



Ranking spreaders by decomposing complex networks



An Zeng^{a,b,*}, Cheng-Jun Zhang^a

^a Department of Physics, University of Fribourg, Chemin du Musée 3, CH-1700 Fribourg, Switzerland

^b Institute of Information Economy, Hangzhou Normal University, Hangzhou 310036, China

ARTICLE INFO

Article history:

Received 19 October 2012

Received in revised form 15 February 2013

Accepted 19 February 2013

Available online 27 February 2013

Communicated by C.R. Doering

Keywords:

Complex networks

Network decomposition

Spreading

ABSTRACT

Ranking the nodes' ability of spreading in networks is crucial for designing efficient strategies to hinder spreading in the case of diseases or accelerate spreading in the case of information dissemination. In the well-known *k*-shell method, nodes are ranked only according to the links between the remaining nodes (residual links) while the links connecting to the removed nodes (exhausted links) are entirely ignored. In this Letter, we propose a mixed degree decomposition (MDD) procedure in which both the residual degree and the exhausted degree are considered. By simulating the epidemic spreading process on real networks, we show that the MDD method can outperform the *k*-shell and degree methods in ranking spreaders.

© 2013 Elsevier B.V. All rights reserved.

1. Introduction

Spreading processes are important in various fields including physics, chemistry, medical science, biology and sociology [1]. For example, reaction diffusion processes [2], pandemics [3], cascading failures in electric power grids [4,5] and information dissemination [6] can be naturally addressed within the framework of spreading process. In particular, spreading in complex networks has been intensively studied in the past decade. Many studies have revealed that the spreading process is strongly influenced by the network topologies [7–10]. With the understanding of spreading pathways on networks, many methods have been developed to manipulate the network structure to control the spreading threshold [11,12]. Moreover, in order to avoid the wide propagation of the disease, various efficient immunization strategies were also proposed [13,14].

Though lots of former works are dedicated to understand and control the spreading process in a macroscopic sense, recently more and more attention has been given to microscopically study the *spreading ability* for each node, i.e., how many nodes will finally be covered when the spreading originates from this single node [15–17]. For example, the knowledge of node spreading ability is crucial for developing efficient methods to either decelerate spreading in the case of diseases, or speed up spreading in the case of information flow. Moreover, it can be helpful for identifying the initial spreader of certain disease or information [18].

There are many classic topology metrics which can be naturally used to rank spreaders, such as degree, betweenness [19], closeness [20], clustering coefficient [21], Katz centrality [22] and so on. Although the most connected nodes (hubs) and the nodes with high betweenness centrality are commonly believed to be the most influential spreaders in networks, the *k*-shell (also called *k*-core) method was found to perform better in identifying the best individual spreaders [15,23]. The *k*-shell method starts by removing all nodes with one connection only (with their links), until no more of such nodes remain, and assign them to the 1-shell. For each remaining node, the number of links connecting to the other remaining nodes is called its *residual degree* and the number of links connecting to the removed nodes is called its *exhausted degree*. After assigning the 1-shell, all nodes with residual degree 2 are recursively removed and the 2-shell is created. This procedure continues as the residual degree increases until all nodes in the networks have been assigned to one of the shells. The nodes with high *k*-shell value tend to locate in the center of the network and the spreading starting from each of these nodes is likely to cover a large part of the network. Actually, the *k*-shell method was originally introduced in [24] to address issues of social network research. Subsequently, it was used to analyze internet graphs [25], transportation [26], biological systems [27], economic systems [28] and network robustness [29]. The method has also been extended to weighted networks [30]. Moreover, a similar idea has also been applied to assign direction to the link of undirected networks and significant improvement in synchronizability can be achieved [31,32].

When the *k*-shell method is used to decompose networks, only the links between the remaining nodes are considered while the links connecting to the removed nodes are completely ignored. In

* Corresponding author.

E-mail address: an.zeng@unifr.ch (A. Zeng).

order to have the same spreading ability, the nodes within the same shell should have the same exhausted degree or at least have homogeneous exhausted degree distribution. However, the exhausted degree of the nodes in real networks indeed has very heterogeneous distribution. Given the same residual degree, even though the nodes with large exhausted degree perform much better than the other nodes in spreading, the k -shell method gives them the same rank. Consequently, ignoring the exhausted degree will lead to some problems in k -shell method. For example, the k -shell method assigns every node to the same shell in the tree network [34] and Barabasi–Albert network [35].

