative number of infected individuals at steady state I and the weight α of the DC index, which reveals that there is an optimal weight α of DC, which makes the infection ability reach the maximum. (i) In the Email, Wiki and Hamsterster networks, the relationship curve between the cumulative number of infected individuals at steady state I and the weight α shows a trend of first decreasing and then stabilizing, and the optimal value of the weight α is generally small ($\alpha < 0.5$). The results show that the propagation ability of vital nodes identified by the DC index is similar to that of the cycle ratio, and the improvement is not obvious. (ii) In the Chongqing-BUS, Beijing-BUS and Chengdu-BUS networks, the relationship curve between the cumulative number of infected individuals at steady state I and the weight α presents a trend of first rising and then declining, and the optimal value of the weight α is generally large ($\alpha > 0.5$). The results show that the propagation ability of vital nodes identified by the DC index is significantly improved compared with other metrics. (iii) Therefore, it is found that for epidemic spreading, in most subways networks, the DC method can accurately identify the most influential spreaders and maximize information dissemination in a short period of time.

3.3. Targeted immunization

The selection of appropriate immunization strategies is extremely important for the prevention and control of infectious diseases. Targeted immunization of vital nodes through the SIS model can quickly contain infection.[28] Targeted immunization hopes to achieve the best possible immunization effect by selectively immunizing a small number of nodes. Once these nodes are immunized, it means that their connected edges can be removed from network, greatly reducing the possible connection routes for disease spreading.[29,30] This paper identifies vital nodes in network based on targeted immunization and proves that disease immunization is an effective method to measure node centrality.

If each susceptible node is connected to one or more infected nodes, its infection rate is β. Infected nodes are cured and become susceptible nodes again with rate γ. This paper does not consider the rate of an individual being removed due to death or acquired immunization, so the individual runs randomly in the cycle of susceptibility $\longrightarrow$ infection $\longrightarrow$ susceptibility, which is the SIS propagation model. The simulation is implemented under a fixed spreading rate of 0.25, and the rate of an infected node recovering to a susceptible node $\gamma = 0.5$. We first infect half of nodes in network, and then update the rules of SIS model in parallel to analyze the stationary characteristics of the infected node density ρ_{p} under different immunization densities P.

Figure 13 shows the sorting of vital nodes by six real-world networks under different weights of α for DC indicators, and the changes of ρ_{p}/ρ_0 with immune density P. In Figs. 13(a)–13(c), the difference between curves is small under different α. In Figs. 13(d)–13(f), there is a big difference between the curves under different α. To quantify and compare the differences between the curves in Fig. 13, we propose observation ψ to describe the performance of metrics affected by immunization. The smaller the ψ, the better the immunization effect and the faster the elimination of disease spreading is. The formula is

$$\psi = \sum_{i \in P} \frac{\rho_i}{\rho_0 N_P}, \tag{7}$$

where ρ_0 is the steady-state infection density without immunization, N_P represents the number of equally spaced intervals

taken by the horizontal coordinate, and ρ_i is the steady-state infection density after immunization of nodes with the proportion $i \in P$ in network. It describes the area formed by the curve of ρ_P/ρ_0 as a function of P and the x-axis and y-axis.

