Modified localized attack on complex network

2016 EPL 113 28002

(http://iopscience.iop.org/0295-5075/113/2/28002)

A Letters Journal Exploring the Frontiers of Physics

January 2016

EPL, **113** (2016) 28002
doi: 10.1209/0295-5075/113/28002

www.epljournal.org

Modified localized attack on complex network

Gaogao Dong, Ruijin Du, Huifang Hao and Lixin Tian

Nonlinear Scientific Research Center, Faculty of Science, Jiangsu University - Zhenjiang, 212013, China

received 10 October 2015; accepted in final form 21 January 2016
published online 8 February 2016

PACS 89.75.Fb – Structures and organization in complex systems
PACS 64.60.aq – General studies of phase transitions: Networks

Abstract – Since a shell structure contains a wealth of information, it is not only very important for understanding the transport properties of the network, but also essential to identify influential spreaders in complex networks. Nodes within each shell can be classified into two categories: protected nodes and unprotected nodes. In this paper, we propose a generalization of the localized attack, modified localized attack, which means that when a randomly chosen node (root node) is under attack, protected nodes will not be removed, but unprotected nodes in the nearest shells will fail. We numerically and analytically study the system robustness under this attack by taking an Erdös-Rényi (ER) network, a regular random (RR) network and a scale-free (SF) network as examples. Moreover, a fraction of nodes belonging to giant component S and a critical threshold q_c, where S approaches to zero, are given. The result implies that increasing connection density has been found to be useful to significantly improve network robustness.

Introduction. – Inspired by empirical studies of networked systems such as the Internet, social networks, and biological networks, networks have also been studied extensively by researchers in physics, mathematics, informatics and computer science etc. [1,2]. A complex network is an abstract consideration for real systems, which not only understand functional behaviors of real networks better, but also provide theoretical guidance for designing real networks [3,4]. The massive and comparative analysis of networks from different fields has produced a series of unexpected and dramatic results [5–9]. In fact, the last decade has witnessed the birth of a new movement of interest and research in the study of complex networks, which involve topological properties of complex networks, spreading of epidemics and rumors, synchronization and link prediction on complex networks [10–16]. Specially, understanding the network robustness can shed light on the connection mechanisms of real networks. Network robustness under deliberate attacks and failures has experienced a growing interest in recent years, which was also one of the first issues to be explored in the literature on complex networks because of obvious practical reasons [10,17–20]. The first numerical studies on the robustness of real networks are reported in refs. [21,22]. Albert *et al.* numerically studied the robustness of scale-free networks under randomly attack, which means nodes are randomly removed from a network [22]. Their findings highlight that scale-free networks display a surprisingly high degree of tolerance against random failures. Besides numerical simulations, a series of analytical approaches to study tolerance to errors and attacks in complex networks has been proposed. Cohen *et al.* [23] numerically and analytically studied the robustness of scale-free networks under randomly and intentional attack by using percolation theory, respectively [10]. Gallos *et al.* studied the tolerance and topology of scale-free networks under targeted attack by applying the probability for high-degree nodes to be attacked is higher than for low-degree nodes. And, they found that little knowledge on the highly connected nodes in an intentional attack reduces the threshold drastically compared with the random case [24]. Shao *et al.* developed a percolation framework to analytically and numerically study the robustness of complex networks against localized attack, which means that all nodes in the nearest shells of the central nodes are removed [25]. Their results can provide an insight into understanding how to facilitate the design of resilient infrastructures. More recently, since real systems are susceptible to geographically localized damage, Berezin *et al.* developed a theoretical and numerical approach and studied a general model of spatially embedded networks with dependences under localized attacks. Additionally, they also predicted the effects of localized attacks on spatially embedded systems with dependences [26]. Besides, different approaches to

study the vulnerability of a network have been proposed by Latora *et al.* [27], Goh *et al.* [28], Kim *et al.* [29], Motter *et al.* [30] and Holme *et al.* [31,32].

