New Journal of Physics
The open access journal at the forefront of physics

Deutsche Physikalische Gesellschaft DPG
IOP Institute of Physics

Published in partnership with: Deutsche Physikalische Gesellschaft and the Institute of Physics

OPEN ACCESS

RECEIVED
1 October 2018

REVISED
27 November 2018

ACCEPTED FOR PUBLICATION
10 December 2018

PUBLISHED
18 January 2019

PAPER

Localized attack on networks with clustering

Gaogao Dong[1,3], Huifang Xiao[1], Fan Wang[1], Ruijin Du[1,3,5], Shuai Shao[3], Lixin Tian[2], H Eugene Stanley[3,5] and Shlomo Havlin[4]

[1] Institute of Applied System Analysis, Faculty of Science, Jiangsu University, Zhenjiang, 212013 Jiangsu, People's Republic of China
[2] School of Mathematical Science, Nanjing Normal University, Jiangsu, 210042, People's Republic of China
[3] Center for Polymer Studies and Department of Physics, Boston University, Boston, MA 02215 United States of America
[4] Department of Physics, Bar-Ilan University, Ramat-Gan 52900, Israel
[5] Authors to whom any correspondence should be addressed.

E-mail: dudo999@126.com **and** hes@bu.edu

Keywords: robustness, percolation, localized attack, clustering network, resilience

Abstract

Network systems with clustering have been given much attention due to their wide occurrence in the real world. One focus of these studies has been on robustness of single clustered networks and interdependent clustered networks under random attack (RA) or hub-targeted attack. However, infrastructure networks could suffer from a damage that is localized, i.e. a group of neighboring nodes attacked or fail, a topic that was not studied earlier on clustered networks. In this paper, we analytically and via simulations study the robustness under localized attack (LA) of single Erdős–Rényi clustered network and interdependent clustered network. For generating networks with clustering we use two models: (i) double Poisson distribution (DPD) and (ii) fixed degree distribution (FDD). For the LA case, the DPD model shows a second order phase transition behavior for a single clustered network, while for dependent networks, the system undergoes a change of percolation phase transition from a first order (abrupt transition) to a second order (continuous) transition when the coupling strength q decreases below a critical value q_c. Our results imply that single networks become significantly more vulnerable with increasing clustering coefficient c with respect to LA. This is in contrast to RA where the robustness is almost independent of c. We obtain similar results when testing different real networks. For LA on dependent networks, we also observe that the system becomes more vulnerable as c increases. This is again in contrast to RA, where for, $q < q_c$, the system robustness is almost unaffected by increasing clustering. We also solved analytically the case of LA on random regular networks which are clustered and interdependent and find that as m (the number of clustered networks that each network depends on) or c increases, the system becomes significantly more vulnerable. We also analyzed via simulations the case of generating clustering in networks for the model of keeping a FDD, and find that the influence of clustering on the robustness of two partially interdependent networks under LA is smaller than for DPD, which is very different from these cases under RA.

1. Introduction

Over the past two decades, the study of complex networks has gained increasing attention. The main reason is that many real systems in our daily life can be described and better understood when represented as complex networks. Examples include the Internet and World Wide Web, food webs, social networks, transportation systems, electricity distribution networks, genetic networks, brain networks and many others [1–11]. An important concern in the study of complex networks is their robustness, which is important for many fields, such as ecology, biology, economics and engineering [12–18]. Network robustness deals usually with the question of the response of the network to random failures and targeted attacks. This question can be analyzed and characterized using percolation theory by studying the critical thresholds or the integrated size of the largest cluster during the attack process [19–23].

Many useful results have been obtained by analyzing the robustness of single isolated networks. However, in many real scenarios, critical infrastructures rarely appear in isolated state but usually depend on other infrastructures for functioning. This has led to the emerging sub-field of research in network science, called interdependent networks or more general, networks of networks (NON). Many constructive conclusions have been obtained which improve our understanding of the robustness of interdependent networks. Buldyrev *et al* [24] developed a framework for understanding the robustness of two fully interdependent networks under random failures, and found that interdependent networks become, due to cascading failures, significantly more vulnerable compared to their single networks counterparts and undergo an abrupt (first order transition) collapse. Subsequently, a system of two partially interdependent networks (where a fraction of q nodes in both networks depend on each other) under random failures has been studied by Parshani *et al* [25]. It was found, both analytically and numerically that reducing the coupling strength below a critical value q_c, yields to a change from a first order to a second order percolation transition. Gao *et al* developed a general framework to study the percolation behavior of n interdependent networks, suffering from random failures [26–28]. The above studies reveal that dependency links between networks make the system highly vulnerable to random failures that may yield cascading failures and understanding their mechanisms might help to design resilient infrastructures and improve existing infrastructures.

Due to the broad degree distribution of real networks, it was proposed to analyze the vulnerability with respect to a targeted attack on the high degree nodes. Such attacks have dramatic structural effects on single networks and can lead easily to network fragmentation [18, 20, 29–33]. By introducing a probability function of node degree to fail, Gallos *et al* found that for the targeted attacks case, even little knowledge of the highly connected nodes can reduce significantly the robustness compared to the random attack (RA) case [34]. By mapping the targeted-attack problem to the RA problem, Huang *et al* [35] studied the robustness of two fully interdependent networks under targeted attack. Later, Dong *et al* [36, 37] studied the robustness of two partially interdependent networks against targeted attack, and further proposed a general theoretical framework for understanding the targeted-attack problem in a NON system.

However, in many real scenarios, attacks are neither random nor targeted, but localized, which means a group of neighboring nodes in a network are attacked or fail due to natural disasters like earthquakes or floods. For example, when an earthquake occurs, it releases energy in the form of seismic waves that spread from the epicenter in all directions. According to local amplification effect, even for low-intensity earthquakes, local geological features can induce high levels of shaking ground surface in a certain radius around the center, which can destroy locally the infrastructures. Only few studies on such localized attacks (LAs) strategy have been reported. Shao *et al* developed a theoretical and numerical approach to study the robustness of complex networks against LA [38]. Berezin *et al* described and predicted the effects of LA on spatially embedded systems with dependencies, and found that a LA can cause substantially more damage compared to an equivalent RA [39]. By mapping the LA problem to a RA problem, Yuan *et al* showed how the broadness of the degree distribution affects the fragility of interdependent networks due to LA [40]. Zhao *et al* finds a mapping between overload failures and dependency links [41]. Dong *et al* proposed a modified partially LA strategy, and studied the network robustness against this attack analytically and numerically [42].

As one of the key issues in complex networks, clustered networks, which is a realistic feature appearing frequently in real network, have attracted much attention in both theoretical research and in various applied fields [43]. However, networks with clustering were studied only with respect to random failures or high degree attacks [44, 45], while the effect of LAs on clustered networks has not been studied earlier.

In this paper we study the percolation behavior due to LAs in two types of clustered networks models in single clustered network, as well as in network of interdependent networks with clustering. The two models are: (a) we generate networks for which the degree distribution of the clustered network follows double Poisson distribution (DPD) [46]. (b) We generate networks with fixed degree distribution having a Poisson distribution (FDD) [45, 47]. The results for single networks and real networks are described in section 2. The robustness of NON with clustering are analyzed in section 3.

2. Single networks with clustering

In a network, the clustering feature can be characterized by specifying the fraction of nodes connected to s single links and having t triangles (clustering). As a special case, we consider an Erdős–Rényi (ER) type network with clustering having a probability density which obeys a DPD [46]

$$P(s, t) = \mathrm{e}^{-\langle s\rangle}\frac{\langle s\rangle^s}{s!}\mathrm{e}^{-\langle t\rangle}\frac{\langle t\rangle^t}{t!}, \tag{1}$$

where $\langle s\rangle$ and $\langle t\rangle$ are the average numbers of single links and triangles per node, respectively. The average degree of a node is thus, $\langle k\rangle = \langle s\rangle + 2\langle t\rangle$. The generating function of the DPD can be expressed as [46],

$$G_0(x, y) = \sum_{s,t=0}^{\infty} P(s, t)x^s y^t = e^{\langle s\rangle(x-1)}e^{\langle t\rangle(y-1)}. \tag{2}$$

The clustering coefficient is defined as $c = \frac{2\langle t\rangle}{2\langle t\rangle + \langle k\rangle^2}$, which implies the average of the clustering coefficient of all nodes by using the probability that two edges share a node in the network. For $c = 0$, the network does not have clustering, which is the limit of the ER network.

