IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received March 19, 2020, accepted April 1, 2020, date of publication April 6, 2020, date of current version April 20, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2985713

Identifying Influential Spreaders Based on Adaptive Weighted Link Model

ZHE LI[1], TAO REN[1,2], YANJIE XU[1], BOYU CHANG[1], DONGMING CHEN[1], AND SHIXIANG SUN[1]

[1]Software College, Northeastern University, Shenyang 110819, China
[2]Liaoning Key Laboratory of Intelligent Diagnosis and Safety for Metallurgical Industry, Northeastern University, Shenyang 110819, China

Corresponding author: Tao Ren (chinarentao@163.com)

This work was supported in part by the National Natural Science Foundation of China under Grant 61473073 and Grant 61433014, and in part by the Fundamental Research Funds for the Central Universities under Grant N2017009, Grant N182608003, Grant N181706001, Grant N2018008, Grant N161702001, and Grant N171706003.

ABSTRACT Identifying influential spreaders in complex networks is crucial in understanding, controlling and accelerating spreading processes for diseases, information, innovations, behaviors, and so on. We proposed a semi-local-information-based algorithm named the adaptive weighted link model (AWLM), which classifies the links in the subgraph made up of the second-order neighbors of nodes and gives them different weights adaptively. The adaptive weighted link model is completely depends on the semi-local topological structure and thus can be calculated not only faster but also under the case where the global topology is not known, especially when the network is sparse, the time complexity is approximate linear. Empirical analyses of the Susceptible-Infected-Recovered (SIR) spreading dynamics on ten real networks show that the adaptive weighted link model always perform the best in comparison with well-known state-of-the-art methods.

INDEX TERMS Complex networks, influential spreaders, semi-local information, adaptive method.

I. INTRODUCTION

Network science is playing an increasingly significant role in many domains [1]. The heterogeneous nature of real networks [2] asks for a crucial question: How to measure a node's importance quantitatively in a dynamical process? A good answer is an efficient algorithm to identify influential spreaders in complex networks, which can help to better control the outbreak of an epidemic [3], optimize the use of limited resources to facilitate the dissemination of information [4], prevent catastrophic disruptions of power grid or the Internet [5], discover the candidates of drug target and essential proteins [6], find the important species for ecosystems [7], [8], and so on.

Till far, most known methods only make use of the structural information [9], which can be roughly classified into neighborhood-based centralities and path-based centralities. Typical representatives of neighborhood-based centralities are degree centrality [10] (DC), H-index [11], k-shell decomposition method [12] (KS) and LocalRank [13] (LR). Two well-studied path-based centralities are closeness centrality [14] (CC) and betweenness centrality [15] (BC).

Recently, some more potential methods that only use the semi-local structural information are proposed and perform much better than the above well-known state-of-the-art methods, such as Quasi-Laplacian centrality [16] (QC) and Local Gravity Model [17] (LGM). QC is defined as the drop of the Laplacian energy of the network with the deletion of the target node from the network. LGM, inspired by the gravity law, takes both neighborhood information and path information into account.

We also proposed a semi-local-information-based algorithm named the adaptive weighted link model (AWLM), which classifies the links in the subgraph made up of the second-order neighbors of nodes and gives them different weights adaptively. Empirical results show that AWLM performs always the best in comparison with the eight methods mentioned above.

The remainder of the paper is organized as follows: In Section II, the related works of methods to identify influential spreaders are introduced. In Section III, our adaptive weighted link model is proposed. The network data description and numerical results based on various centrality

The associate editor coordinating the review of this manuscript and approving it for publication was Chao-Yang Chen.

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

measures applying on real networks are shown in Section IV. Finally, conclusions are made in Section V.

II. RELATED WORKS

An undirected and unweighted network is represented by $G = (N, M)$ with N nodes and M links, and its structure can be described by an adjacent matrix $A = (a_{ij})_{N\times N}$ where $a_{ij} = 1$ if node i links to node j, and $a_{ij} = 0$ otherwise.

Many centrality measures have been proposed to rank nodes. A simple one is DC [10], which is defined as

$$DC(i) = \sum_{j} a_{ij}. \tag{1}$$

Another two famous methods just considering the local information are H-index and LR.