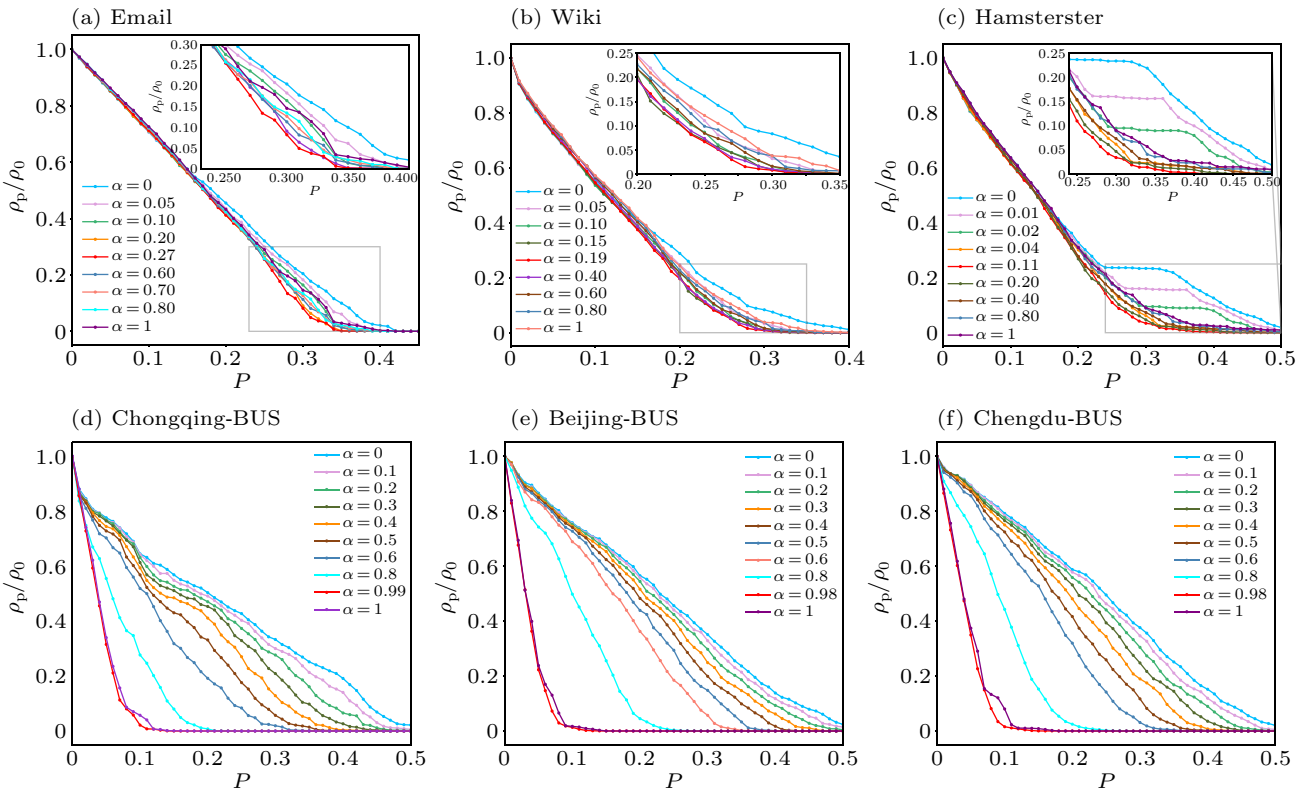


Fig. 13. The SIS propagation dynamics model is simulated in networks according to DC metrics for different α, the steady-state infection density ρ_p/ρ_0 is a function of the immunization density P. The first $P \cdot N$ nodes selected based on DC are set as immune nodes, then half of the remaining nodes are randomly selected as infected nodes, and the other part is set as susceptible nodes. Infection rate $\beta = 0.25$, the recovery rate $\gamma = 0.5$.

Based on the premise that the steady-state infection density reaches 0, this paper explores the changes of ψ in a steady state under different weight α based on the networks in Table 2. As shown in Figs. 14(a)–14(c), with the increase of α, the observation ψ of most networks shows a trend of first decreasing and then increasing. This non-monotonicity characteristic proves that there is an optimal value of α, ensuring that the extent of immunization of the vital nodes identified by DC is maximized. As shown in Figs. 14(d)–14(f), with the increase of α, the observation ψ of the network shows a trend of decreasing first and then stabilizing. This nonlinear relationship shows that there is an optimal value for the weight α, and immunity achieves the best effect when α is larger ($\alpha > 0.5$). We take the α corresponding to the bottom of the curve as the weight of DC.

As shown in Fig. 15, under different immune indexes, steady-state infection density gradually decreased with the increase of immune density P. In Figs. 15(a)–15(c), the steady-state infection density decreased faster and the corresponding immune density threshold was smaller than other indicators based on the immunization strategy of identifying important nodes in DC. The immune effect of local clustering coefficient was the worst. These results show that the vital nodes identified by the DC indicators have a strong effect on the immune prevention and control of the cycle-dominant network, and can eliminate the spread of the virus faster than other influence indicators. In Figs. 15(d)–15(f), the immune effect of DC is the best, the speed of eliminating disease transmission is faster, and the immune density critical value is smaller. The second is degree, and the index of cycle ratio has the worst effect. This means that in the degree-dominant network, the vital nodes identified according to the cycle ratio are not effective in controlling information dissemination and solving congestion. However, only a smaller number of nodes need to be immunized to rapidly control the spread of the epidemic and ensure the maximum immunization effect under limited resources.

In summary, the non-monotonicity relationship between the weight α of the DC index and the observation ψ reveals that there is an optimal α in DC, so that the network can minimize the speed and scope of disease spreading. (i) Generally, for Email, Wiki and Hamsterster networks, the optimal α is small ($\alpha < 0.5$); that is, the weight of the cycle ratio is large, and DC is inclined to identify high cycle ratio nodes connecting different communities in high-clustering networks, which are usually located in the center of tightly connected clusters. Immunizing these nodes can significantly reduce the speed and extent of the spread of disease, rumors, etc. (ii) Generally, for Chongqing-BUS, Beijing-BUS and Chengdu-BUS, the opti-

mal α is large ($\alpha > 0.5$), that is, the DC index is focused on the node degree. Protected high-degree nodes (large interchange stations, transportation hubs, power stations or substations) can ensure the normal operation of traffic flow during peak periods or emergencies. (iii) In conclusion, for targeted immunization, the DC method shows an effective effect in cycle-dominant networks, and also proves that observation ψ is an effective indicator to measure node centrality.