The LA is performed as follows. We randomly choose a node as a 'root' node and denote all nodes distances from this root, shell by shell according to increasing distance. Next, the LA is performed by the following two stages. We first remove around the root node all nodes shell by shell according to increasing distance, and remove all the links connecting all pairs of the removed nodes until a fraction of $1 - p$ nodes from the whole network is removed. In this stage we keep the links between the removed nodes and the remaining nodes. The distribution of nodes with s single links and t triangles in the remaining network is [35, 38]

$$P_p(s, t) = \frac{A_p(s, t)}{pN}, \tag{3}$$

where $A_p(s, t)$ denotes the number of nodes with s single links and t triangles. When one more node is being removed, we get

$$A_{(p-1/N)}(s, t) = A_p(s, t) - \frac{P_p(s, t)s}{\langle k\rangle_p} - \frac{2P_p(s, t)t}{\langle k\rangle_p}, \tag{4}$$

where $\langle k\rangle_p \equiv \sum P_p(s, t)(s + 2t)$. As $N \to \infty$, equation (4) can be presented by differentiating $A_p(s, t)$ with respect to p,

$$\frac{dA_p(s, t)}{dp} = N\frac{P_p(s, t)(s + 2t)}{\langle k\rangle_p}. \tag{5}$$

By differentiating equation (3) with respect to p and substituting it in equation (5), we get,

$$p\frac{dP_p(s, t)}{dp} + P_p(s, t) - \frac{P_p(s, t)(s + 2t)}{\langle k\rangle_p} = 0. \tag{6}$$

The solution of equation (6) can be written as,

$$P_p(s, t) = P(s, t)\frac{h^s h^{2t}}{G_0(h, h^2)}, \tag{7}$$

where $G_0(h, h^2) = p$. The generating function of the residual network is

$$G_a(x, y) = \sum_{s,t=0} P_p(s, t)x^s y^t = \frac{G_0(hx, h^2y)}{G_0(h, h^2)} = \frac{e^{\langle s\rangle(hx-1)}e^{\langle t\rangle(h^2y-1)}}{e^{\langle s\rangle(h-1)}e^{\langle t\rangle(h^2-1)}}. \tag{8}$$

The probability of a link to end at an unremoved node in the remaining network can be expressed as [38],

$$\tilde{p} = \frac{G_0'(h, h^2)}{G_0'(1, 1)h} = \frac{p}{h}. \tag{9}$$

In the second stage we remove all remaining links from the removed nodes, which are connected to the remaining non removed nodes. The generating function of the remaining network is

$$G_0^p(x, y) = G_a(1 - \tilde{p} + \tilde{p}x, 1 - \tilde{p} + \tilde{p}y). \tag{10}$$

If we find a clustering network $\tilde{A}$ of generating function $\tilde{G}_0(x, y)$, after a RA of removing $1 - p$ fraction of nodes than the generation function of the remaining network becomes $\tilde{G}_0^p = \tilde{G}_0(1 - p + px, 1 - p + py)$ [45, 47]. Next, we map the LA problem on network A to a RA problem on network $\tilde{A}$. By using $\tilde{G}_0^p = G_0^p(x, y)$ and equation (10), we deduce

$$\tilde{G}_0(x, y) = e^{\langle s\rangle(x-1)}e^{\langle t\rangle h(y-1)}. \tag{11}$$

Thus, the fraction of the giant component of the remaining clustered network is

$$g(p) = 1 - \tilde{G}_0[1 - p(1 - f(p)), (1 - p(1 - f(p)))^2] = 1 - e^{\langle s\rangle[p(f(p)-1)]}e^{\langle t\rangle h\{[1-p(1-f(p))]^2-1\}}, \tag{12}$$

where $f(p)$ satisfies $f(p) = \tilde{G}_1[1 - p(1 - f(p), (1 - p(1 - f(p))^2)] = e^{\langle s\rangle[p(f(p)-1)]}e^{\langle t\rangle h\{[1-p(1-f(p))]^2-1\}}$ and $\tilde{G}_1(x, y) = \frac{\tilde{G}'_0(x,y)}{\tilde{G}'_0(1,1)} = \tilde{G}_0(x, y)$. The fraction of the giant component with respect to the original clustered network is

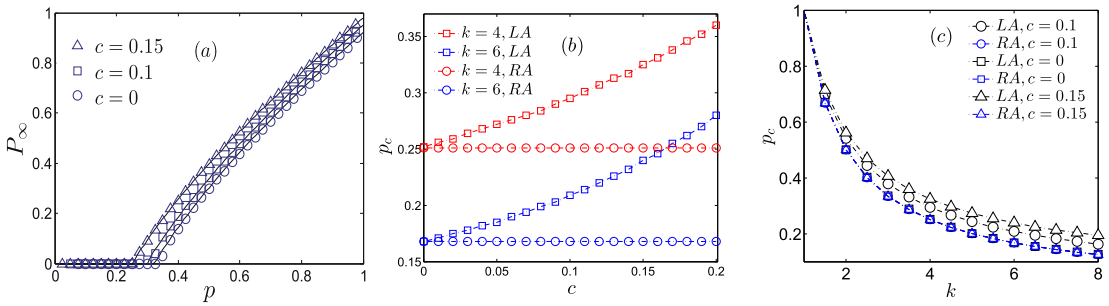

Figure 1. (a) Size of giant component P_∞, as a function of p for $\langle k\rangle = 4$ and different values of c, where solid lines are from equations (12) and (13) and symbols are from simulations with a network of number of nodes $N = 10^5$. All simulation results are obtained by averaging over 100 realizations. (b) p_c as a function of c for different $\langle k\rangle = 4,\ 6$ from equation (14). Results are also compared to RA. (c) p_c as a function of k with different $c = 0, 0.1$ and 0.15 from equation (14). Results are also compared to RA.

$$P_\infty = pg(p). \tag{13}$$

Figure 1(a) presents both the numerical solution of P_∞ in equation (13) and simulations for several values of c, which support well the theory. In addition, when comparing different c, we find that p_c increases with c and P_∞ increases continuously from zero at the critical threshold p_c to a finite value, which means that the system undergoes a second order phase transition. As $f(p_c) \to 1$, the critical threshold of the second order phase transition p_c can be found as [38],

$$p_c = \frac{1}{\langle s\rangle + 2h\langle t\rangle}. \tag{14}$$

Moreover, the dependence of p_c on c and $\langle k\rangle$ are shown in figures 1(b) and (c). As can be seen, p_c increases as c increases and $\langle k\rangle$ decreases. Thus, the network becomes more vulnerable with increasing clustering coefficient. This is in marked contrast to RA, where the system robustness does not change with increasing clustering coefficient, as seen in figures 1(b) and (c).

Clustering has an important effect in epidemic processes, information spread, network resilience, and dynamical systems on the networks. For example, in a social network, two friends of an individual have a high probability to become friends. From above, we analyze and compare the robustness of several real networks under LA and RA, and analyze them while changing the actual clustering in the real network using rewiring algorithm [48–52]. Figure 2 demonstrates that S as a function of p for several real networks with different clustering c under LA and RA. For real networks, we use a small S_{cutoff} to find p_c. Simulation results suggest that p_c is almost unchanged for RA but p_c dramatically increases for LA, as seen in figure 3. Thus again, clustering coefficient almost have no effect on system robustness for the case of RA, but it has a significant effect on system robustness for LA. As clustering coefficient increases, the system becomes more and more vulnerable and significantly more difficult to protect for LA, similar with above theoretical results. Note that the effect of clustering in the real networks (figure 3) on their vulnerability is significantly more than found in our theory (figure 1(b)). The reason is that the model we solved analytically is for Poisson degree distribution (equation (1)) but the real networks shown in figure 3 are of scale free type, where we expect a stronger effect in LA. This since neighbors of nodes in scale free networks are usually high degree and a LA will remove them in the first stages [53] while RA removes mostly low degree nodes.