The H-index [11] of node i, denoted by $H(i)$, is defined as the maximal integer satisfying that there are at least $H(i)$ neighbors of node i whose degrees are all no less than $H(i)$.

LR [13] of node i is defined as

$$LR(i) = \sum_{j\in\Gamma_i} Q(j), \tag{2}$$

$$Q(j) = \sum_{k\in\Gamma_j} N(k), \tag{3}$$

where Γ_i is the set of the nearest neighbors of node i and $N(k)$ is the number of the nearest and the next nearest neighbors of node k.

There are also many well-known state-of-the-art methods considering the global information, such as KS, CC and BC.

KS [12] assigns a k-shell index to each node based on its topological location, where nodes closer to the core of the network will get higher k-shell indices, and nodes in the periphery will get lower k-shell indices.

CC [14] of node i is defined as

$$CC(i) = \frac{N-1}{\sum_{j\neq i} d_{ij}}, \tag{4}$$

where d_{ij} is the shortest distance between node i and node j.

BC [15] of node i is defined as

$$BC(i) = \sum_{s\neq i, s\neq t, i\neq t} \frac{g_{st}(i)}{g_{st}}, \tag{5}$$

where g_{st} is the number of shortest paths between nodes s and t, and $g_{st}(i)$ is the number of shortest paths between nodes s and t that pass through node i.

Furthermore, to identify the influential spreaders more effectively, some more potential methods are proposed and perform better than the well-known state-of-the-art methods, such as QC and LGM.

QC [16] of node i is defined as

$$QC(i) = k_i^2 + k_i + 2\sum_{j\in\Gamma_i} k_j, \tag{6}$$

where k_i is the degree of node i.

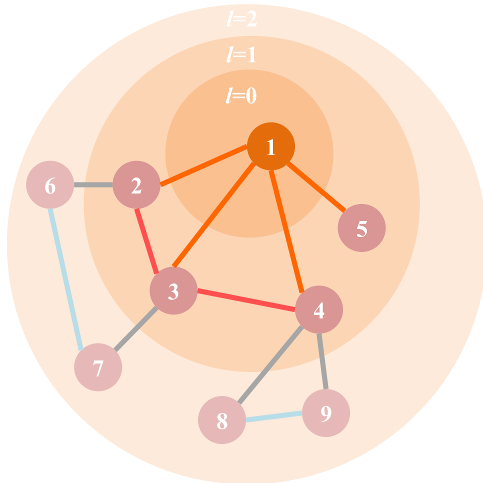


FIGURE 1. **A simple network to illustrate the proposed method.**

LGM [17] of node i is defined as

$$LGM(i) = \sum_{d_{ij}\leq R, j\neq i} \frac{k_i k_j}{d_{ij}^2}, \tag{7}$$

where R is the truncation radius. Reference [17] reveals that the optimal truncation radius, denoted by R^*, approximately scales linearly with the average distance, denoted by $\langle d\rangle$, as

$$R^* \approx \frac{1}{2}\langle d\rangle \tag{8}$$

at $\beta = \beta_c$. Such approximately linear relation also holds for other values of β not so far from β_c, where β is the infection rate and β_c is the epidemic threshold of the SIR model [18].

III. ALGORITHMS

If some node is selected as the center of the network, the other nodes are arranged in different layers according to their distance from the center node. As is shown in Figure 1, node 1 is the center node, its nearest neighbors (node 2, 3, 4, 5) occupy 1-layer, and its next nearest neighbors (node 6, 7, 8, 9) occupy 2-layer.

