



OPEN

Identifying influential spreaders in complex networks for disease spread and control

Xiang Wei^{1✉}, Junchan Zhao², Shuai Liu¹ & Yisi Wang³

Identifying influential spreaders is an important task in controlling the spread of information and epidemic diseases in complex networks. Many recent studies have indicated that the identification of influential spreaders is dependent on the spreading dynamics. Finding a general optimal order of node importance ranking is difficult because of the complexity of network structures and the physical background of dynamics. In this paper, we use four metrics, namely, betweenness, degree, H-index, and coreness, to measure the central attributes of nodes for constructing the disease spreading models and target immunization strategies. Numerical simulations show that spreading processes based on betweenness centrality lead to the widest range of propagation and the smallest epidemic threshold for all six networks (including four real networks and two BA scale-free networks generated according to Barabasi–Albert algorithm). The target immunization strategy based on the betweenness centrality of nodes is the most effective for BA scale-free networks but displays poor immune effect for real networks in identifying the most important spreaders for disease control. The immunization strategy based on node degrees is the most effective for the four real networks. Findings show that the target immune strategy based on the betweenness centrality of nodes works best for standard scale-free networks, whereas that based on node degrees works best for other nonstandard scale-free networks. The results can provide insights into understanding the different metrics of measuring node importance in disease transmission and control.

An interesting challenge in network science is understanding the relationship between the system structure and its dynamics. This condition makes it important to find the decisive network structure factors that will allow better control of the system's functionality^{1–4}. Epidemic spreading in complex networks has attracted much attention for a long time, the SIS model for the dynamics of malaria in human populations is studied⁵, it is found that whenever the number of reproduction is greater than 1, the time of disease extinction is exponentially distributed, so it becomes very large in a large population. María et al. analyzed the disease dynamics in highly clustered networks with short average path length⁶, they revealed that there was an exponent $p \neq 1$ for which $\langle S/N \rangle = p = R_0^{-1}$. Mayra et al.⁷ constructed a metapopulation model through the network to study the population migration in the case of dengue. The model can determine the flow pattern of the disease in the large region of southeast Mexico. Li et al.⁸ developed the Colombian ZIKV transmission model on the complex network. The model considers the sexual communication and media communication between people in the transmission process, and displays the intrusion area of ZIKV in these two parameter spaces. Many research results are concentrated on determining which network topological attributes are closely related to information and virus transmission^{9–11}. In particular, many results focus on the centrality of a topological feature, including betweenness, degree, H-index, and core centralities. These agents are expected to be the most influential expanders because most central nodes can spread their effect to the entire network faster than others. Recently, Kitsak et al.¹² used different locations of the spreading origin and found evidence to support this hypothesis in the case of an epidemic outbreak. The most influential spreader can be predicted through K-core decomposition analysis. These agents are located at the core of the network and do not need to be closely connected. Silva et al.¹³ studied the correlation between the heterogeneous propagation and central properties of nodes that originally transmitted disease and found that a strong correlation exists between node degree and propagation accessibility. For different dynamic processes, such as the standard rumor models, identifying the most influential spreaders with the same metrics¹⁴ is impossible. Nine centrality metrics correlating with the rumor and disease spreading

¹Department of Engineering, Honghe University, Honghe 661100, People's Republic of China. ²School of Science, Hunan University of Technology and Business, Changsha 410205, People's Republic of China. ³School of Big Data Science and Application, Chongqing Wenli University, Chongqing 402160, People's Republic of China. ✉email: weixiangwx2003@163.com

capabilities of the nodes were investigated¹⁵, and the generalized reachability is more related to the propagation process in spatial networks than that in nonspatial networks.

Many studies of disease transmission in complex networks mainly use mean-field method^{16,17}, which is simple and easy to understand. The mean-field method is widely used in the analysis and prediction of epidemiology propagation dynamics^{18–20}, voting model²¹, and synchronization phenomenon^{22,23}. The mean-field method is a method of collectively treating the effects at the level of classes. Assume that the nodes of the same degree behave equally. For homogeneous²⁴ and uniform networks (degree distribution obeys Normal or Poisson distribution), scholars applied the homogeneous mean-field method to model analysis. For heterogeneous Barabási-Albert (BA) networks²⁵, such as scale-free networks, scholars used heterogeneous mean-field methods by assuming that all nodes with equal degree k follow the same dynamic process. Many mean-field methods are used to model the spreading process on the basis of the node degree distribution, and the attributes used to characterize the importance of nodes include the metrics of node degree and the metrics of betweenness, H-index, coreness^{26,27}. The relationship between degree, H-index, and coreness was displayed in²⁶, where degree, H-index, and coreness are the initial, intermediate, and steady states of these sequences, respectively. Since the mean-field method used to describe propagation dynamics assumes that nodes with the same degree have the same dynamics, can the mean-field method be constructed and assume the nodes with the same H-index or coreness or betweenness have the same dynamics?

Although many studies have provided evidence for the presence of influential spreaders in disease transmission, the conclusions are not universal. No consensus is reached under the definition of network centrality because different metrics are used to quantify the centrality of nodes for specific physical background²⁸. Controlling the spread of infectious diseases through complex networks, such as the spread of rumors in social interaction and the spread of infection in the population, has attracted increasing attention. Sk et al. proposed a novel mathematical model to forecast COVID-19 and assess control strategies, they found that investing on the quarantined was more effective than isolated people in reducing cases²⁹. Li et al. developed a SEIQR difference-equation model to investigate the dynamics of COVID-19. The advantage of this model is that it does not need to estimate the initial value of the model³⁰. From the statistical and mathematical analysis of COVID-19 case reports, human mobility and temporal and spatial changes of transmission control measures, they concluded that China's control measures made it possible to avoid hundreds of thousands of cases³¹. Considering the impact of blockade and medical resources on the spread of COVID-19 in Wuhan, Sun et al. put forward a dynamic model of epidemic transmission and found that the more abundant medical resources, the smaller the final scale³². Immunizing a small number of the highest degree nodes can effectively eliminate virus spreading in scale-free networks³³. An improved mean-field model to study immunization was proposed by utilizing the degree centrality before and after the immunization³⁴, showing that the epidemic threshold for infectious diseases is low. A new measure of centrality that utilizes the coreness and eigenvector centralities was proposed to identify the influential spreaders³⁵. The results displayed that this approach is more influential than other benchmarks. Many metrics of node centrality are used to select the nodes for target immunization.

In this study, heterogeneous mean-field propagation models based on node betweenness centrality, degree, H-index, and coreness are constructed. The numerical simulations show that the spreading processes based on betweenness centrality lead to the widest range of propagation processes. With the target immunity for disease transmission, the numerical simulations of uniform and target immunizations based on node betweenness centrality, degree, H-index, and coreness display that the immunization strategy based on betweenness centrality performs competitively for BA scale-free networks, and the immunization strategy based on degree centrality performs competitively in comparison with other strategies for real networks.