Nextly, we consider the FDD model for single clustered networks, which preserves the total degree distribution $P(k)$ for different c using the method proposed by Hackett *et al* as given by equation (15) below, i.e. changing c but keeping a FDD [45]

$$\begin{aligned} P(s,\ t) &= P(k)\delta_{k,s+2t}[(1-f)\delta_{t,0} + f\delta_{t,\lfloor (s+2t)/2\rfloor}], \\ c &= f\frac{\sum_k k[P(2k) + P(2k+1)]}{\sum_k \binom{k}{2} P(k)}, \end{aligned} \tag{15}$$

where $P(k) = \langle k\rangle^k e^{-\langle k\rangle}/k!$, $f \in [0,\ 1]$ and $\lfloor . \rfloor$ is the floor function [45, 47, 54]. According to above expression, equation (15), clustered network are assumed to have joint distribution P_{st} from a given degree distribution $P(k)$ by randomly choosing a fraction f of nodes to be connected to maximum possible number of triangles while the

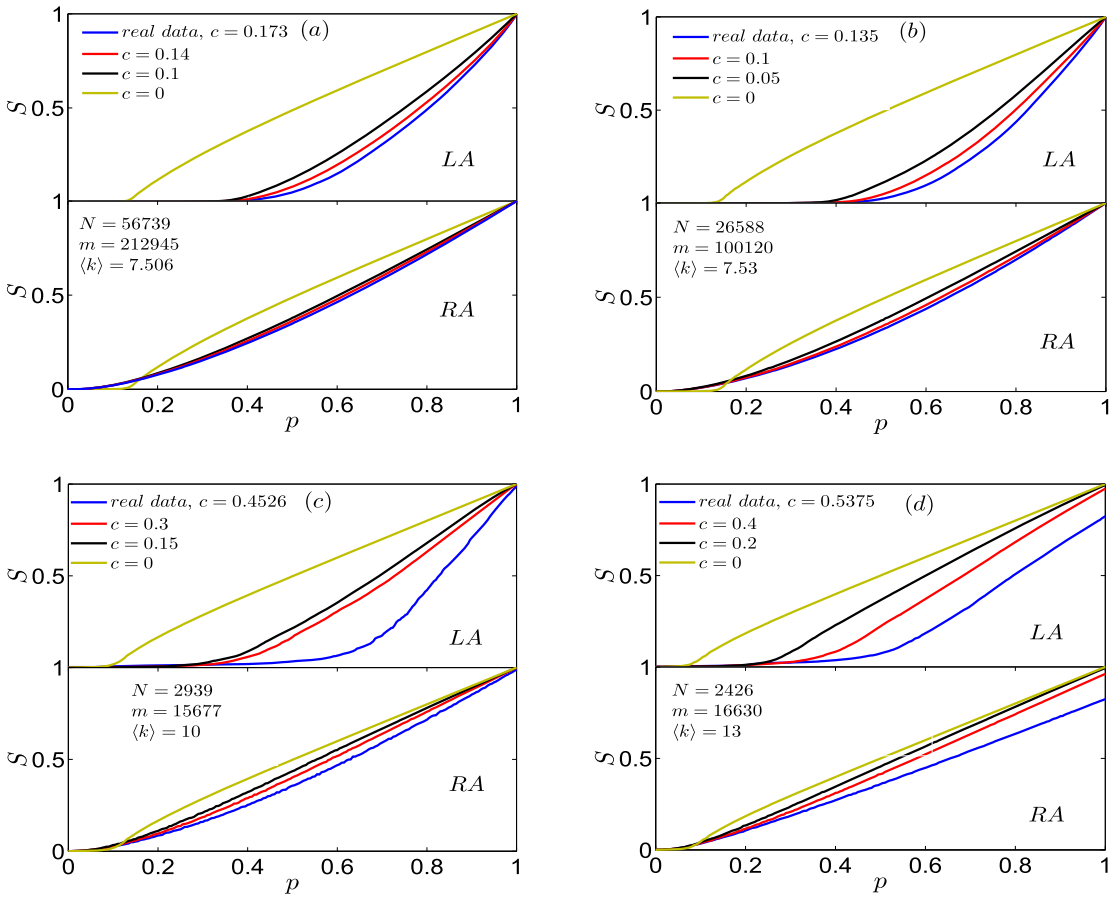

Figure 2. Comparison of S as a function of p under LA and RA with different parameters c for real networks. (a) Undirected Brightkite network [49], (b) undirected Epinions network [50], (c) undirected Openflight network [51], (d) undirected HAMSTERster network [51]. The parameters N, M, $\langle k\rangle$ denote the number of nodes and links, average degree, respectively.

remaining fraction $1-f$ of nodes are attached to single edges only. From this definition, we can get the above equation (15) for the clustering coefficient c as a function of f [45].

From figure 4, one can observe the peaks of the second largest cluster, $P_{\infty,2}$ for different c corresponding to the phase transition point from simulation results. It is also seen that p_c under LA is almost unchanged with increasing c, but increases for RA as seen in both figures 4(b) and (c). This means that changing clustering coefficient for a single FDD has a little effect for LA but increasing clustering coefficient can make single network under FDD more vulnerable, which is in marked contrast to the case of DPD.

3. NON with clustering

3.1. Two interdependent clustered networks

In this subsection, we study the robustness under LAs of two partially interdependent clustered networks A and B, which obey the DPD $P(s, t)$, with parameters $\langle s\rangle_A = \langle s\rangle_B = \langle s\rangle$, $\langle t\rangle_A = \langle t\rangle_B = \langle t\rangle$ and $N_A = N_B = N$, respectively. We assume a fraction q_A (q_B) of nodes in network A (B) depends on nodes in network B (A). This means that a node in network B which depends on a failed node in network A, will also fail, and vice versa. We start by removing a fraction $1-p$ of nodes from network A and B separately through LA, cascading failures occur, until the system reaches a steady state. At this time, the remaining fraction of nodes in network A and network B are equal to X and Y [25, 36],

$$\begin{aligned} X &= p[1-q_A(1-g_Bp)],\\ Y &= p[1-q_B(1-g_Ap)]. \end{aligned} \tag{16}$$

The size of the giant components of networks A and B can be expressed as $P_{\infty,A}$ and $P_{\infty,B}$

$$P_{\infty,A} = Xg_A,\ P_{\infty,B} = Yg_B, \tag{17}$$

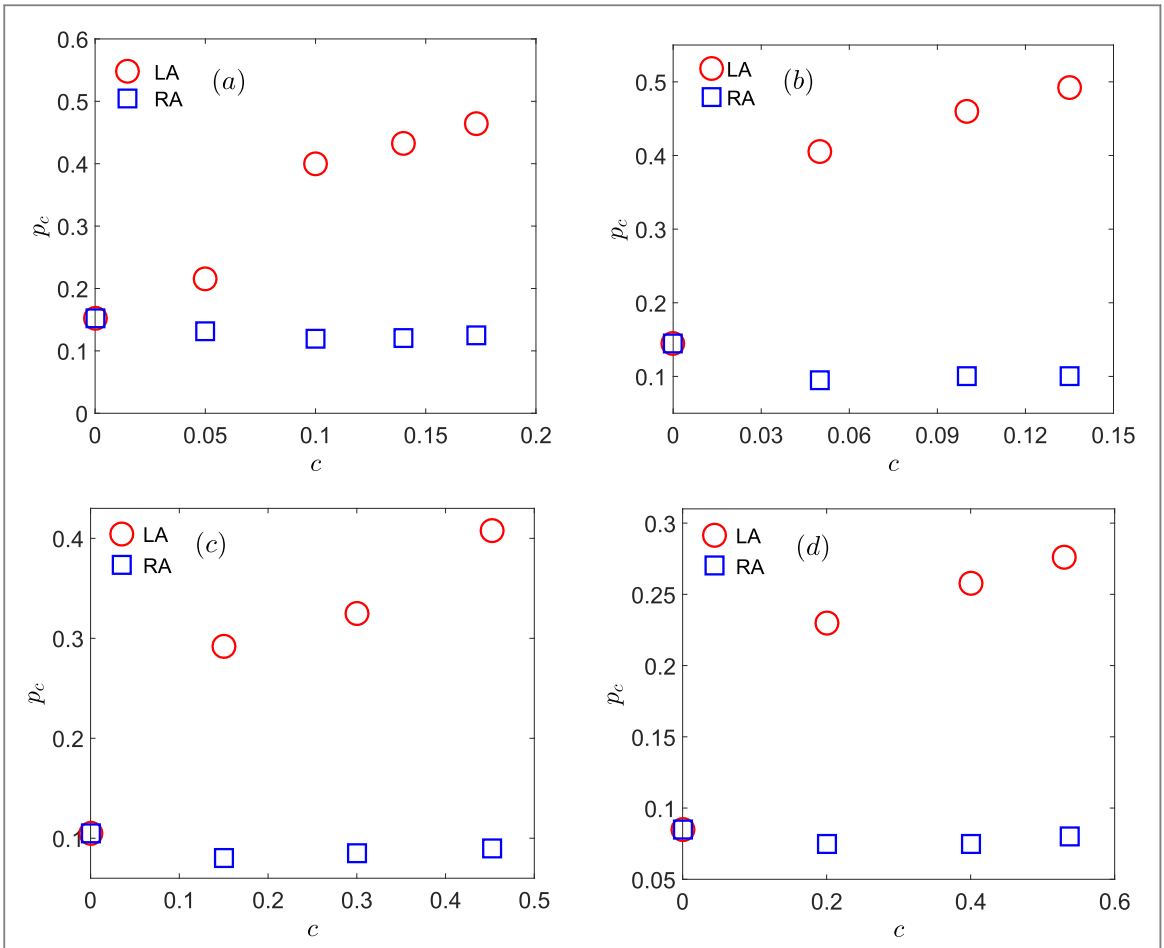

Figure 3. Comparison of p_c as a function of c under LA(circles) and RA(squares) with different parameters c in real networks for $S_{\text{cutoff}} = 0.001$. (a) Brightkite network, (b) Epinions network, (c) Openflight network, (d) HAMSTERster network.