An important characteristic of a network is the structure of its shells, where it is defined as the set of nodes that are at some distance from a randomly chosen root node. In fact, since the shell structure in a complex network contains a wealth of information, it is not only very important for understanding the transport properties of the network such as the epidemic spread and information diffusion, but also essential to identify and rank influential spreaders in complex networks [8,33]. Additionally, because many real systems show the characteristic of a shell structure, cascading failures caused by previous attacks modes, random attack and targeted attack, cannot adequately and clearly describe real-world scenarios. For instance, strong earthquakes destroy infrastructures, shell after shell, noxious pathogens infect the surrounding cells and make them lose the normal functions, and some dangerous trojan virus can even cause the whole computer network paralysis by spreading on the shell structure, etc. Because of this, one takes protective measures to avoid damages or reduce risks, such as improving the earthquake resistance of the infrastructures, injecting vaccines to increase immunity, and installing or updating Anti-Virus Software and so on. However, these protective measures sometimes take a great cost, and not all the nodes in the system have safe protective measures.

Inspired by these motivations, nodes within each shell are classified into two categories: protected nodes and unprotected nodes. Here we propose a generalization of the localized attack, the modified localized attack model, which means that when a randomly chosen root node is under attack and malfunction, only a fraction of nodes of each shell (unprotected nodes) will be removed because of the connection relationship between shells (as shown in fig. 1).

Model. – For a network with degree distribution $P(k)$, the generating function and generation function of this branching process are given by $G_0 = \sum_{k=0}^{\infty} P(k)x^k$ and $G_1 = \frac{G_0'(x)}{G_0'(1)}$, respectively [19,20]. The averaged fraction of the l-th nodes z_l satisfies the following recursion relation [8]:

$$z_l = [G_1'(1)]^{l-1} G_0'(1) = \left[\frac{z_2}{z_1}\right]^{l-1} z_1, \tag{1}$$

where $z_1 = G_0'(1)$ and $z_2 = G_0'(1)G_1'(1)$.

Let p_l denote the fraction of randomly chosen unprotect nodes within the l-shell for a randomly selected root node. Thus, the fraction of removed nodes from the network is

$$q = \frac{1 + \sum_1^l p_l N_l}{N}, \tag{2}$$

where N_l is the number of nodes in the l-shell. Without loss of generality, we begin to focus on the changes of the

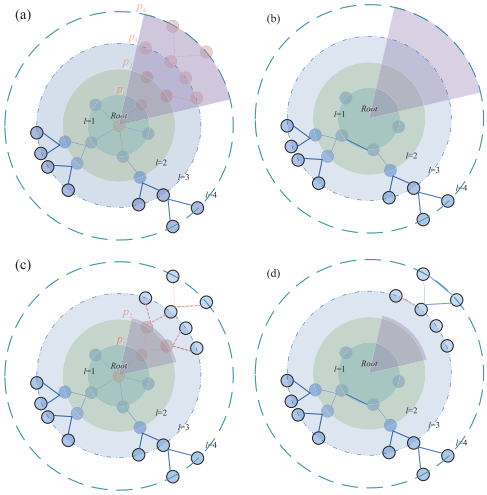

Fig. 1: (Colour online) Schematic representation of modified localized attack progress. p_l (red nodes) and $1 - p_l$ denote the fraction of unprotected and protected nodes at each shell l ($l = 1, 2, 3, 4$) in (a) and (c), respectively. (b) and (d): root node (attacked node), unprotected nodes p_l in shell l and the corresponding links are removed. (c) There exists unprotected nodes in the whole shell $l = 1, 2, 3, 4$. (d) Unprotected nodes only exist in the partial shell $l = 1, 2$.

generating function after the shell attack. When randomly removing the q fraction of nodes shell by shell but keeping the links connecting the removed nodes to the remaining nodes, the degree distribution $P_{1-q}(k)$ can be expressed as [25,34,35]

$$P_{1-q}(k) = P(k)\frac{f^k}{G_0(f)}, \tag{3}$$

where $f = G_0^{-1}(1-q)$.

