

The fastest spreader in SIS epidemics on networks

Zhidong He^a and Piet Van Mieghem

Faculty of Electrical Engineering, Mathematics and Computer Science, Delft University of Technology,
2628 CD Delft, The Netherlands

Received 20 December 2017 / Received in final form 13 February 2018

Published online 2 May 2018

© The Author(s) 2018. This article is published with open access at [Springerlink.com](https://www.springerlink.com)

Abstract. Identifying the fastest spreaders in epidemics on a network helps to ensure an efficient spreading. By ranking the average spreading time for different spreaders, we show that the fastest spreader may change with the effective infection rate of a SIS epidemic process, which means that the time-dependent influence of a node is usually strongly coupled to the dynamic process and the underlying network. With increasing effective infection rate, we illustrate that the fastest spreader changes from the node with the largest degree to the node with the shortest flooding time. (The flooding time is the minimum time needed to reach all other nodes if the process is reduced to a flooding process.) Furthermore, by taking the local topology around the spreader and the average flooding time into account, we propose *the spreading efficiency* as a metric to quantify the efficiency of a spreader and identify the fastest spreader, which is adaptive to different infection rates in general networks.

1 Introduction

Identifying the most influential initial spreaders in a network constitutes a basic endeavor in network science, which helps to optimize the utility of resources and to ensure an efficient diffusion [1]. Injecting information in the fastest spreaders results in the most efficient spreading performance. The knowledge of the fastest spreader can be applied in direct marketing [2] or idea spreading [3], where the resources are limited to start the spreading with a small number of spreaders.

Many topological metrics have been proposed to measure the influence of nodes in networks [4], such as degree, betweenness, closeness [5], eigenvector centrality [6] and the square eigenvector component [7]. Kitsak et al. [8] suggest that coreness constitutes a better topological descriptor to identify influential spreaders in epidemics [8]. However, many nodes performing differently in a spreading process may have the same k -core value. Therefore, new metrics based on the existing centrality are proposed to improve the identification of the influential nodes by coreness [9,10]. Considering removing the nodes causing the biggest drop in the energy function, Morone and Makse [11] propose the metric of collective influence through optimal percolation, which performs well in locally tree-like networks. Van Mieghem et al. [12] propose that the best conduction node in a resistor network is the minimizer of the diagonal elements of the pseudoinverse matrix Q^\dagger of the weighted Laplacian matrix of the graph.

In the Susceptible-Infected-Removed (SIR) model [13], Šikić et al. [14] show that the ranking of nodal influences is sensitive to the spreading dynamics, which depends on the

infection rate and the curing rate. Measured by the cumulative infection probabilities of nodes, the degree centrality can better identify influential spreaders when the spreading rate is very small. However, the eigenvector centrality performs better when the spreading rate is close to the epidemic threshold [15]. Holme [16] discovers similar results and proposes an exact method to identify the best spreaders for influence maximization (the expected outbreak size) in the SIR model, but the method is only tractable in small graphs. In the Susceptible-Infected-Susceptible (SIS) model, Qu et al. [17] unveil that the ranking of nodal metastable infection probability also changes with the effective infection rate.

The “influence” of the spreader in the SIS model is not well defined. In this paper, we confine ourselves to the *spreading time* $T_m(i)$, defined as the time [18] when the number of infected nodes in the metastable state is first reached, started with one initially infected node i . The spreading time of an epidemic process generally determines the preferred period to take immunization actions to eradicate the spreading [19]. We investigate the average spreading time $E[T_m(i)]$ to identify *the fastest spreader* in an SIS epidemic on a general network.

This paper is organized as follows. Section 2 introduces the spreading time and shows that the average spreading time depends on the topological metrics in an ER random graph. Section 3 shows that the fastest spreader changes with the dynamic process in SIS epidemics. Further, we propose the spreading efficiency to identify the fastest spreader. We show the performance in four artificial and real networks in Section 4. Finally, we conclude our results in Section 5.