Methods

Epidemic model construction strategies. The classical susceptible-infected-susceptible (SIS) model¹⁶ is used to describe the spreading dynamics on heterogeneous network at the level of classes of nodes. Assume that all nodes of the same degree (betweenness, H-index, and k -coreness) behave equally. Define the partial prevalence $\rho_k(t)$ ($\rho_b(t)$, $\rho_h(t)$ and $\rho_s(t)$) as the fraction of infected nodes with a given degree k degree (betweenness b , H-index h , and coreness s). The goal is to understand the propagation process under the measurement of four different node importance metrics on the epidemic processes. We compare the prevalence of the four different propagation dynamics. In the heterogeneous network, let $P(k)$ denote the fraction of infected nodes that a node with a given degree k and $P(k'|k)$ be the conditional probability that a node of degree k is connected to a node of degree k' . The normalization conditions $\sum_k P(k) = 1$ and $\sum_k P(k'|k) = 1$ hold. The average number of links connecting a node of degree k to some nodes of degree k' is $kP(k'|k)$. Thus, the dynamic evolution processes can be written as

$$\frac{d\rho_k(t)}{dt} = -\rho_k(t) + \lambda k(1 - \rho_k(t))\Theta_k(t), \quad (1)$$

where

$$\Theta_k(t) = \frac{1}{\langle k \rangle} \sum_{k'} k' P(k'|k) \rho_{k'}(t). \quad (2)$$

The first term on the right-hand side denotes that infected nodes of degree k can be recovered. The second term indicates that susceptible nodes of degree k are infected by their infected neighbors, where $1 - \rho_k(t)$ represents the fraction of susceptible nodes of degree k , and $\Theta_k(t)$ is the probability that a link emanating from the

nodes of degree k points to an infected node in the complex network. $\langle k \rangle$ is the first moment of degree distributions. When the right-hand side of Eq. (1) is zero, we obtain

$$\rho_k = \frac{k\lambda\Theta_k(t)}{1 + k\lambda\Theta_k(t)}, \quad (3)$$

Based on Eqs. (1) and (2), then one obtains

$$\Theta_k(t) = \frac{1}{\langle k \rangle} \sum_k kP(k) \frac{k\lambda\Theta_k(t)}{1 + k\lambda\Theta_k(t)}, \quad (4)$$

The equilibrium point of endemic diseases in Eq. (1) must meet the following conditions:

$$\frac{d}{d\Theta_k(t)} = \left(\frac{1}{\langle k \rangle} \sum_k kP(k) \frac{k\lambda\Theta_k(t)}{1 + k\lambda\Theta_k(t)} \right) \Big|_{\Theta_k=0} \geq 1, \quad (5)$$

one gets

$$\sum_k \frac{kP(k)\lambda k}{\langle k \rangle} = \frac{\langle k^2 \rangle}{\langle k \rangle} \lambda \geq 1, \quad (6)$$

where $\langle k^2 \rangle$ is the second moment of degree distributions.

The epidemic threshold of heterogeneous network based on degree is¹⁶

$$\lambda_c^k = \frac{\langle k \rangle}{\langle k^2 \rangle}, \quad (7)$$

For the metrics of H-index, coreness, and betweenness centrality, the dynamic processes are constructed as follows:

$$\frac{d\rho_h(t)}{dt} = -\rho_h(t) + \lambda h(1 - \rho_h(t))\Theta_h(t), \quad (8)$$

where $\Theta_h(t) = \frac{1}{\langle h \rangle} \sum_{h'} h' P(h') \rho_{h'}(t)$, h is the metric for the nodes of H-index h .

$$\frac{d\rho_s(t)}{dt} = -\rho_s(t) + \lambda s(1 - \rho_s(t))\Theta_s(t), \quad (9)$$

where $\Theta_s(t) = \frac{1}{\langle s \rangle} \sum_{s'} s' P(s') \rho_{s'}(t)$, s is the metric for the nodes of coreness s .

$$\frac{d\rho_b(t)}{dt} = -\rho_b(t) + \lambda b(1 - \rho_b(t))\Theta_b(t), \quad (10)$$

where $\Theta_b(t) = \frac{1}{\langle b \rangle} \sum_{b'} b' P(b') \rho_{b'}(t)$, b represents the number of shortest paths through node i . Node betweenness is defined as the ratio of the number of shortest paths through node i to the total number of shortest paths in the network. We use the number of shortest paths through node i as metrics for modeling.

For the metrics of H-index, coreness, and betweenness centrality, by the same methods, we can conclude that the models have the following epidemic thresholds:

$$\lambda_c^h = \frac{\langle h \rangle}{\langle h^2 \rangle}, \quad (11)$$

where $\langle h \rangle$ is the first moment and $\langle h^2 \rangle$ is the second moment of H-index distributions.

$$\lambda_c^s = \frac{\langle s \rangle}{\langle s^2 \rangle}, \quad (12)$$

where $\langle s \rangle$ is the first moment and $\langle s^2 \rangle$ is the second moment of coreness distributions.

$$\lambda_c^b = \frac{\langle b \rangle}{\langle b^2 \rangle}, \quad (13)$$

where $\langle b \rangle$ is the first moment and $\langle b^2 \rangle$ is the second moment of betweenness distributions.

Optimized immunization strategies. For epidemic spreading in complex networks, the immunization procedure can effectively prevent the spread of disease. Given that the proportional immunization schemes can effectively increase the immunity threshold, the special network of scale-free networks allows for efficient strategies on the basis of the hierarchy of nodes. Scholars have designed targeted immunization schemes where the highly connected nodes (i.e., nodes are likely to spread disease) are gradually immunized. Choosing the most effective metric to identify the most efficient spreaders for targeted immunization in a network is an important step. In contrast to common belief, the best spreaders correspond to the most highly connected or the most central people. In this study, four metrics (betweenness, degree, H-index, and coreness) are used to identify the most

Networks	N	E	$\langle k \rangle$	$\langle d \rangle$	r
Email	1005	4855	9.622	3.61	0.078
Polblogs	1222	16,714	27.35	2.740	- 0.200
USAir	332	2126	12.81	2.738	- 0.208
Ego-facebook	4039	88,234	43.6	3.670	- 0.182
Scale-free network 1	5,000	150,000	60	2.670	- 0.01
Scale-free network 2	10,000	1,000,000	200	3.24	- 0.011

Table 1. The topological features of the six networks.

