



Identifying influential spreaders based on indirect spreading in neighborhood

Senbin Yu^a, Liang Gao^{a,b,*}, Lida Xu^c, Zi-You Gao^a

^a Key Laboratory of Transport Industry of Big Data Application Technologies for Comprehensive Transport (Beijing Jiaotong University), Beijing 100044, China

^b Yinchuan Municipal Bureau of Big Data Management and Service, Yinchuan 750011, China

^c Beijing Key Laboratory of Bioprocess, College of Life Science and Technology, Beijing University of Chemical Technology, Beijing 100029, China

HIGHLIGHTS

- We concentrate on how to improve the ranking of the spreading influence of nodes.
- A topological measure for ranking the spreading influence of a node is proposed.
- The proposed measure relies on the indirect spread in the neighborhood of a node.
- The proposed measure outperforms the previous measures in accuracy.

ARTICLE INFO

Article history:

Received 1 April 2018

Received in revised form 3 October 2018

Available online 12 February 2019

Keywords:

Indirect spread
Spreading strength
Influential spreader
SIR model
Complex network

ABSTRACT

Identifying influential spreaders plays a crucial role in understanding and controlling the spreading processes on complex networks. Previous works mainly focus on the direct spread via edge e_{ij} from node i to node j . However, the indirect spread through an intermediate node k , i.e. the infection is spread successively via edge e_{ik} and edge e_{kj} , is a ubiquitous phenomenon in a spreading process. Considering the spreading influence of a node, an asymmetric connection strength c_{ij} is proposed, which combines the indirect infections in i 's neighborhood with the traditional direct infections. Then node i 's spreading influence s_i^c , is defined as the sum of c_{ij} among i 's neighbors. We investigate the Susceptible–Infected–Removed (SIR) model on nine real-world networks to evaluate the accuracy of s_i^c in ranking the spreading influence of nodes. The results show that s_i^c is more accurate and more robust ranking of nodes' spreading influence in general compared with the node strength s and s -shell index s_s . Our research sheds light on the mechanism that dominates the spreading strength of nodes. The indirect spread among the neighborhood effectively catches more details for ranking the node influence in the spreading process.

© 2019 Published by Elsevier B.V.

1. Introduction

Identifying influential spreaders which is a basic question in spreading process [1], such as the cascading failure in power grid [2], the disease propagation in human society [3], and the advertisement of products [4], is full of practical and theoretical significance in the studies of complex networks. Because it is crucial for developing efficient strategies to control epidemic

* Corresponding author at: Key Laboratory of Transport Industry of Big Data Application Technologies for Comprehensive Transport (Beijing Jiaotong University), Beijing 100044, China.

E-mail address: lianggao@bjtu.edu.cn (L. Gao).

transmission [5], accelerate information diffusion, promote new product and so on. From the perspective of the severity of a spreading process, the most influential spreaders are those resulting in a much larger final infected proportion of the whole system when the spread originates from them than that from other nodes [6].

Many centralities had been used to solve this problem, like degree [7], betweenness [8], closeness [9], and Katz centrality [10], etc. It is worth mentioning that Kitsak et al. argued that the topology of the network organization plays an important role, and the most efficient spreaders are those located within the core of the network as identified by the k -shell decomposition analysis [11]. However, the nodes in the same shell layer often have different influences, or worse yet, k -shell index fails in identifying influential spreaders in some networks without core structure like BA scale-free network [12]. To deal with the two conditions, some improved measures of k -shell index were proposed [13–15], such as considering the mixed degree decomposition (MDD) [16], the shortest distance to the highest k -shell [17], neighborhood coreness centrality [18], the iteration information in k -shell decomposition [19], an operator \mathcal{H} on the neighbor's degree called the H-index [20] which was originally used to measure the citation impact of a scholar or journal [21] and so on.

Besides the centrality-based algorithms, many other methods had been proposed to evaluate the influential importance of a node. By counting the steps from the original infected node to each susceptible node, a walk counting approach was proposed to identify the influential spreaders [22]. Heuristic method was also introduced to identify the influential spreaders [23–25]. For temporal networks, a dynamic-sensitive method was applied [26]. The community-based method considering both the number and the size of communities was defined. To measure the influence of nodes, it is commonly believed that the global knowledge of the network is inevitably required. Hu et al. find that any node's global influence can be accurately measured by a small characteristic number of nodes [27].

In an unweighted and undirected network, each edge is treated as equivalent in measuring node influence [28]. However, edges show different importance in measuring node influence, which should be described by weight and direction [29]. Recent research suggested that two unequal weights w_{ij} and w_{ji} should be assigned to edge e_{ij} and e_{ji} respectively, which may express different importance for the spreading from i to j and for the spreading from j to i [6]. And the strength (s_i) of node i was defined as the sum of the weight of its edges, which measured the direct contribution of i 's neighbor on i 's spreading influence. Actually, indirect infection via another organism (either a vector or an intermediate host) is a ubiquitous phenomenon in epidemic transmission. Researchers always consider the direct and indirect spreadings through environment in epidemic models to describe the propagation of disease [30–34]. In this paper, to measure the spreading influence of node more accurately, an asymmetric connection strength c_{ij} is proposed, which combines the indirect infection with the direct spreading effect in i 's neighborhood. Then node i 's spreading influence s_i^c , is defined as the sum of c_{ij} between node i and i 's neighbor. To evaluate the accuracy of s^c in ranking the spreading influence of nodes, we investigate the Susceptible–Infected–Removed (SIR) model on nine real-world networks.