In this Letter, we propose a mixed degree decomposition (MDD) procedure combining both the residual degree and the exhausted degree. The new method can effectively remove the degeneracy of the k -shell method. We remark that there are many different kinds of spreading processes on networks such as epidemic, information, rumor spreading and so on [1]. According to a recent study [33], though the k -shell method is effective in identifying the most influential spreader for disease, it may not be a good predictor for the spreading ability of nodes when the information or rumor propagation is considered. In this Letter, we focus on the epidemic spreading process as Ref. [15] and employ the well-known SIR model in our simulation [1]. By simulating the epidemic process on 15 real networks, we show that the MDD performs more accurately than the k -shell and the degree-based methods in ranking the spreading ability of nodes. Finally, we discuss in detail the ranking of hub nodes in the different methods.

2. Mixed degree decomposition method

The k -shell method is a dynamical network decomposition procedure in which the residual degree of nodes is updated in each step while all the information of the removed nodes is dropped. In the Mixed Degree Decomposition (MDD) method, not only the residual degree but also the exhausted degree of the nodes are recorded, and the decomposition is based on both of them. For a node i , we denote the residual degree (number of links connecting to the remaining nodes) and the exhausted degree (number of links connecting to the removed nodes) as $k_i^{(r)}$ and $k_i^{(e)}$, respectively. In each step of the MDD procedure, the nodes are removed according the mixed degree

$$k^{(m)} = k^{(r)} + \lambda * k_i^{(e)}, \quad (1)$$

where λ is a tunable parameter between 0 and 1. The detailed decomposition is done with the following procedure:

1. Initially, $k^{(m)}$ of each node is equal to $k^{(r)}$ since there is no removed node in the network.
2. Remove all the nodes with the smallest $k^{(m)}$ (denoted as M) and assign them to the M -shell.
3. Update $k^{(m)}$ of all the remaining nodes by $k^{(m)} = k^{(r)} + \lambda * k_i^{(e)}$. Then, remove all the nodes with $k^{(m)}$ smaller than or equal to M and assign them to the M -shell too. This step is recursively carried on until $k^{(m)}$ of all remaining nodes are larger than M .
4. Repeat step 2 and 3 as M value increases until all nodes in the network have been assigned to one of the shells.

When $\lambda = 0$, the MDD method coincides with the k -shell method in Refs. [15,23]. When $\lambda = 1$, the MDD method is equivalent to the degree centrality method. Different from the original k -shell method, note that the shell values in MDD method are no longer integer since $k^{(m)}$ can be decimal when λ is between 0 and 1. To better illustrate the procedure of MDD, a simple example is shown in Fig. 1 in which parameter λ for the MDD method is set as 0.7.

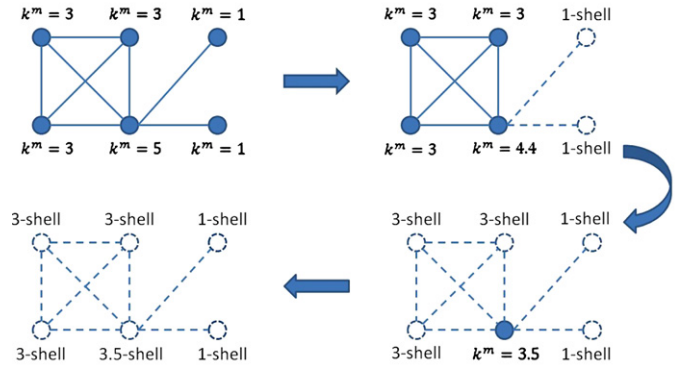


Fig. 1. (Color online.) A simple example to illustrate the procedure of the Mixed Degree Decomposition (MDD). The nodes represented with dashed line are the removed nodes at the current step and the links with dashed line are the exhausted links. Here, the parameter λ in MDD is set as 0.7. Note that $k^{(m)}$ of each remaining node has to be updated in each step.

