Efficient calculation of the robustness measure R for complex networks

Chen Hong[a], Ning He[a], Oriol Lordan[b,*], Bo-Yuan Liang[c,d,e], Nai-Yu Yin[c,d,e]

[a] *College of Information Technology, Beijing Union University, Beijing 100101, P.R.China*
[b] *Universitat Politècnica de Catalunya-BarcelonaTech, Colom 11, Terrassa 08222, Spain*
[c] *School of Electronic and Information Engineering, Beihang University, Beijing 100191, P.R.China*
[d] *Beijing Key Laboratory for Network-based Cooperative Air Traffic Management, Beijing 100191, P.R.China*
[e] *Beijing Laboratory for General Aviation Technology, Beijing 100191, P.R.China*

Abstract

In a recent work [Proc. Natl. Acad. Sci. USA, 108 (2011) 3838], Schneider et al. proposed a new measure R for network robustness, where the value of R is calculated within the entire process of malicious node attacks. In this paper, we present an approach to improve the calculation efficiency of R, in which a computationally efficient robustness measure R' is introduced when the fraction of failed nodes reaches to a critical threshold q_c. Simulation results on three different types of network models and three real networks show that these networks all exhibit a computationally efficient robustness measure R'. The relationships between R' and the network size N and the network average degree $\langle k \rangle$ are also explored. It is found that the value of R' decreases with N while increases with $\langle k \rangle$. Our results would be useful for improving the calculation efficiency of network robustness measure R for complex networks.

Keywords: Network robustness, Robustness measure, Malicious attack, Complex networks

*Corresponding author
Email address: oriol.lordan@upc.edu (Oriol Lordan)

1. Introduction

A wide range of systems in nature and society can be described as complex networks, such as the World Wide Web, neural networks and air transportation networks, etc. In the past decades, the study of complex networks has given rise to great achievements in many fields [1, 2, 3, 4], such as network modeling [5, 6, 7], cascading failures [8, 9, 10, 11, 12], evolutionary games [13, 14, 15, 16], optimization [17, 18, 19] and traffic dynamics [20, 21, 22] and so on.

Large infrastructure networks such as the Internet, power grids and transportation systems [23, 24] play a significant role in the modern world. As the robustness of infrastructure networks is becoming more and more important, the robustness of complex networks has attracted many researchers in recent years [25, 26, 27, 28, 29]. Albert et al [30] found that complex networks with scale-free character are robust to random failures but vulnerable under malicious attacks. Cohen et al [31] explored the robustness of the Internet and proposed an analytical approach to find the critical percolation threshold on random networks. Holme et al [32] investigated the effect of four attacking strategies: removal by descending order of betweenness and degree, calculated for either the current network during the removal process or the initial network. It is found that adaptive attack strategies are more effective than attack strategies based on the initial network.

Recently, Schneider et al. [33] proposed a new measure R for network robustness and investigated optimal network structure against high-degree node removal with respect to this measure. They found that the final robust networks exhibit an onion structure in which highly connected nodes form a core surrounded by rings of nodes with decreasing degree. Following the pioneering work of Schneider et al, many researchers have used this new robustness measure and onion-like structure to explore the robustness of networks [34, 35, 36, 37]. Wu et al. [38] proposed a generative algorithm to efficiently produce synthetic scale-free networks with onion structure and validated the robustness of their generated networks against malicious attacks and random failures. Complementary to the node-robustness measure, Zeng et al. [39] proposed a link-robustness index and designed a hybrid greedy algorithm to against both node and link attacks. The results show that network robustness can be significantly improved. In previous works, the value of network robustness measure R is calculated within the whole process of malicious attacks, resulting in a time-consuming calculation process. In modern so-

ciety, there are many large networks such as the World Wide Web and the Internet, hence the computation cost of network robustness measure on large networks needs to be considered. This calls for a quicker, smarter method to calculate the value of network robustness measure R. In this paper, we propose a computationally efficient robustness measure corresponding to the critical fraction of attacked nodes and confirm its reasonability on three types of network models and three real complex networks.