where $g_A(X)$ and $g_B(Y)$ satisfy

$$\begin{aligned} g_A &= 1 - f_A, \\ g_B &= 1 - f_B, \\ f_A &= e^{\langle s\rangle\{1-[p(1-q_A(1-(1-f_B)p))](1-f_A)-1\}} e^{\langle t\rangle h\{\{1-[p(1-q_A(1-(1-f_B)p))](1-f_A)\}^2-1\}}, \\ f_B &= e^{\langle s\rangle\{1-[p(1-q_B(1-(1-f_A)p))](1-f_B)-1\}} e^{\langle t\rangle h\{\{1-[p(1-q_B(1-(1-f_A)p))](1-f_B)\}^2-1\}}, \end{aligned} \tag{18}$$

and $p = G_0(h, h^2)$. Figure 5(a) shows that simulation results for several values of c, are in good agreement with the theoretical results obtained from equations (16)–(18) We can see in figure 5(a) that $P_{\infty,A}$ continuously increases from zero to finite value at the second order critical threshold p_c^{II} for $q = 0.2$, but abruptly jumps from zero to a finite value at the first order critical threshold p_c^{I} for $q = 0.8$. These results suggest that the phase transition nature changes from second order to first order at a critical coupling strength q_c. By combining equations (16)–(18), we can obtain p_c as a function of q, as shown in figure 5(b). The upper panel in figure 5(b) shows that p_c increases with increasing c, which means that increasing either clustering coefficient or coupling strength makes the network more vulnerable to LA. In contrast, as seen in the lower panel, for RA, the system robustness almost remains the same for $q < q_c$ and increasing c and becomes more vulnerable for $q > q_c$. In figures 5(c) and (d), we see also that increasing c or/and decreasing $\langle k\rangle$ will enlarge p_c, which means increasing clustering coefficient make networks more vulnerable. From comparing the two kinds of attacking strategies in figure 5(b), we can also see that q_c is almost constant when changing clustering for RA, but increases as c increases for LA.

Next, we perform a LA by simulated removing a fraction $1 - p$ of nodes from network A. Let $P_{\infty,1}$ denote the size of the giant component of network A at the steady state. Figures 6(a) and (b) compare simulation results of FDD with theoretical results of DPD from equations (16)–(17) of [45]. Figures 6(a) and (b) indicate that for LA, second and first order phase transition behaviors can be observed, respectively, for weak ($q = 0.2$) and strong ($q = 0.8$) coupling strength with different c. Note that p_c increases with increasing c for both, FDD and DPD. This suggests that robustness of both cases of clustered networks decreases with increasing clustering coefficient. Note, however, that, p_c of DPD is larger than that of FDD except for $c = 0$. As seen in figures 6(c) and (d), that the

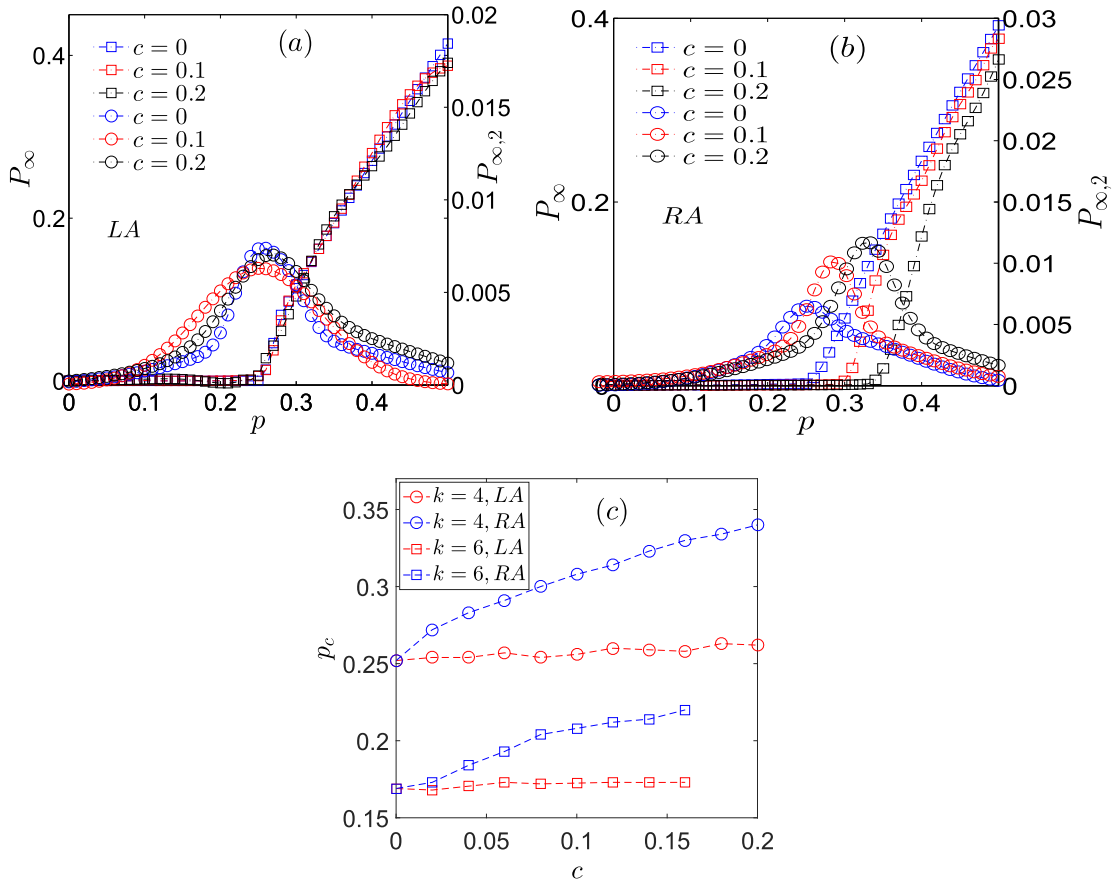

Figure 4. (a) and (b) show that P_∞ (circles) and $P_{\infty,2}$ (squares) as functions of p with different c under LA and RA, respectively, for a single FDD network, where $P_{\infty,2}$ is the size of the second largest component of the single clustered network after attack. (c) p_c as a function of c with $k = 4$ and 6 under LA and RA for single clustered network. Simulation results are obtained by averaging over 100 realizations with $N = 10^5$.

behavior in LA is in contrast with RA [45]. Additionally, for clustered networks of types FDD and DPD, figures 6(c) and (d) compare the change of p_c with q under LA and RA. The left panel of figure 6(c) for FDD under LA illustrates that increasing clustering coefficient has almost no effect on robustness for weak coupling strength. But for DPD, we notice that the system become more vulnerable as c increases for all coupling strengths as seen in the right panel of figure 6(c). In marked contrast, if the FDD system suffers from RA, it becomes more vulnerable as c increases for both weak and strong coupling strength as shown in the left panel of figure 6(d). However for DPD, the system gradually becomes vulnerable as c increases only for strong coupling strength as seen in right panel of figure 6(d).

3.2. Star-like NON of ER networks with clustering

We generalize our results for two interdependent networks with clustering analyzed in above subsection, to a system whose dependence structure is a network and each node is a clustered network, i.e. NON [26]. For simplicity, we assume that all clustered networks satisfy a joint degree distribution with the same $\langle s \rangle$ and $\langle t \rangle$. Here, we adopt the non-feedback condition [27] for dependency structure like in the above section. In this subsection, we study the cases of a star-like NON formed of ER networks with clustering (as demonstrated in figure 7(a)) and random regular (RR) of ER networks with clustering (as shown in figure 7(b)).