^a e-mail: Z.He@tudelft.nl

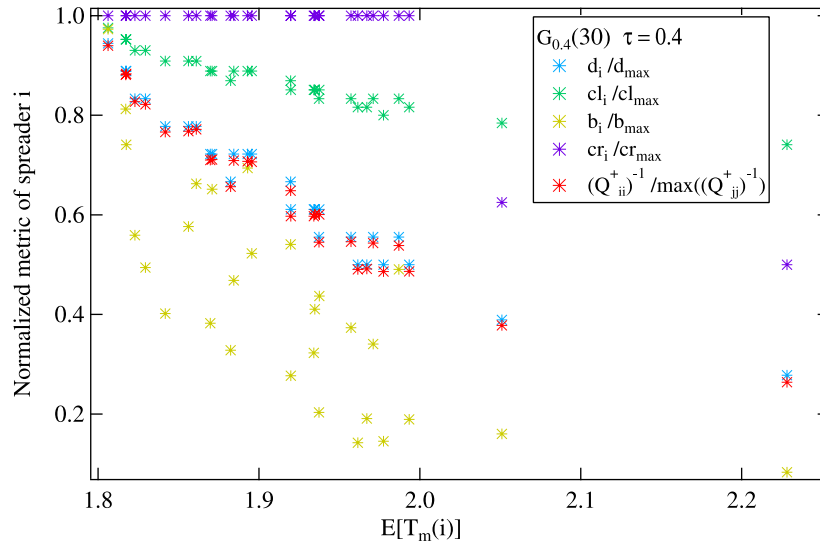


Fig. 1. The normalized topological metrics of the initial spreader i , e.g., the degree d_i , the betweenness b_i , the closeness c_l_i , the coreness cr_i and the reciprocal of the diagonal element (Q_{ii}^+) of the pseudoinverse matrix Q^+ , versus the average spreading time $E[T_m(i)]$ in an connected ER random graph $G_{0.4}(30)$ with $N = 30$ nodes and link density $p = 0.4$. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

2 The spreading time in epidemics on networks

We concentrate on the Markovian SIS epidemics [20] on networks, where both the curing and infection processes are Poisson processes. In the SIS epidemics model on a network G with N nodes and L links, the ratio between the infection rate β and the curing rate δ is called the effective infection rate $\tau = \beta/\delta$. The SIS model features a phase transition [21] around the epidemic threshold τ_c . Viruses with an effective infection rate τ above the epidemic threshold τ_c can infect a sizable portion of the population and stay for a long time in the network. A first-order mean-field approximation of the epidemic threshold $\tau_c^{(1)} = 1/\lambda_1$, where λ_1 is the spectral radius of the adjacency matrix A of the network G , was shown to be a lower bound for the epidemic threshold [20]. We denote by $x = \tau/\tau_c^{(1)}$ the normalized effective infection rate.

The spreading time $T_m(i)$ of the Markovian SIS process resembles a lognormal-like distribution with deep tails [18]. The average spreading time $E[T_m(i)]$ approximates the average hitting time when the average fraction y_∞ of infected nodes in the metastable state is reached. Physically, the spreading time $T_m(i)$ describes the spreading velocity in the early stage of the spreading process, which depends on the local topology around the initial spreader i . The analytic expression of the spreading time in a general graph is hard to derive in closed form [19]. Due to the limitation of the analytical methods, an event-driven simulator SSIS for the SIS spreading process based on the Gillespie algorithm is implemented to determine the spreading time [18].

A faster initial spreader speeds up the spreading in the outbreak period and leads to a shorter average spreading time, which measures the efficiency of the spreader. We

can identify the fastest nodes by ranking the average spreading time. We first show the effect of the topological properties of the spreader i on the average spreading time $E[T_m(i)]$ in a SIS epidemics on an Erdős-Rényi (ER) random network. Figure 1 shows the normalized topological metrics of node i versus the average spreading time $E[T_m(i)]$, which demonstrates that the average spreading time $E[T_m(i)]$ depends on the topological properties of initial spreader i . Specifically, the degree and the closeness of the initial spreader seem to have a similar behavior as the average spreading time in the ER random graph, while the betweenness of the initial spreader has a weaker correlation with the average spreading time. The reciprocal of the diagonal element (Q_{ii}^+) of the pseudoinverse matrix Q^+ also performs well in ranking the fastest spreaders and behaves similarly as the degree in the ER random graph [12]. Figure 1 illustrates that the nodes with the same coreness may occupy a large proportion of the network so that the fastest spreader cannot be identified well by their coreness.

3 The fastest spreader in SIS epidemics

In this section, we further investigate the fastest spreader in the SIS epidemics. The change of the fastest spreader with the effective infection rate τ is presented in an exemplified barbell-like graph. Then, we propose a new metric to identify the fastest spreader.

3.1 Change of the fastest spreader with τ in a barbell-like graph

We generate an asymmetric barbell-like graph G_{20} where a path graph L_2 connects an ER random graph $G_{0.5}(10)$ and a star graph $K_{1,7}$, as shown in Figure 2. The barbell-like