Initially, node 1 in the network is in the infected state and the others are in the susceptible state. According to the rules of SIR model, node 1 can infect its susceptible neighbors (node 2, 3, 4, 5) with probability β. And in next step, node 1 changes to be recovered and will never participate in the dynamics with probability λ. The spreading process repeats until there are no more infected nodes in the network. The first spread is completed by the four orange links between 0-layer and 1-layer, and the second spread is completed by the two red links in 1-layer and the four grey links between 1-layer and 2-layer. If node 1 fails to infect node 4 and there is no red link between node 3 and node 4, node 8 and node 9 can never be infected. If node 1 fails to infect node 2 and there is no red link between node 2 and node 3, the number of infected paths of node 6 is reduced from two (i.e., 1-3-2-6 and 1-3-7-6) to one (i.e., 1-3-7-6). It is well known that the farther the distance between the infected node and the susceptible node, the more difficult the infection is, and multiple infections can alleviate the problem to some extent. Based on the above two examples, it is not hard to see that the red links are extraordinarily important compared with the grey links and they should be given a greater weight. The different color links play different roles in the spread process. Therefore, the different color links should be given different weights. Based on the above analysis, this paper proposes the adaptive weighted link model (AWLM) only utilizing semi-local information. The adaptive weighted link of node i is defined as

$$awl(i)=\alpha k_i + (1-\alpha)\left[(\sum_{j\in\Gamma_i} k_j - k_i - 2R_i) + R_i(1+\frac{1}{k_i})\right], \tag{9}$$

where $R_i = (\sum_{j\in\Gamma_i, k\in\Gamma_i, j\neq k} a_{jk})/2$, Γ_i is the set of the nearest neighbor of node i, and $\alpha \in [0, 1]$. Here, we define orange links as type I links, grey links as type II links and red links as type III links, respectively. Obviously, k_i is the number of type I links, $\sum_{j\in\Gamma_i} k_j - k_i - 2R_i$ is the number of type II links, and R_i is the number of type III links. In particular, $1 + 1/k_i$ is used to amplify the contribution of type III links. When the number of type III links of two nodes is equal, compared with the neighbors of the node with large degrees, the neighbors of the node with small degrees are more closely connected, therefore, the amplification effect of the node with small degrees is more obvious. The adaptive coefficient α is used to adjust the weight of the first spread (i.e., type I links) and second spread (i.e., type II links and type III links). For a specific network, α can be obtained by traversing it (more details see Section IV). In conclusion, the weight of type I links is α, the weight of type II links is $1-\alpha$ and the weight of type III links is $(1-\alpha)(1+1/k_i)$. Finally, a node i's influence can be estimated as

$$AWLM(i) = \sum_{j\in\Gamma_i} awl(j). \tag{10}$$

Taking node 1 in Figure 1 as an example, set $\alpha = 0.3$, $awl(2) = 0.3*k_2+0.7*(k_1+k_3+k_6-k_2-2*(a_{13}+a_{31})/2+(a_{13}+a_{31})/2*(1+1/k_2)) = 5.3333$, $awl(3) = 6.4500$, $awl(4) = 5.7500$, $awl(5) = 2.4000$, therefore, $AWLM(1) = awl(2)+awl(3)+awl(4)+awl(5) = 19.9333$.

The most time-consuming operation in AWLM is to calculate R_i and its time complexity is $\langle k\rangle^2$, where $\langle k\rangle$ is the average degree. So the time complexity of AWLM is $N\langle k\rangle^2$. Especially when the network is sparse, the time complexity of AWLM is approximate linear. Therefore, AWLM is a promising method for large-scale networks to identify the influential spreaders.

IV. RESULTS

A. DATA DESCRIPTION

In this paper, ten real networks from disparate fields are used to test the performance of AWLM, including two collaboration networks (Jazz and NS), four social networks (PB, Facebook, WV and Sex), one transportation network (USAir), one communication network (Email), one infrastructure network (Power) and one technological network (Router). Jazz [19] is a collaboration network of jazz musicians. NS [20] is a co-authorship network of scientists working on network science. PB [21] is a network of US political blogs. Facebook [22] describes social circles from Facebook. WV [23] is a network of Wikipedia who-votes-on-whom. Sex [24] is a bipartite network in which nodes are females (sex sellers) and males (sex buyers) and links between them are established when males write posts indicating sexual encounters with females. USAir [25] is the United States airports transportation network. Email [26] describes email interchanges between users including faculty, researchers, technicians, managers, administrators, and graduate students of the Rovira i Virgili University. Power [27] is the power grid of the western United States. Router [28] is a symmetrized snapshot of the structure of the Internet at the level of autonomous systems. These networks' topological features are shown in Table 1, including the number of nodes, the number of links, the average degree, the clustering coefficient [27], the assortative coefficient [29], the degree heterogeneity [30] and the epidemic threshold [31] of the SIR model.