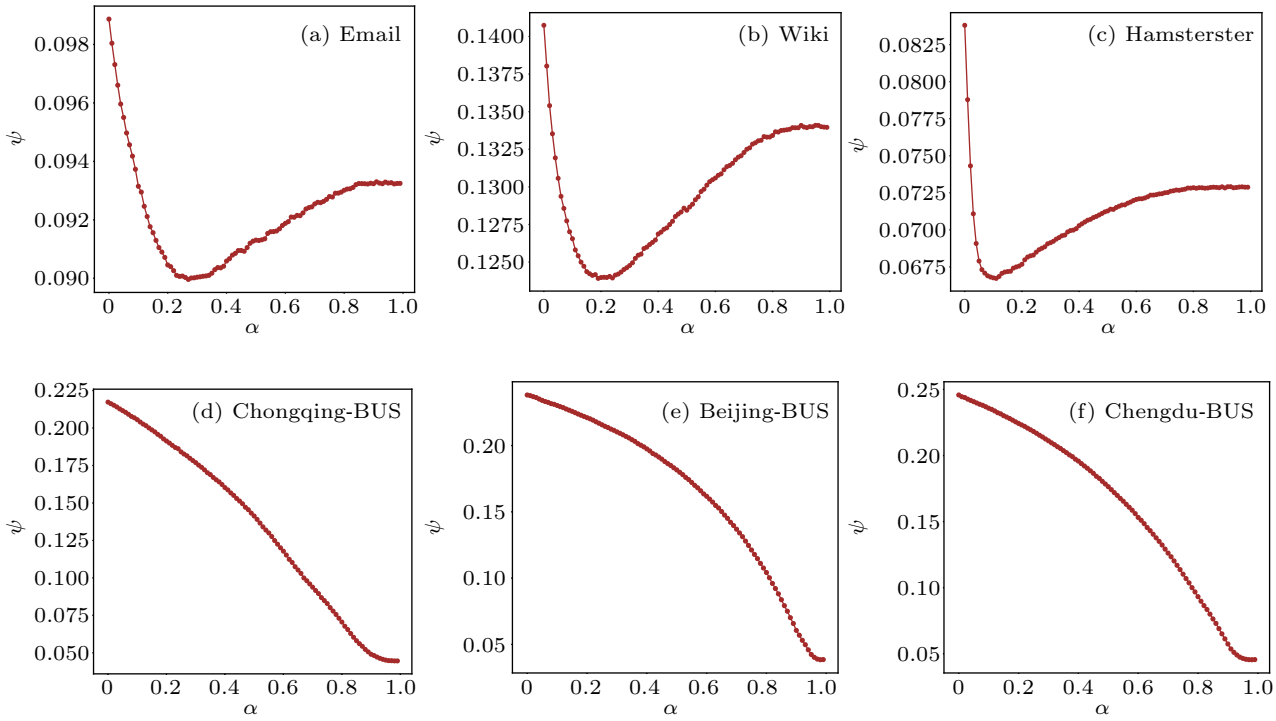


Fig. 14. The SIS propagation dynamics model is simulated in networks. Based on DC metric, the observation ψ is a function of the weight α.

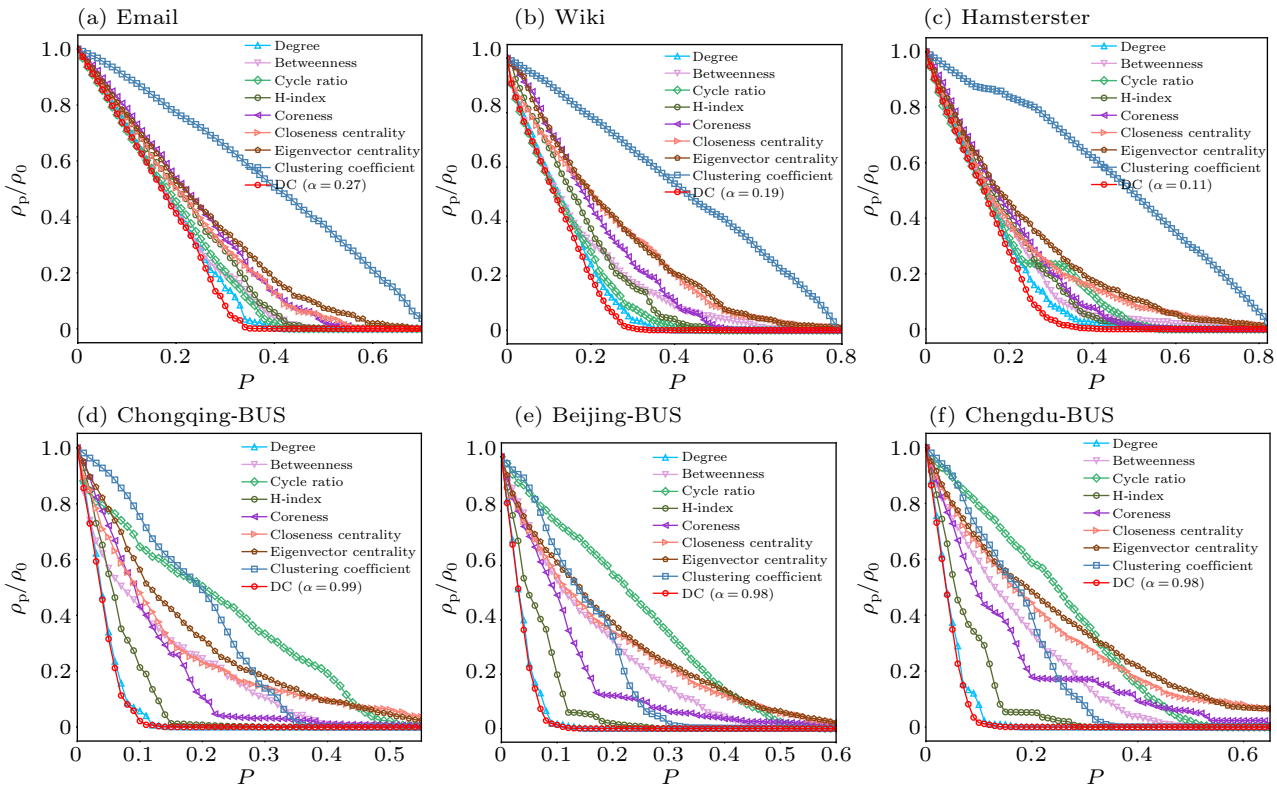


Fig. 15. The SIS propagation dynamics model is simulated in networks. Based on different metrics, the steady-state infection density ρ_P/ρ_0 is a function of the immunization density P.

3.4. LFR synthetic network

Lancichinetti *et al.*[44] introduced a class of synthetic networks with a power-law distribution of both degree distribution and community size distribution, i.e., LFR benchmark, including the following parameters: the number of nodes N; the power law exponent for the degree distribution v; the power law exponent for the community size distribution δ; the mixing parameter μ; the average degree $\langle k \rangle$. The network construction steps are as follows: first, the LFR network assigns degrees to the nodes using a power-law distribution with an exponent of v, with an average degree of $\langle k \rangle$. Then, a configuration model is used to connect the nodes to maintain the degree order of the nodes. Select the community size using a power-law distribution with an exponent of δ so that the sum of all community sizes equals the number of nodes N of the graph, and ensure that nodes of any degree belong to at least one community. At the same time, the location of the nodes within the community is adjusted through the iterative process and rewiring steps, so that each node shares a part of the link $1-\mu$ with other nodes in its community and a part of the link μ with other nodes in the network. Such networks more closely resemble real-world networks with community structures.