Networks	Betweenness	Degree	H-idx	Core
Email	44,013	140	75	34
Polblogs	92,228	143	73	36
USAir	11,376	58	32	23
Ego-facebook	778,993	223	148	97
Scale-free network 1	82,724	343	131	78
Scale-free network 2	203,277	741	270	152

Table 2. The maximum value of the four metrics.

efficient spreaders for targeted immunization. For a fixed infection rate λ , the control parameter is the immunity proportion g , which is defined as part of the immune nodes that exist in the network. For the mean-field model, the presence of immunity can reduce the infection rate λ by the factor $(1 - g)$. By substituting $\lambda \rightarrow \lambda(1 - g)$ in Eq. (3), the prevalence behavior of immunization rates is increasing. Consider that a proportion g of the nodes with the highest connectivity are immunized. This process corresponds to introducing an upper threshold k_t , such that all nodes with connectivity $k > k_t$ are immunized. The proportion of immunized nodes is given by Eq. (3) at the mean-field level, and the presence of immunity will reduce the infection rate λ by the factor.

$$g = \sum_{k > k_t} P(k). \quad (14)$$

For the metrics of H-index, coreness, and betweenness, the fraction g of the immunized nodes is given as follows:

$$g = \sum_{h > h_t} P(h), g = \sum_{s > s_t} P(s), g = \sum_{b > b_t} P(b). \quad (15)$$

Data description. Four real networks and two BA scale-free networks from different fields are used to simulate the performance of the four models in Eqs. (3), (8), (9) and (10), including two communication networks (Email, ego-Facebook), one social networks (Political blogs), one transportation network (USAir), and two BA scale-free networks. Email^{36,37} describes the email interchanges between various users. Political blogs^{38,39} is a network between weblogs on US politics. ego-Facebook^{40,41} is collected from survey participants using the Facebook app. USAir⁴² is a network that describes US air transportation. Two scale-free networks are formed using the algorithm proposed by Barabási and Albert²⁵ with different parameters. The topological features of these networks, including the number of nodes and links, average degree, average distance, and assortative coefficient, are shown in Table 1.

The features of these networks, including the max betweenness, max degree, max H-index, and max coreness, are shown in Table 2. For each network, the max betweenness of nodes is the largest, followed by max degree, max H-index, and max coreness. Therefore, the node importance fluctuations based on betweenness are the largest, followed by those fluctuations based on degree, H-index, and coreness. This finding implies that the most fluctuant metric is betweenness, followed by degree, H-index, and coreness.

Results

Numerical simulations for epidemic model. To study the present mean-field methods based on the four metrics, the numerical simulations of the SIS model on the six networks are used to display the spread dynamic process. The infection prevalence in the stationary state as a function of spreading rate λ was simulated. The initial fraction of infected nodes is set to 0.05 (without lack of generality, we set the fixed probability of curing $\mu = 1$). The comparisons of infection prevalence ρ of the infected nodes among the four models and Monte Carlo simulations for six networks are displayed in Figs. 1, 2 and 3. Different disease spreading models exhibit diverse spreading processes. The model based on betweenness centrality exhibits the highest infection

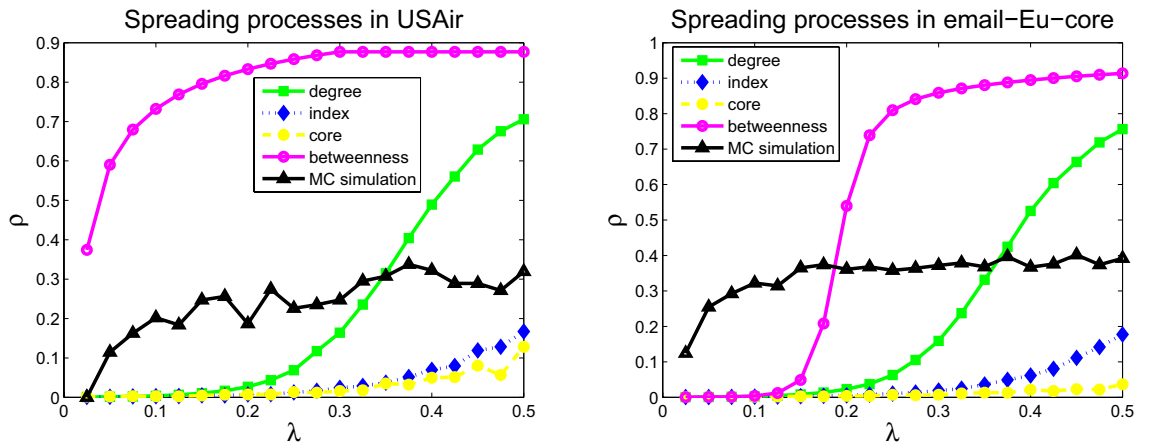


Figure 1. Infections prevalence ρ as a function of spreading rate λ for USAir (left) and email-Eu-core (right).

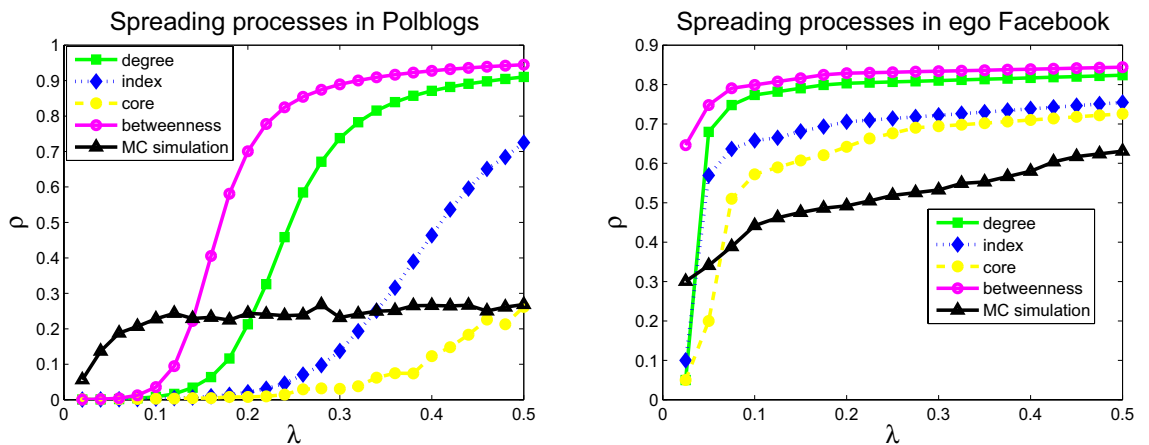


Figure 2. Infections prevalence ρ as a function of spreading rate λ for Polblogs (left) and ego-Facebook (right).