B. EMPIRICAL RESULTS

The well-known SIR model is used to compare the rankings of influences produced by algorithms and simulations. Initially, one node in the network is in the infected state (I) and the others are in the susceptible state (S). Each of the infected nodes can infect its susceptible neighbors with probability β. And in each step, every infected node changes to be recovered state (R) and will never participate in the dynamics with probability λ. The spreading process repeats until there are no more infected nodes in the network. The influence of any node i can be estimated by

$$F(i) = N_r/N, \tag{11}$$

where N_r is the number of recovered nodes at the end of the dynamics. For simplicity, λ is set to 1, and the corresponding epidemic threshold [31] is

$$\beta_c \approx \frac{\langle k\rangle}{\langle k^2\rangle - \langle k\rangle}, \tag{12}$$

where $\langle k^2\rangle$ is the second-order moment of the degree distribution.

Given a network and the infection rate β, to obtain the standard ranking of nodes' influences (i.e., the ranking of nodes' influences calculated by Eq. 11), we use 1000 independent implementations for averaging, in each implementation every node is selected once as the seed once. The accuracy of an

TABLE 1. The basic topological features of the ten real networks. $\langle k\rangle$ is the average degree. C is the clustering coefficient. r is the assortative coefficient. H is the degree heterogeneity. β_c is the epidemic threshold of the SIR model.

Networks	N	M	$\langle k\rangle$	C	r	H	β_c
Jazz	198	2742	27.6970	0.6334	0.0202	1.3951	0.0266
NS	379	914	4.8232	0.7981	-0.0817	1.6630	0.1424
PB	1222	16714	27.3552	0.3600	-0.2213	2.9707	0.0125
Facebook	4039	88234	43.6910	0.6170	0.0636	2.4392	0.0095
WV	7115	100762	28.3238	0.2089	-0.0831	5.1319	0.0069
Sex	15810	38540	4.8754	0.0000	-0.1145	5.8276	0.0365
USAir	332	2126	12.8072	0.7494	-0.2079	3.4639	0.0231
Email	1133	5451	9.6222	0.2540	0.0782	1.9421	0.0565
Power	4941	6594	2.6691	0.1065	0.0035	1.4504	0.3483
Router	5022	6258	2.4922	0.0329	-0.1384	5.5031	0.0786

TABLE 2. The algorithms' accuracies for $\beta = \beta_c$, measured by the Kendall's Tau (τ). The best performed algorithm for each network is emphasized by bold.

Networks	BC	CC	DC	H-index	KS	LR	QC	LGM	AWLM
Jazz	0.4720	0.7158	0.8174	0.8470	0.7541	0.8610	0.8738	0.8715	**0.9354**
NS	0.3004	0.3386	0.5777	0.5629	0.5118	0.8104	0.8073	0.8281	**0.9051**
PB	0.6742	0.7806	0.8511	0.8692	0.8597	0.8775	0.9001	0.9011	**0.9174**
Facebook	0.4559	0.3942	0.6819	0.7085	0.7088	0.7536	0.7907	0.8265	**0.8398**
WV	0.6971	0.8170	0.7606	0.7651	0.7648	0.8288	0.8347	0.8312	**0.8415**
Sex	0.4268	0.6061	0.4785	0.4890	0.4936	0.6513	0.6692	0.6699	**0.6717**
USAir	0.5243	0.8002	0.7355	0.7581	0.7532	0.8707	0.9062	0.8850	**0.9065**
Email	0.6199	0.8107	0.7637	0.7881	0.7695	0.8963	0.8703	0.8646	**0.9255**
Power	0.3250	0.3836	0.4261	0.3994	0.3114	0.7535	0.6218	0.7295	**0.7817**
Router	0.3101	0.6377	0.3146	0.1941	0.1814	0.7542	0.6919	0.7946	**0.8123**