The paper is organized as follows. In the next section we demonstrate the computationally efficient robustness measure and attacking strategies in detail. In Section 3, simulation results and correspondent theoretical analysis are provided. Finally, the work is summarized in Section 4.

2. The Model

The unique network robustness measure proposed by Schneider et al. is defined as [33]

$$R = \frac{1}{N} \sum_{q=1/N}^{1} s(q), \tag{1}$$

where N is the number of nodes in the network, q is the fraction of removed nodes and $s(q)$ is the fraction of nodes in the largest connected component after removing qN largest degree nodes. The normalization factor $1/N$ ensures the comparability of network robustness of different sizes. The range of possible values of R is between 0 and 0.5, where $R = 0$ corresponds to a star network, in which all nodes in the network are isolated after removing the hub node. If $R = 0.5$, the original network is a fully connected network and the largest connected component decreases only one node at each node attack step [34]. Obviously, networks with higher value of R are of stronger resistance to targeted node attacks.

Since the robustness measure R captures the effects on the network over the entire attack sequence, it is especially time-consuming when the size of the network is huge. From the definition of R, we can see that R is strongly correlated with the size of the largest connected component. It is known that network robustness can also be measured by the critical percolation threshold q_c, which is the minimum value of the remaining node fraction required for a unique giant component to be of the order of the entire network under attacks [30, 31, 32]. For huge networks, since the change of $s(q)$ is relatively small after the giant component completely collapses, it is reasonable that

the calculation efficiency of R will be efficiently improved if the calculation process is stopped at $q = q_c$.

To estimate the calculation efficiency of R, we define a cost-based function

$$R(t) = \frac{1}{N}\sum_{q=1/N}^{t} s(q) \approx t - \frac{t(tN+1)}{2N}, \tag{2}$$

where t $(1/N \leq t \leq 1)$ is the cost indicator of the calculation process and $R = R(1)$. Obviously, the smaller the value of t, the lower the calculation cost of $R(t)$. If the network is fully connected and the largest connected component decreases only one node at each node attack step, we can get $R(t) \approx \frac{1}{N}(\frac{N-1}{N} + \frac{N-2}{N} + \cdots + \frac{N-tN}{N}) = t - \frac{t(tN+1)}{2N}$. Consequently, $R(1) \approx (N-1)/2N \approx 0.5$, $R(1/N) \approx (N-1)/N^2$ and $R(1/N) \approx 0$ for large N values, indicating that the range of $R(t)$ values is the same as that for R.

Based on above analyses, we propose a computationally efficient robustness measure which is defined as

$$R' = R(q_c) = \frac{1}{N}\sum_{q=1/N}^{q_c} s(q) \approx q_c(1 - \frac{1}{2N}) - \frac{q_c^2}{2}, \tag{3}$$

where q_c is the critical threshold at which the giant component is completely collapsed. Since q_c is usually smaller than one, the calculation efficiency of R' is higher than that of R. Obviously, the higher the value of R' is, the stronger the network robustness is.

To investigate the robustness of networks, different attacking strategies have been intensively studied [40, 41], including node attacking strategies [30, 31, 42] and edge attacking strategies [32]. Node attacking strategies are usually based on node centrality measures [43], such as node degree [44], betweenness [45] and closeness [46], etc. In our model, we adopt two commonly used node attacking strategies: the high-degree adaptive attack (HDA) [32, 36, 47] and the high-betweenness adaptive attack (HBA) [32]. For HDA, the node with the highest degree is deleted at each attack step and the highest degree of the network is recalculated before each attack. In the case of HBA, the node with the highest betweenness will be removed at each attack step and we recalculate the highest betweenness of the network before each attack. It is noteworthy that the edges linking the removed node are deleted simultaneously.

To validate the reasonability of the computationally efficient robustness measure R', we will apply it to the investigation of the robustness of three

different types of network models and three real-world networks under HDA and HBA.