Here we study star-like NON formed of n clustered networks, in which a central network is linked via dependency links with other $n - 1$ networks. I.e. the $n - 1$ networks are mutually dependent on the central network but do not depend on each other, see figure 7(a). We assume that a fraction $q_{i,1}$ ($i = 2, 3, \ldots, n$) of nodes in network A_i and vice versa. depends on nodes in the central network A_1. If one of a pair of interdependent nodes fail, the other node that depends on it also fail to function. The initial attack is exerted on each network by removing locally a fraction $1 - p$ of nodes and this damage spreads in this system back and forth until no node depends on a disabled node, and the remaining network is stable or fully collapsed. For simplicity, but without loss of generality, we set $q_{2,1} = q_{3,1} = \ldots = q_{n,1} = q$. The fraction of left nodes in A is equal to X and in the other $n - 1$ networks A_i ($i = 2, 3, \ldots, n$) is equal to Y, following the expressions [27, 37],

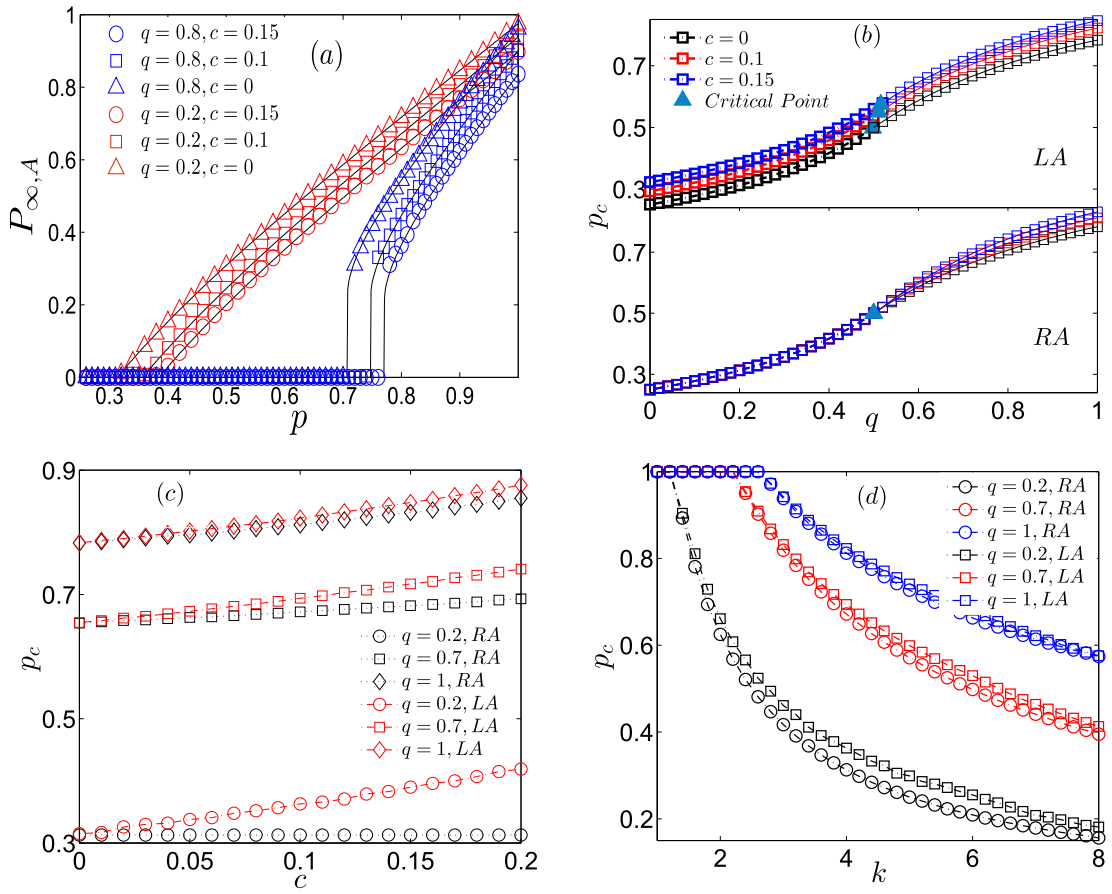

Figure 5. (a) $P_{\infty,A}$ as a function of p for $\langle k\rangle = \langle k\rangle_A = \langle k\rangle_B = 4$, where solid lines drawn from equations (16)–(18) and symbols represent simulation results with $N = 10^5$ for several q and c values. The simulation results are obtained by averaging over 100 realizations. (b) Critical threshold p_c as a function of q ($q = q_A = q_B$) for $c = 0, 0.1$ and 0.15 with $\langle k\rangle = 4$. Blue full triangles indicate the critical value q_c from equations (16)–(18). (c) Critical threshold p_c as a function of c for $q = 0.2, 0.7$ and 1 for $\langle k\rangle = \langle k\rangle_A = \langle k\rangle_B = 4$ under LA and RA obtained from equations (16)–(18). (d) p_c as a function of $\langle k\rangle$ for different $q = 0.2, 0.7$ and 1 with $c = 0.1$ under LA and RA obtained from equations (16)–(18).

$$X = p(1 - q + pqg_2)^{n-1},$$
$$Y = p[1 - q + pqg_1(1 - q + pqg_2)^{n-2}]. \tag{19}$$

The size of the giant components of network A and the other $(n-1)$ networks A_i, can be expressed as $P_{\infty,1}$ and $P_{\infty,2}$

$$P_{\infty,1} = Xg_1,\ P_{\infty,2} = Yg_2, \tag{20}$$

where g_1 and g_2 satisfy

$$\begin{aligned} g_1 &= 1 - f_1,\\ g_2 &= 1 - f_2,\\ f_1 &= e^{-\langle s\rangle p[1-q+pq(1-f_2)]^{n-1}(1-f_1)} e^{\langle t\rangle h[\{1-p[1-q+pq(1-f_2)]^{n-1}(1-f_1)\}^2-1]},\\ f_2 &= e^{-\langle s\rangle p\{1-q+pq(1-f_1)[1-q+pq(1-f_2)]^{n-2}\}(1-f_2)} e^{\langle t\rangle h\{[1-p\{1-q+pq(1-f_1)[1-q+pq(1-f_2)]^{n-2}\}(1-f_2)]^2-1\}}, \end{aligned} \tag{21}$$

and $p = G_0(h, h^2)$.

Simplifying equations (19)–(21), we get

$$\begin{aligned} P_{\infty,1} &= p\{1 - q + pq[1 - e^{\langle t\rangle hP_{\infty,2}^2 - (\langle s\rangle + 2\langle t\rangle h)P_{\infty,2}}]\}^{n-1}[1 - e^{\langle t\rangle hP_{\infty,1}^2 - (\langle s\rangle + 2\langle t\rangle h)P_{\infty,1}}],\\ P_{\infty,2} &= p\{1 - q + pq[1 - e^{\langle t\rangle hP_{\infty,1}^2 - (\langle s\rangle + 2\langle t\rangle h)P_{\infty,1}}]\}\{1 - q + pq[1 - e^{\langle t\rangle hP_{\infty,2}^2 - (\langle s\rangle + 2\langle t\rangle h)P_{\infty,2}}]\}^{n-2}\}\\ &\quad \times [1 - e^{\langle t\rangle hP_{\infty,2}^2 - (\langle s\rangle + 2\langle t\rangle h)P_{\infty,2}}]. \end{aligned} \tag{22}$$

Figure 8(a) shows that simulation results agree well with theoretical predictions obtained from equation (22). Additionally, we observe in figures 8(a) and (b) that as q decreases, the system changes from an abrupt first order phase transition to a continuous second order transition at a critical coupling strength q_c. And, the results show that the system becomes more vulnerable for larger c. Moreover, it is seen that for RA as q