The remainder of this paper is organized as follows: In Section 2, we briefly review several centrality indices which are compared in this work, and the detailed description of our method is presented. Then nine real networks are illustrated and analyzed to show accuracy and robustness for our method in Section 3. And in Section 4, a simple conclusion is summarized.

2. Methods

2.1. Measures

Given a network $G = (V, E)$, $N = |V|$ is the number of nodes, and $M = |E|$ is the number of edges. Let e_{ij} represent the edge connecting node i and node j , and Γ_i denotes the set of neighbor nodes of node i . The degree k_i is the number of links node i carries.

k -shell index (k_s), which is also called as *coreness*, can be obtained by k -core decomposition [35]. Kitsak et al. [11] argued that k -shell index (k_s) is more significant than degree (k) in identifying influential spreaders under single seed SIR spreading process.

Recently, Liu et al. [6] proposed s -shell index (s_s), which is obtained by weighted core (s -core) decomposition. They define the weight w_{ij} of an edge e_{ij} as

$$w_{ij} = 1 + (k_i k_j^{\text{out}})^{\alpha}, \quad (1)$$

where k_j^{out} is the number of edges $\{e_{jl} | l \notin \Gamma_i \text{ and } l \neq i\}$, and the parameter α serves to tune the contribution of $k_i k_j^{\text{out}}$ to the importance of edge e_{ij} . Node l lies on the outside of Γ_i , i.e. node l is a blank circle node in Fig. 1. As shown in Fig. 1, $w_{ij} = 1 + (2 \times 3)^{\alpha}$. This definition of edge weight describes to what extent the infection spreading from i to j would reach out to the nodes on the outside of the neighbors Γ_i . And the strength s_i of node i is defined as

$$s_i = \sum_{j \in \Gamma_i} w_{ij}. \quad (2)$$

The s_n -core of a weighted network consists of the subset of nodes i with strength $s_i > s$. The process iteratively prunes all nodes i with strength $s_i \leq s$. To achieve s_{n+1} -core, $s \equiv s_m = \min_{i \in s_n\text{-core}} s_i$. Then the above pruning process is repeated until all rest nodes with strength $s_i > s (= s_m)$. A node belongs to the s_n -core but not to the s_{n+1} -core is assigned an s -shell index $s_s = s (= s_m)$. They show that node strength s and s -shell index s_s generally outperform the rankings based on the degree k and the k -shell index k_s .

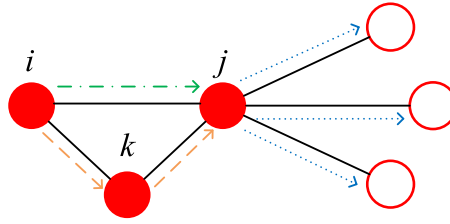


Fig. 1. Sample network. The solid circle nodes are the neighborhood of node i . The blank circle nodes are the nodes outside i 's neighborhood.

To rank the spreading influence of nodes more accurate by measures, the infection spreads among i 's neighborhood is an important step at the beginning of a spreading process. Actually, the infection spreading from node i to node j will include two main types: (1) direct spread from node i to node j by edge e_{ij} (green dash dot arrow in Fig. 1); (2) indirect spread from node i to node j by i 's neighbor node k , which means the spreading by path $i \rightarrow k \rightarrow j$ (orange dash arrow path in Fig. 1). There can be longer spreading path from node i to node j by more than one i 's neighbor nodes, but the influence on the spreading strength of node i will be too weak to take into account [36,37]. Therefore, in this paper, we will only consider the two-step indirect spreading to improve the accurate in ranking nodes' importance for spreading.

Considering a spreading process originating from node i , once node j is infected, the infection could be spread to the blank circle nodes, which lie on the outside of i 's neighborhood (Fig. 1). The number of j 's out-linking edges (blue dot line arrows) is denoted by k_j^{out} [6]. For example, in Fig. 1, $k_j^{\text{out}} = 3$ and $k_k^{\text{out}} = 0$.

To take the direct and indirect infection spreading into account, we define spreading strength c_{ij} as

$$c_{ij} = 1 + k_j^{\text{out}} \left[1 + \frac{|D_{ij,2}|}{2^2} \right]^\alpha, \quad (3)$$

where $|D_{ij,2}|$ is the number of paths from node i and node j whose length are 2, and the parameter α serves to tune the contribution of $1 + \frac{|D_{ij,2}|}{2^2}$ to the direct and indirect infection spreading from node i to node j . In Fig. 1, $|D_{ij,2}| = 1$ and $c_{ij} = 1 + 3 \times 1.25^\alpha$. The c_{ij} measures the strength of the direct and indirect spreading, which originates from node i and spreads via node j to other parts of the network, the outside of Γ_i . Note that c_{ij} and c_{ji} are describing the spreading strength of different directions, and they are generally unequal. Even if c_{ij} and c_{ji} sometimes have the same value, they may be still different meanings. The spreading strength of node i is defined as

$$s_i^c = \sum_{j \in \Gamma_i} c_{ij}. \quad (4)$$

2.2. Single seed SIR model

The single seed SIR model is applied to study the spreading process. In the single seed SIR model, all nodes are initially in the susceptible state (S) except for the only one seed node in the infectious state (I). At each time step, the I state nodes infect their susceptible neighbors with probability β and then enter the recovered state (R), where they become immunized and cannot be infected again. This process is repeated until there are no infected nodes in the network. The number of recovered nodes gives the final infected scope of the seed node, which is adopted to measure the spreading influence of the seed node. The higher β , the larger population infected, no matter from where it originates. In this paper, the infection probability β is set slightly larger than the critical infection probability $\beta_c \sim \langle k \rangle / \langle k^2 \rangle$ [38].