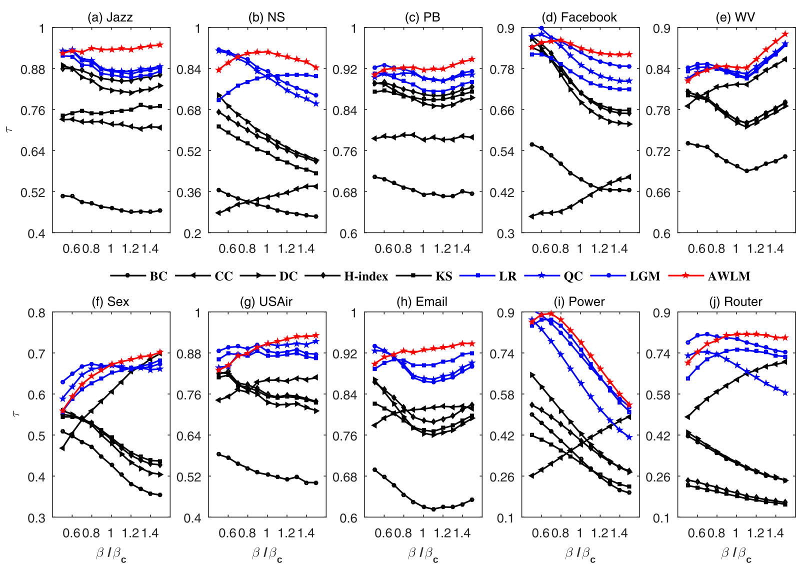


FIGURE 2. The algorithms' accuracies for different β, measured by the Kendall's Tau (τ).

algorithm is measured by the Kendall's Tau (τ) [32] between the standard ranking and the ranking by the algorithm. The Kendall's Tau is an index measuring the correlation strength between two sequences. Considering two sequences with N elements, $X = (x_1, x_2, \ldots, x_N)$ and $Y = (y_1, y_2, \ldots, y_N)$. Any pair of two-tuples (x_i, y_i) and (x_j, y_j) $(i \neq j)$ are concordant if both $x_i > x_j$ and $y_i > y_j$ or both $x_i < x_j$ and $y_i < y_j$. They are discordant if $x_i > x_j$ and $y_i < y_j$ or $x_i < x_j$ and $y_i > y_j$. If $x_i = x_j$ or $y_i = y_j$, the pair is neither concordant nor discordant. The Kendall's Tau of two sequences X and Y can be calculated as

$$\tau = \frac{2(n_+ - n_-)}{N(N-1)}, \tag{13}$$

where n_+ and n_- denote the number of concordant and discordant pairs, respectively. It can be seen that the extent to which τ exceeds zero indicates the strength of the correlation. A larger value of τ means a stronger correlation between the two sequences and thus a better performance. Table 2 compares the accuracies of the proposed algorithms (i.e., AWLM) and eight benchmark algorithms. The infection rate for each case is fixed as $\beta = \beta_c$ (for more values of β, see Figure 2) and the parameters in relevant algorithms are all adjusted to their optimal values subject to the largest τ.

As shown in Table 2, AWLM always performs the best among the nine algorithms. The results reported in Table 2 demonstrate the advantage of AWLM and show that a semi-local index (AWLM) can outperform global indices

TABLE 3. The Kendall's Tau between the standard ranking and the two rankings by AWLM with $\alpha = (1 - r)/2$ and $\alpha = 0.5$. The best results for each network is emphasized by bold.

Networks	Jazz	NS	PB	Facebook	WV	Sex	USAir	Email	Power	Router
$\alpha = 0.5$	0.9348	0.9043	0.9162	0.8393	0.8412	0.6611	0.9033	0.9251	0.7773	0.7946
$\alpha = (1 - r)/2$	**0.9348**	**0.9048**	**0.9164**	**0.8394**	**0.8413**	**0.6619**	**0.9037**	**0.9252**	**0.7782**	**0.8023**

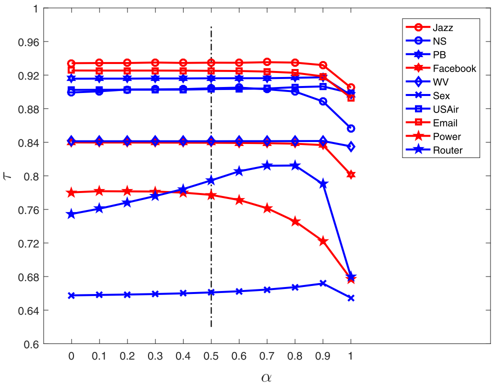


FIGURE 3. The Kendall's Tau between the standard ranking and the eleven rankings by AWLM with different α.