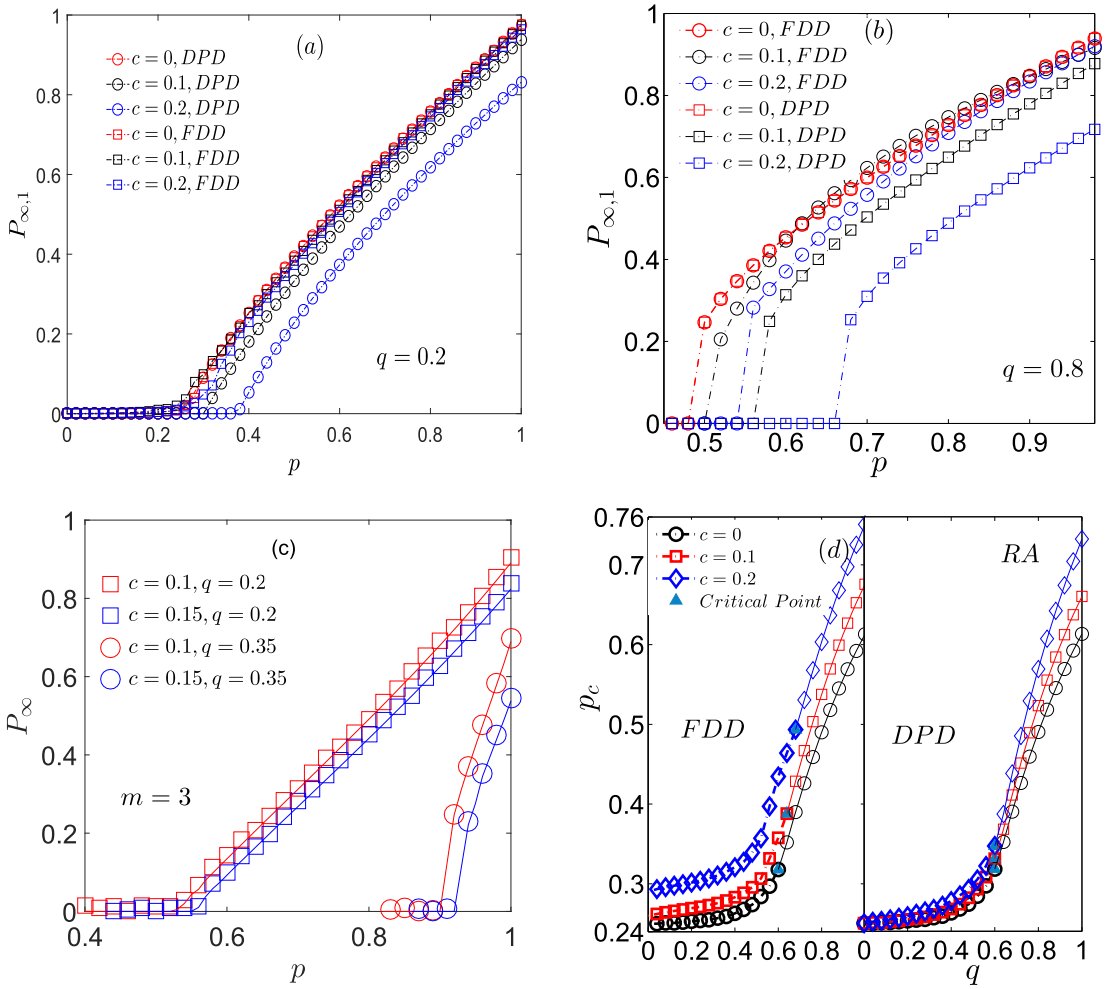

Figure 6. Parameters in all panels include different $c = 0, 0.1$ and 0.2, $N = 10^5$ and $\langle k\rangle = \langle k\rangle_A = \langle k\rangle_B = 4$. All the simulation results are obtained by averaging over 100 realizations. (a) and (b) $P_{\infty,1}$ as a function of p for $q = 0.2$ and $q = 0.8$, respectively, comparing FDD and DPD. (c) p_c as a function of q under LA for both FDD and DPD. (d) p_c as a function of q under RA for both FDD and DPD.

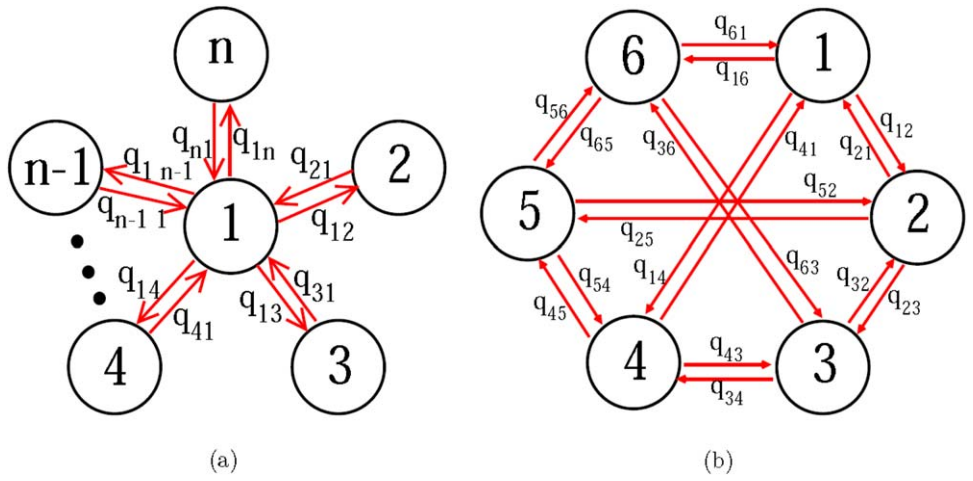

Figure 7. (a) Star-like NON where each network is clustered. Here, network 1 is the central network, which has interdependent relationships with other $n - 1$ networks. (b) Random regular network of clustered networks consist of m ER clustered networks which are mutually interdependent [45].

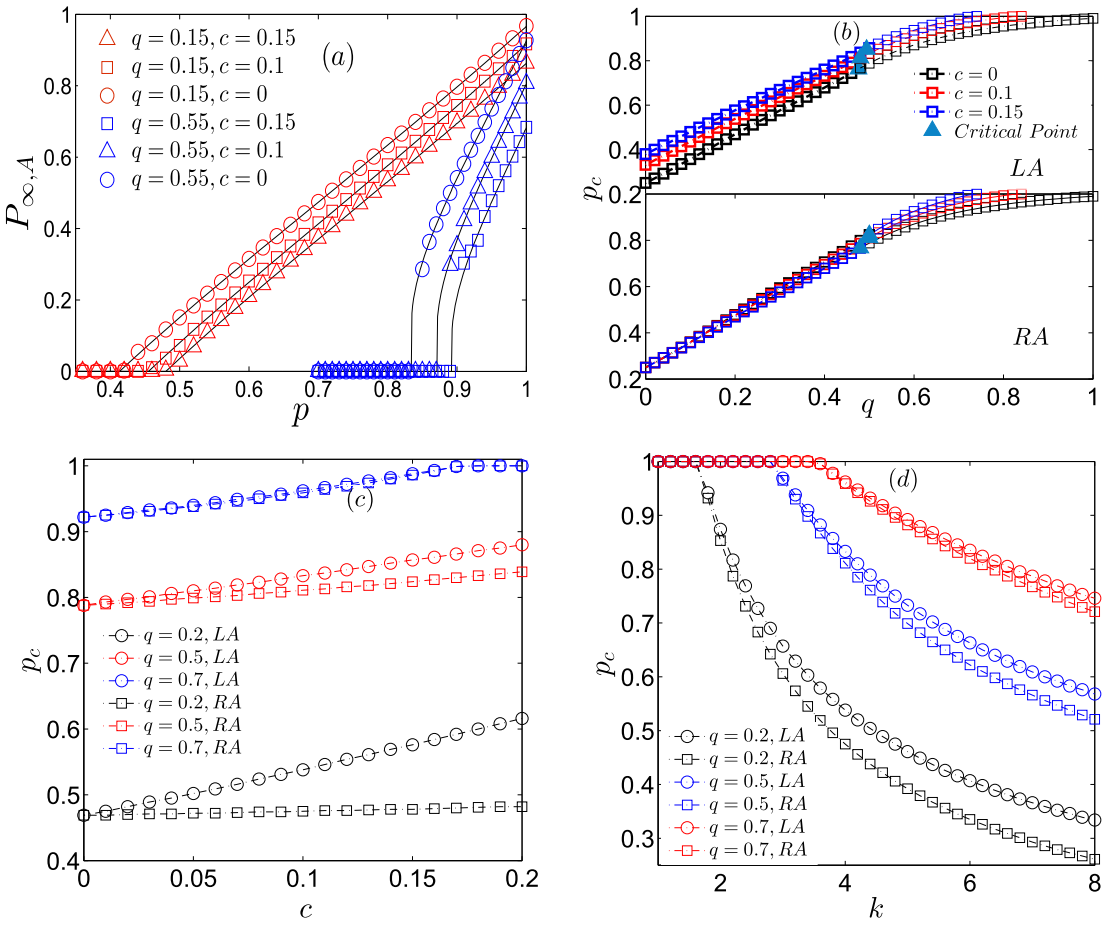

Figure 8. (a) $P_{\infty,A}$ as a function of p for a star-like NON with $\langle k\rangle = \langle k\rangle_A = \langle k\rangle_{A_i} = 4$ and $n = 5$, where solid lines are obtained from equation (22) and symbols are simulation results with $N = 10^5$ for different q and c. All results are obtained by averaging over 100 realizations. (b) p_c as a function of q for $c = 0, 0.1$ and 0.15 with $\langle k\rangle = 4$ and $n = 5$ obtained from equation (22). (c) p_c as a function of c for different q with $\langle k\rangle = 4$ and $n = 5$ obtained from equation (22). (d) p_c as a function of k for different q with $c = 0.1$ and $n = 5$ obtained from equation (22).

approaching to 0, p_c lines for different c gradually become closer, while for LA, as c increases, p_c becomes larger, suggesting that increasing clustering within networks makes the system more vulnerable to LA. This is seen even more clearly in figures 8(c) and (d) comparing p_c as functions of c and $\langle k\rangle$ for different q under LA and RA. As seen in figure 8(c) the system under LA and RA shows very different robustness as c increases for weak coupling strength $q = 0.2, 0.5$ and $\langle k\rangle = 4$. The system is significantly more robust under RA compared to LA for weak coupling strength. But for strong coupling strength $q = 0.7$, the system shows similar robustness for both attacking strategies. From figure 8(d), one can see that increasing $\langle k\rangle$ enhances the robustness for both attack strategies, but LA has a more destructive power, that makes the system more vulnerable compared to RA. Note that as $\langle k\rangle$ increases, the robustness of system gradually shows differences for different attacking strategies even for strong coupling strength, as shown in figure 8(d).