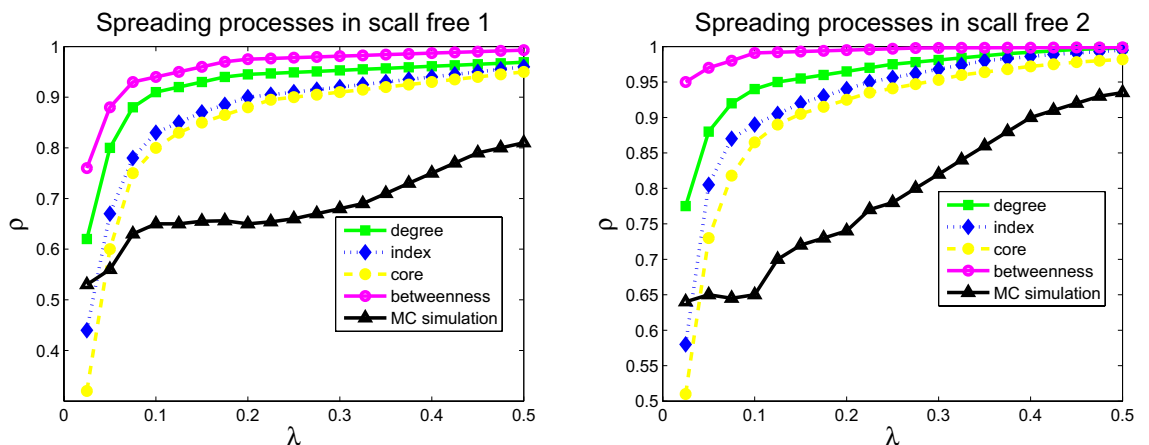


Figure 3. Infections prevalence ρ as a function of spreading rate λ for scale-free 1 (left) and scale-free 2 (right).

prevalence for all six networks, followed by the model based on degree, H-index, and coreness. Monte Carlo simulations results (black solid curve and triangle) were displayed in Figs. 1, 2 and 3. In the two networks (USAir and Email) with small number of nodes and small average degree, the variance between Monte Carlo simulation results and simulation results based on degree distribution is the smallest. In the other four networks (polblogs, ego-Facebook, scale-free network 1 and scale-free network 2) with large number of nodes and large average

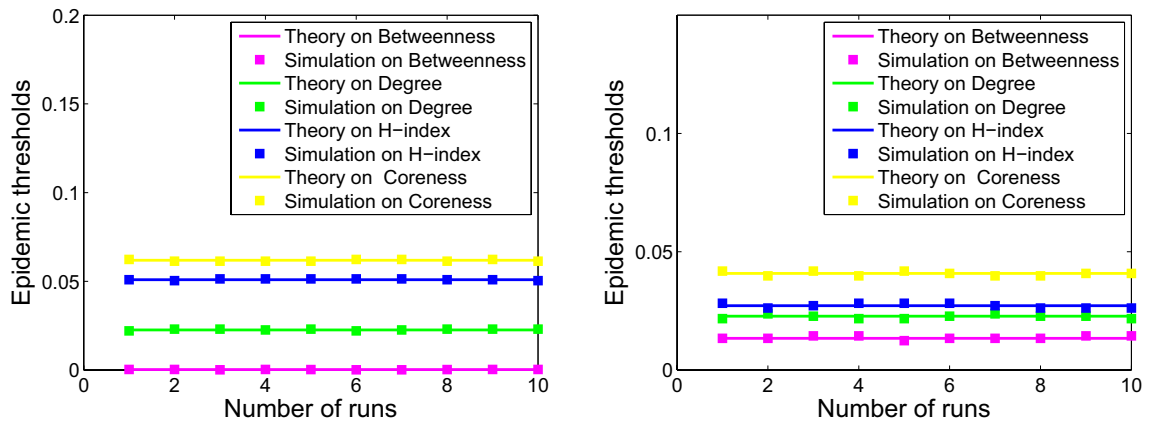


Figure 4. Numerical simulation (squares) and theoretical (solid curve) results of epidemic thresholds for USAir (left) and email-Eu-core (right).

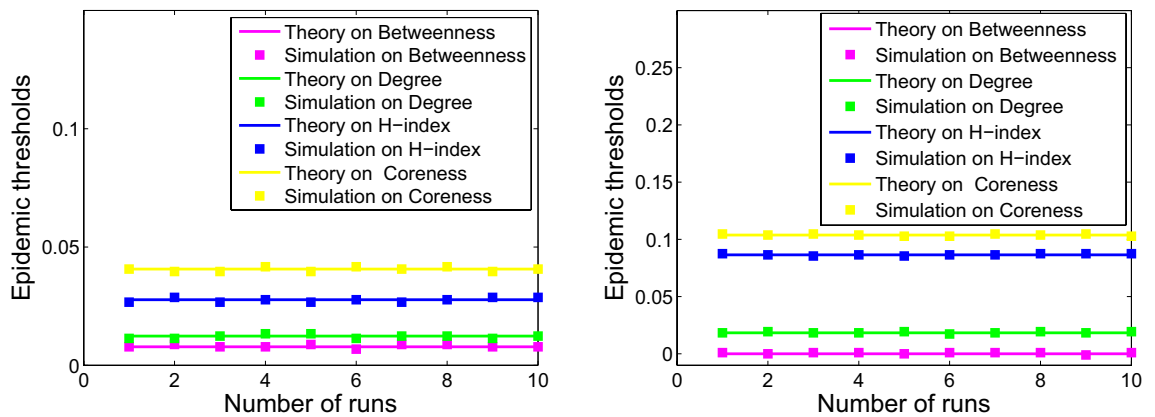


Figure 5. Numerical simulation (squares) and theoretical (solid curve) results of epidemic thresholds for Polblogs (left) and ego-Facebook (right).

degree, the variance between Monte Carlo simulation results and simulation results based on coreness distribution is the smallest. Note that in the networks (ego-Facebook, scale-free network 1 and scale-free network 2) with large number of nodes, infections prevalence of Monte Carlo simulation is smallest compared with the four other models.

Comparing the epidemic thresholds of epidemic models. Since more fluctuant metric lead to larger second moment, one can obtain:

$$\langle b^2 \rangle > \langle k^2 \rangle > \langle h^2 \rangle > \langle s^2 \rangle, \tag{16}$$

which imply that the model based on betweenness centrality exhibits the smallest epidemic threshold for all six networks, followed by the model based on degree, H-index, and coreness. Therefore, we can conclude that:

$$\lambda_c^b < \lambda_c^k < \lambda_c^h < \lambda_c^s. \tag{17}$$

The comparison of the epidemic thresholds for the networks from theoretical analysis described in Eqs. (7), (11), (12), and (13) with that from numerical simulations are displayed in the Figs. 4, 5 and 6. It shows that the theoretical analysis agrees well with numerical simulations with some minor deviation. Disease spreading models based on different metrics exhibit diverse spreading processes. Moreover, the simulation results are consistent with that of the previous analysis in Eq. (17).