And the generating function becomes

$$G_a(x) = \sum_{k=0} P_{1-q}(k)x^k = \frac{G_0(fx)}{G_0(f)}. \tag{4}$$

Next, the links that connected to the remaining nodes and to other nodes on the same shell are removed. And the number of links of the remaining nodes is [25,33]

$$\tilde{L}(f) = \frac{L(f)(1-q)N\langle k(f)\rangle}{(1-q)N\langle k(f)\rangle + L(f)} = N\left[fG_0'(f) - \frac{G_0'(f)^2}{G_0'(1)}\right], \tag{5}$$

where $L(f) = N(G_0'(1)f^2 - G_0'(f)f)$ denotes the open links belonging to the aggregate (root and all nodes already connected to the root) [33] and the average degree of the remaining network is $\langle k(f)\rangle = \frac{fG_0'(f)}{G_0(f)}$.

Thus, the fraction of the original links that connect to the remaining nodes satisfies

$$1 - \tilde{q} = 1 - \frac{\tilde{L}(f)}{(1-q)N\langle k(f)\rangle} = \frac{G_0'(f)}{G_0'(1)f}. \tag{6}$$

At this stage, the generation function becomes

$$G_b(x) = G_a(\tilde{q} + x - \tilde{q}x) = \frac{1}{G_0(f)} G_0\left[f + \frac{G_0'(f)}{G_0'(1)}(x-1)\right]. \tag{7}$$

Since the random connection process within the network can be modeled as a branching process, the outgoing links of any node have a probability to be connected to a node with degree k. Accordingly, the generation function of this branching process is expressed as

$$H(x) = \frac{\sum_{k=0}^{\infty} P_{1-q} k x^{k-1}}{\langle k \rangle} = \frac{G_b'(x)}{G_b'(1)}. \tag{8}$$

Let R be the probability that the node, arriving by following a randomly chosen connectivity link, belongs to the giant component of the final network. R satisfies the transcendental equation [8]

$$R = H(R). \tag{9}$$

Furthermore, the probability that a randomly chosen node does belong to a giant component is given by

$$g(R) = 1 - G_b(R). \tag{10}$$

Accordingly, S_1 is equal to the probability that a randomly chosen node belongs to the giant component is

$$S_1 = (1-q)g(R) = (1-q)(1 - G_b(R)). \tag{11}$$

Moreover, as q increases, the nontrivial solution R gradually approaches to the trivial solution $H(1) = 1$, which means that the network collapses and the size of the giant component becomes zero at the critical value $1 - q_c$. Thus, the critical value can be obtained from

$$\begin{cases} q_c = \dfrac{1 + \sum_1^l p_l [G_1'(1)^{l-1} G_0'(1)]}{N}, \\ \dfrac{G_0''(f_c)}{G_0'(1)} = 1, \\ f_c = G_0^{-1}(1 - q_c). \end{cases} \tag{12}$$

Applications. – In this part, we study the robustness of three different topological networks, ER, RR and SF networks, under modified localized attack. For the ER network with Poisson degree distribution $P_k = \frac{e^{-\widetilde{k}}\widetilde{k}^k}{k!}$, the corresponding generation functions are be written as $G_0 = G_1 = e^{\widetilde{k}(x-1)}$, respectively, where $\widetilde{k}$ is the average degree of the network [8,35]. Thus, from eqs. (7) and (11), we express the size of the giant component under the attack as follows:

$$S_1 = \frac{N - 1 - \sum_1^l p_l N_l}{N}(1 - e^{-\widetilde{k}S}), \tag{13}$$

where N_l can be obtained from eq. (1) (standard result of the networks). Figures 2(a)–(c) suggest that the simulation results agree well with the theoretical predictions.