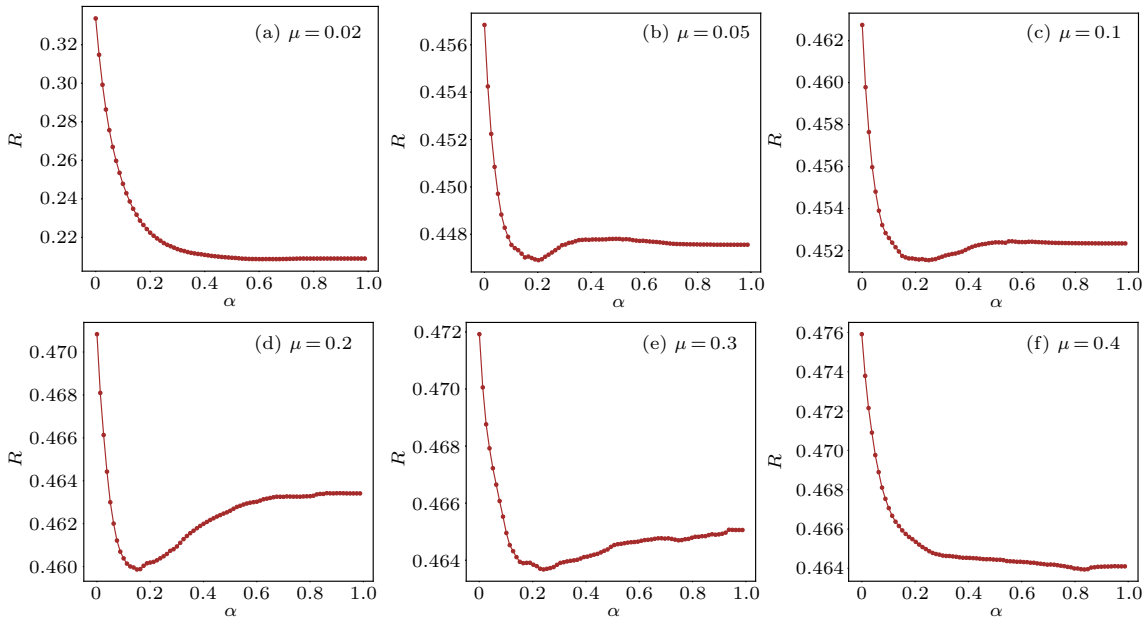


Fig. 16. The observation R as a function of the coefficient α for LFR networks. The curve is the result of averaging 50 LFR networks. $N = 3000$, $v = 3$, $\delta = 1.5$, $\langle k \rangle = 20$.

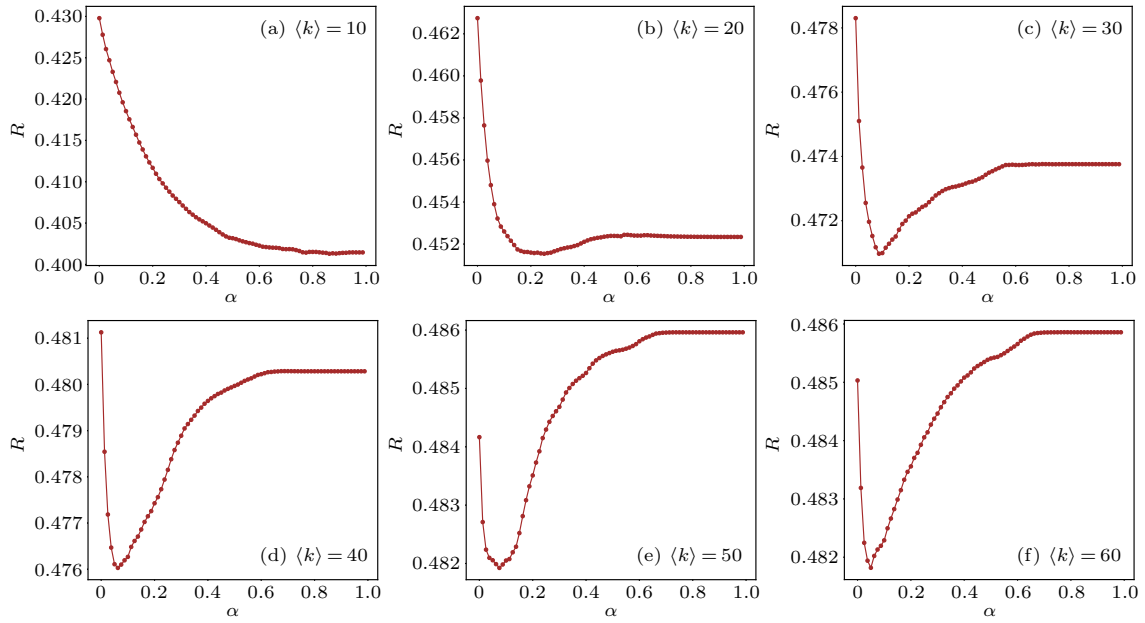


Fig. 17. The observation R as a function of the coefficient α for LFR networks. The curve is the result of averaging 50 LFR networks. $N = 3000$, $v = 3$, $\delta = 1.5$, $\mu = 0.1$.