3. Result on real networks

To validate the effectiveness of the MDD method, we then apply it to 15 real networks which are from social and nonsocial systems. Social networks are: Dolphins (friendship) [36], Jazz (musical collaboration) [37], Netsci (collaboration network of network scientists) [38], Email (communication) [39], HEP (collaboration network of high-energy physicists) [40], PGP (an encrypted communication network) [41], Astro Phys (collaboration network of astrophysics scientists) [40], Cond Matt (collaboration network of condensed matter scientists) [40]. Nonsocial networks are: Word (adjacency relation in English text) [38], E. coli (metabolic) [42], C. elegans (neural) [43], TAP (yeast protein–protein binding network generated by tandem affinity purification experiments) [44], Y2H (yeast protein–protein binding network generated using yeast two hybridization) [45], Power (connections between power stations) [46], Internet (router level) [47]. To better illustrate the performance of the MDD method, we select four relatively large networks (email, PGP, Astro Phys and Cond Matt) as representative examples. The results of the other networks are detailedly reported in Table 1.

We begin our analysis by investigating the distribution of the exhausted degree in real networks. We first use the k -shell method to decompose the above-mentioned four networks. The nodes within the same shell are selected, and their exhausted degree can be easily calculated as the number of links from each of them to the nodes with lower shells. The exhausted degree distribution of those nodes with the same shell are shown in Fig. 2. As expected, in all of the four real networks, the exhausted degree distribution shows a long tail, which indicates the distribution is inhomogeneous. If a virus originates from a node with large exhausted degree, not only it has the same possibility as the other nodes in the same shell to infect the nodes in the higher shells, but also it has a bigger branch of nodes in the lower shells to infect, so that this virus will end up covering much more nodes at the end. Therefore, the information of the exhausted degree cannot be overlooked when ranking nodes.

In Fig. 3, we show the frequency of the appearance of different ranks in k -shell method, degree centrality method and MDD method. Obviously, the k -shell only has limited number of ranks and the number of nodes in each rank is quite high, which implies that node differences are not well distinguished in the k -shell method. By using the degree centrality to rank the nodes, larger number of ranks will be obtained. In the MDD method, the degeneracy is further removed and nodes are in more detailed ranked than the previous two methods. Specifically, we obtained 203 ranks in Email network, 415 ranks in PGP network, 1389 ranks in

Table 1

Structural properties and ranking results of the different real networks. Structural properties include network size (N), edge number (E), degree heterogeneity ($H = \langle k^2 \rangle / \langle k \rangle^2$), degree assortativity (r), clustering coefficient ($\langle C \rangle$) and average shortest path length ($\langle d \rangle$). For the ranking results, all nodes are considered when calculating $\langle \tau \rangle$. In $\langle \bar{\tau} \rangle$, only $N \cdot L$ nodes with the largest degree in the networks are taken into account (here, $L = 10\%$).

Network	N	E	H	r	$\langle C \rangle$	$\langle d \rangle$	$\langle \tau \rangle_{ks}$	$\langle \tau \rangle_k$	$\langle \tau \rangle_{MDD}^*$	$\langle \bar{\tau} \rangle_{ks}$	$\langle \bar{\tau} \rangle_k$	$\langle \bar{\tau} \rangle_{MDD}^*$
Dolphins	62	159	1.327	−0.044	0.259	3.357	0.555	0.742	0.754	0.000	0.533	0.543
Word	112	425	1.815	−0.129	0.173	2.536	0.655	0.750	0.760	0.000	0.720	0.720
Jazz	198	2742	1.395	0.020	0.618	2.235	0.709	0.817	0.835	0.316	0.471	0.499
E. coli	230	695	2.365	−0.015	0.224	3.784	0.665	0.689	0.708	0.330	0.488	0.595
C. elegans	297	2148	1.801	−0.163	0.292	2.455	0.638	0.717	0.745	0.000	0.600	0.707
Netsci	379	914	1.663	−0.082	0.741	6.042	0.547	0.655	0.665	0.245	0.551	0.618
Email	1133	5451	1.942	0.078	0.220	3.606	0.733	0.750	0.773	0.295	0.532	0.623
TAP	1373	6833	1.644	0.579	0.529	5.224	0.673	0.723	0.742	0.454	0.546	0.598
Y2H	1458	1948	2.667	−0.210	0.071	6.812	0.428	0.553	0.563	0.146	0.638	0.686
Power	4941	6594	1.450	0.004	0.080	18.989	0.374	0.613	0.623	0.133	0.474	0.530
HEP	5835	13 815	1.926	0.185	0.506	7.026	0.586	0.652	0.667	0.433	0.545	0.663
PGP	10 680	24 316	4.147	0.238	0.266	7.463	0.465	0.515	0.525	0.521	0.494	0.608
Astro Phys	14 845	119 652	2.820	0.228	0.670	4.847	0.697	0.690	0.714	0.324	0.479	0.564
Internet	22 963	48 436	61.978	−0.198	0.230	3.850	0.400	0.382	0.401	0.643	0.381	0.677
Cond Matt	36 458	171 736	2.960	0.177	0.657	5.476	0.611	0.602	0.634	0.654	0.445	0.735