Comparing immunization effect of different immunization strategies. The immunization behaviors for the SIS model on six networks are studied through numerical simulations for different immunization schemes. The targeted immunization scheme and uniform immunization are tested by immunizing the gN nodes. For a network of fixed size N , the uniform immunization is conducted by randomly selecting and immunizing gN nodes, and the targeted immunizations are performed by choosing the most central gN nodes to immunize. Node centrality is identified from betweenness, degree, H-index, and coreness centralities. The nodes at the top of the lists will be maintained. Infection prevalence ρ_g in the stationary state as a function of

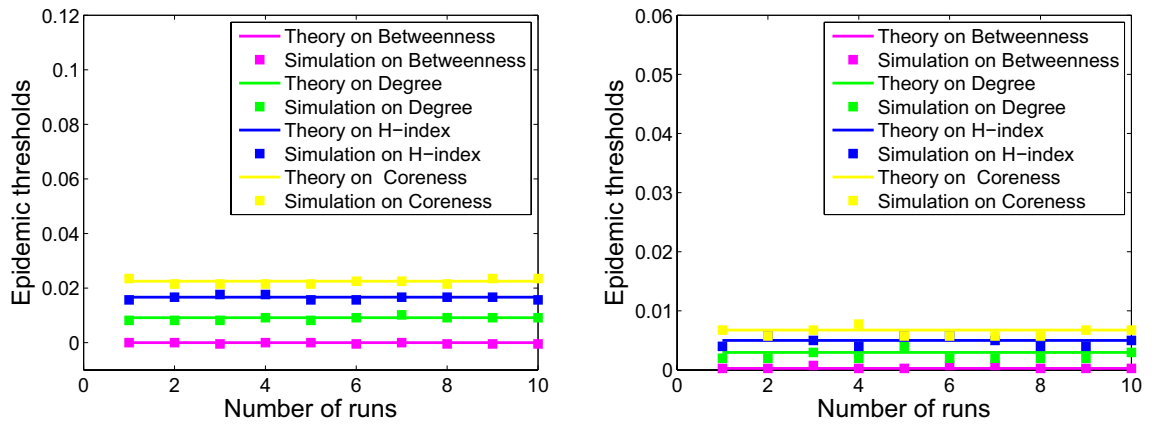


Figure 6. Numerical simulation (squares) and theoretical (solid curve) results of epidemic thresholds for scale-free 1 (left) and scale-free 2 (right).

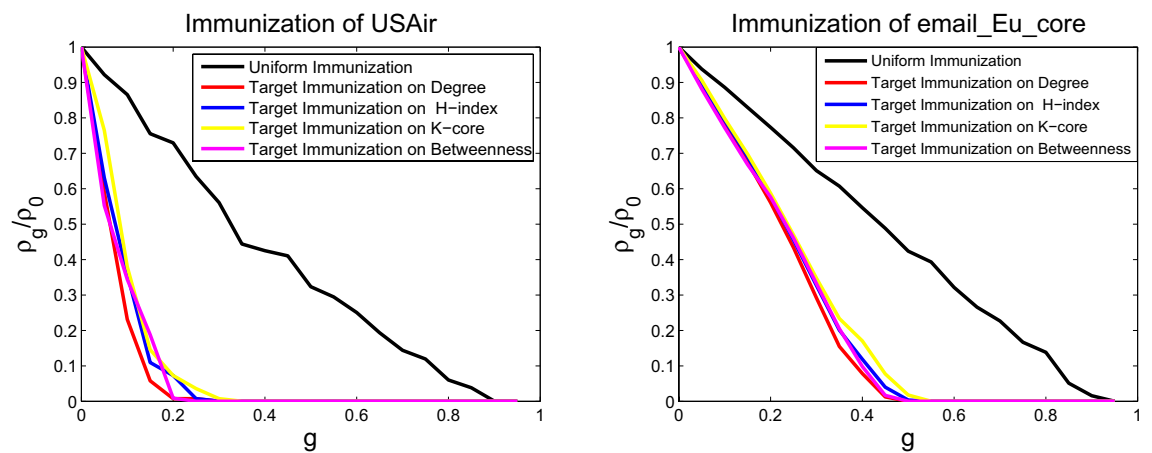


Figure 7. Comparison of the prevalence ρ_g/ρ_0 as a function of immunity proportion g with uniform and four targeted immunization for USAir (left) and email-Eu-core (right), at a fixed spreading rate $\lambda = 0.15$.

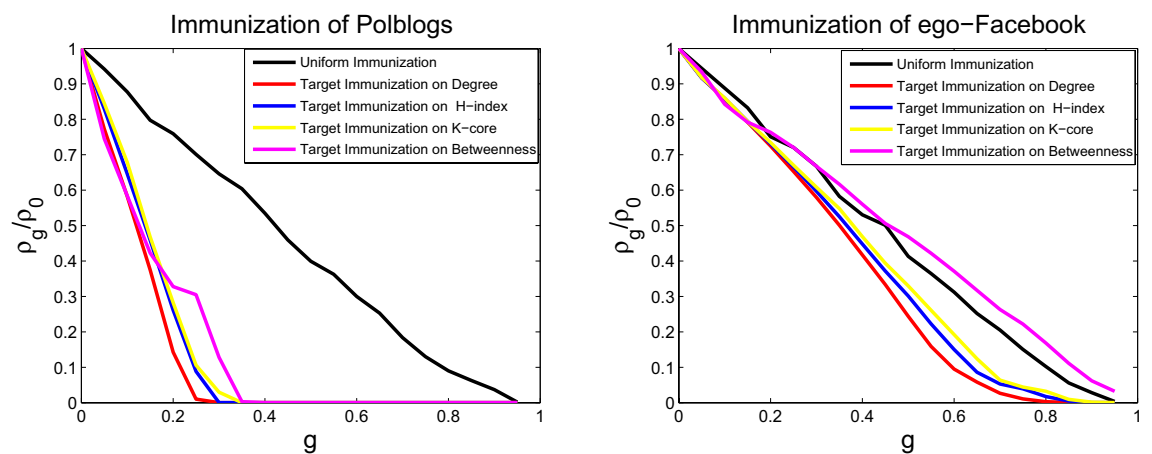


Figure 8. Comparison of the prevalence ρ_g/ρ_0 as a function of immunity proportion g with uniform and four targeted immunization for Polblogs (left) and ego-Facebook (right), at a fixed spreading rate $\lambda = 0.15$.