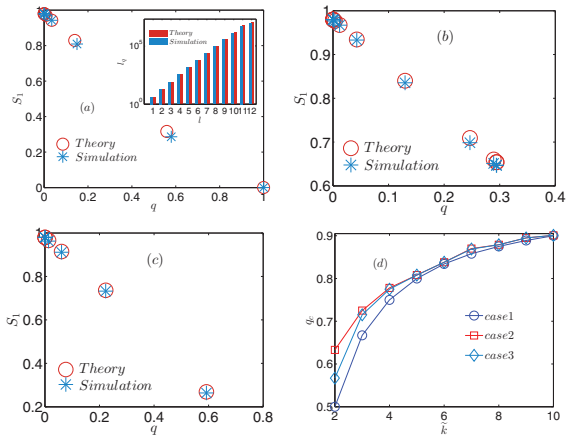

Fig. 2: (Colour online) (a)–(c) Comparison between theory and simulations of the size of the giant component S_1 *vs.* q with different parameters for the ER network. (a) $\widetilde{k} = 4$, $p_l = 1$, $l = 12$ and the number of nodes $N = 10^7$. Otherwise, comparison between eq. (1) and simulations of the size of each shell L_q as a function of l are shown in the inset. This inset verifies the eq. (1) validity, which is considered as a standard result in ref. [8]. For (b), (c), assuming $N \to \infty$ in eq. (1) and simulation complexity for N_l, we use simulated results N_l to substitute the theoretical values N_l of eq. (1) to verify analytical formula (13). (b) $\widetilde{k} = 4$, $p_1 = p_2 = p_3 = p_4 = p_5 = 0.5$, $p_6 = p_2 = \cdots = p_L = 0.3$ and $N = 10^6$, where L denotes the furthest shell from the root node. (c) $\widetilde{k} = 4$, $p_1 = p_3 = p_5 = p_7 = p_8 = p_9 = p_{10} = 1$, the fraction of unprotected nodes in other shells is equal to zero and $N = 10^6$. (d) q_c as a function of $\widetilde{k}$ for three different cases. The parameters are, respectively, $p_1 = p_2 = \cdots = p_L = q$ (the first case), $p_1 = p_2 = \cdots = p_{L-4} = 0$, $p_{L-3} = \cdots = p_L = q$ (the second case) and $p_1 = p_2 = \cdots = p_{L-3} = 0$ and $p_{L-2} = \cdots = p_L = q$ (the third case). Additionally, for the fraction of nodes in L shell nodes, we use $N - \frac{1+\sum_{L-1}^{l} N_{L-1}}{N}$ to express it. The simulation results are averaged over $N = 10^7$ realizations for (a), the others are 10^3 realizations in the simulations.

One can find that the system undergoes a second-order phase transition as the attacking strength q increases. The critical threshold q_c can be obtained according to eq. (12),

$$q_c = \frac{1 + \sum_1^l p_l N_l}{N} = \frac{\bar{k} - 1}{\bar{k}}. \tag{14}$$

From eq. (14), we can find that q_c does not depend on p_l, but on the average degree of network. And we also observe that q_c is the same as that obtained in ref. [25] for $1 - p_c$ from eq. (14). As shown in fig. 2(d), q_c as a function of $\widetilde{k}$ for different cases either attack all shells or attack partial shells. Figure 2(d) suggests that the critical threshold gradually increases as $\widetilde{k}$ increases, which means that the system becomes more robust as the average degree increases. And, for the low connection density network, the critical threshold becomes bigger as the number of attacked shells decreases. However, the critical threshold remains almost constant for large $\widetilde{k}$, which implies that the change of the attacking region (whole shell or partial