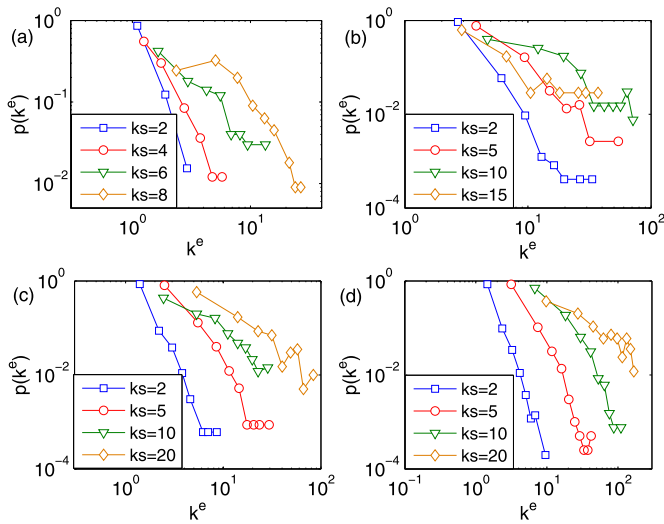


Fig. 2. (Color online.) The distribution of the exhausted degree in real networks. The networks are: (a) Email, (b) PGP, (c) Astro Phys and (d) Cond Matt.

Astro Phys network, 1323 ranks in Cond Matt network. As we can see from Fig. 3, the number of ranks can be even ten times larger than the degree method. More importantly, the number of nodes in each high rank is almost 1 which suggests that these nodes are well separated. We also check the performance of MDD method on the tree network and BA model in which k -shell method is not valid, we find that the MDD can effectively detect the difference between nodes in these networks.

The ranking generated by k -shell, degree and MDD methods are all obtained by analyzing network topology. In principle, an effective topology-based ranking should be as close as possible to the ranking by the real spreading coverage. In this Letter, we employ the SIR model [1] to simulate the spreading process on networks. The number of nodes that are finally infected when the infection starts from a given node i is denoted as its spreading ability s_i^p where p is the infection rate in the SIR model. For all the methods mentioned above, we generate the final ranking of nodes. We then use the Kendall's tau rank correlation coefficient (τ) [48] to estimate how a certain topology-based ranking is correlated to the ranking by the true spreading ability s of the nodes.

The Kendall's tau coefficient considers a set of observations of the joint variables X and Y (in our case, X can be ranking results and Y can be the spreading results of all nodes). Any pair of observations (x_i, y_i) and (x_j, y_j) are said to be concordant if the ranks for both elements agree: that is, if both $x_i > x_j$ and $y_i > y_j$ or if

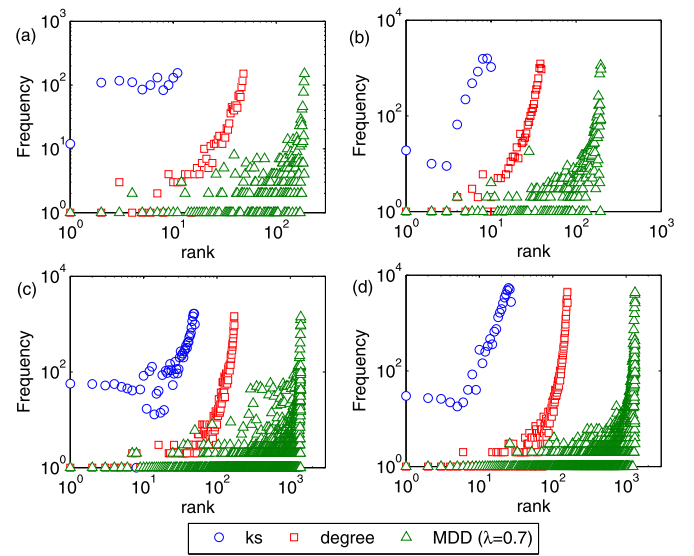


Fig. 3. (Color online.) The frequency of appearance of different ranks in k -shell method, degree centrality method and MDD method ($\lambda = 0.7$). The networks are: (a) Email, (b) PGP, (c) Astro Phys and (d) Cond Matt.