Figure 9 shows the dependence of the critical coupling strength q_c on n from equation (22) for several c. The results show that q_c monotonically decreases as n increases, implying that the NON system more easily shows first order phase transition, when n increases. Additionally, q_c increases as c increases for different n, which imply that the region of occurring second phase transition becomes larger as clustering coefficient increases.

3.3. RR NON of clustered networks

Next we study RR NON formed of ER networks with clustering. The ER networks are clustered networks while the NON is a RR network with same degree m for all the nodes, which means that every ER clustered network is partially dependent (with q dependency nodes) on other m clustered networks, see figure 7(b) [55]. Assuming that the initial attack is exerted on each ER network through locally removing a fraction $1 - p$ of nodes of the network by removing all nodes shell by shell around a randomly chosen node. For simplicity, we let each clustered network to have the same average number of single links $\langle s\rangle$, average number of triangles $\langle t\rangle$ and coupling strength q. When the cascading failures process reaches a stable state, the remaining fraction of nodes in any clustered network is X [55]

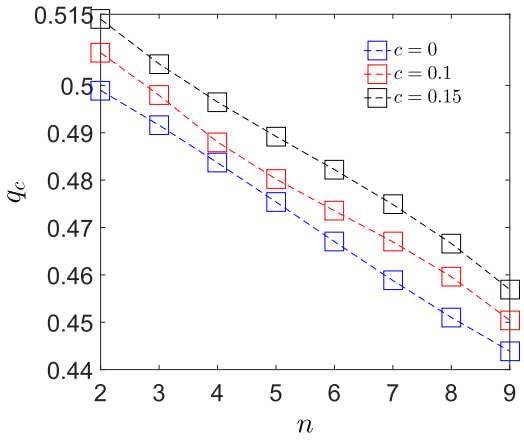

Figure 9. The critical coupling, q_c, as a function of n with $\langle k\rangle = 4$ for different c values under LA from equation (21).

$$\begin{aligned} X &= p[qYg(X) - q + 1]^m, \\ Y &= p[qYg(X) - q + 1]^{m-1}, \end{aligned} \tag{23}$$

where

$$\begin{aligned} g(X) &= 1 - f(X), \\ f(X) &= e^{\langle s\rangle\{1-p[qYg(X)-q+1]^m(1-f)-1\}} e^{\langle t\rangle h\{\{1-p[qYg(X)-q+1]^m(1-f)\}^2-1\}}, \\ p &= G_0(h, h^2). \end{aligned} \tag{24}$$

The size of giant component in each network is

$$P_\infty = Xg(X). \tag{25}$$

Simplifying equations (23)–(25), we obtain,

$$P_\infty = p\left[1 - e^{\langle t\rangle hP_\infty^2 - (\langle s\rangle + 2\langle t\rangle h)P_\infty}\right]\left[\frac{1 - q + \sqrt{(q-1)^2 + 4qP_\infty}}{2}\right]^m. \tag{26}$$

The critical threshold of the second order phase transition p_c^{II} can be obtained from $P_\infty(p_c^{\mathrm{II}}) \to 0$ in equation (26). And, by equating the first derivative of both sides of equation (26) with respect to P_∞, we obtain the critical threshold at q_c. Thus, the critical coupling q_c, where the first order phase transition changes to a second order phase transition can be obtained analytically as follows

$$\begin{aligned} q_c &= \frac{\{[\langle s\rangle + 2\langle t\rangle h]^2 + 2\langle t\rangle h\}(1 - q_c)^2}{2m[\langle s\rangle + 2\langle t\rangle h]}, \\ p_c^{\mathrm{II}} &= \frac{1}{(1 - q_c)^m[\langle s\rangle + 2\langle t\rangle h]}, \\ G_0(h, h^2) &= p_c^{\mathrm{II}}. \end{aligned} \tag{27}$$

Figures 10(a)–(c) show that simulation results and theoretical predictions obtained from equation (26) with $\langle k\rangle = 4$ agree well. In addition, we observe in figures 10(a)–(c) that P_∞ continuously increases at a critical threshold p_c^{II} for $q = 0.2$ and $m = 2, 3$ (second order phase transition). However, as seen in figures 10(a) and (b) P_∞ changes abruptly at p_c^{I} (first order phase transition) for $q = 0.5$, and in figure 10(c) for $q = 0.35$. Figures 10(a)–(c) indicate that increasing coupling strength changes the phase transition from second order to first order, and p_c increases as clustering coefficient increases, as can be seen in figure 10(d). Note that the system becomes more vulnerable as m increases since the p_c value increases. And, q_c gradually increases with m as seen in figure 10(d).

Figure 11(a) shows that increasing m induce decrease of q_c and implies that it is more easy to have a first order phase transition. The lower panel in figure 11(b) shows that for RA, the p_c values for different c are almost constant for small q, which means changing clustering coefficient has little influence for RA, but, as seen in the top panel, there exists obvious differences under LA. This illustrates that for LA, increasing clustering coefficient makes the network more vulnerable. Figures 11(c) and (d) show p_c as a function of c and $\langle k\rangle$ for different q.