3. Simulation Results

Firstly, we discuss the robustness of three well-known network models: Erdös-Rényi (ER) random networks [5], Watts-Strogatz (WS) small-world networks [6] and Barabási-Albert (BA) scale-free networks [7]. For above network models, the networks size N is 3000. Here we set $m = m_0 = 3$ for BA networks, the rewiring probability $p = 0.01$ for WS networks and the connection probability $p = 0.002$ for ER networks. Figure 1(a) shows $R(t)$ and $s(q)$ as a function of q under HDA on ER networks. One can see that the value of $s(q)$ decreases with the increment of q and $s(q) \approx 0$ when $q = q_c$. On the other hand, the value of $R(t)$ increases with q and $R = R(1) \approx R(q_c) = R'$ when $q > q_c$. This means that the calculating procedure of R can be effectively stopped at $q = q_c$, where the giant component is completely destroyed. Meanwhile, we can see that $q_c \ll 1$, which indicates that the computational cost of R can be significantly decreased. In the case of WS networks (Figure 1(b)), the increment of $R(t)$ can also be ignored when $q > q_c$. For BA networks (Figure 1(c)), although the value of q_c is quite small, we can see that $R \approx R'$ when $q > q_c$.

For attacking strategy of HBA, the computationally efficient robustness measure R' are displayed as well (Fig. 1(d)-(f)), indicating that the computational cost of R can also be lowered under HBA for three different network models. Meanwhile, due to the lowest q_c value of BA networks, the calculating efficiency of R' for BA networks is highest in three network models.

To further confirm the reasonability of R', we also investigate $R(t)$ and $s(q)$ as a function of q on three real networks: coauthorship network, power grid and the Internet. Here, the coauthorship network is a network of coauthorships between scientists working on network theory and experiment [48]; the power grid is an undirected unweighted representation of the topology of the western states power grid of the United States [6]; the Internet is a symmetrized snapshot of the Internet structure at the level of autonomous systems. From Figure 2, one can see that $R \approx R'$ when $q > q_c$ in all three real networks whatever the attacking strategy is, which is in good accordance with the result of three network models.

To measure the computational efficiency for R' against R, we investigate the reduction of node attack steps for R' against R on different networks

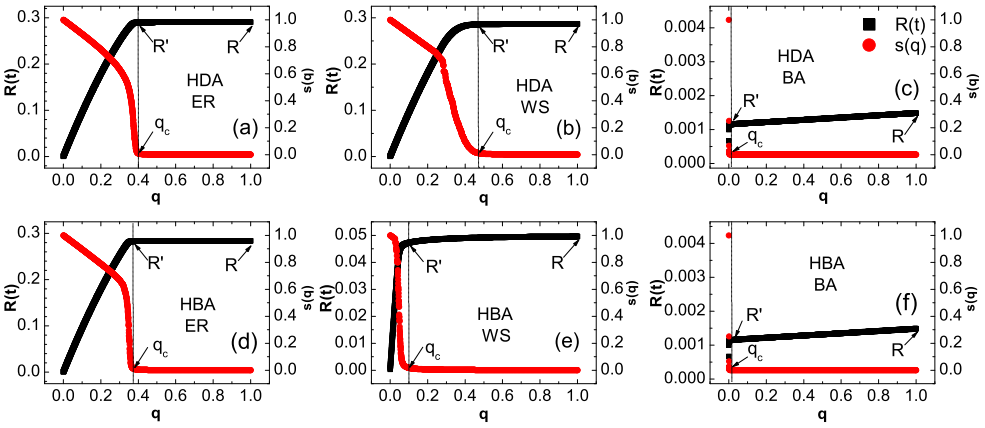

Figure 1: $R(t)$ and $s(q)$ as a function of q under HDA on different network models: (a) ER; (b) WS; (c) BA. $R(t)$ and $s(q)$ as a function of q under HBA on three network models: (d) ER; (e) WS; (f) BA. For these network models, the network size $N = 3000$ and the average degree $\langle k \rangle = 6$. Here we set $m = m_0 = 3$ for BA networks, the rewiring probability $p = 0.01$ for WS networks and the connection probability $p = 0.002$ for ER networks. Each figure is averaged over 10 independent realizations.