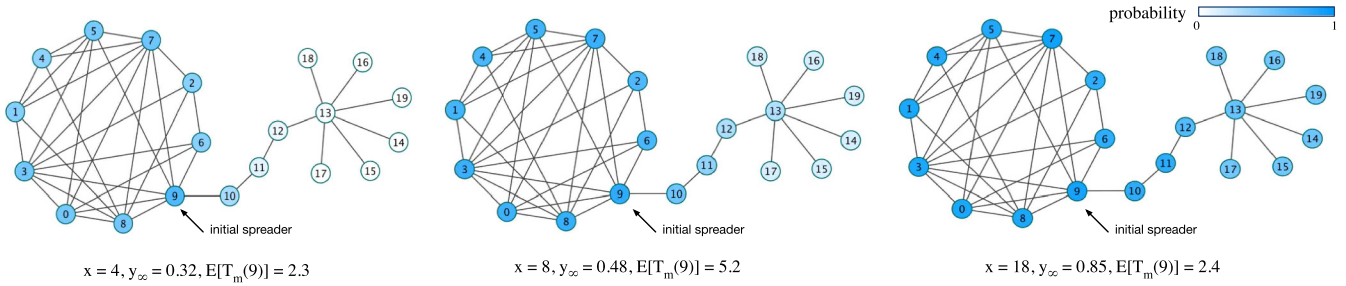


Fig. 2. The probability that the nodes is infected at the spreading time for different normalized effective infection rate $x = \tau/\tau_c$. Node 9 is the initially infected spreader. The darkness of the nodes represents the probability. The results is based on 10^5 realizations.

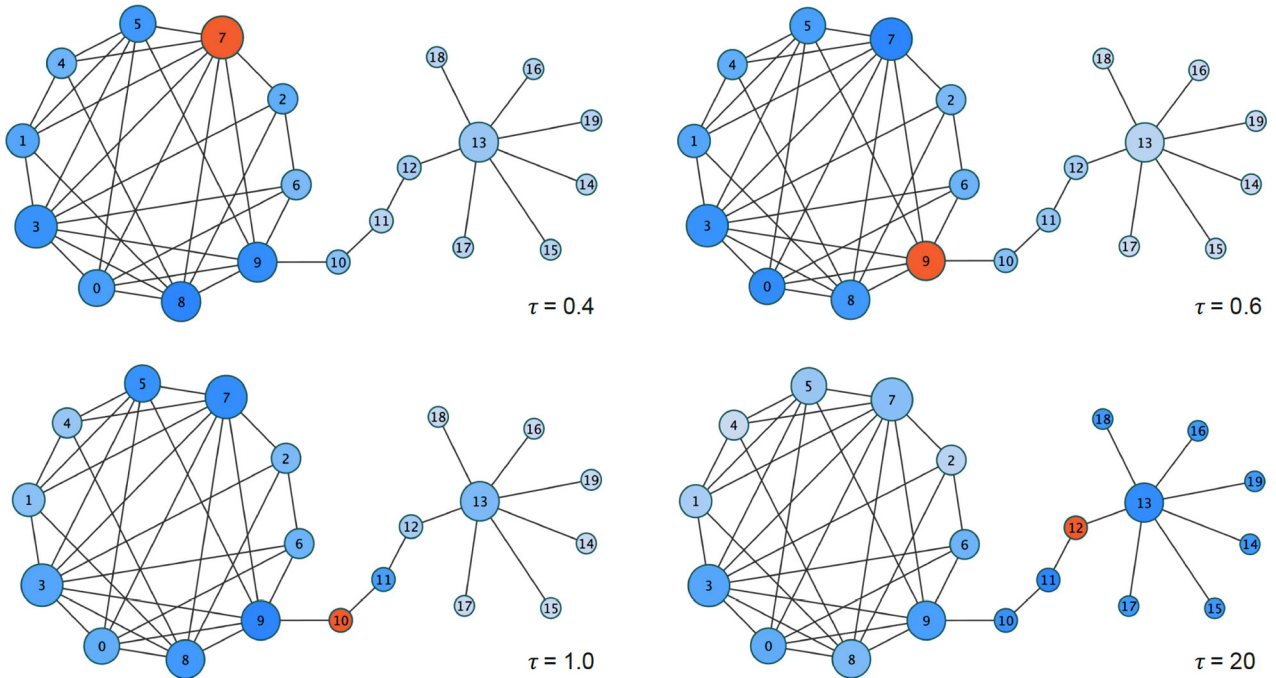


Fig. 3. Illustration of the changing of the fastest spreader with different τ . The size of the nodes represents the degree, and the darker node represents the faster spreader. The orange node is the fastest initial spreader.

graph helps us to trace the fastest spreader if the effective infection rate τ changes. Figure 2 illustrates the probability that the nodes is infected at the spreading time. Figure 2 shows that the infected nodes are usually localized around the initial spreader at the spreading time, e.g., the viruses seldom reach node 14 for a small normalized effective infection rate $x = 4$.

Figure 3 exemplifies that the fastest spreader changes with the effective infection rate τ in G_{20} . The fastest spreader changes dramatically from the highest degree node to the lowest degree node with increasing effective infection rate τ . Specifically, we observe three different cases in Figure 3. If the effective infection rate τ is relatively small, the fastest spreader tends to be located in the dense part (the ER random subgraph) of the network. With the increasing the effective infection rate τ , the fastest spreader transits to nodes with a larger closeness in the path subgraph. At last, the process approximates a flooding process if the effective infection rate τ is large

enough. Since the average time to infect all nodes in the star subgraph is larger than that in the ER random subgraph,¹ the fastest spreader should be closer to the star subgraph.