We discuss the non-monotonicity relationship between the weight α and R on the LFR networks. First setting $N = 3000$, $v = 3$, $\delta = 1.5$, $\langle k \rangle = 20$, and 50 LFR networks are randomly generated for testing under each μ. As shown in Fig. 16, the observation R varies with the weight α when nodes are sorted according to the DC index at $\mu = 0.02, 0.05, 0.1, 0.2, 0.3, 0.4$. It is found that the existence of optimal α can only be observed in a certain interval of μ ($0.05 < \mu < 0.3$).

Next, we set $N = 3000$, $v = 3$, $\delta = 1.5$, $\mu = 0.1$, and 50 LFR networks are randomly generated for testing under each $\langle k \rangle$. As shown in Fig. 17, the observation R varies with the weight α when nodes are sorted according to the DC index at $\langle k \rangle = 10, 20, 30, 40, 50, 60$. With the gradual increase of $\langle k \rangle$, the non-monotonicity of the curve becomes stronger and stronger, and the optimal α gradually decreases.

4. Conclusion

In this paper, combining degree (weight α) and cycle ratio (weight $1 - \alpha$), a new vital node identification method, the DC method (degree and cycle ratio), is proposed to quantify the importance of a single node. This method not only makes up for the defect of vital node identification under a single degree index, but also overcomes the limitation caused by the omission of cycle-based nodes without direct correlation. In this paper, networks are divided into cycle-dominant networks and degree-dominant networks, and three dynamic processes, node percolation, epidemic spreading, and targeted immunization, are used to test the effectiveness of the DC method.

First, in node percolation, there is a non-monotonicity relationship between the observation R and the weight α of the DC index, which reveals that the existence of the optimal α of the DC index can cause the network to collapse quickly. In general, for cycle-dominant networks, the optimal α is small ($\alpha < 0.5$), and the effect of the cycle ratio is more obvious, which occupies a dominant position in identifying vital nodes. For degree-dominant networks, the optimal α is large ($\alpha > 0.5$), the degree value effect is significant, and the DC index focuses on the degree of nodes.

Secondly, in epidemic spreading, there is a non-monotonicity relationship between the cumulative number of infected individuals at steady state I and the weight α of the DC index, which reveals that there is an optimal weight α of DC index to maximize the infection capacity. Generally, for cycle-dominant networks, the relationship curve between the cumulative number of infected individuals I and α shows a trend of first decreasing and then stabilizing, and the optimal value of α is generally small ($\alpha < 0.5$). The propagation ability of vital nodes identified by the DC index is similar to the effect of cycle ratio, and the improvement is not obvious. While, for degree-dominant networks, the relationship curve between the cumulative number of infected individuals at steady state I and α shows a trend of first rising and then declining, and the optimal value of α is generally large ($\alpha > 0.5$). The propagation ability of vital nodes identified by the DC index is significantly improved compared with the other eight indicators.

Next, in targeted immunization, the non-monotonicity relationship between the weight α of DC index and the observation ψ reveals that there is an optimal α of DC so that the network can minimize the speed and range of disease spreading. Generally, for cycle-dominant networks, the optimal α is small ($\alpha < 0.5$), and the cycle ratio is dominant in identifying vital nodes. The immunization effect of vital nodes identified by the DC index was significantly improved compared with the other eight indicators. However, for degree-dominant networks, the optimal α is large ($\alpha > 0.5$), and the degree is dominant in identifying vital nodes. The propagation ability of vital nodes identified by the DC index is similar to the effect of cycle ratio, and the improvement is not obvious.

Finally, this paper uses the LFR synthetic network model to explore the relationship between the weight α and observation R, and finds that the change of mixed parameters μ or $\langle k \rangle$ has a great influence on the non-monotonicity of R vs. α curve, and the optimal α is usually small.

In conclusion, this paper uses three dynamic processes, namely, node percolation, epidemic spreading, and targeted immunization, to test the effectiveness of the DC vital node identification method, and verifies that, in most networks, DC performs better than a single metric.

Appendix A: Node percolation

In order to test the effect of the vital nodes identified by DC on more networks, we studied more datasets, as shown below. The cycle-dominant network includes: Diseasome (Human disease network), KKI (Label network),[37] Crime (Criminal network),[45] PolBlogs (Political blog network),[46] PH (Petster-Hamster online social network),[45] Socfb-Reed98 (Facebook network), Collaration (Collaborative Network of Physicists),[40] Grid-plant (Biological Networks), Wm2 (Economic network), Pages-food and Pages-tvshow (Facebook social network), and Corecipient (Democratic email corecipient network).[37] The degree-dominant network includes: the power grid[37] (Grid, Power-bcspwr09, Power-bcspwr10 and Power-662-BUS), road network (Beijing-Gaosu (Beijing-highway), euroroad (European road network, Road-Minnesota (a road network of Minnesota)) and subway network (Chongqing-SUB, Chengdu-SUB, Beijing-SUB, Shanghai-SUB, Shenzhen-SUB). Their basic topological features are shown in Table A1.

Table A1. The topological properties of networks. Average degree $\langle k\rangle$; average path length $\langle L\rangle$; average clustering coefficient $\langle C\rangle$.