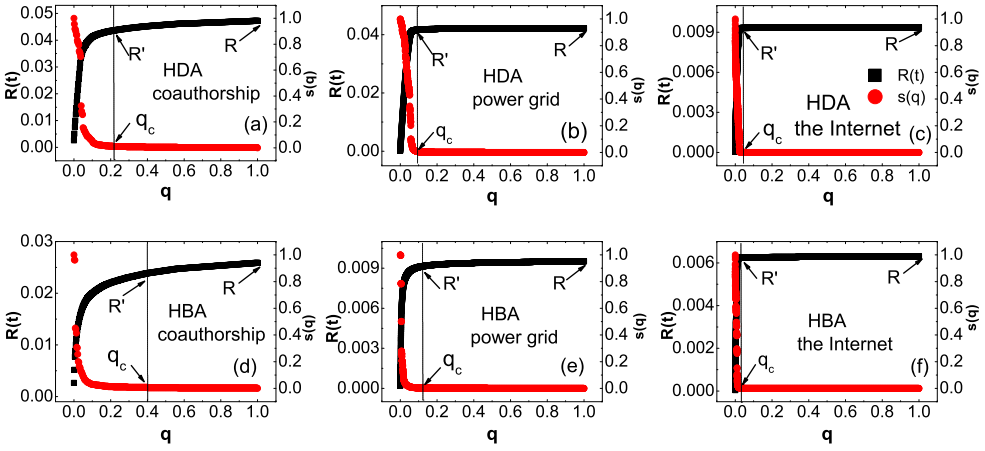

Figure 2: $R(t)$ and $s(q)$ as a function of q under HDA on three real networks: (a) coauthorship network; (b) power grid; (c) the Internet. $R(t)$ and $s(q)$ as a function of q under HBA on three real networks: (d) coauthorship network; (e) power grid; (f) the Internet. Here the size of three real networks is $N_{coauthor} = 379$, $N_{power} = 4941$ and $N_{internet} = 22963$, respectively.

(Table 1). In our approach, the reduction of node attack steps $N_r = N - q_c N = (1 - q_c)N$, where N is the number of nodes in the network.

Table 1: The reduction of node attack steps N_r for different networks. Here the value of q_c and N are come from data of Figures 1 and 2.

	ER		WS		BA		coauthorship		power grid		the Internet	
	$1-q_c$	N_r	$1-q_c$	N_r	$1-q_c$	N_r	$1-q_c$	N_r	$1-q_c$	N_r	$1-q_c$	N_r
HDA	0.60	1800	0.61	1830	0.96	2880	0.78	295	0.90	4447	0.96	22044
HBA	0.62	1860	0.90	2700	0.97	2910	0.60	227	0.88	4348	0.97	22274

From Table 1, one can see that the reduction of node attack steps is considerable for all networks. More than 60% of attack steps are unnecessary, especially for BA networks (97%) and the Internet (97%). As we do not need to calculate the largest connected component and search for the node of largest degree in it in the omitted steps, calculation costs decrease considerably. Therefore, the computational efficiency for R' against R is greatly improved.