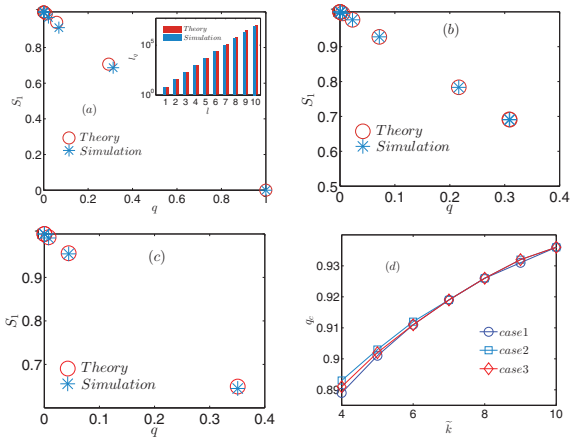

Fig. 3: (Colour online) (a)–(c) Comparison between theory and simulations of the size of the giant component S_1 *vs.* q with different parameters for the RR network. (a) $\widetilde{k} = 6$, $p_l = 1$, $l = 1, \cdots, 10$ and the number of nodes $N = 10^7$. The comparison between eq. (1) and simulations of the size of each shell L_q as a function of l is shown in the inset. (b) $\widetilde{k} = 6$, $p_1 = p_2 = p_3 = p_4 = p_5 = 0.5$, $p_6 = p_2 = \cdots = p_L = 0.3$ and $N = 10^6$. (c) $\widetilde{k} = 6$, $p_2 = p_4 = p_5 = p_8 = 1$, the other $p_l = 0$ and $N = 10^6$. (d) q_c as a function of $\widetilde{k}$ for three different cases. The parameters are, respectively, $p_1 = p_2 = \cdots = p_L = q$ (the first case), $p_1 = p_2 = \cdots = p_{L-5} = 0$, $p_{L-4} = \cdots = p_L = q$ (the second case) and $p_1 = p_2 = \cdots = p_{L-5} = 0.5$, $p_{L-2} = \cdots = p_L = q$ (the third case). The simulation results agree well with the theoretical predictions. In (a)–(c), processing methods of N_l and L are identical to those in fig. 2. The simulation results are averaged over $N = 10^7$ realizations for (a), the others are 10^3 realizations in the simulations.

shell) has no influence on the network robustness for the high connection density network. And the same results also are found for the RR network as shown in fig. 3. For the RR network with generation functions $G_0 = x^{\widetilde{k}}$ and $G_1 = x^{\widetilde{k}-1}$ [33], according to eqs. (7), (11) and (12), the analytical solutions of the giant component and critical threshold are expressed as follows:

$$\left(\frac{N-1-\sum_1^l p_l N_l}{N}-S_1\right)^{\frac{1}{\widetilde{k}}}\left[1-\left(\frac{N-1-\sum_1^l p_l N_l}{N}-S_1\right)^{\widetilde{k}-1}\right]$$
$$-\left(\frac{N-1-\sum_1^l p_l N_l}{N}\right)^{\frac{1}{\widetilde{k}}}\left[1-\left(\frac{N-1-\sum_1^l p_l N_l}{N}\right)^{\widetilde{k}-1}\right]=0, \tag{15}$$

$$q_c = \frac{1+\sum_1^l p_l N_l}{N} = 1-(\widetilde{k}-1)^{-\frac{\widetilde{k}}{\widetilde{k}-2}}. \tag{16}$$

From eq. (16), one can notice that q_c does not depend on p_l but on $\bar{k}$. And p_c within eq. (16) is the same as that obtained in ref. [25] for $1 - p_c$. Figures 3(a)–(c) show the comparison between theory and simulations of the size of the giant component S *vs.* q. In fig. 3(d) q_c is shown as a function of $\widetilde{k}$ for three different cases for $\widetilde{k} > 2$.