Experiments of the single seed SIR model are carried on nine real networks (Table 1). The spreading influence of each node is measured by averaging over 200 simulations. And we get the rank vector of SIR model R_{SIR} .

2.3. Evaluation methods

We employ the single seed SIR model to simulate the spreading process, and evaluate the performance of measures, where the influence of nodes is denoted by R_{SIR} . We apply the Kendall τ_b correlation coefficient to quantify the accuracy of each measure.

Kendall τ_b correlation coefficient [39] is adopted to evaluate the consistency between two rankings. Given R_μ , the rank vector of a measure μ , and R_{SIR} , that of the single seed SIR model, the Kendall τ_b correlation coefficient is defined as

$$\tau_b(R_\mu, R_{\text{SIR}}) = \frac{n_c - n_d}{\sqrt{(n_0 - n_1)(n_0 - n_2)}}, \quad (5)$$

where n_c is the number of concordant pairs, n_d is the number of discordant pairs. And $n_0 = n(n-1)/2$, $n_1 = \sum_i t_i(t_i-1)/2$, $n_2 = \sum_j u_j(u_j-1)/2$, where n is the size of rank vectors and t_i and u_j are the number of tied values in the i th and j th group of ties, respectively. Since all measures are evaluated by R_{SIR} , $\tau_b(R_\mu, R_{\text{SIR}})$ will be denoted by $\tau_b(\mu)$ for short.

Table 1

Properties of the nine real-world networks studied in this work. Structural properties of different networks include number of nodes (N), number of edges (M), epidemic threshold ($\beta_c \sim \langle k \rangle / \langle k^2 \rangle$), infection probability (β), average degree ($\langle k \rangle$), maximum degree (k_{max}), optimal α for node strength and s -shell index (α_{opt}) [6] and optimal α for spreading strength (α_{opt}^{sc}).

Name	N	M	β_c	β	$\langle k \rangle$	k_{max}	α_{opt}	α_{opt}^{sc}
Facebook	4039	88234	0.00938	0.02	43.6910	1045	1.1	0.2
PGP	10680	24316	0.0529	0.16	4.5536	205	2.0	1.0
Protein interaction	1870	2203	0.1546	0.30	2.3561	56	1.5	1.7
Guntella08	6301	20777	0.0566	0.08	6.5948	97	2.2	0.0
GrQc	5242	14496	0.0593	0.20	5.5307	81	1.6	1.0
CondMat	23133	93497	0.0453	0.05	8.0830	279	1.5	1.0
HepTh	9877	25998	0.0798	0.16	5.2644	65	1.5	0.7
US Air	332	2126	0.0225	0.04	12.8072	139	1.0	1.6
Power grid	4941	6594	0.2583	0.30	2.6691	19	1.7	2.0

The **imprecision function** [11] is also employed to quantify the accuracy of each measure in identifying the most influential spreaders. Given a fraction $p \in [0, 1]$, $V_{eff}(p)$ is the pN most efficient spreaders, and $V_\mu(p)$ is the pN nodes with the highest value of measure μ . Their average spreading scope is denoted by $S_{eff}(p)$ and $S_\mu(p)$, respectively. Then the imprecision function is defined as

$$\epsilon_\mu(p) = 1 - \frac{S_\mu(p)}{S_{eff}(p)}. \quad (6)$$

The smaller ϵ_μ is, the higher the accuracy of measure μ in identifying the most influential spreaders will be. The measure μ , discussed in this work, could be node strength s , s -shell index s_s [6], or spreading strength s^c .

3. Results

We evaluate the accuracy and the robustness of three measures (node strength s , s -shell index s_s [6], and spreading strength s^c) on nine real-world networks. They are two social networks (Facebook [40] and PGP [41]), one protein interaction network (Protein [42]), one file sharing network (Gnutella08 [43,44]), three collaboration networks (GrQc [43], CondMat [43], HepTh [43]), one transportation network (US Air [45]), and one power grid network (Power Grid [46]). And their basic topological features are summarized in Table 1.

3.1. Accuracy

In Fig. 2, when $\alpha = 0$, $s_i = 2k_i$, s_i is functionally equivalent to degree k_i , and the s -shell of node i is functionally equivalent to k -shell index k_s . s_s works better than s in identifying influential spreaders, which can be validated in Fig. 2 except for Power Grid. When $\alpha > 0$, τ_b of s , s_s and s^c are all larger than τ_b at $\alpha = 0$. Note that $\tau_b(s^c)$ is larger than $\tau_b(s)$ and $\tau_b(s_s)$ for $\alpha > 0$, which means s^c ranks nodes more accurate and outperforms other measures. In the US Air and Power Grid, s^c outperforms the other two measures for different α in ranking nodes' spreading importance. As α increases, τ_b of the three measures approach to the same value in the GrQc, Facebook, HepTh, PGP, and CondMat. Fig. 2 also shows that an optimal α value of each network exists, where τ_b gets its maximum for s , s_s and s^c , respectively. At the optimal α values, s^c also performs better than s and s_s in all nine networks.