Network	N	E	$\langle k\rangle$	$\langle L\rangle$	$\langle C\rangle$	Network	N	E	$\langle k\rangle$	$\langle L\rangle$	$\langle C\rangle$
Diseasome	516	1188	4.60	6.51	0.64	Grid	4941	6594	2.67	18.99	0.08
KKI	2238	4019	3.59	6.35	0.39	Power-bcspwr09	1723	4117	4.78	5.49	0.076
Crime	754	2127	5.64	6.82	0.73	Power-bcspwr10	5300	8271	5.12	20.85	0.09
PolBlogs	1490	16715	22.44	2.74	0.26	Power-662-BUS	662	1568	4.74	10.24	0.046
PH	2000	16098	16.10	3.59	0.54	Beijing-highway	355	498	2.81	11.99	0.079
Socfb-Reed98	962	18812	39.11	2.46	0.32	Euroroad	1174	1417	2.41	18.39	0.016
Collaration	4158	13422	6.46	6.05	0.56	Road-minnesota	2642	3303	2.50	35.35	0.02
Grid-plant	1745	3098	3.61	8.29	0.15	Chongqing-SUB	226	244	2.16	15.63	0.003
Wm2	260	2942	19.56	2.62	0.25	Chengdu-SUB	334	375	2.25	18.98	0.001
Pages-food	620	2102	6.78	5.09	0.33	Beijing-SUB	342	388	2.27	16.29	0.014
Pages-tvshow	3892	17262	8.87	6.28	0.37	Shanghai-SUB	419	495	2.36	15.85	0.013
Corecipient	906	10429	23.02	2.76	0.49	Shenzhen-SUB	240	275	2.29	13.70	0.004

A1. Cycle-dominant network

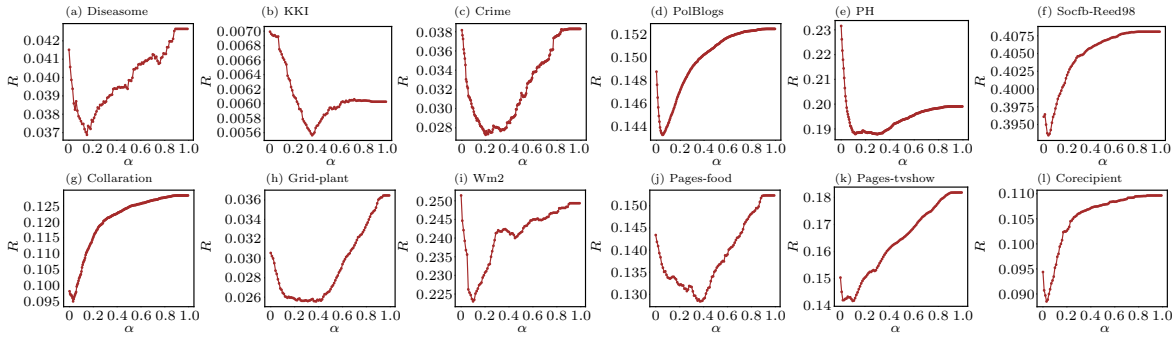


Fig. A1. The observation R as a function of the coefficient α for cycle-dominant networks.

A2. Degree-dominant network

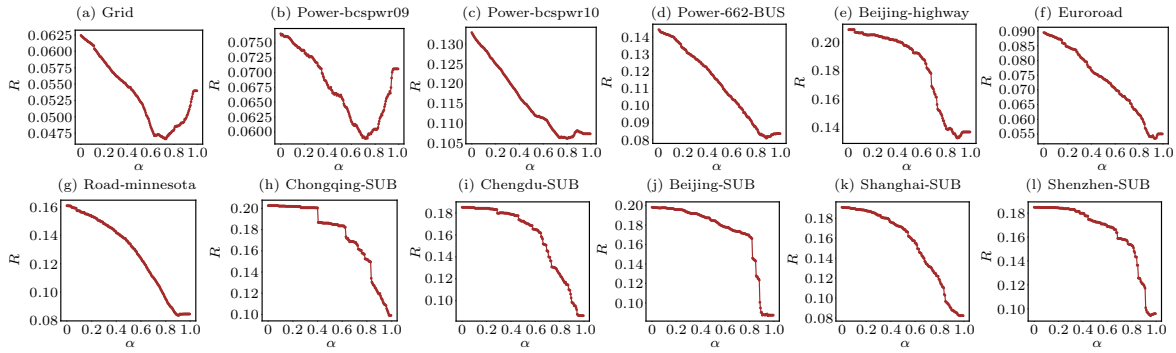


Fig. A2. The observation R as a function of the coefficient α for degree-dominant networks.

Appendix B: Epidemic spreading

B1. Cycle-dominant network

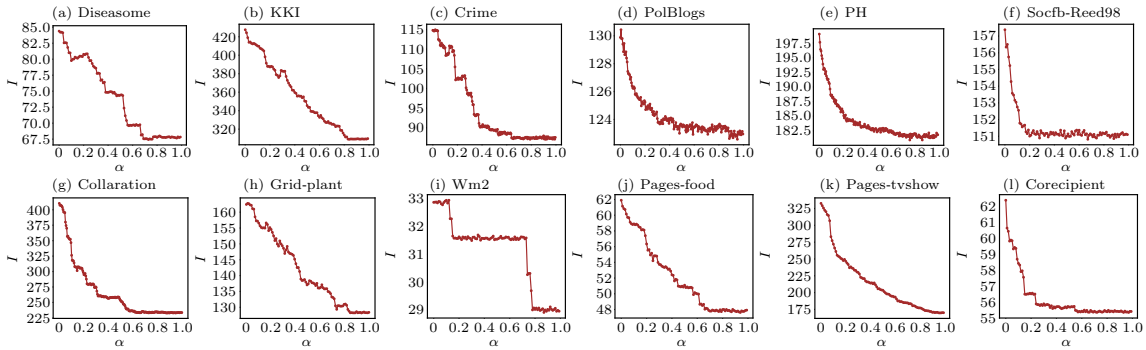


Fig. B1. The SIR propagation dynamics model is simulated in cycle-dominant networks, and the observation I is a function of the coefficient α. The first $0.1N$ nodes under the DC index are set as infected nodes, and the other nodes are set as susceptible nodes. The value of infection rate β is shown in formula (5), and the recovery rate $\gamma = 1$.