In Figure 4, the crossings of the average spreading time $E[T_m(i)]$ with the effective infection rate τ for different initial spreaders demonstrate that not only the fastest spreader but also the ranking of spreaders is not fixed for different effective infection rates τ . Therefore, we conclude that the fastest initial spreader in SIS model, only inferred by its location in the underlying graph of the network, cannot be determined. Our finding implies that

¹ The average time to infect all nodes [22] in an ER random graph $G_p(N)$ is estimated to be $\frac{1}{\beta} \sum_{n=1}^{N-1} \frac{1}{np(N-n)} \sim \frac{2 \log(N-1)}{\beta N}$. The average time to infect all nodes in a star graph $K_{1,N}$ from the center is estimated to be the maximum of N exponentially distributed random variables with mean $1/\beta$, which approximates $\sum_{n=1}^N \frac{1}{\beta n} \sim \frac{\log N}{\beta}$.

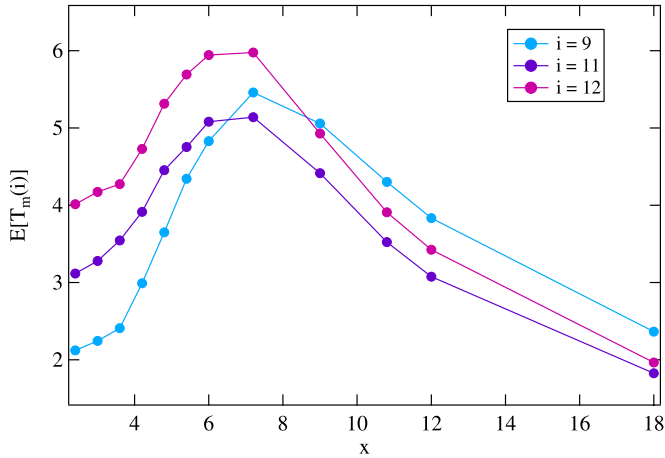


Fig. 4. The average spreading time $E[T_m(i)]$ as a function of the normalized effective infection rate x in G_{20} , started from node 9, 11 and 12, in the barbell-like graph G_{20} . (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

time-dependent “importance or centrality” of a node is usually strongly coupled to the dynamic process and the underlying graph itself.

3.2 A heuristic topological metric for the fastest spreader

In this section, we discuss the topological property of the fastest spreader throughout the increase of the effective infection rate τ , i.e., $\tau \downarrow \tau_c$, $\tau > \tau_c$ and very large τ .

3.2.1 Case: $\tau \downarrow \tau_c$

Invoking the infection probability vector $V(t) = (v_1(t), v_2(t), \dots, v_N(t))^T$, we approximate the spreading dynamics in the early stage of the spreading [23] and obtain

$$\frac{dV(t)}{dt} \approx \beta AV(t) - \delta IV(t). \quad (1)$$

The average fraction $y(t, \tau)$ of infected nodes at the spreading time t_m with τ obeys that

$$y(t_m, \tau) = \frac{1}{N} u^T V(t_m) \approx \frac{1}{N} u^T e^{(\beta A - \delta I)t_m} V(0), \quad (2)$$

where $u^T = (1, 1, \dots, 1)$. If the effective infection rate $\tau = \frac{\beta}{\delta}$ approaches the first order mean-field approximation of the epidemic threshold $\tau_c^{(1)} = \frac{1}{\lambda_1}$, only a very small proportion $y(t_m, \tau)$ of nodes will be infected in the metastable state. The spreading time t_m , defined as the the first hitting time when $Ny(t_m, \tau)$ nodes are infected in the Markovian SIS process without extinction [18], is finite. Figure 4 also exemplifies that the average spreading time is relatively small if $\tau \downarrow \tau_c^{(1)}$. The matrix $(\beta A - \delta I)t_m$ in (2) is dominated by the largest eigenvalue $\delta(\tau\lambda_1 - 1)t_m$, which tends to be 0 if the effective infection rate $\tau \downarrow \tau_c^{(1)}$

(by Perron-Frobenius Theorem [5]). Simplified, invoking the degree vector $d = Au$ and $V(0) = e_i$, we arrive at

$$\begin{aligned} \lim_{\tau \downarrow \tau_c^{(1)}} y(t_m, \tau) &\approx \frac{1}{N} u^T (I + (\beta A - \delta I)t_m) V(0) \\ &\approx \frac{1}{N} (u^T + (\beta d^T - \delta u^T)t_m) V(0) \\ &= \frac{1}{N} (1 - t_m \delta + t_m \beta d_i). \end{aligned} \quad (3)$$

Relation (3) exhibits that the degree of the spreader dominates the spreading time t_m for the unaltered rates β , δ and their corresponding $y(t_m) = y_\infty$. This result is different from the result that the eigenvector of the adjacent matrix A belonging to the largest eigenvalue determines the infection probability vector in the metastable state [5]. We here exemplify an extreme case: if the effective infection rate τ approaches τ_c , and there is only one infected node in the metastable state, i.e., $y(t_m) = \frac{1}{N}$, the spreading time t_m equals the minimum time when any one of the neighbors of the spreader i is infected. Then, the average spreading time $E[T_m]$ is the minimum of the d_i exponential distributed random variables with a mean $1/\beta$, where d_i is the degree of the spreader i . Thus, the average spreading time follows $E[T_m] = \frac{1}{\beta d_i}$, which is determined by the degree of the initial spreader.