The imprecision functions of s , s_s , and s^c are compared in Fig. 3. Recall that the lower the imprecision is, the higher the accuracy in identifying the most influential spreaders will be. The imprecision of s , s_s , and s^c is less than 0.1 for all fraction p in nearly all cases. The s and s_s always have the same performance in the measure of imprecision function. More noticeable is that the s^c performs even better than the other two measures in most cases, except at some smaller values of p in Protein and Facebook.

There are two possible reasons to answer why the s^c is superior to s and s_s . First, s_i^c considers the indirect spreading inside the neighborhood of a node i before the infection be spread out of i 's neighborhood, which has taken more detailed spreading structure information into consideration. Second, we use the spreading paths in the Eq. (3) to identify the spreading strength. So the summation of spreading strength can be more objective and realistic to reflect the spreading process.

3.2. Robustness

In the preliminary analysis, we utilize the optimal value of α to evaluate s , s_s and s^c , and compare ranking results in SIR model with a probability β by imprecision function. However, the optimal value of α and the spreading probability β are not often known precisely in real applications. So it will be useful to examine the performance of s , s_s and s^c for some arbitrarily chosen value of α and various spreading probability β . First, we set $\alpha = 0.5$ so that the term $(k_i k_j^{out})^\alpha$ in w_{ij} represents a geometric mean which tested in Ref. [6]. s^c obtain a lower imprecision than the other two measures in Fig. 4.

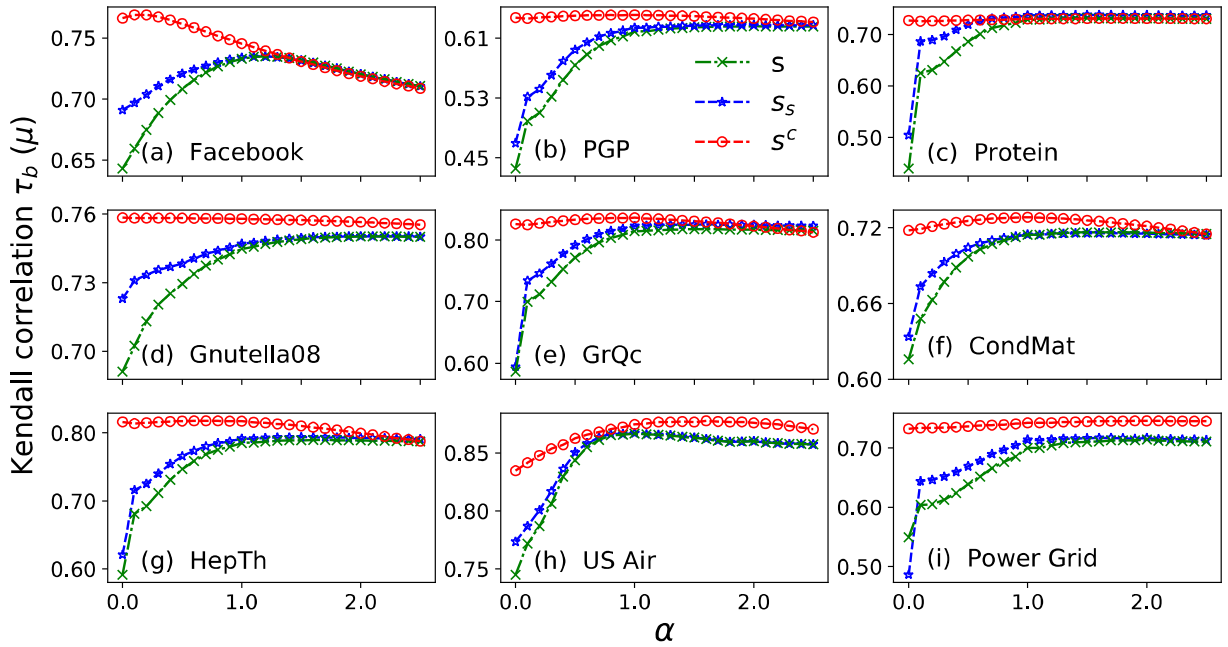


Fig. 2. Kendall correlation coefficient $\tau_b(\mu)$ as a function of α for real-world networks. μ : node strength (s) in green cross, s -shell index (s_s) in blue star, and spreading strength (s^c) in red circle.