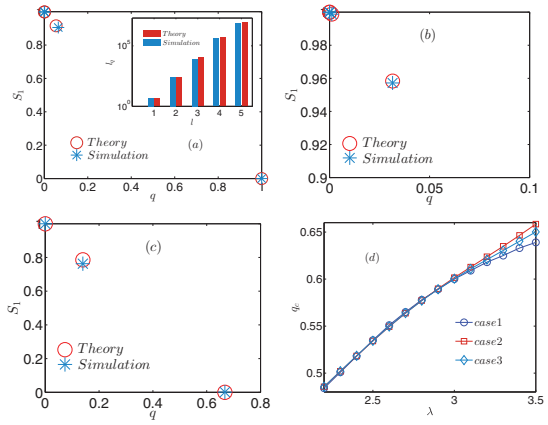

Fig. 4: (Colour online) Comparisons between theory and simulations of the size of the giant component S_1 as a function of q with different parameters for the SF network. (a) $\lambda = 2.5$, $p_l = 1$, $l = 1, \cdots, 5$ and the number of nodes $N = 10^7$. The comparison between eq. (1) and simulations of the size of each shell L_q as a function of l is shown in the inset. (b) $\lambda = 2.5$, $p_1 = p_2 = 0.5$, $p_3 = p_2 = \cdots = p_L = 0.3$ and $N = 10^6$. (c) $\lambda = 2.5$, $p_1 = p_3 = 0$, the other $p_l = 1$ and $N = 10^6$. (d) q_c as a function of λ for three different cases. The parameters are, respectively, $p_1 = p_2 = \cdots = p_L = q$ (the first case), $p_1 = p_2 = \cdots = p_{L-3} = 0$, $p_{L-2} = \cdots = p_L = q$ (the second case) and $p_1 = p_2 = \cdots = p_{L-3} = 0.3$, $p_{L-2} = \cdots = p_L = q$ (the third case). In (a)–(c), processing methods of N_l and L are similar to those in fig. 2. The simulation results are averaged over $N = 10^7$ realizations for (a), the others are 10^3 realizations in the simulations.

Additionally, from eqs. (14) and (16), we notice that the localized attack with a fraction of protected nodes does not change q_c.

For the SF network the two generation functions are respectively [35–37]

$$\begin{cases} G_0(x) = \sum_m^M \left[\left(\frac{m}{k}\right)^{\lambda-1} - \left(\frac{m}{k+1}\right)^{\lambda-1}\right] x^k, \\ G_1(x) = \dfrac{\sum_m^M \left[\left(\frac{m}{k}\right)^{\lambda-1} - \left(\frac{m}{k+1}\right)^{\lambda-1}\right] k x^{k-1}}{\sum_m^M \left[\left(\frac{m}{k}\right)^{\lambda-1} - \left(\frac{m}{k+1}\right)^{\lambda-1}\right] k}. \end{cases} \tag{17}$$

Thus, From eqs. (7), (11) and (12), the size of the giant component and that of the critical threshold are graphically shown in fig. 4. The theoretical predictions are in good agreement with the simulation results from figs. 4(a)–(c). And, for networks with power-law degree distribution, the system becomes more robust as λ increases from fig. 4(d). Moreover, the network robustness is almost unchangeable as the number of attacked shell decreases for small λ.

Conclusion. – In this paper, we propose a modified localized attack, based on the shell structure of complex networks. And for the shell attack, we can attack whole shells or partial shells for a randomly chosen root node by

adjusting parameter p_l. Additionally, we also analytically and numerically study the robustness of complex networks under this attack by using the percolation theory. Taking ER, RR, SF networks as examples, we find that the system undergoes a second-order phase transition behaviors as q increases. And the exact analytical solutions S and p_c are also expressed for three different topological networks in this paper. Results can provide a theoretical guidance and a reference value for the design of resilient infrastructures.

* * *

This work was supported by the National Natural Science Foundation of China (Grant Nos. 61403171, 71403105, 71303095), the Major research breeding project of the National Natural Science Foundation of China (Grant No. 91546118), the National Natural Science Foundation of Jiangsu Province (Grant Nos. 14KJB120001, 13KJB120001), the Jiangsu Postdoctoral Science Foundation (Grant Nos. 1402077B, 1501100B), the China Postdoctoral Science Foundation (Grant No. 2015M581738) and the Senior talents Foundation of Jiangsu University (Grant Nos. 14JDG143, 14JDG144).