3.2.2 Case: increasing τ

We then investigate the case for the increasing effective infection rate τ . Inspired by the illustration in Section 3.1, we postulate that the fastest spreader depends on the local topology around itself, i.e., the number of nodes and the connectivity of nodes around the spreader. We first consider the number of nodes around the initial spreader and regard that the efficiency of the initial spreader is related to the expansion [5] of the subgraph centered at the spreader. Specifically, assuming that the hop count h is the farthest distance from the initial spreader i that the viruses can reach before the spreading time, the expansion of the subgraph is the number of nodes $|C_i(h)|$ within h hops from the initial spreader i .

We then consider the connectivity of the nodes around the initial spreader. An epidemic behaves like a continuous time Markov branching process in the early stage [24]. For a branching process, we obtain that the number of infected nodes follows

$$Ny(t) \approx u^T e^{\beta A t} V(0) \leq e^{\beta \lambda_1 t} Ny(0) \quad (4)$$

which implies that the lower bound of the time to infect $Ny(t)$ nodes around the initial spreader follows that $t \geq \frac{\log(Ny(t))}{\beta \lambda_1}$. Inspired by (4), we propose $\frac{\lambda_i(h)}{\log |C_i(h)|}$ as an indication of the connectivity of the local topology around the spreader i for a fixed infection rate β , where $\lambda_i(h)$ is the largest eigenvalue of the subgraph within h hops around the initial spreader i . A larger $\frac{\lambda_i(h)}{\log |C_i(h)|}$ implies a higher connectivity that leads to a faster spreading in the local network within h hops.

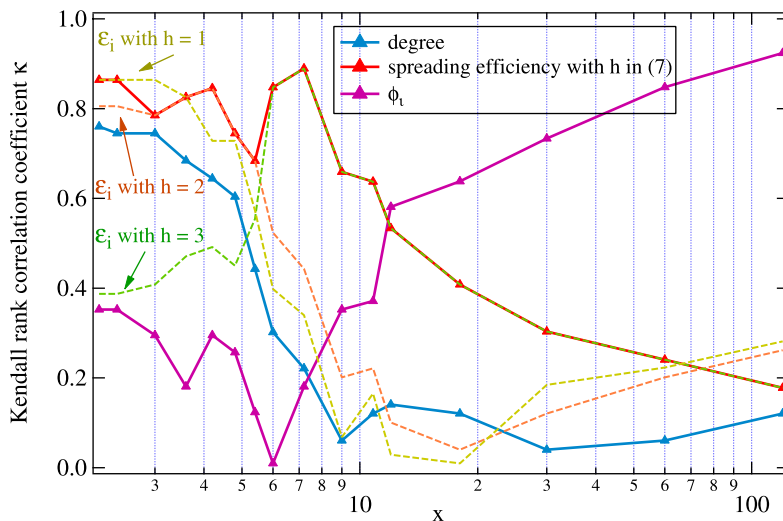


Fig. 5. The Kendall rank correlation coefficient κ between the average spreading time $E[T_m(i)]$ and the metrics: the degree d_i , the spreading efficiency \mathcal{E}_i and the reciprocal of the average flooding time ϕ_i in the barbell graph G_{20} .

Table 1. The topological properties of the giant component of the four experimental networks.

	N	L	ρ	λ_1	C_G	ρ_D
Les Misérables	77	254	5	12.01	0.57	-0.17
SmallWCitation	233	994	4	20.96	0.56	-0.30
Barbell G_{20}	20	39	7	6.14	0.33	-0.11
NetScience	379	914	17	10.38	0.74	-0.08

Considering the above two factors including the expansion $|C_i(h)|$ of the subgraph and the connectivity indication $\frac{\lambda_i(h)}{\log |C_i(h)|}$ within the subgraph, we propose the spreading efficiency as a new metric to measure the efficiency of the initial spreader in the SIS model. The spreading efficiency of node i is defined as

$$\mathcal{E}_i = \frac{\lambda_i(h)}{\log |C_i(h)|} |C_i(h)|. \tag{5}$$

In case that the sub-graph expansion $|C_i(h)|$ of the initial spreaders are the same, a larger sub-eigenvalue $\lambda_i(h)$ leads a higher spreading efficiency in the subgraph due to a higher connectivity of nodes.