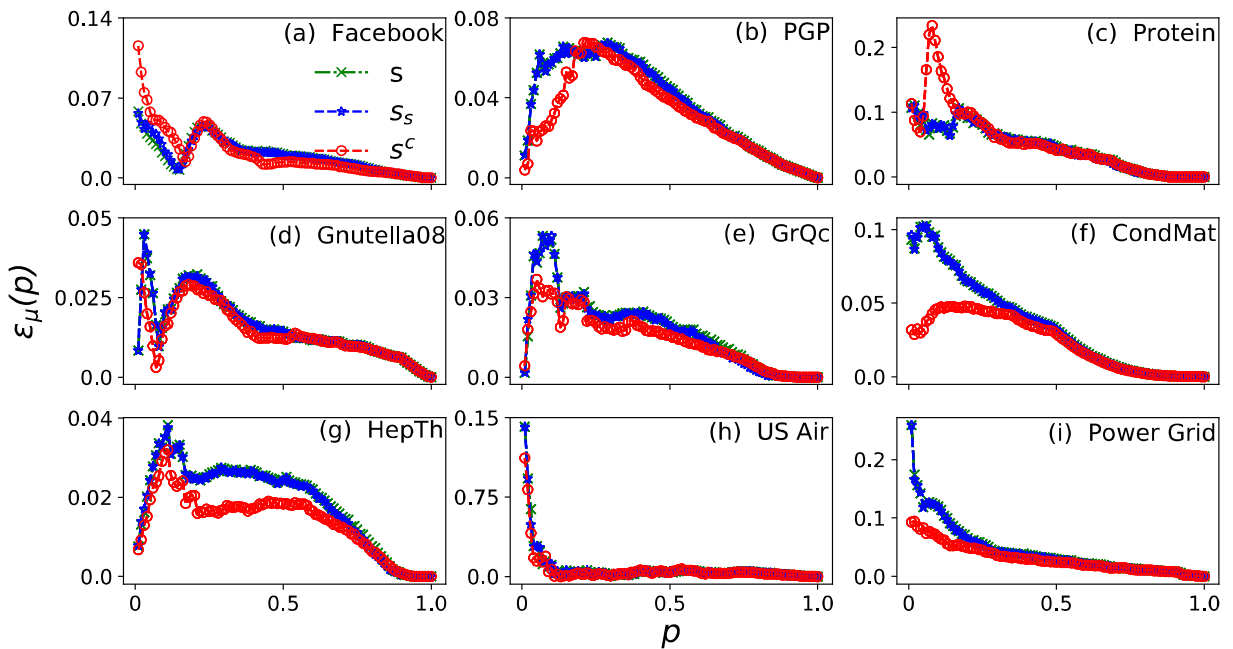


Fig. 3. The imprecision function $\epsilon_\mu(p)$ as a function of node fraction $p \in [0.01, 1.0]$ at α_{opt}^μ (optimal alpha) for each real-world network.

Fig. 5 compares the effectiveness of s , s_s and s^c for $\alpha = 1$. s^c works better again in most cases, which is similar to those results in Figs. 3 ($\alpha = \alpha_{opt}^\mu$) and 4 ($\alpha = 0.5$). The results indicate that s^c does not need α_{opt}^μ to find the most influential spreaders, and consolidate that the spreading strength is a better measure for identifying influential spreaders.

Next, we investigate the Kendall correlation coefficient $\tau_b(\mu)$ in three real-world networks. The calculations, to evaluate the effect of infection probability β , are shown in Fig. 6. The s^c exhibits obviously correctness on a wide range of probabilities β in three real networks, especially when $\beta > \beta_c$. When β is very small, the spreading is typically confined to the neighborhood of the initially infected node, hence the node with larger degree can infect more nodes [47]. The s and s_s

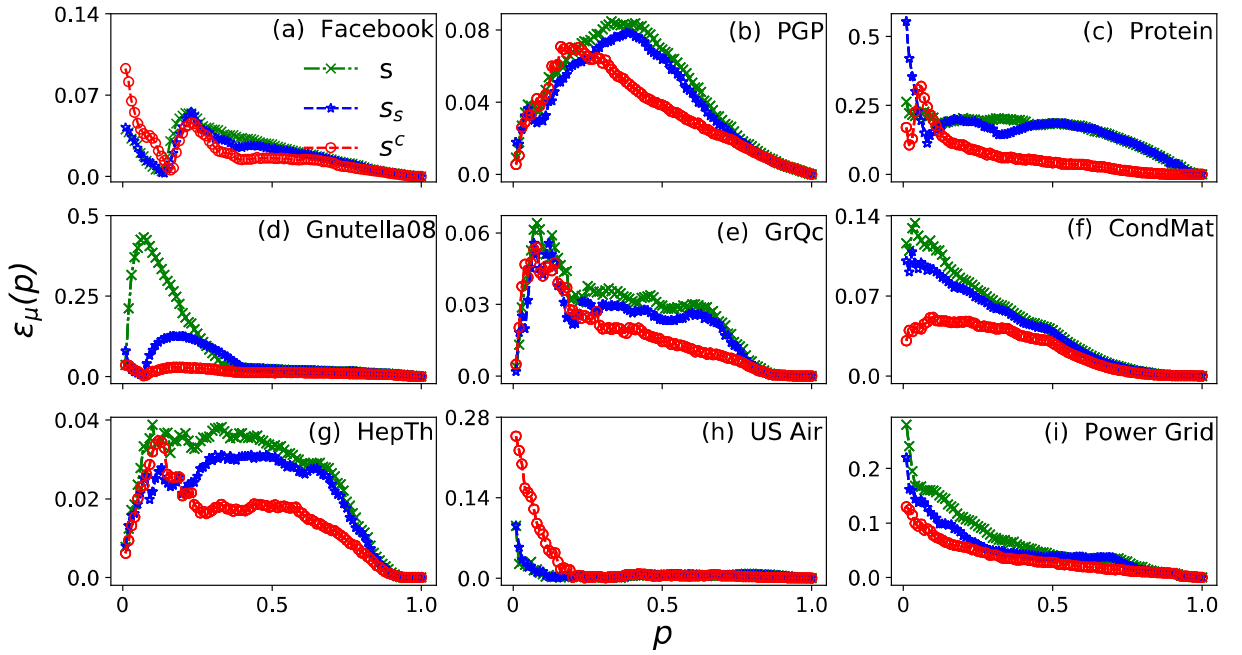


Fig. 4. The imprecision function $\epsilon_{\mu}(p)$ as a function of node fraction $p \in [0.01, 1.0]$ at $\alpha = 0.5$ for each real-world network.