different immunization proportion g is obtained through numerical simulation. The initial fraction of infected nodes is set to 0.05, the fixed spreading rate is $\lambda = 0.15$, and the fixed probability of curing is $\mu = 1$. The infection prevalence is computed by averaging over 100 runs of each model. Figures 7, 8 and 9 display the behavior of infection prevalence ρ_g/ρ_0 (where ρ_0 is the prevalence without immunization) as a function of immunization proportion g . For uniform immunization, the prevalence experiences a passive drop and exhibits the onset of

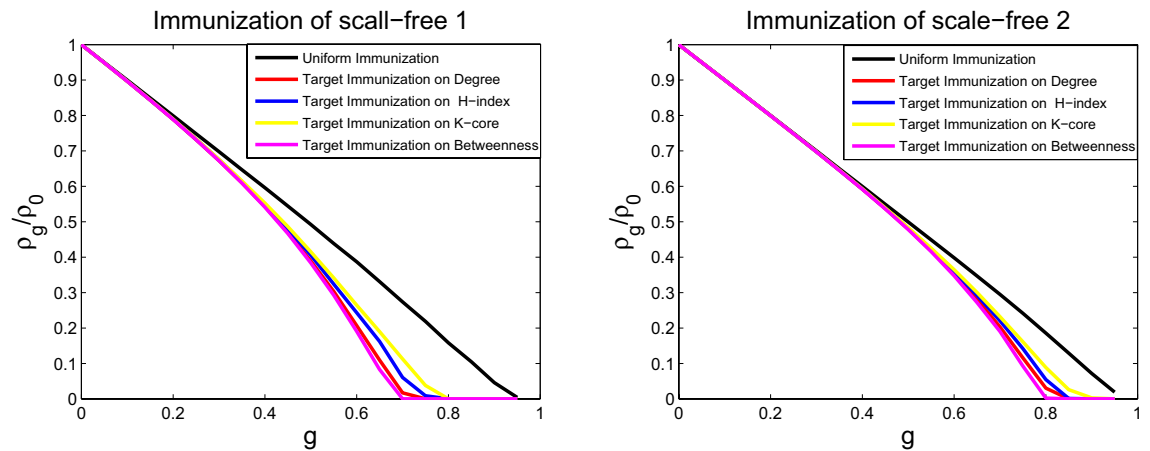


Figure 9. Comparison of the prevalence ρ_g/ρ_0 as a function of immunity proportion g with uniform and four targeted immunization for scale-free 1 (left) and scale-free 2 (right), at a fixed spreading rate $\lambda = 0.15$.

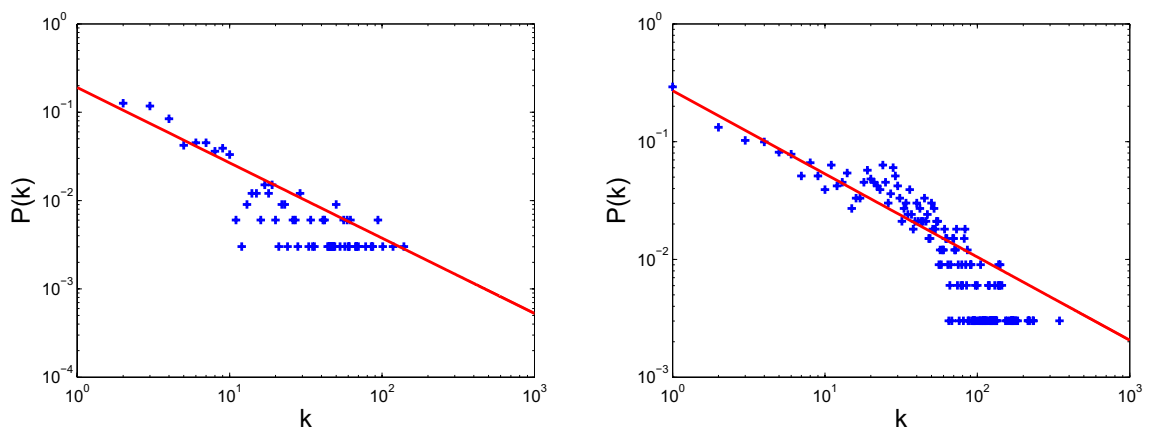


Figure 10. Degree distribution of USAir (left) and email-Eu-core (right).

large immunization threshold. On the contrary, the infection prevalence experiences an extremely sharp drop and exhibits the onset of small immunization threshold for targeted immunization (immunization threshold denotes that ρ_g/ρ_0 is 0 in Figs. 7, 8 and 9). The targeted immunization strategy using betweenness centrality to identify the most important nodes displays different immune effect for real networks, it displays the second best effect for USAir and email-Eu-core networks, as shown in Fig. 7, and displays the worst target immune effect among the four strategies, as shown in Fig. 8 for Political blogs and ego-Facebook networks. But it displays the best target immune effect for the two BA scale-free networks in Fig. 9. Excluding the target immunity based on betweenness centrality. For all networks, the target immunity strategy using degree centrality shows the most effective immune effect and obtains the smallest immunization threshold, followed by the model using H-index centrality, and the third is the model using coreness centrality.

Discussion

The models for disease propagation based on betweenness exhibit the best propagation effect, followed by those models using the metrics of degree, H-index, and coreness. The models for targeted immunization based on betweenness exhibit the best immunization effect on BA scale-free networks, followed by those models using degree, H-index, and coreness. The targeted immunizations based on betweenness exhibit inconsistent effect for real networks because the node degrees of these real networks do not conform to the standard power-law distribution. The node degree distributions of these real networks and two BA scale-free networks are displayed in Figs. 10, 11 and 12. The right panels of Figs. 8 and 11 display that ego-Facebook whose node degrees are the least consistent with the power-law distribution shows the worst target immunity effect, and Figs. 9 and 12 show that the two BA scale-free networks whose node degrees are the most consistent with the power-law distribution show the best target immunity effect. The target immunization based on the node betweenness exhibits the best immunization effect on the network with standard power-law distribution (that is, the node degree distribution is a straight line in the double logarithmic scales). The models for disease propagation and targeted immunization based on degree exhibit the best propagation and immunization effects, followed by those models using H-index and coreness when the optimal of point betweenness is excluded. The key to obtaining these conclusions is the heterogeneity of node centrality. The greater the centrality disturbance of the nodes is, the more favorable it is