both $x_i < x_j$ and $y_i < y_j$. They are said to be discordant, if $x_i > x_j$ and $y_i < y_j$ or if $x_i < x_j$ and $y_i > y_j$. If $x_i = x_j$ or $y_i = y_j$, the pair is neither concordant nor discordant. Denoting n_1 as the number of concordant pairs, n_2 as the number of discordant pairs, and n as the number of nodes in the network, the Kendall tau coefficient is defined as:

$$\tau = \frac{n_1 - n_2}{0.5n(n-1)}. \quad (2)$$

In the most ideal case where $\tau = 1$, for each two nodes i and j , if i is ranked before j by the topology-based method, i will have stronger spreading ability than j . In Fig. 4, we show the value of τ of the k -shell, degree centrality and MDD methods under different p . In this Letter, we use relatively small values for p , namely $p \in (0, 0.1]$, so that the infected percentage of the nodes is not so large. When $p = 0.1$, the average percentage of infected nodes is 0.11 in Email network, 0.0059 in PGP network, 0.18 in Astro Phys network, 0.074 in Cond Matt network. In the case of large p values, where spreading can cover almost all the network, the role of individual nodes is no longer important since the final coverage of virus is independent of where it originated from. Interestingly, though the k -shell method is claimed to be able to identify the most influential node, its τ value is not significantly higher than

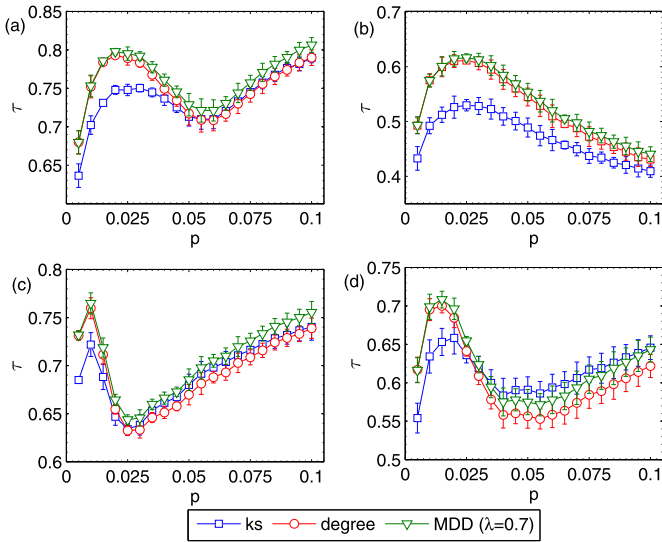


Fig. 4. (Color online.) The value of τ of the k -shell, degree centrality and MDD methods ($\lambda = 0.7$) under different infection rate p in the SIR model. The networks are: (a) Email, (b) PGP, (c) Astro Phys and (d) Cond Matt. The results are averaged over 100 independent realizations.

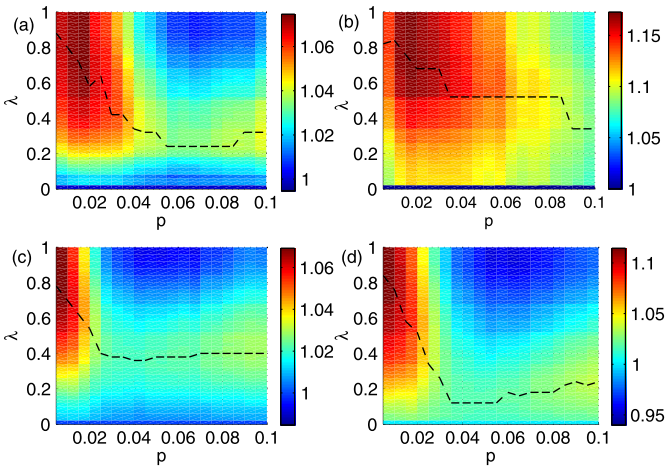


Fig. 5. (Color online.) The value of τ' under different parameter λ and p . The networks are: (a) Email, (b) PGP, (c) Astro Phys and (d) Cond Matt. The black lines mark the optimal λ^* under different p . The results are averaged over 100 independent realizations.

that of the degree centrality method. We again set $\lambda = 0.7$ in the MDD as an example and show its τ value in Fig. 4. As we can see, for this value of λ , the MDD method can have a slightly higher τ than the other two methods in Fig. 4(a), (b) and (c). However, the τ value of the MDD method in Fig. 4(d) is lower than that of k -shell method. These results indicate that different real systems may ask for a different λ for the MDD method to achieve its best ranking of the spreaders, which we will discuss more detailedly in Fig. 5.