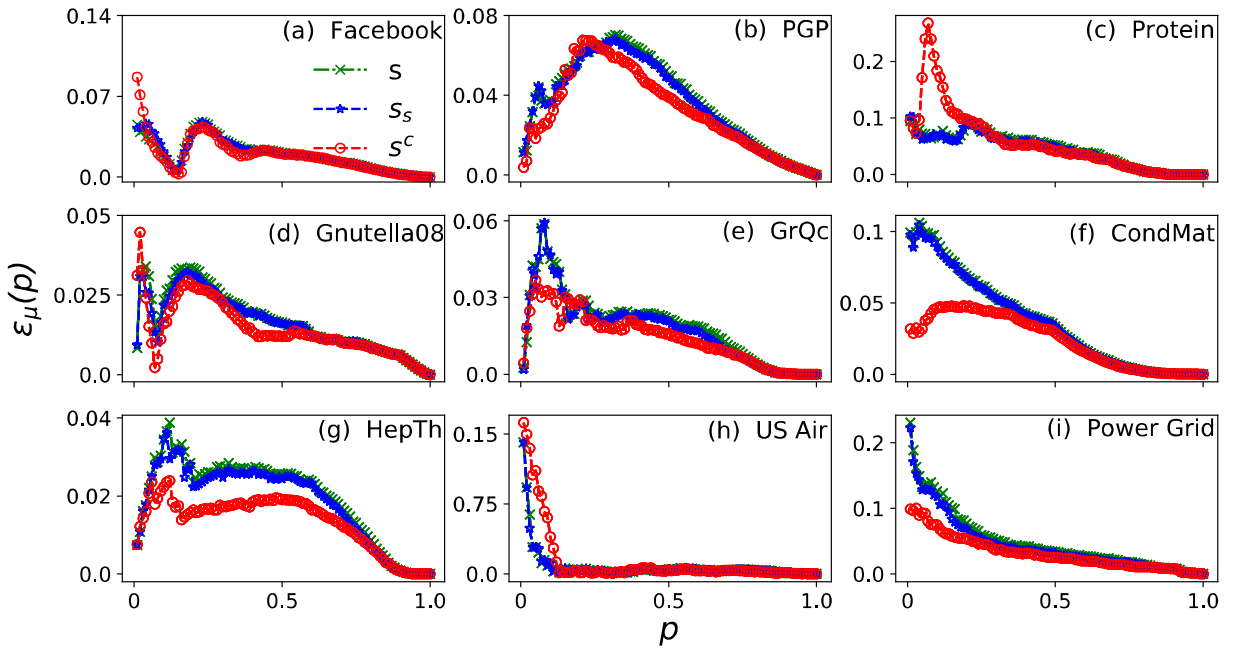


Fig. 5. The imprecision function $\epsilon_{\mu}(p)$ as a function of node fraction $p \in [0.01, 1.0]$ at $\alpha = 1.0$ for each real-world network.

are based on degree of each node, so they achieve the largest τ values when $\beta < \beta_c$ in Protein, GrQc and Power grid. When the value of β becomes larger, spreading strength begins to show its better performance. It is worth to be mentioned that the Kendall correlation coefficient becomes nearer with the increase of α . So, the above-mentioned results demonstrated that the s^c is a better indicator to identify the spreading influence in complex networks.