(i.e., BC, CC and KS). As shown in Figure 2, results for other values of β not too far from the threshold are consistent to the one at β_c, suggesting the robustness of our findings.

Notice that, the optimization process for α is as follows: α is increased from 0 to 1 by 0.1 each step and eleven rankings of nodes' influences can be obtained, then, the Kendall's Tau between the standard ranking and the eleven rankings by AWLM with different α can be calculated, finally, the α that makes the Kendall's Tau largest is selected as the optimal α, denoted by α^*. However, when network is particularly large, the calculation of the Kendall's Tau will take a lot of time. Therefore, a heuristics approach is proposed to set α. As is shown in Figure 3, we discover that α^* of the assortative network (denoted by red) is more inclined to be less than 0.5, and α^* of the disassortative network (denoted by blue) is more inclined to be more than 0.5. Therefore, α^* can be calculated by $(1 - r)/2$ approximately. In order to verify the validity of the heuristic approach, the Kendall's Tau between the standard ranking and the two rankings by AWLM with $\alpha = (1 - r)/2$ and $\alpha = 0.5$ is compared. As is shown in Table 3, it is reported that the heuristic approach indeed works. Even though it is not necessarily taking the optimal α, it is better than the fifty-fifty mode.

V. CONCLUSION

To measure influences of nodes in a certain networked dynamics, we propose a semi-local-information-based method, namely the adaptive weighted link model (AWLM), in which the links in the subgraph made up of the second-order neighbors of nodes are classified into three categories: type I links, type II links and type III links. The type I links working in the first spread are given the weight of α. The type II and type III links working in the second spread are given the weight of $1 - \alpha$. In particular, we hold that the type III links are more important than the type II links, so the weight of type III links are amplified $1 + 1/k_i$ times. In conclusion, these three link are given different weights adaptively by optimizing α. Empirical results show that AWLM performs always the best in comparison with the eight benchmark methods.

Furthermore, to improve the efficiency of the algorithm, a heuristic approach is proposed to set α instead of traversing α. Even though the heuristic approach is not necessarily taking the optimal α, it is better than the fifty-fifty mode. When network is particularly large, it will save a lot of time and obtain a competitive result.

AWLM completely depends on the semi-local topological structure and thus can be calculated not only faster but also under the case where the global topology is not known, especially when the network is sparse, the time complexity of AWLM is approximate linear. Therefore, AWLM is an extraordinary promising algorithm for real application.

Because our method is mainly for the undirected and unweighted networks, we only need to calculate the number of different types of links. If we want to extend it to the weighted networks, we need to consider the inherent weight of each link. The simplest way is to calculate the sum of the weight of the three type links, respectively, furthermore, set the three type links different weights. However, in weighted complex networks, the heterogeneity of the links greatly change their importance [33]. Our method may not work as well as it does on the undirected and unweighted network. Therefore, it is still a challenge to extend our model to the weighted complex networks.