In order to systematically study how the parameter λ affects the performance of the MDD method, we study the performance of MDD method from $\lambda = 0$ to $\lambda = 1$. In order to see more clearly the advantage of the MDD method, we calculate the ratio of τ_{MDD} and τ_{ks} as $\tau' = \tau_{\text{MDD}}/\tau_{\text{ks}}$. In Fig. 5, we show the dependence of τ' on the infection rate p and the MDD parameter λ . One can immediately notice that τ' can be larger than 1 in different p , which indicates that the MDD method can constantly improve k -shell method under different infection rate. However, there is also some region where $\tau' < 1$ (i.e. τ_{MDD} is smaller than τ_{ks}). It indi-

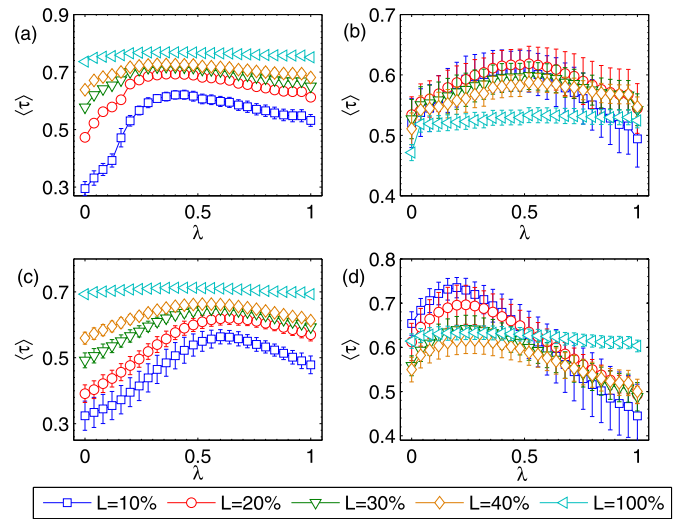


Fig. 6. (Color online.) The value of $\langle \tilde{\tau} \rangle$ of the MDD methods under different parameter λ . L denotes the percentage of nodes with the largest degree considered when calculating $\langle \tilde{\tau} \rangle$. The networks are: (a) Email, (b) PGP, (c) Astro Phys and (d) Cond Matt. The results are averaged over 100 independent realizations.

cates that the performance of MDD method depends a lot on the parameter λ . Actually, given a certain p , there is always an optimal parameter λ^* which can achieve the highest τ' . The change of the optimal λ^* with p is marked by the black curves in Fig. 5. As we can see, λ^* decreases with p . When p is small, relatively large λ works better. Under this situation, the virus only spread a few steps and the local structure property (degree information) mainly determines the spreading ability of nodes. On the other hand, with a large p the virus can cover more nodes, hence the global property (information from the decomposition) plays a more important role in determining the spreading abilities. Consequently, larger λ generally works better in this case.

We then calculate $\langle \tau \rangle$ by averaging all the τ values under different infection rates p we considered. In this way, we can investigate whether the MDD method can reflect the overall spreading ability of nodes more accurately than the other two methods. We report the results of this measure in all 15 real networks in Table 1. We see that the MDD method can outperform both k -shell method and degree method in $\langle \tau \rangle$ in all real networks we considered.

Usually, people are more interested in the most influential spreaders [15]. In our work, since the measurement $\langle \tau \rangle$ takes all the nodes into account, another interesting aspect is to investigate $\langle \tilde{\tau} \rangle$, which only considers the hubs (i.e. nodes with large degree). Therefore, we further move to investigate whether the MDD method can improve $\langle \tilde{\tau} \rangle$ as well. In $\langle \tilde{\tau} \rangle$, we first denote a percentage L and only consider the $N \cdot L$ nodes with the largest degree in the network. Except for this, the calculation of $\langle \tilde{\tau} \rangle$ is exactly the same as for $\langle \tau \rangle$. As shown in Fig. 6, when we only consider the hub nodes, the MDD can remarkably improve $\langle \tilde{\tau} \rangle$ when adjusting λ . Actually, we can see that the improvement of MDD method is larger when L is smaller. It suggests that MDD works more effectively in ranking the most influential spreaders. The results of $\langle \tilde{\tau} \rangle$ in the other real networks are investigated, and similar results are observed. The value of $\langle \tilde{\tau} \rangle$ in the other networks can be seen in Table 1.