B2. Degree-dominant network

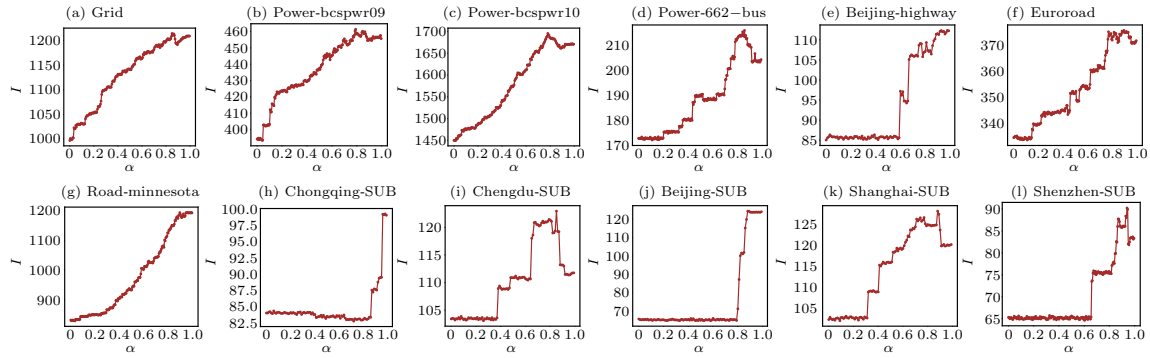


Fig. B2. The SIR propagation dynamics model is simulated in degree-dominant networks, and the observation I is a function of the coefficient α.

Appendix C: Targeted immunization

C1. Cycle-dominant network

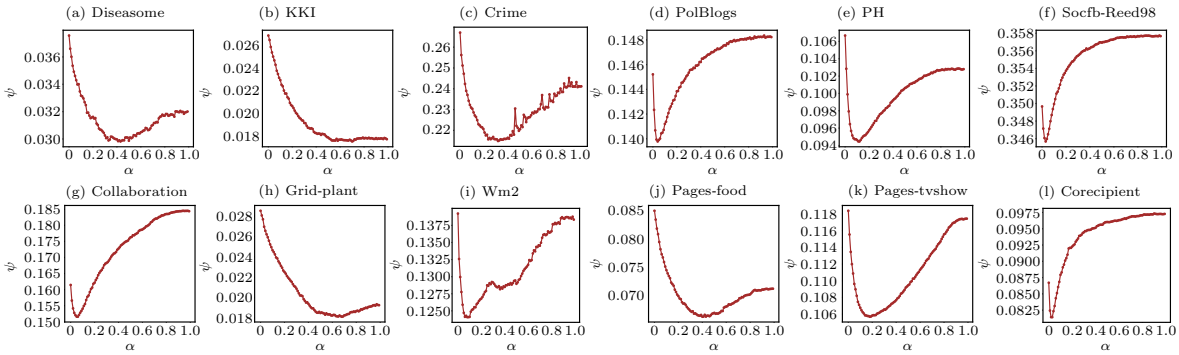


Fig. C1. The SIS propagation dynamics model is simulated in cycle-dominant networks. Based on DC metric, the observation ψ is a function of the coefficient α. Infection rate $\beta = 0.25$, the recovery rate $\gamma = 0.5$.

C2. Degree-dominant network

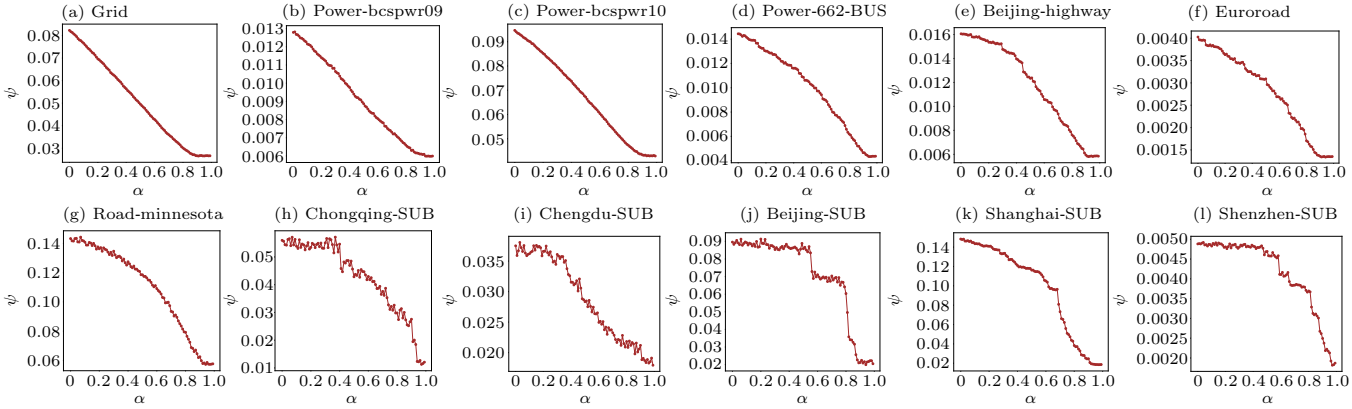


Fig. C2. The SIS propagation dynamics model is simulated in degree-dominant networks. Based on DC metric, the observation ψ is a function of the weight α.

Program availability

The code for the program has been uploaded to GitHub at https://github.com/kust-yangbo/Degree-and-Cycle-ratio and is also openly available in Science Data Bank at https://www.doi.org/10.57760/sciencedb.j00113.00158, which are freely available to academic users. The program was written based on the windows operating platform, which is commonly used for Python software, making it easy to use with other software. For more detail, please read the README.md file in the Degree-and-Cycle-ratio program.

Acknowledgment

Project supported by Yunnan Fundamental Research Projects (Grant No. 202401AT070359).