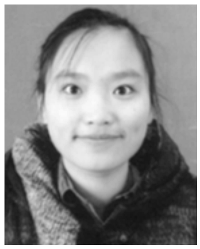
REFERENCES

[1] M. E. J. Newman, *Networks*. London, U.K.: Oxford Univ. Press, 2018, pp. 1–12.
[2] G. Caldarelli, *Scale-Free Networks: Complex Webs in Nature and Technology*. London, U.K.: Oxford Univ. Press, 2007, pp. 1–10.
[3] R. Pastor-Satorras and A. Vespignani, "Immunization of complex networks," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 65, no. 3, Feb. 2002, Art. no. 036104.
[4] F. Morone and H. A. Makse, "Influence maximization in complex networks through optimal percolation," *Nature*, vol. 524, no. 7563, pp. 65–68, Aug. 2015.
[5] R. Albert, I. Albert, and G. L. Nakarado, "Structural vulnerability of the north American power grid," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 69, no. 2, Feb. 2004, Art. no. 025103.
[6] P. Csermely, T. Korcsmaros, H. J. M. Kiss, G. London, and R. Nussinov, "Structure and dynamics of molecular networks: A novel paradigm of drug discovery: A comprehensive review," *Pharmacol. Therapeutics*, vol. 138, no. 3, pp. 333–408, Jun. 2013.
[7] M. Bellingeri and A. Bodini, "Food Web's backbones and energy delivery in ecosystems," *Oikos*, vol. 125, no. 4, pp. 586–594, Apr. 2016.
[8] M. Bellingeri, D. Cassi, and S. Vincenzi, "Increasing the extinction risk of highly connected species causes a sharp robust-to-fragile transition in empirical food Webs," *Ecological Model.*, vol. 251, pp. 1–8, Feb. 2013.
[9] L. Lu, D. Chen, X.-L. Ren, Q.-M. Zhang, Y.-C. Zhang, and T. Zhou, "Vital nodes identification in complex networks," *Phys. Rep.*, vol. 650, pp. 1–63, Sep. 2016.

[10] P. Bonacich, "Factoring and weighting approaches to status scores and clique identification," *J. Math. Sociol.*, vol. 2, no. 1, pp. 113–120, Jan. 1972.
[11] L. Lu, T. Zhou, Q.-M. Zhang, and H. E. Stanley, "The H-index of a network node and its relation to degree and coreness," *Nature Commun.*, vol. 7, no. 1, Apr. 2016, Art. no. 10168.
[12] M. Kitsak, L. K. Gallos, S. Havlin, F. Liljeros, L. Muchnik, H. E. Stanley, and H. A. Makse, "Identification of influential spreaders in complex networks," *Nature Phys.*, vol. 6, no. 11, pp. 888–893, Nov. 2010.
[13] D. Chen, L. Lü, M.-S. Shang, Y.-C. Zhang, and T. Zhou, "Identifying influential nodes in complex networks," *Phys. A, Stat. Mech. Appl.*, vol. 391, pp. 1777–1787, Feb. 2012.
[14] L. C. Freeman, "Centrality in social networks conceptual clarification," *Social Netw.*, vol. 1, no. 3, pp. 215–239, 1979.
[15] L. C. Freeman, "A set of measures of centrality based on betweenness," *Sociometry*, vol. 40, no. 1, pp. 35–41, Mar. 1977.
[16] Y. Ma, Z. Cao, and X. Qi, "Quasi-laplacian centrality: A new vertex centrality measurement based on quasi-Laplacian energy of networks," *Phys. A, Stat. Mech. Appl.*, vol. 527, Aug. 2019, Art. no. 121130.
[17] Z. Li, T. Ren, X. Ma, S. Liu, Y. Zhang, and T. Zhou, "Identifying influential spreaders by gravity model," *Sci. Rep.*, vol. 9, no. 1, Dec. 2019, Art. no. 8387.
[18] H. W. Hethcote, "The mathematics of infectious diseases," *SIAM Rev.*, vol. 42, no. 4, pp. 599–653, Dec. 2000.
[19] P. M. Gleiser and L. Danon, "Community structure in Jazz," *Adv. Complex Syst.*, vol. 6, no. 4, pp. 565–573, Dec. 2003.
[20] M. E. J. Newman, "Finding community structure in networks using the eigenvectors of matrices," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 74, no. 3, Sep. 2006, Art. no. 036104.
[21] L. A. Adamic and N. Glance, "The political blogosphere and the 2004 U.S. Election: Divided they blog," in *Proc. 3rd Int. Workshop Link Discovery (LinkKDD)*, 2005, pp. 36–43.
[22] J. McAuley and J. Leskovec, "Learning to discover social circles in ego networks," in *Proc. Neural Inf. Process. Syst. Conf. (NIPS)*, 2012, p. 1–9.
[23] J. Leskovec, D. Huttenlocher, and J. Kleinberg, "Predicting positive and negative links in online social networks," in *Proc. 19th Int. Conf. World Wide Web (WWW)*, 2010, pp. 641–650.
[24] L. E. C. Rocha, F. Liljeros, and P. Holme, "Simulated epidemics in an empirical spatiotemporal network of 50,185 sexual contacts," *PLoS Comput. Biol.*, vol. 7, no. 3, Mar. 2011, Art. no. e1001109.
[25] V. Batagelj and A. Mrvar. (2006). *Pajek Datasets*. [Online]. Available: http://vlado.fmf.uni-lj.si/pub/networks/data/
[26] R. Guimera, L. Danon, A. Díaz-Guilera, F. Giralt, and A. Arenas, "Self-similar community structure in a network of human interactions," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 68, no. 6, Dec. 2003, Art. no. 065103.
[27] D. J. Watts and S. H. Strogatz, "Collective dynamics of 'small-world' networks," *Nature*, vol. 393, no. 6684, pp. 440–442, Jun. 1998.
[28] N. Spring, R. Mahajan, D. Wetherall, and T. Anderson, "Measuring ISP topologies with rocketfuel," *IEEE/ACM Trans. Netw.*, vol. 12, no. 1, pp. 2–16, Feb. 2004.
[29] M. E. J. Newman, "Assortative mixing in networks," *Phys. Rev. Lett.*, vol. 89, no. 20, Oct. 2002, Art. no. 208701.
[30] H.-B. Hu and X.-F. Wang, "Unified index to quantifying heterogeneity of complex networks," *Phys. A, Stat. Mech. Appl.*, vol. 387, no. 14, pp. 3769–3780, Jun. 2008.
[31] C. Castellano and R. Pastor-Satorras, "Thresholds for epidemic spreading in networks," *Phys. Rev. Lett.*, vol. 105, no. 21, Nov. 2010, Art. no. 218701.
[32] M. G. Kendall, "A new measure of rank correlation," *Biometrika*, vol. 30, nos. 1–2, pp. 81–93, Jun. 1938.
[33] M. Bellingeri, D. Bevacqua, F. Scotognella, and D. Cassi, "The heterogeneity in link weights may decrease the robustness of real-world complex weighted networks," *Sci. Rep.*, vol. 9, no. 1, Dec. 2019, Art. no. 10692.