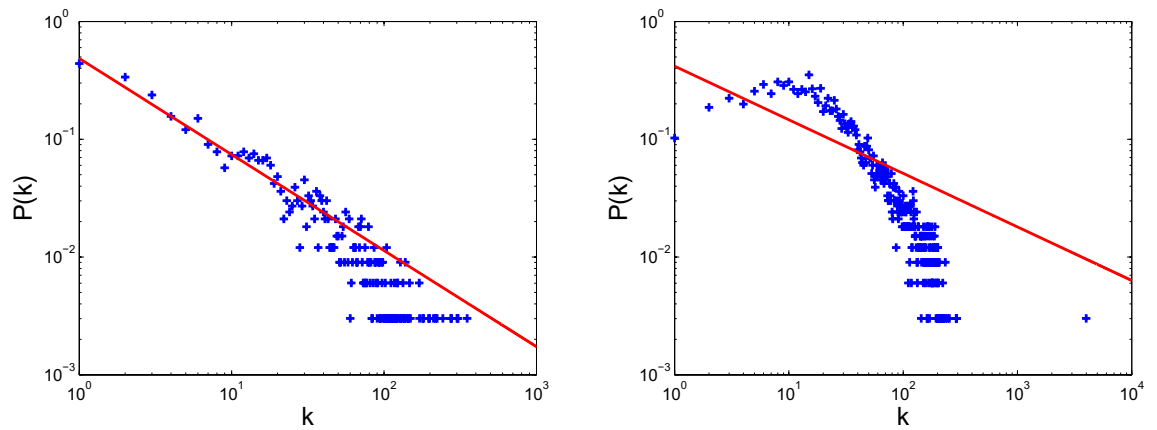


Figure 11. Degree distribution of Polblogs (left) and ego-Facebook (right).

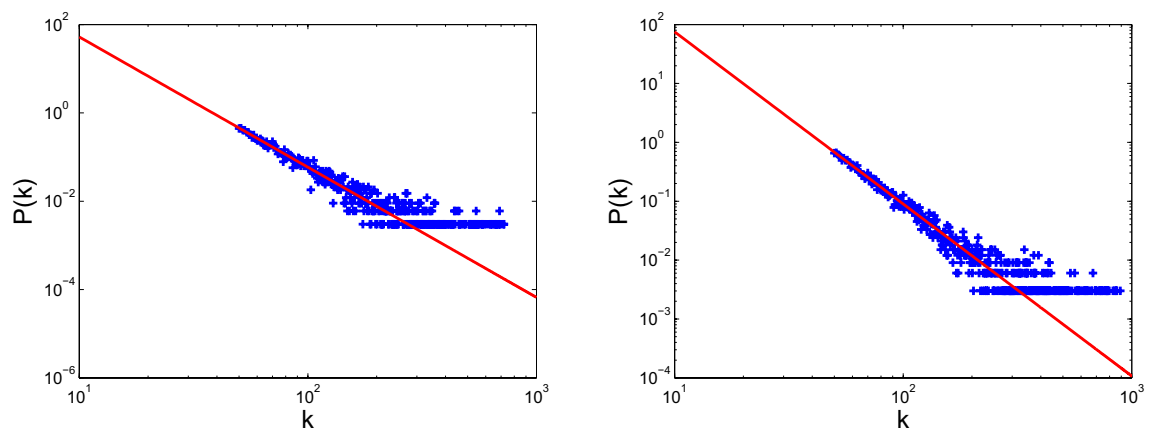


Figure 12. Degree distribution of scale-free 1 (left) and scale-free 2 (right).

to propagate and find the target immune nodes. The node importance fluctuations based on betweenness are the largest, followed by those fluctuations based on degree, H-index, and coreness. The results of this study are consistent with the conclusions in paper²⁶, where degree, H-index, and coreness are the initial, intermediate, and steady states of the sequences operated by operator \mathcal{H} , respectively. Operator \mathcal{H} makes strong fluctuations in the degree distribution to moderate fluctuations in the H-index distribution and abstains weak fluctuations in the coreness distribution. The simulations confirm that the most fluctuant metric is effective in disease propagation and identifying the spreading influences of nodes for target immunization. These results reveal that the largest fluctuant metric can be used to construct the fastest disease spreading model and identify the influential spreaders for target immunization. For networks with standard power law distribution, the best immune effect can be obtained by finding the most influential spreaders using the betweenness centrality as the immune target. However, for many real networks with nonstandard power law distribution, the best immune effect can be obtained using the degree centrality to find the most influential spreaders as the immune target. The 2019 Coronavirus Disease (COVID-19) has caused an outbreak on a global scale, so it is necessary to investigate control strategies to develop health care plans. The findings of this paper can guide people to accurately find the targets that need immunization, so as to effectively control the spread of the epidemic.

Conclusion

The models for disease propagation and target immunization strategies based on node betweenness, degree, H-index, and coreness metrics are proposed in this study. The models for disease propagation based on betweenness always exhibit the best propagation effect, followed by those using the metrics of degree, H-index, and coreness. The models for targeted immunization based on betweenness exhibit the best immunization effect in BA scale-free networks, followed by those using degree, H-index, and coreness. However, the models for targeted immunization in four real networks based on node degrees exhibit the best immunization effect, while that based on betweenness for different real networks exhibit inconsistent effects (second-best or worst effects) because the node degrees of these real networks do not conform to the standard power-law distribution. The results provide reference for the public health sector to put forward effective measures for disease prevention and control.

Data availability

All relevant data can be downloaded from their respective web pages.