We also check the performance of MDD method on two modeled networks, the classic Watts–Strogatz network model [21] and a variant of Barabasi–Albert network model [49]. In the small-world network model, we observe that $\langle \tau \rangle$ of MDD is 0 in regular network and it increases with the rewiring probability. In the scale-free network model, we find that a larger degree power-law

exponent yields a higher $\langle \tau \rangle$ of MDD. This is because the degeneracy among the small degree nodes is less serious when the exponent is larger. In both of these two network models, the MDD method outperforms the degree and k -shell methods in $\langle \tau \rangle$.

4. Conclusion

The well-known k -shell method has been shown to outperform degree and betweenness methods in ranking spreaders in networks. In this Letter, we show that the k -shell method can be further improved by considering the exhausted degree (i.e., links connecting the remaining nodes to the removed nodes) during the decomposition process. We propose a Mixed Degree Decomposition (MDD) procedure with a tunable parameter λ to rank the spreading ability for nodes in networks. We apply our method in 15 real networks from both social and nonsocial systems. By using Kendall's tau rank correlation coefficient (τ) to measure the correlation between the MDD ranking and the real spreading ability from the SIR model, we find that our method can achieve a higher τ than the degree method and k -shell method. After designing a local $\tilde{\tau}$ which only considers hub nodes, we further show the MDD works also effectively in ranking the most influential spreaders.

In the k -shell method, some nodes with large degree are seriously underestimated. Compared to the other nodes in the same shell, these nodes have much stronger spreading ability. By considering the exhausted degree in the network decomposition, the ranking of these nodes are properly improved by the MDD method. That's why the MDD method outperforms the k -shell method. Though the MDD method can improve the k -shell method in ranking spreaders, it may still not be the optimal way to address this problem. For example, directly considering the number of possible spread paths and weighting them with some proper damping factor might obtain an even more accurate ranking for spreading ability. However, this method can be with much higher computational complexity than the network decomposition-based methods. Therefore, some more effective and efficient methods are still asked for further investigation.

Acknowledgements

We would like to thank Yi-Cheng Zhang, Stanislao Gualdi, Giulio Cimini and Seung-Woo Son for helpful suggestions. This work is supported by the Swiss National Science Foundation (No. 200020-132253). AZ acknowledges the support from the opening foundation of Institute of Information Economy in Hangzhou Normal University (Grant No. PD12001003002002).

References

- [1] S.N. Dorogovtsev, A.V. Goltsev, J.F.F. Mendes, *Rev. Mod. Phys.* 80 (2008) 1275.