ZHE LI received the B.S. degree in information security and the M.S. degree in communication and information system from Northeastern University, Shenyang, China, in 2011 and 2015, respectively, where he is currently pursuing the Ph.D. degree. His current research interests include network science and machine learning.

TAO REN received the B.S. degree in automatic control and the M.S. and Ph.D. degrees in control theory and control engineering from Northeastern University, Shenyang, China, in 2003, 2005, and 2007, respectively. He is currently a Professor with Northeastern University. He is in charge of 14 projects, such as the National Natural Science Foundation of China, and so on. He has published more than 30 high-qualified academic articles in several high-ranking journals or conferences. His current research interests include complex networks and intelligent optimization algorithm. He has received the Chinese People's Liberation Army Science and Technology Progress Award and several other academic awards.



YANJIE XU received the B.S. degree in software engineering from Nanchang University, Nanchang, China, in 2015. She is currently pursuing the Ph.D. degree in software engineering with Northeastern University, Shenyang, China. Her current research interests include complex networks and machine learning.

BOYU CHANG is currently pursuing the bachelor's degree in software engineering with Northeastern University, Shenyang, China. Her current research interests include network science and machine learning.

DONGMING CHEN received the Ph.D. degree, in 2006. From December 2008 to December 2009, he worked as a Visiting Scholar with the Department of Information Science, University of Arkansas at Little Rock, USA. He is currently a Professor and a Ph.D. Supervisor with the Software College, Northeastern University, China. He has undertaken the National Natural Science Foundation of China, the Natural Science foundation of Liaoning Province, and other projects. His research interests include complex networks, social network analysis, and deep learning. He is a member of ACM and the IEEE Computer Society. He is also a Senior Member of the China Computer Federation (CCF) and the China Institute of Communications (CIC). He is an editor of some international journals and a program committee of some international conferences on computer and software areas.

SHIXIANG SUN received the B.S. and M.S. degrees from the School of Mathematics and Statistics Science, Ludong University, Yantai, China, in 2014 and 2018, respectively. She is currently a Doctor with Northeastern University. She has published five qualified academic articles in high-ranking journals or conferences. Her current research interests include intelligent optimization algorithm, complex networks, and stochastic control.