The hop count h describes the average farthest distance of the infected nodes from the spreader at the spreading time for the effective infection rate τ , which is difficult to be determined precisely in a general network. Morone and Makse [11] identify the influential spreaders by the Ball (subgraph) centered at the spreader, where the optimal radius of the Ball is 3 or 4.

The optimal hop $h = f(\tau)$ in our method is more flexible, which is a function of the effective infection rate τ . We hereby proceed with an approximation. First, the average fraction of infected nodes y_∞ in the metastable state can be estimated by the NIMFA approach for a determined τ . The number N_C of nodes in a branch process follows $N_C \approx \frac{\mu^{H+1}-1}{\mu-1}$, where H is the largest hop count from the

root and $\mu = E[D] - 1$ is the mean degree minus 1 in this graph [5]. In that case, we have the largest hop count $H \approx \frac{\log(N_C(\mu-1)+1)}{\log \mu} - 1 \approx \frac{\log N_C}{\log \mu}$ if $N_C \gg \mu$. Invoking the fact that a spreading process approximates a branching process in the early stage, we can estimate the hop count h in a sparse, large graph by

$$h = \left\lceil \frac{\log N y_\infty}{\log \mu} \right\rceil. \tag{6}$$

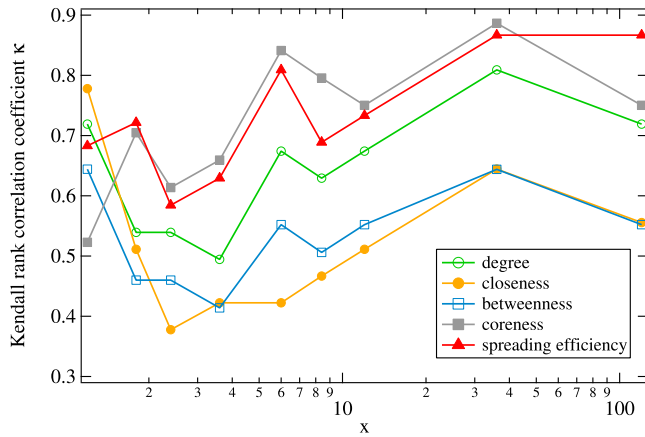
3.2.3 Case: large τ

With the increase of the effective infection rate τ and the average fraction y_∞ of infected nodes in the metastable state, the nodes that need relatively more time to be reached gradually dominate the spreading time. Thus, the fastest spreader could be closer to the sparser subgraph of the network. Finally, if the effective infection rate τ is large enough, the SIS process is reduced to be a flooding process [22]. The average flooding time $E[T_N(i)]$ of an initial spreader i is the average minimum time for the virus to reach all other nodes in a flooding process. Therefore, we could regard the reciprocal of the average flooding time $\phi_i = \frac{1}{E[T_N(i)]}$ determines the fastest spreader if τ is very large.

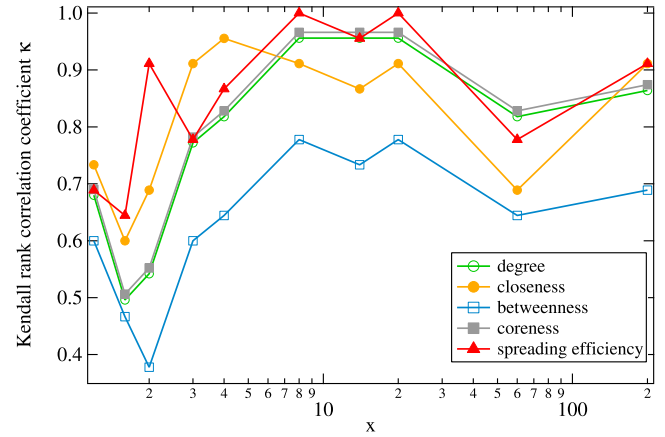
Assuming that $\lambda_1(0) = 1$ and $|C_i(0)| = d_i$, the spreading efficiency in (5) with $h < 1$ follows the same rank as the degree d_i . In summary, we simplify and propose the overall metric “spreading efficiency” to identify the fastest initial spreader in an SIS epidemics as

$$\mathcal{E}_i = \begin{cases} \frac{\lambda_1(h)|C_i(h)|}{\log |C_i(h)|} & y_\infty \leq y_\infty^* \\ \phi_i & y_\infty > y_\infty^*, \end{cases} \tag{7}$$