- [2] V. Colizza, R. Pastor-Satorras, A. Vespignani, *Nature Phys.* 3 (2007) 276.
- [3] R. Pastor-Satorras, A. Vespignani, *Phys. Rev. Lett.* 86 (2001) 3200.
- [4] A.E. Motter, *Phys. Rev. Lett.* 93 (2000) 098701.
- [5] S.V. Buldyrev, R. Parshani, G. Paul, H.E. Stanley, S. Havlin, *Nature* 464 (2010) 1025.
- [6] M. Medo, Y.-C. Zhang, T. Zhou, *Europhys. Lett.* 88 (2009) 38005.
- [7] V.M. Eguiluz, K. Klemm, *Phys. Rev. Lett.* 89 (2002) 108701.
- [8] M. Boguna, R. Pastor-Satorras, A. Vespignani, *Phys. Rev. Lett.* 90 (2003) 028701.
- [9] M.A. Serrano, M. Boguna, *Phys. Rev. Lett.* 97 (2006) 088701.
- [10] C. Castellano, R. Pastor-Satorras, *Phys. Rev. Lett.* 105 (2010) 218701.
- [11] M. Schlapfer, L. Buzna, *Phys. Rev. E* 85 (2012) 015101(R).
- [12] A.N. Bishop, I. Shames, *Europhys. Lett.* 95 (2011) 18005.
- [13] R. Cohen, S. Havlin, D. Ben-Avraham, *Phys. Rev. Lett.* 91 (2003) 247901.
- [14] Y. Chen, G. Paul, S. Havlin, F. Liljeros, H.E. Stanley, *Phys. Rev. Lett.* 101 (2008) 058701.
- [15] M. Kitsak, L.K. Gallos, S. Havlin, F. Liljeros, L. Muchnik, H.E. Stanley, H.A. Makse, *Nature Phys.* 6 (2010) 888.
- [16] D. Chen, L. Lu, M.-S. Shang, Y.-C. Zhang, T. Zhou, *Physica A* 391 (2012) 1777.
- [17] R.A.P. Da Silva, M.P. Viana, L. Da, F. Costa, *arXiv:1202.0024v1*, 2012.
- [18] C.H. Comin, L. Da, F. Costa, *Phys. Rev. E* 84 (2011) 056105.
- [19] L. Freeman, *Sociometry* 40 (1977) 35.
- [20] G. Sabidussi, *Psychometrika* 31 (1966) 581.
- [21] D.J. Watts, S. Strogatz, *Nature* 393 (1998) 440.
- [22] L. Katz, *Psychometrika* 18 (1953) 39.
- [23] S. Carmi, S. Havlin, S. Kirkpatrick, Y. Shavitt, E. Shir, *Proc. Natl. Acad. Sci. USA* 104 (2007) 11150.
- [24] S.B. Seidman, *Social Networks* 5 (1983) 269.
- [25] J.I. Alvarez-Hamelin, L. Dall'Asta, A. Barra, A. Vespignani, *Networks and Heterogeneous Media* 3 (2008) 371.
- [26] Z. Torozkai, K.E. Bassler, *Nature* 428 (2004) 716.
- [27] S. Wuchty, E. Almaas, *Proteomics* 5 (2005) 444.
- [28] A. Garas, P. Argyrakis, C. Rozenblat, M. Tomassini, S. Havlin, *New J. Phys.* 12 (2010) 113043.
- [29] S.N. Dorogovtsev, A.V. Goltsev, J.F.F. Mendes, *Phys. Rev. Lett.* 96 (2006) 040601.
- [30] A. Garas, F. Schweitzer, S. Havlin, *New J. Phys.* 14 (2012) 083030.
- [31] S.-W. Son, B.J. Kim, H. Hong, H. Jeong, *Phys. Rev. Lett.* 103 (2009) 228702.
- [32] A. Zeng, S.-W. Son, C.H. Yeung, Y. Fan, Z. Di, *Phys. Rev. E* 83 (2011) 045101(R).
- [33] J. Borge-Holthoefer, Y. Moreno, *Phys. Rev. E* 85 (2012) 026116.
- [34] A. Zeng, Y. Hu, Z. Di, *Europhys. Lett.* 87 (2009) 48002.
- [35] A.-L. Barabasi, R. Albert, *Science* 286 (1999) 509.
- [36] D. Lusseau, et al., *Behav. Ecol. Sociobiol.* 54 (2003) 396.
- [37] P.M. Gleiser, L. Danon, *Adv. Complex Syst.* 6 (2003) 565.
- [38] M.E.J. Newman, *Phys. Rev. E* 74 (2006) 036104.
- [39] R. Guimera, L. Danon, A. Diaz-Guilera, F. Giralt, A. Arenas, *Phys. Rev. E* 68 (2003) 065103.
- [40] M.E.J. Newman, *Proc. Natl. Acad. Sci. USA* 98 (2001) 404.
- [41] M. Boguna, R. Pastor-Satorras, A. Diaz-Guilera, A. Arenas, *Phys. Rev. E* 70 (2004) 056122.
- [42] H. Jeong, B. Tombor, R. Albert, Z.N. Oltvai, A. Barabasi, *Nature* 407 (2000) 651.
- [43] J. Duch, A. Arenas, *Phys. Rev. E* 72 (2005) 027104.
- [44] A.C. Gavin, et al., *Nature* 415 (2002) 141.
- [45] H. Jeong, S.P. Mason, A. Barabasi, Z.N. Oltvai, *Nature* 411 (2001) 41.
- [46] D.J. Watts, S.H. Strogatz, *Nature* 393 (1998) 440.
- [47] M.E.J. Newman, <http://www-personal.umich.edu/~mejn/netdata/>.
- [48] M.G. Kendall, *Biometrika* 30 (1938) 81.
- [49] M. Chavez, D.U. Hwang, A. Amann, H.G. Hentschel, S. Boccaletti, *Phys. Rev. Lett.* 94 (2005) 218701.