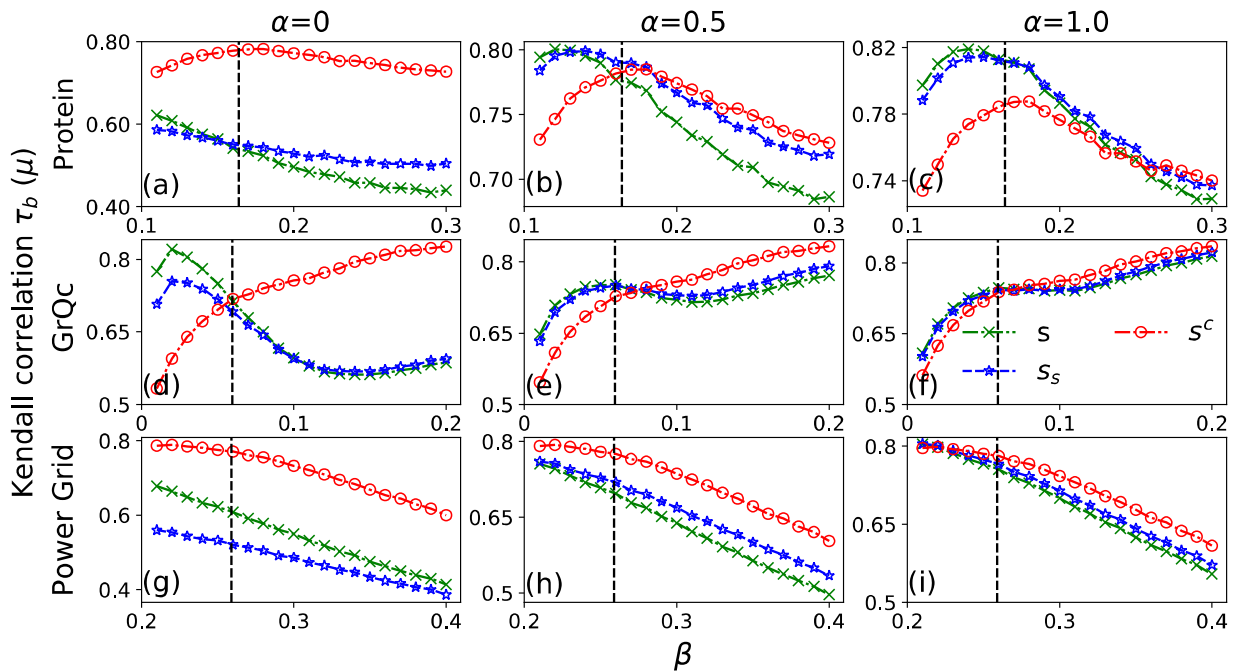


Fig. 6. Kendall correlation coefficient $\tau_b(\mu)$ as a function of the infection probability β for Protein (upper row), GrQc (middle row) and Power Grid (bottom row). α : 0 (left column), 0.5 (middle column), and 1.0 (right column). The dash line in each sub-figure is showing the critical infection rate β_c .

4. Conclusions

In this paper, we define spreading strength s^c , which simultaneously takes the direct and the indirect infection spreading from infected node to all its neighbor nodes into account. The present analyses provide convincing evidence that the spreading strength s^c is an excellent predictor for identifying influential spreaders. The novel defined s^c gets not only more accurate but also more robust ranking of nodes' spreading influence, since it brings the indirect spreading in the two-step neighborhood of a node into consideration. Although we focus on the indirect spread in the two-step neighborhood, it is worth considering the indirect spread in the higher step neighborhood in the future, which maybe a direction for improving the accuracy of ranking nodes' importance in epidemic spreading.

In addition, it should be noticed that in our current work, we have chosen the single seed SIR model to describe the spreading process. This is due to the consideration that it fits well to identify the influential spreaders. Nevertheless, it would also be interesting to examine the spreading strength s^c in identifying the top K influential spreaders under multiple seeds SIR model. It would be a more complex and more challenging topic in the future.

Acknowledgments

The authors thank for support from the National Natural Science Foundation of China (No. 71571017, No. 91646124, No. 71621001, and No. 91746201), and support from the Fundamental Research Funds for the Central Universities, China (2017YJS102, 2018JBM026).

References

- [1] G.F. de Arruda, F.A. Rodrigues, Y. Moreno, Fundamentals of spreading processes in single and multilayer complex networks, *Phys. Rep.* (2018) <https://doi.org/10.1016/j.physrep.201806007>.
- [2] R. Albert, I. Albert, G.L. Nakarado, Structural vulnerability of the north american power grid, *Phys. Rev. E* 69 (2004) 025103.
- [3] K.I. Goh, M.E. Cusick, D. Valle, B. Childs, M. Vidal, A. Barabási, The human disease network, *Proc. Natl. Acad. Sci. USA* 104 (2007) 8685–8690.
- [4] J. Goldenberg, B. Libai, E. Muller, Talk of the network: A complex systems look at the underlying process of word-of-mouth, *Mark. Lett.* 12 (2001) 211–223.
- [5] L.L. Ma, C. Ma, H.F. Zhang, B.H. Wang, Identifying influential spreaders in complex networks based on gravity formula, *Physica A* 451 (2016) 205–212.
- [6] Y. Liu, M. Tang, Y. Do, P.M. Hui, Accurate ranking of influential spreaders in networks based on dynamically asymmetric link weights, *Phys. Rev. E* 96 (2017) 022323.
- [7] L. Freeman, Centrality in social networks conceptual clarification, *Social Networks* 1 (1978) 215–239.
- [8] L.C. Freeman, A set of measures of centrality based on betweenness, *Sociometry* 40 (1977) 35–41.
- [9] L. Katz, A new status index derived from sociometric analysis, *Psychometrika* 18 (1953) 39–43.
- [10] G. Sabidussi, The centrality index of a graph, *Psychometrika* 31 (1966) 581–603.