Received: 3 July 2021; Accepted: 23 February 2022

Published online: 01 April 2022

References

1. Aguirre, J., Sevilla-Escoboza, R., Gutiérrez, R., Papo, D. & Buldú, J. Synchronization of interconnected networks: The role of connector nodes. *Phys. Rev. Lett.* **112**, 248701 (2014).
2. Li, Y., Wu, X., Lu, J.-A. & Lü, J. Synchronizability of duplex networks. *IEEE Trans. Circuits Syst. II Express Briefs* **63**, 206–210 (2017).
3. Feng, Y. & Wang, B. C. A unified framework of epidemic spreading prediction by empirical mode decomposition based ensemble learning techniques. *IEEE Trans. Comput. Soc. Syst.* **6**, 660–669 (2019).
4. Kang, H., Sun, M., Yu, Y., Fu, X. & Bao, B. Spreading dynamics of an SEIR model with delay on scale-free networks. *IEEE Trans. Netw. Sci. Eng.* **7**, 489–496 (2020).
5. Fombohforba, M. Y. & Ngwa, G. A. Stochastic analysis of a closed SIS malaria model. *J. Cameroon Acad. Sci.* **12**, 87–113 (2015).
6. del Valle Rafo, M., Mauro, J. P. D. & Aparicio, J. P. Disease dynamics and mean field models for clustered networks. *J. Theor. Biol.* **526**, 110554. <https://doi.org/10.1016/j.jtbi.2020.110554> (2021).
7. López, M., Ramos, L. A. & Velasco-Hernandez, J. X. Migration rate estimation in an epidemic network. *Appl. Math. Model.* **82**, 1949–1964 (2020).
8. Li, L. *et al.* Analysis of transmission dynamics for zika virus on networks. *Appl. Math. Comput.* **347**, 566–577 (2019).
9. Zhan, *et al.* Coupling dynamics of epidemic spreading and information diffusion on complex networks. *Appl. Math. Comput.* **332**, 437–448 (2018).
10. Zhang, D., Wang, Y. & Zhang, Z. Identifying and quantifying potential super-spreaders in social networks. *Sci. Rep.* **9**, 14811 (2019).
11. Chin, W. C. B. & Bouffanais, R. Spatial super-spreaders and super-susceptibles in human movement networks. *Sci. Rep.* **10**, 18642 (2020).
12. Kitsak, M. *et al.* Identification of influential spreaders in complex networks. *Nat. Phys.* **6**, 888–893 (2010).
13. Silva, R. A. P. D., Viana, M. P. & Costa, L. D. F. Predicting epidemic outbreak from individual features of the spreaders. *J. Stat. Mech.* **2012**, P07005 (2012).
14. Borge Holthoef, J. & Moreno, Y. Absence of influential spreaders in rumor dynamics. *Phys. Rev. E* **85**, 026116 (2012).
15. Arruda, G. F. D. *et al.* Role of centrality for the identification of influential spreaders in complex networks. *Phys. Rev. E* **90**, 032812 (2014).
16. Pastor-Satorras, R. & Vespignani, A. Epidemic dynamics and endemic states in complex networks. *Phys. Rev. E* **63**, 066117 (2001).
17. Newman, M. E. J. Spread of epidemic disease on networks. *Phys. Rev. E* **66**, 016128 (2002).
18. Beniamino, G. & Jesús, G. G. Annealed and mean-field formulations of disease dynamics on static and adaptive networks. *Phys. Rev. E* **82**, 035101 (2010).
19. Piao, Z. & Fu, X. Epidemic spread on one-way circular-coupled networks. *Acta Mathematica Scientia (English Series)* **39**(6), 1713–32 (2019).
20. Muthukumar, S., Muthukrishnan, S. & Chinnadurai, V. Dynamic behaviour of competing memes' spread with alert influence in multiplex social-networks. *Computing* **2019**, 1177–1197 (2019).
21. Sood, V. & Redner, S. Voter model on heterogeneous graphs. *Phys. Rev. Lett.* **94**, 178701 (2004).
22. Guo, W., Sun, W. & Lu, X. Distributed input observer-based finite-time consensus for the leader-following multi-agent systems. *J. Frankl. Inst.* **357**, 5556–5569 (2020).
23. Oh, E., Lee, D. S., Kahng, B. & Kim, D. Synchronization transition of heterogeneously coupled oscillators on scale-free networks. *Phys. Rev. E* **75**, 011104 (2007).
24. Watts, D. J. & Strogatz, S. H. Collective dynamics of small-world networks. *Nature* **6684**, 440–442 (1998).
25. Barabási, A. L. & Albert, R. Emergence of scaling in random networks. *Science* **286**, 509–512 (1999).
26. Lü, L., Zhou, T., Zhang, Q. M. & Stanley, H. E. The H-index of a network node and its relation to degree and coreness. *Nat. Commun.* **7**, 10168 (2016).
27. Wu, X., Wei, W., Tang, L., Lu, J.-A. & Lü, J. Coreness and H-index for weighted networks. *IEEE Trans. Circuits Syst. I Regul. Pap.* **66**, 3113–3122 (2019).
28. Liao, H., Mariani, M. S., Medo, M., Zhang, Y. C. & Zhou, M. Y. Ranking in evolving complex networks. *Phys. Rep.* **689**, 1–54 (2017).
29. Nadim, S. S., Ghosh, I. & Chattopadhyay, J. Short-term predictions and prevention strategies for COVID-19: A model-based study. *Appl. Math. Comput.* **404**, 126251 (2021).
30. Li, M. T., Sun, G. Q., Zhang, J., Zhao, Y. & Jin, Z. Analysis of COVID-19 transmission in Shanxi province with discrete time imported cases. *Math. Biosci. Eng.* **17**, 3710–3720 (2020).
31. Tian, H. *et al.* An investigation of transmission control measures during the first 50 days of the COVID-19 epidemic in China. *Science* **368**(6491), 638–642 (2020).
32. Sun, G. Q., Wang, S. F., Li, M. T., Li, L. & Feng, G. L. Transmission dynamics of COVID-19 in Wuhan, China: Effects of lockdown and medical resources. *Nonlinear Dyn.* **101**, 1981–93 (2020).
33. Pastor, S. & Vespignani, A. Immunization of complex networks. *Phys. Rev. E* **2001**, 036134 (2001).
34. Wu, Q. & Fu, X. Immunization and epidemic threshold of an SIS model in complex networks. *Phys. A* **444**, 576–581 (2016).
35. Sara, A. *et al.* Identifying influential spreaders by gravity model. *Sci. Rep.* **2019**, 8387 (2019).
36. Leskovec, J. Snap. <http://snap.stanford.edu/data/email-Eu-core.html> (2007).
37. Guimerà, R., Danon, L., Díaz-Guilera, A., Giralt, F. & Arenas, L. Self-similar community structure in a network of human interactions. *Phys. Rev. E* **68**, 065103 (2004).
38. Adamic, L. A. & Glance, N. The political blogosphere and the 2004 U.S. election: divided they blog. In *In Proceedings of the 3rd International Workshop on Link Discovery*, 36–43 (ACM Press, 2005).
39. Newman, M. Network data. <http://www-personal.umich.edu/~mejn/netdata/> (2005).
40. Leskovec, J. ego-networks. <http://snap.stanford.edu/data/ego-Facebook.html> (2012).
41. McAuley, J. J. & Leskovec, J. Learning to discover social circles in ego networks. *Adv. Neural Inf. Process. Syst.* **25**, 548–556 (2012).
42. Batageli, V. & Mrvar, A. Pajek Datasets. <http://vlado.fmf.uni-lj.si/pub/networks/data/> (2003).

Acknowledgements

This work is partially supported by National Natural Science Foundation of China (41761079), and Local Undergraduate Colleges and Universities Joint Special Foundation of Yunnan Province (2017FH001-062, 2018FH001-046), National Social Science Foundation of China (18BTJ025), Research Fund of Yunnan Provincial Department of Education (2020J0675), Scientific Research Fund of Yunnan Provincial Department of Education (2022J0893), and Research Fund Project of Honghe University (XJ18B01).

Author contributions

X.W., conceived, performed the analyses and wrote the paper; J.Z., performed the analyses and wrote the paper; S.L., performed the research and simulations; Y.W., analyzed the data.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to X.W.

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

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



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

© The Author(s) 2022