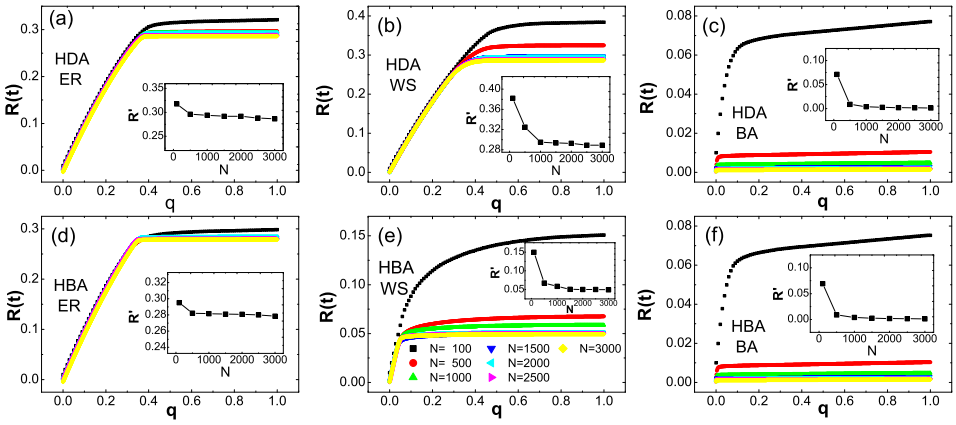

Figure 3: $R(t)$ as a function of q under HDA and HBA with different network sizes ($N = 100, 500, \cdots, 2500$) on three network models: (a) ER, HDA; (b) WS, HDA; (c) BA, HDA; (d) ER, HBA; (e) WS, HBA; (f) BA, HBA. The inset shows the relationship between the computationally efficient robustness measure R' and network size N. Here the average degree $\langle k \rangle = 6$, and each datum is averaged over 10 independent realizations.

In our approach, the time requirement for R' and R is $O(N^3)$. To explore the calculation efficiency of R' against R in more detail, we investigate the real execution time for R' and R on different networks (Table 2).

Table 2: The real execution time (in seconds) for R' and R on a specific environment, where CPU: i5-4590; Memory: 8G; OS: Ubuntu 16.04 LTS; Programming language: R 3.3.2. Here the network parameters are the same as that in Figures 1 and 2.

	ER		WS		BA		coauthorship		power grid		the Internet	
	R'	R	R'	R	R'	R	R'	R	R'	R	R'	R
HDA	8.70	14.18	7.85	13.57	0.58	10.55	0.25	0.93	4.99	25.06	33.26	377.58
HBA	430.05	435.72	60.40	69.98	1.51	11.56	0.41	0.92	41.49	60.84	2754.98	3096.76

From Table 2, we can see that the real execution time for R' is significantly smaller than execution time of R on different types of networks. Especially, more than 94% execution time is saved for BA networks under HDA.

Next, we will discuss the effect of the network size N on the computationally efficient robustness measure R'. Figure 3 shows $R(t)$ as a function of q with different network sizes ($N = 100, 500, \cdots, 2500$), and the relationship between R' and N is depicted in the inset. This shows that, for three network models, $R \approx R'$ when the value of q goes beyond a critical value whatever the value of N is, reflecting that the calculation efficiency of R can be improved under different network sizes. One can see that the value of R' decreases with the increment of N for three network models. On the other hand, the value of R' is almost the same when N is large. From Eq. (3), we can get $R' \approx q_c - q_c^2/2$ for large N values, thus an approximate value of R' is displayed with respect to large N values.

The average degree is an important topology parameter in complex networks. It is known that network robustness is heavily affected by the network average degree [42]. To explore the impact of the average degree $\langle k \rangle$ on R', we plot the relationship between $R(t)$ and q on three network models with different average degrees ($\langle k \rangle = 6, 8, \cdots, 16$), and the inset shows R' as a function of $\langle k \rangle$ (Figure 4). This shows that, for three network models, $R \approx R'$ when the value of q goes beyond a certain threshold whatever the value of $\langle k \rangle$ is, indicating that the calculation efficiency of R can be improved under different average degrees. Besides, one can see that the value of R' increases with $\langle k \rangle$, regardless of the network models and attacking strategies, indicating improved network robustness for more compact structures.

4. Conclusion

To summarize, we have proposed a computationally efficient robustness measure R', where the time-consuming calculation process of robustness measure R can be effectively stopped at a critical threshold q_c, at which the giant