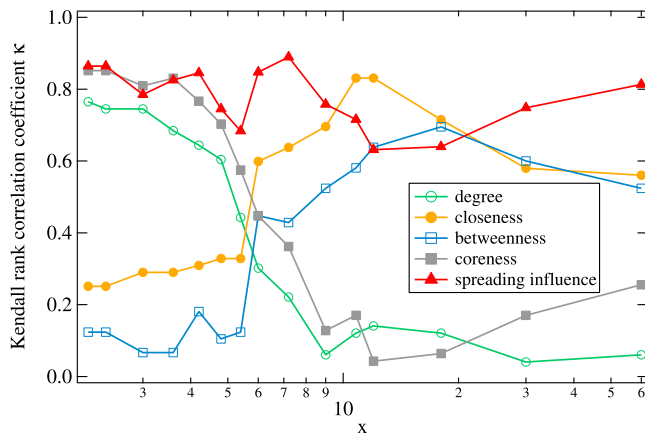
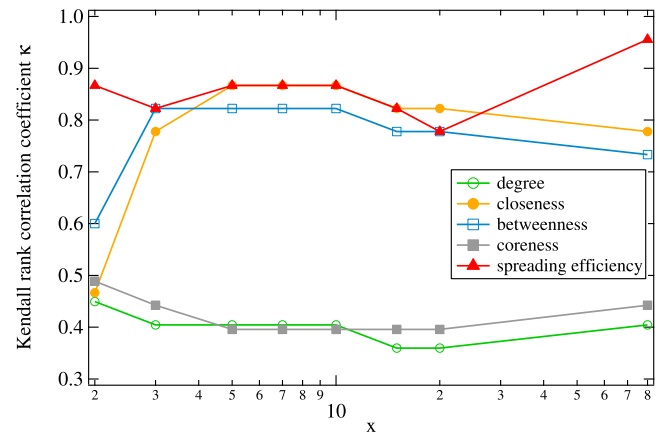
where y_∞^* is a prescribed parameter indicating that the process approximates a flooding process if $y_\infty > y_\infty^*$. We set $y_\infty^* = 0.8$ in the paper for the simulation.



(a) Characters network in Les Misérables



(b) SmallW citation network

(c) Barbell network G_{20} 

(d) NetScience network

Fig. 6. The Kendall rank correlation coefficient κ between the average spreading time $E[T_m(i)]$ and the metrics including degree d_i , closeness cl_i , betweenness b_i , coreness cr_i and the spreading efficiency \mathcal{E}_i in four networks.

Figure 5 shows the Kendall rank correlation coefficient κ between the average spreading time and the above discussed metrics, including the degree, the spreading efficiency in (5) and the reciprocal of the average flooding time ϕ_i via Monte-Carlo estimation. If the effective infection τ is close to the epidemics threshold $\tau_c \approx 0.17$, the degree centrality could be a better metric. We then observe that the best hop count h increases with the effective infection rate τ , and the spreading efficiency \mathcal{E}_i with the proposed hop count h in (6) can lead to the maximum correlation coefficient κ in a wide range of τ . At last, the reciprocal of the average flooding time shows the advantage when τ is large enough.

4 Numerical results

We evaluate the performance by identifying the ranking of the fastest initial spreaders in four, artificial and real, networks with different sizes and topologies: co-appearances of characters in Les Misérables [25], small world citation network (SmallWCitation) [26], the

artificial barbell network G_{20} and Co-authorship network of scientists (Net-Science) [27]. Table 1 shows some properties of the giant component of the four networks including the number of nodes N , the number of links L , the diameter ρ , the largest eigenvalue λ_1 , the clustering coefficient C_G , the Pearson degree correlation coefficient ρ_D .

We extract the giant component of the above network and select 10 nodes randomly in each network. In each implementation, only one of the selected nodes is infected initially, and then the virus spreads in the network according to the Markovian SIS model. After obtaining the average spreading time via SSIS started from different initial spreaders, we compare the Kendall rank correlation coefficients κ between the average spreading time and some other metrics including degree, closeness, betweenness, coreness and the proposed spreading efficiency in (7). Physically, the identification of the fastest spreaders in a flooding process is a 1-center problem [28] in a graph, where the weights of links in the graph are exponentially distributed random variables with mean $1/\beta$. Thus, we estimate the average flooding time

$E[T_N(i)]$ by Monte-Carlo approach and the efficiency shortest path algorithm [29].

Figure 6 shows the performance of the several centrality metrics for ranking the fastest initial spreader in four networks. We observe that, in the networks with a small diameter (e.g., Les Misérables and SmallW citation network), the spreading efficiency performs similarly with the coreness, both of which are better than other centrality metrics. In addition, the spreading efficiency shows its advantage over the coreness if the effective infection rate τ is relatively large because the reciprocal of the average flooding time determines the fastest spreader in that case.

However, the degree and coreness show the vulnerability in the community networks with a large diameter (e.g., Barbell and NetScience network). Meanwhile, the closeness becomes a better metric, which considers the average length of the path between the spreader and all other nodes. Especially, in the Barbell G_{20} , we observe the changing of the performance of the centrality metrics with the increasing effective infection rate τ . When τ is small, the degree and the coreness perform better, but the closeness and the betweenness become better if τ is large enough, which further convinces us that a single existing centrality metric fails to identify the fastest spreader in the SIS model. The results suggest that, in the real world, the viruses or information may spread more efficiency starting from the spreader with a large degree within the community for a small τ , but it is better to choose the spreader with a high closeness for a large τ .