REFERENCES

[1] Watts D. J. and Strogatz S. H., *Nature*, **393** (1998) 440.
[2] Albert R. and Barabási A.-L., *Rev. Mod. Phys.*, **74** (2002) 47.
[3] Dorogovtesev S. N. and Mendes J. F. F., *Adv. Phys.*, **51** (2002) 1079.
[4] Newman M. E. J., *SIAM Rev.*, **45** (2003) 167.
[5] Boccaletti S., Latora V., Moreno Y., Chavez M. and Hwang D. U., *Phys. Rep.*, **424** (2006) 175.
[6] Barabási A.-L. and Albert R., *Science*, **286** (1999) 509.
[7] Newman M. E. J., *Phys. Rev. Lett.*, **89** (2002) 208701.
[8] Newman M. E. J., *Phys. Rev. E*, **67** (2003) 026126.
[9] Marchiori M. and Latora V., *Physica A*, **285** (2000) 539.
[10] Cohen R., Erez K., BenAvraham D. and Havlin S., *Phys. Rev. Lett.*, **85** (2000) 4626.
[11] Song C., Havlin S. and Makse H. A., *Nature*, **433** (2005) 392.
[12] Gao J., Buldyrev S. V., Havlin S. and Stanley H. E., *Phys. Rev. E*, **85** (2012) 066134.
[13] Bashan A., Bartsch R. P., Kantelhardt J. W., Havlin S. and Ivanov P. C., *Nat. Commun.*, **3** (2012) 702.
[14] Buldyrev, Parshani R., Paul G., Stanley H. E. and Havlin S., *Nature*, **464** (2010) 1025.
[15] Parshani R., Buldyrev S. V. and Havlin S., *Phys. Rev. Lett.*, **105** (2010) 048701.
[16] Gao J., Buldyrev S. V., Havlin S. and Stanley H. E., *Phys. Rev. Lett.*, **107** (2011) 195701.
[17] Cohen R., Avraham D. B. and Havlin S., *Phys. Rev. E*, **66** (2002) 36113.
[18] Vázquez A., Moreno Y. and Havlin S., *Phys. Rev. E*, **67** (2003) 015101(R).
[19] Dong G., Tian L., Du R., Xiao J., Zhou D. and Stanley H. E., *EPL*, **102** (2013) 68004.
[20] Dong G., Gao J., Tian L., Du R. and He Y., *Phys. Rev. E*, **85** (2012) 016112.
[21] Broder A., Kumar R., Maghoul F., Raghavan P., Rajagopalan S., Stata R., Tomkins A. and Wiene J., *Comput. Netw.*, **33** (2000) 309.
[22] Albert R., Jeong H. and Barabási A.-L., *Nature*, **406** (2000) 378.
[23] Callaway D. S., Newman M. E. J., Strogatz S. H. and Watts D. J., *Phys. Rev. Lett.*, **85** (2000) 5468.
[24] Gallos L. K., Cohen R., Argyrakis P., Bunde A. and Havlin S., *Phys. Rev. Lett.*, **94** (2005) 188701.
[25] Shao S., Huang X., Stnaley H. E. and Havlin S., *New J. Phys.*, **17** (2015) 023049.
[26] Berezin Y., Bashan A., Danziger M. M., Li D. and Havlin S., *Sci. Rep.*, **5** (2015) 8394.
[27] Latora V. and Marchiori M., *Phys. Rev. E*, **71** (2005) 015103(R).
[28] Goh K. I., Oh E. S., Jeong H., Kahng B. and Kim D., *Proc. Natl. Acad. Sci. U.S.A.*, **99** (2002) 125832.
[29] Kim J.-H., Goh K.-I., Kahng B. and Kim D., *Phys. Rev. Lett.*, **91** (2003) 058701.
[30] Motter A. E., Nishikawa T. and Lai Y., *Phys. Rev. E*, **66** (2002) 065103.
[31] Holme P. and Kim B. J., *Phys. Rev. E*, **65** (2002) 066109.
[32] Holme P., *Phys. Rev. E*, **66** (2002) 036119.
[33] Shao J., Buldyrev S. V., Braunstein L. A. and Havlin S., *Phys. Rev. E*, **80** (2009) 036105.
[34] Huang X., Gao J., Buldyrev S. V., Havlin S. and Stanley H. E., *Phys. Rev. E*, **83** (2011) 065101(R).
[35] Dong G., Gao J., Du R., Tian L., Stanley H. E. and Havlin S., *Phys. Rev. E*, **87** (2013) 052804.
[36] Dong G., Tian L., Du R., Fu M. and Stanley H. E., *Phys. Rev. E*, **394** (2014) 370.
[37] Zhou D., Gao J., Stanley H. E. and Havlin S., *Phys. Rev. E*, **87** (2013) 052812.