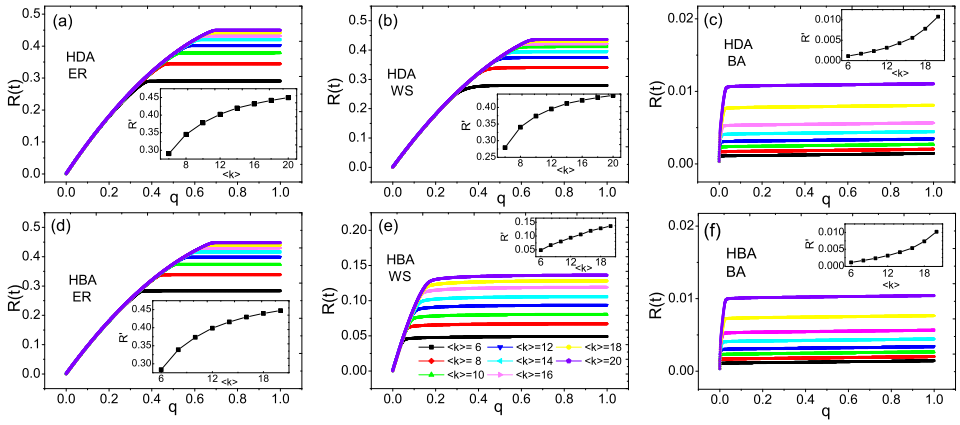

Figure 4: $R(t)$ as a function of q under HDA and HBA with different average degrees ($\langle k\rangle = 6, 8, \cdots, 16$) on three network models: (a) ER, HDA; (b) WS, HDA; (c) BA, HDA; (d) ER, HBA; (e) WS, HBA; (f) BA, HBA. The inset shows the relationship between the computationally efficient robustness measure R' and the average degree $\langle k\rangle$. Here the network size $N = 3000$, and each datum is averaged over 10 independent realizations.

component is completely collapsed. We confirm the reasonability of R' on three different types of network models and three real-world networks. The results show that R' is effective on all these networks. Furthermore, the effects of the network size and the average degree on R' are investigated. It is found that the value of R' increases with the average degree of networks, yet an opposite effect is observed for the network size.

Acknowledgements

The authors would like to thank Wenbo Du for useful conversations. This paper is supported by the National Natural Science Foundation of China (Grant Nos. 91538204, 61425014, 61521091, 61370138, 61572077), National Key Research and Development Program of China (Grant No. 2016YFB1200100), National Key Technology R&D Program of China (Grant No. 2015BAG15B01), and Beijing Municipal Natural Science Foundation (Nos. 4152017, 4162027).

References

[1] M. E. J. Newman, SIAM Rev. **45** (2003) 167.

[2] S. Boccaletti, G. Bianconi, R. Criado, C. I. del Genio, J. Gómez-Gardeñes, M. Romance, I. Sendiña-Nadal, Z. Wang and M. Zanin, Phys. Rep. **544** (2014) 1.

[3] D.-J. Wei, X.-Y. Deng, X.-G. Zhang, Y. Deng and S. Mahadevan, Physica A **392** (2013) 2564.

[4] C.-R. Cai, Z.-X. Wu, M. Z. Q. Chen, P. Holme and J.-Y. Guan, Phys. Rev. Lett. **116** (2016) 258301.

[5] P. Erdös and A. Rényi, Publ. Math. Inst. Hung. Acad. Sci. **5** (1960) 17.

[6] D. J. Watts and S. H. Strogatz, Nature **393** (1998) 440.

[7] A. L. Barabási and R. Albert, Science **286** (1999) 509.

[8] A. E. Motter and Y.-C. Lai, Phys. Rev. E **66** (2002) 065102(R).

[9] D. J. Watts, Proc. Natl. Acad. Sci. U.S.A. **99** (2002) 5766.

[10] R.-R. Liu, W.-X. Wang, Y.-C. Lai and B.-H. Wang, Phys. Rev. E **85** (2012) 026110.

[11] F. Tan, Y.-X. Xia, W.-P. Zhang and X.-Y. Jin, EPL **102** (2013) 28009.

[12] R.-R. Liu, M. Li, C.-X. Jia and B.-H. Wang, Sci. Rep. **6** (2016) 25294.

[13] M. Perc and A. Szolnoki, BioSystems **99** (2010) 109.

[14] W.-B. Du, X.-B. Cao, M.-B. Hu and W.-X. Wang, EPL **87** (2009) 60004.

[15] Z. Wang, A. Szolnoki and M. Perc, New J. Phys. **16** (2014) 033041.

[16] W.-B. Du, X.-B. Cao, L. Zhao and M.-B. Hu, Physica A **388** (2009) 4509.

[17] W.-B. Du, W. Ying, G. Yan, Y.-B. Zhu and X.-B. Cao, IEEE Trans. Circuits Syst. II, doi: 10.1109/TCSII.2016.2595597.