In summary, we can observe that the proposed spreading efficiency performs better than the compared topological metrics in general, which is adaptive to different topologies and different dynamic process. We find that the accuracy of the spreading efficiency drops a little around the effective infection rate corresponding to the transition parameter y_∞^* . We also expect a better transition method and a better estimation of hop $h = f(\tau)$ that can improve the performance.

5 Conclusion

We investigated the properties of the fastest initial spreader with the shortest average spreading time in the SIS model. We showed that the fastest spreader changes from the node with the largest degree to the node with the shortest flooding time for the increasing effective infection rate, which implies that the fastest spreader is coupled to not only the underlying graph but also the dynamic process.

By considering the expansion and the largest eigenvalue of the subgraph around the spreader, we proposed the spreading efficiency as a metric to rank the fastest spreaders. The spreading efficiency depends on the effective infection rate τ , and reduces to the reciprocal of the flooding time for a large τ . The simulation results on four networks show that the spreading efficiency can better rank the fastest spreaders than some existing topological metrics including degree, closeness, betweenness, and coreness, in different topologies and dynamic processes.

This research is supported by the China Scholarship Council (CSC). We are grateful to H. Cetinay for useful comments.

Author contribution statement

Z. He and P. Van Mieghem derived the results and wrote the manuscript. Z. He performed the numerical results.

Open Access This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

References

1. L. Lü, D. Chen, X. Ren, Q. Zhang, Y. Zhang, T. Zhou, Phys. Rep. **650**, 1 (2016)
2. T. Zhou, J. Liu, W. Bai, G. Chen, B. Wang, Phys. Rev. E **74**, 056109 (2006)
3. M. Nekovee, Y. Moreno, G. Bianconi, M. Marsili, Physica A **374**, 457 (2007)
4. J.M. Hernández, P. Van Mieghem, Delft University of Technology, report 20111111 (2011)
5. P. Van Mieghem *Performance analysis of complex networks and systems* (Cambridge University Press, 2014)
6. T. Martin, X. Zhang, M.E.J. Newman, Phys. Rev. E **90**, 052808 (2014)
7. P. Van Mieghem, preprint [arXiv:1401.4580](https://arxiv.org/abs/1401.4580) (2014)
8. M. Kitsak, L.K. Gallos, S. Havlin, F. Liljeros, L. Muchnik, H.E. Stanley, H.A. Makse, Nat. Phys. **6**, 888 (2010)
9. A. Zeng, C. Zhang, Phys. Lett. A **377**, 1031 (2013)
10. J. Liu, Z. Ren, Q. Guo, Physica A **392**, 4154 (2013)
11. F. Morone, H.A. Makse, Nature **524**, 65 (2015)
12. P. Van Mieghem, K. Devriendt, H. Cetinay, Phys. Rev. E **96**, 032311 (2017)
13. R. Pastor Satorras, C. Castellano, P. Van Mieghem, A. Vespignani, Rev. Modern Phys. **87**, 925 (2015)
14. M. Šikić, A. Lančić, N. Antulov-Fantulin, H. Štefančić, Eur. Phys. J. B **86**, 440 (2013)
15. J. Liu, J. Lin, Q. Guo, T. Zhou, Sci. Rep. **6**, 21380 (2016)
16. P. Holme, Phys. Rev. E **96**, 062305 (2017)
17. B. Qu, C. Li, P. Van Mieghem, H. Wang, Sci. Rep. **7**, 9233 (2017)
18. Z. He, P. Van Mieghem, Phys. A: Stat. Mech. Appl. **494**, 317 (2018)
19. R. van de Bovenkamp, P. Van Mieghem, in *Tenth International Conference on Signal-Image Technology and Internet-Based Systems (SITIS)* (IEEE, 2014), pp. 347–354
20. P. Van Mieghem, J. Omic, R. Kooij, IEEE/ACM Trans. Netw. **17**, 1 (2009)
21. C. Castellano, R. Pastor-Satorras, Phys. Rev. Lett. **105**, 218701 (2010)
22. R. van der Hofstad, G. Hooghiemstra, P. Van Mieghem, Extremes **5**, 111 (2002)
23. P. Van Mieghem, R. van de Bovenkamp, Phys. Rev. Lett. **110**, 108701 (2013)
24. F. Ball, P. Donnelly, Theory Probab. Appl. **37**, 119 (1993)

25. D.E. Knuth, *The Stanford GraphBase: a platform for combinatorial computing* (Addison-Wesley, Reading, 1993), Vol. 37
26. S. Milgram, *Psychol. Today* **1**, 61 (1967)
27. M.E.J. Newman, *Phys. Rev. E* **74**, 036104 (2006)
28. O. Kariv, S.L. Hakimi, *SIAM J. Appl. Math.* **37**, 513 (1979)
29. T.M. Chan, *SIAM J. Comput.* **39**, 2075 (2010)