- [11] M. Kitsak, L.K. Gallos, S. Havlin, F. Liljeros, L. Muchnik, H.E. Stanley, H.A. Makse, Identification of influential spreaders in complex networks, *Nat. Phys.* 6 (2010) 888–893.
- [12] A.L. Barabási, R. Albert, Emergence of scaling in random networks, *Science* 286 (1999) 509–512.
- [13] Y. Liu, M. Tang, T. Zhou, Y. Do, Improving the accuracy of the k-shell method by removing redundant links: From a perspective of spreading dynamics, *Sci. Rep.* 5 (2015) 13172.
- [14] Y. Liu, M. Tang, T. Zhou, Y. Do, Core-like groups result in invalidation of identifying super-spreader by k-shell decomposition, *Sci. Rep.* 5 (2015) 9602.
- [15] Z.X. Wang, C.J. Du, J.P. Fan, Y. Xing, Ranking influential nodes in social networks based on node position and neighborhood, *Neurocomputing* 260 (2017) 466–477.
- [16] A. Zeng, C.J. Zhang, Ranking spreaders by decomposing complex networks, *Phys. Lett. A* 377 (2013) 1031–1035.
- [17] J.G. Liu, Z.M. Ren, Q. Guo, Ranking the spreading influence in complex networks, *Physica A* 392 (2013) 4154–4159.
- [18] J. Bae, S. Kim, Identifying and ranking influential spreaders in complex networks by neighborhood coreness, *Physica A* 395 (2014) 549–559.
- [19] Z.X. Wang, Y. Zhao, J.K. Xi, C.J. Du, Fast ranking influential nodes in complex networks using a k-shell iteration factor, *Physica A* 461 (2016) 171–181.
- [20] L. Lu, T. Zhou, Q.M. Zhang, H.E. Stanley, The H-index of a network node and its relation to degree and coreness, *Nature Commun.* 7 (2016) 10168.
- [21] J.E. Hirsch, An index to quantify an individual's scientific research output, *Proc. Natl. Acad. Sci. USA* 102 (2005) 16569–16572.
- [22] F. Bauer, J.T. Lizier, Identifying influential spreaders and efficiently estimating infection numbers in epidemic models: A walk counting approach, *Epl* 99 (2012) 367–372.
- [23] X. Wang, Y. Su, C. Zhao, D. Yi, Effective identification of multiple influential spreaders by DegreePunishment, *Physica A* 461 (2016) 238–247.
- [24] J.X. Zhang, D.B. Chen, Q. Dong, Z.D. Zhao, Identifying a set of influential spreaders in complex networks, *Sci. Rep.* 6 (2016) 27823.
- [25] Z.K. Bao, J.G. Liu, H.F. Zhang, Identifying multiple influential spreaders by a heuristic clustering algorithm, *Phys. Lett. A* 381 (2017) 976–983.
- [26] Z. Zhao, X. Wang, W. Zhang, Z. Zhu, A community-based approach to identifying influential spreaders, *Entropy* 17 (2015) 2228–2252.
- [27] Y. Hu, S. Ji, Y. Jin, L. Feng, H.E. Stanley, S. Havlin, Local structure can identify and quantify influential global spreaders in large scale social networks, *Proc. Natl. Acad. Sci. USA* 115 (2018) 7468–7472.
- [28] L.Y. Lu, D.B. Chen, X.L. Ren, Q.M. Zhang, Y.C. Zhang, T. Zhou, Vital nodes identification in complex networks, *Phys. Rep.* 650 (2016) 1–63.
- [29] J. Garcia-Algarra, J.M. Pastor, J.M. Iriondo, J. Galeano, Ranking of critical species to preserve the functionality of mutualistic networks using the k-core decomposition, *PeerJ* 5 (2017) e3321.
- [30] C. Wolf, M. Langlais, F. Sauvage, D. Pontier, A multi-patch epidemic model with periodic demography, direct and indirect transmission and variable maturation rate, *Math. Popul. Stud.* 13 (2006) 153–177.
- [31] Y.N. Xiao, N.P. French, R.G. Bowers, D. Clancy, Pair approximations and the inclusion of indirect transmission: Theory and application to between farm transmission of Salmonella, *J. Theoret. Biol.* 244 (2007) 532–540.
- [32] M.W. Allerson, C.J. Cardona, M. Torremorell, Indirect transmission of influenza a virus between pig populations under two different biosecurity settings, *PLoS One* 8 (2013) e67293.
- [33] F. Brauer, A new epidemic model with indirect transmission, *J. Biol. Dyn.* 11 (2017) 285–293.
- [34] N. Ganguly, T. Krueger, A. Mukherjee, S. Saha, Epidemic spreading through direct and indirect interactions, *Phys. Rev. E* 90 (2014) 032808.
- [35] S.B. Seidman, Network structure and minimum degree, *Social Networks* 5 (1983) 269–287.
- [36] Y. Liu, M. Tang, T. Zhou, Y. Do, Identify influential spreaders in complex networks, the role of neighborhood, *Physica A* 452 (2016) 289–298.
- [37] D.B. Chen, L.Y. Lu, M.S. Shang, Y.C. Zhang, T. Zhou, Identifying influential nodes in complex networks, *Physica A* 391 (2012) 1777–1787.
- [38] C. Castellano, R. Pastor-Satorras, Thresholds for epidemic spreading in networks, *Phys. Rev. Lett.* 105 (2010) 218701.
- [39] M.G. Kendall, A new measure of rank correlation, *Biometrika* 30 (1938) 81–93.
- [40] J. McAuley, J. Leskovec, Learning to discover social circles in ego networks, in: *Proceedings of the 26th International Conference on Neural Information Processing Systems*, Lake Tahoe, USA, 2012, pp. 539–547.
- [41] M. Boguna, R. Pastor-Satorras, A. Diaz-Guilera, A. Arenas, Models of social networks based on social distance attachment, *Phys. Rev. E* 70 (2004) 056122.
- [42] H. Jeong, S.P. Mason, A.L. Barabási, Z.N. Oltvai, Lethality and centrality in protein networks, *Nature* 411 (2001) 41–42.
- [43] J. Leskovec, J. Kleinberg, C. Faloutsos, Graph evolution: Densification and shrinking diameters, *ACM Trans. Knowl. Discov. Data (TKDD)* 1 (2007) 2.
- [44] M. Ripeanu, I. Foster, A. Iamnitchi, Mapping the Gnutella network: Properties of large-scale Peer-to-Peer systems and implications for system design, *Comput. Sci.* 6 (2002) 2002.
- [45] V.M. Batageli, A. Pajek, Datasets, available at: <http://vlado.fmf.uni-lj.si/pub/networks/data/> (accessed 04.05.17).
- [46] D.J. Watts, S.H. Strogatz, Collective dynamics of small-world networks, *Nature* 393 (1998) 440–442.
- [47] P. Holme, Three faces of node importance in network epidemiology: Exact results for small graphs, *Phys. Rev. E* 96 (2017) 062305.