[18] W.-B. Du, Y. Gao, C. Liu, Z. Zheng and Z. Wang, Appl. Math. Comput. **268** (2015) 832.

[19] Y. Gao, W.-B. Du and G. Yan, Sci. Rep. **5** (2015) 9295.

[20] G. Yan, T. Zhou, B. Hu, Z.-Q. Fu and B.-H. Wang, Phys. Rev. E **73** (2006) 046108.

[21] Y.-X. Xia, N.-J Liu and H. H. C. Iu, Chaos Soliton. Fract. **42** (2009) 1700.

[22] Z.-X. Wu, W.-X. Wang and K. H. Yeung, New J. Phys. **10** (2008) 023025.

[23] J. Zhang, X.-B. Cao, W.-B. Du and K.-Q. Cai, Physica A **389** (2010) 3922.

[24] W.-B. Du, X.-L. Zhou, O. Lordan, Z. Wang, Z. Chen and Y.-B. Zhu, Transport. Res. Part E **89** (2016) 108.

[25] S. V. Buldyrev, R. Parshani, G. Paul, H. E. Stanley and S. Havlin, Nature **464** (2010) 1025.

[26] C. Hong, J. Zhang, X.-B. Cao and W.-B. Du, Chaos Soliton. Fract. **86** (2016) 28.

[27] O. Lordan, J. M. Sallan, P. Simo and D. Gonzalez-Prieto, Commun. Nonlinear. Sci. Numer. Simulat. **22** (2015) 587.

[28] F. Tan, Y.-X. Xia and Z. Wei, Phys. Rev. E **91** (2015) 052809.

[29] S. Trajanovski, S. Scellato and I. Leontiadis, Phys. Rev. E **85** (2012) 066105.

[30] R. Albert, H. Jeong and A.-L. Barabási, Nature **406** (2000) 378.

[31] R. Cohen, K. Erez, D. ben-Avraham and S. Havlin, Phys. Rev. Lett. **85** (2000) 4626.

[32] P. Holme, B. J. Kim, C. N. Yoon and S. K. Han, Phys. Rev. E **65** (2002) 056109.

[33] C. M. Schneider, A. A. Moreira, J. S. Andrade Jr, S. Havlin and H. J. Herrmann, Proc. Natl. Acad. Sci. USA **108** (2011) 3838.

[34] L. Bai, Y.-D. Xiao, L.-L. Hou and S.-Y. Lao, Chin. Phys. Lett. **32** (2015) 078901.

[35] L.-L. Ma, J. Liu, B.-P. Duan and M.-X. Zhou, EPL **111** (2015) 28003.

[36] H. J. Herrmann, C. M. Schneider, A. A. Moreira, J. S. Andrade Jr and S. Havlin, J. Stat. Mech. (2011) P01027.

[37] T. Tanizawa, S. Havlin and H. E. Stanley, Phys. Rev. E **85** (2012) 046109.

[38] Z.-X. Wu and P. Holme, Phys. Rev. E **84** (2011) 026106.

[39] A. Zeng and W.-P. Liu, Phys. Rev. E **85** (2012) 066130.

[40] R. Cohen, K. Erez, D. ben-Avraham and S. Havlin, Phys. Rev. Lett. **86** (2001) 3682.

[41] J.-X. Gao, S. V. Buldyrev, S. Havlin and H. E. Stanley, Phys. Rev. Lett. **107** (2011) 195701.

[42] C. Hong, X.-B. Cao, W.-B. Du and J. Zhang, Phys. Scr. **87** (2013) 055801.

[43] P. Crucitti, V. Latora and S. Porta, Phys. Rev. E **73** (2006) 036125.

[44] J. Nieminen, Scand. J. Psych. **15** (1974) 322.

[45] L. C. Freeman, Sociometry **40** (1977) 35.

[46] G. Sabidussi, Psychometrika **31** (1966) 581.

[47] W.-B. Du, B.-Y. Liang, G. Yan, O. Lordan and X.-B. Cao, [arXiv:1608.00142v1] (2016).

[48] M. E. J. Newman, Phys. Rev. E **74** (2006) 036104.