References

[1] Lü L Y, Chen D C, Ren X L, Zhang Q M, Zhang Y C and Zhou T 2016 *Physics Reports* **650** 1

[2] Huang Y M, Wang H, Ren X L and Lü L Y 2024 *Commun. Phys.* **7** 19
[3] Zhao N, Wang H, Wen J J, L J, Jing M and Wang J 2023 *New J. Phys.* **25** 083020
[4] Ma P J, Ren X Z, Zhu J F and Jang Y Q 2024 *Chin. Phys. B* **33** 088903
[5] Yang Y Z, Hu M and Huang T Y 2020 *Chin. Phys. B* **29** 088903
[6] Kempe D, Kleinberg J and Tardos é 2003 *Proc. of Acm Sigkdd Intl Conf. on Knowledge Discovery & Data Mining* **11** 105
[7] Kabir K A, Kuga K and Tanimoto J 2020 *Chaos Soliton. Fract.* **132** 109548
[8] Barabási A L and Albert R 1999 *Science* **286** 509
[9] Freeman L M 1977 *Sociometry* **40** 35
[10] Sabidussi G 1966 *Psychometrika* **31** 581
[11] Chen D B, Lü L Y, Shang M S, Zhang Y C and Zhou T 2012 *Physica A* **391** 1777
[12] Kitsak M, Gallos L K, Havlin S, Liljeros F, Muchnik L, Stanley H E and Makse H A 2010 *Nat. Phys.* **6** 888
[13] Dorogovtsev S N, Goltsev A V and Mendes J F F 2006 *Phys. Rev. Lett.* **96** 40601
[14] Fan T L, Li H, Ren X L, Xu A Q, Gou Y Z and Lü L Y 2021 *Int. J. Mod. Phys. C* **32** 2150121
[15] Lü L Y, Zhou T, Zhang Q M and Stanley H E 2015 *Nat. Commun.* **7** 10168
[16] Brin S and Page L 1998 *Comput Networks ISDN Syst.* **30** 107
[17] Lin Y and Zhang Z Z 2019 *The Computer Journal* **62** 63
[18] Hu J T, Du Y X, Mo G M, Wei D J and Deng Y 2016 *Physica A* **444** 73
[19] Fan T L, Lü L Y, Shi D H and Zhou T 2021 *Commun. Phys.* **4** 272
[20] Zhao Y, Li C, Shi D H, Chen G R and Li X 2021 *Chaos* **33** 073139
[21] Jiang S Y, Zhou J, Small M, Lu J A and Zhang Y Q 2023 *Phys. Rev. Lett.* **130** 187402
[22] Shi W F, Xu S Q, Fan T L and Lü L Y 2023 *Sci. China Inf. Sci.* **66** 192203
[23] Callaway D S, Newman M E J, Strogatz S H and Watts D J 2000 *Phys. Rev. Lett.* **85** 5468
[24] Braunstein A, Dall'Asta L, Semerjian G and Zdeborová L 2016 *Proc. Natl. Acad. Sci. USA* **113** 12368
[25] Schneider C M, Moreira A A, Andrade J S, Jr, Havlin S and Herrmann H J 2011 *Proc. Natl. Acad. Sci. USA* **108** 3838
[26] Goh K I, Cusick M E, Valle D, Childs B, Vidal M and Barabási A L 2007 *Proc. Natl. Acad. Sci. USA* **104** 8685
[27] Zino L and Cao M 2021 *IEEE Circ. Syst. Mag.* **21** 4
[28] Cohen R, Havlin S and ben-Avraham D 2003 *Phys. Rev. Lett.* **91** 247901
[29] Gallos L K, Liljeros F, Argyrakis P, Bunde A and Havlin S 2007 *Phys. Rev. E* **75** 045104
[30] Pastor-Satorras R and Vespignani A 2002 *Phys. Rev. E* **65** 036104
[31] Bonneau H, Hassid A, Biham O, Kühn R and Katzav E 2017 *Phys. Rev. E* **96** 062307
[32] Lizier J T, Atay F M and Jost J 2012 *Phys. Rev. E* **86** 026110
[33] Shi D, Chen G, Thong W W K and Yan X 2013 *IEEE Circuits Syst. Mag.* **13** 66
[34] Ruths J and Ruths D 2014 *Science* **343** 1373
[35] Zhang W, Li W and Deng W 2021 *Commun. Nonlinear Sci. Numer. Simul.* **99** 105804
[36] Wilfried I and Peter F S 1998 *Working Papers* **26** 233
[37] Rossi R A and Ahmed N K 2015 *Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence* **29** 4292
[38] Guimerà R, Danon L, Díaz-Guilera A, Giralt F and Arenas A 2003 *Phys. Rev. E* **68** 065103
[39] Yang R, Shi J M, Xiao X K, Yang Y, Liu J C and Bhowmick S S 2020 *Proc. VLDB Endow.* **14** 37
[40] Leskovec J, Kleinberg J and Faloutsos C 2007 *ACM Trans. Knowl. Discov. Data* **1** 2
[41] Gao Y L, Yu H B, Zhou J, Zhou Y Z and Chen S M 2023 *Chin. Phys. B* **32** 098902
[42] Cui D X, He J L, Xiao Z F and Ren W P 2023 *Chin. Phys. B* **32** 098904
[43] Huang S, Chen J, Li M Y, Xu Y H and Hu M L 2024 *Chin. Phys. B* **33** 030205
[44] Lancichinetti A, Fortunato S and Radicchi F 2008 *Phys. Rev. E* **78** 046110
[45] Kunegis J 2013 *Proceedings of the 22nd International Conference on World Wide Web*
[46] Adamic L A and Glance N 2005 *Proceedings of the 3rd International Workshop on Link Discovery* **8** 36