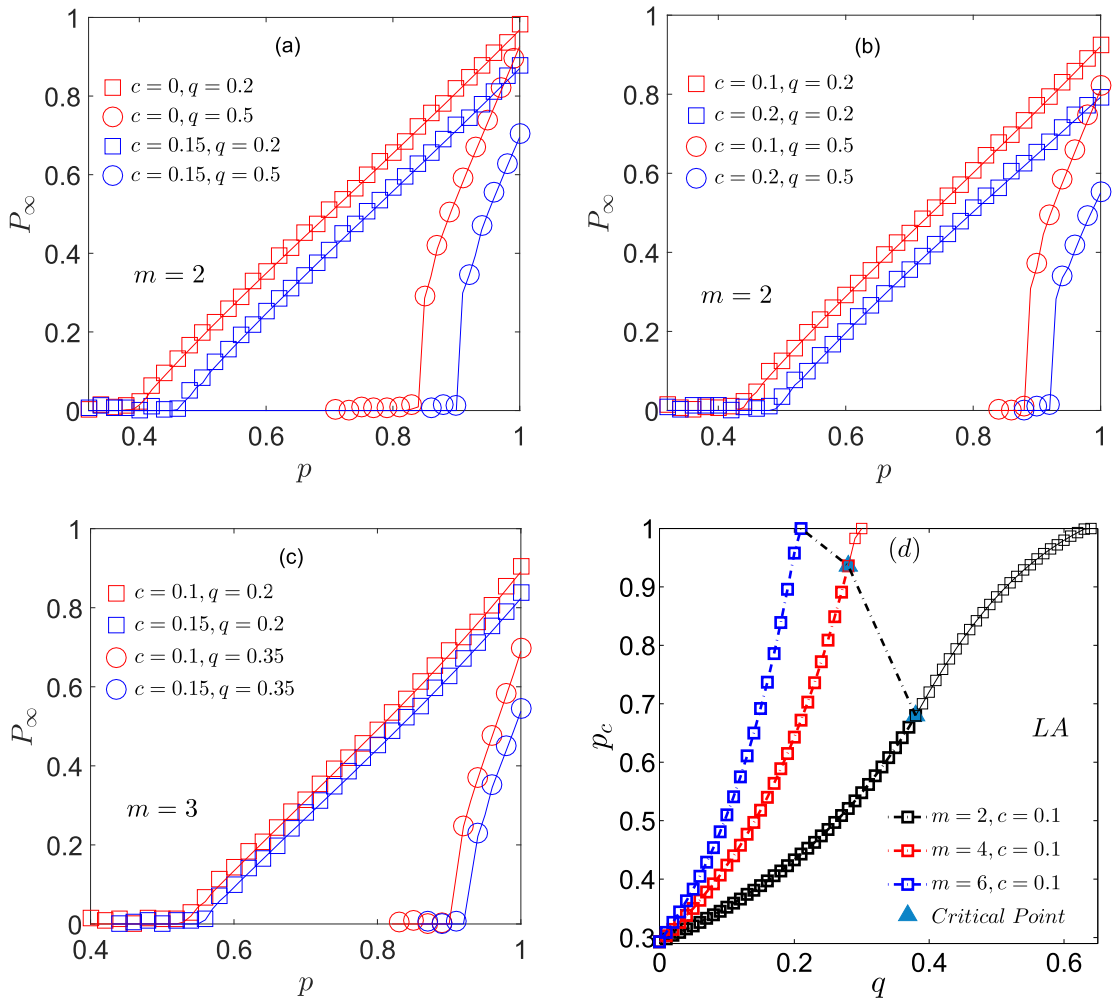

Figure 10. (a) P_∞ as a function of p for $m = 2$ and $n = 6$ for different $q = 0.2, 0.5$ and $c = 0, 0.15$. (b) P_∞ as a function of p for $m = 2$, $c = 0.1, 0.2$ and $q = 0.2, 0.5$. (c) P_∞ as a function of p for $m = 3, c = 0.1, 0.15$ and $q = 0.2, q = 0.35$. In (a)–(c), solid lines represent the analytical results from equation (26) with $\langle k \rangle = 4$, and symbols denote the simulation results with $N = 10^5$. The simulation results are obtained by averaging over 100 realizations. (d) p_c as a function of q for $m = 2, 4, 6$ and $c = 0.1$. Blue full triangles indicate the critical value q_c. Dashed lines and thick symbols denote second order phase transition, while solid lines and thin symbols denote first order phase transition.

4. Conclusions

Here we study both theoretically and numerically the robustness of clustered networks under LA of three cases, single clustered network, two interdependent clustered networks and NON (star-like NON and RR of clustered networks). We also studied, two types of clustering models. One is the DPD and the second is fixed degree distribution obeying Poisson distribution (FDD). For the DPD case, the LA on a single network shows a second order percolation behavior. When considering several real networks, we also find that system becomes more vulnerable and significantly more difficult to protect against LA, but the system robustness does not almost change in RA as clustering coefficient increases. For two interdependent clustered networks and RR NON of clustered networks, increasing q leads to a change from a second order to a first order phase transition. For the case of a single clustered network, p_c increases as c increases and k decreases. This is in marked contrast to RA, where p_c almost keep constant. For the case of dependent networks, the results demonstrate that p_c becomes larger as c and q increase. Besides, for star-like NON, as n increases, p_c becomes larger and q_c decreases gradually, which indicates that the system becomes more vulnerable and is easier to show a first order phase transition. Furthermore, we studied two generalized cases for NONs and found that higher clustering coefficient causes the system to be more vulnerable. For FDD case, by taking two interdependent networks as an example, we found that p_c is almost unaffected for $q < q_c$ when c increases, but for $q > q_c$, p_c increases as c increases, which is in marked contrast to RA.

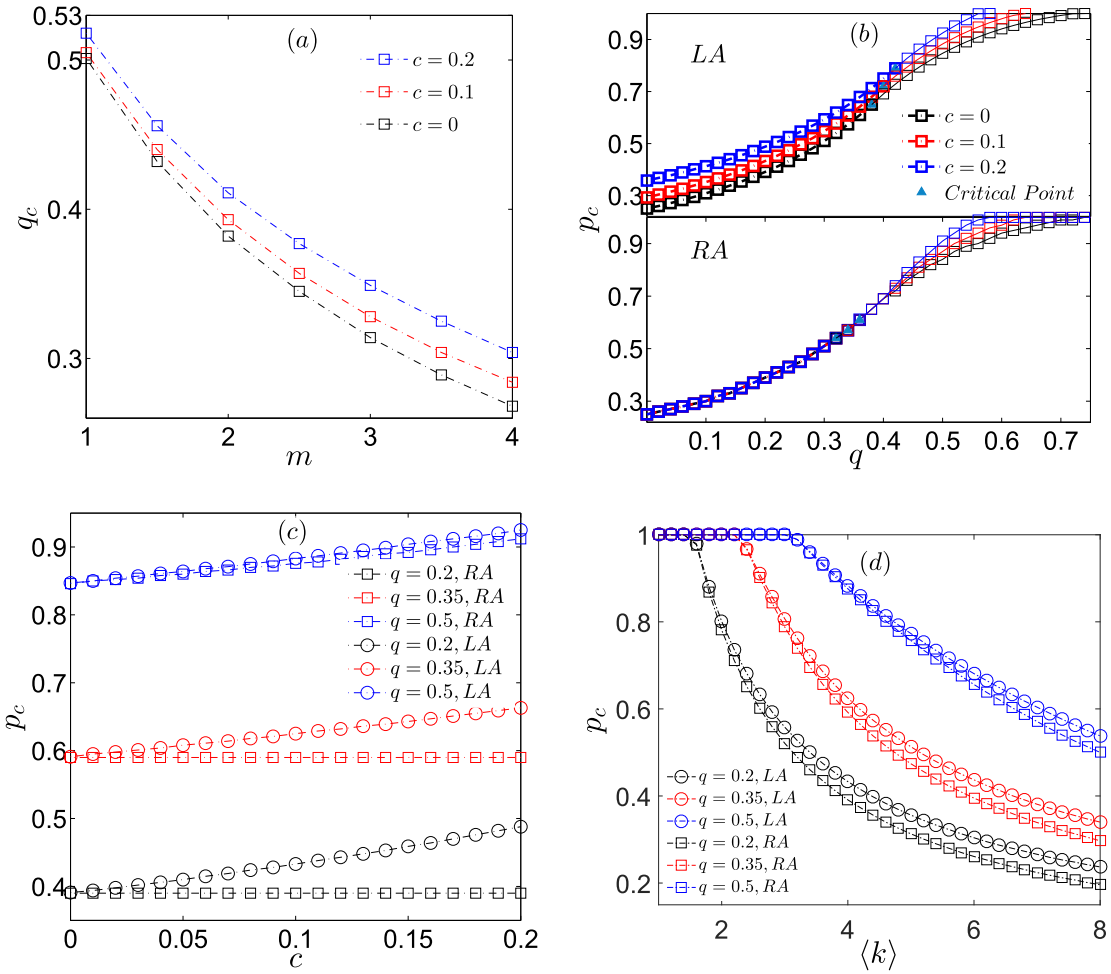

Figure 11. (a) For RR NON, plotted is q_c as a function of m for different c values and $\langle k \rangle = 4$ from equation (27). (b) p_c as a function of q for $c = 0, 0.1, 0.15$, $\langle k \rangle = 4$ and $m = 2$ for both LA and RA. (c) p_c as a function of c for $q = 0.2, 0.35$ and 0.5 for the same parameters as (a). (d) p_c as a function of $\langle k \rangle$ for different $q = 0.2, 0.35$ and 0.5 with $c = 0.1$ and $m = 2$.

Acknowledgments

SH acknowledges support of the Israel–Italian collaborative project NECST, Israel Science Foundation, ONR, Japan Science Foundation, BSF-NSF, the BIU Center for Research in Applied Cryptography and Cyber Security in conjunction with the Israel National Cyber Bureau in the Prime Minister's Office, and DTRA (Grant no. HDTRA-1-10-1- 0014) for financial support. The Boston University Center for Polymer Studies is supported by NSF Grant PHY-1505000 and by DTRA Grant HDTRA1-14-1-0017. We also thank National Natural Science Foundation of China (Grant Nos. 61403171, 71403105, 71303095), the Jiangsu Postdoctoral Science Foundation (Grant No. 1501100B), the China Postdoctoral Science Foundation (Grant No. 2015M581738), the Senior talents Foundation of Jiangsu University (Grant Nos. 14JDG143, 14JDG144) and Scientific research project of Jiangsu University (Grant No. 17A293) for support.

References

[1] Barabási A L and Albert R 1999 Emergence of scaling in random networks *Science* **286** 509–12
[2] Cohen R and Havlin S 2010 *Complex Networks: Structure, Robustness and Function* (Cambridge: Cambridge University Press)
[3] Boccaletti S *et al* 2006 Complex networks: structure and dynamics *Phys. Rep.* **424** 175–308
[4] Albert R and Barabási A L 2002 Statistical mechanics of complex networks *Rev. Mod. Phys.* **74** 47
[5] Newman M E J 2003 The structure and function of complex networks *SIAM Rev.* **45** 167–256
[6] Watts D J and Strogatz S H 1998 Collective dynamics of small-world networks *Nature* **393** 440–2
[7] Saberi A A 2015 Recent advances in percolation theory and its applications *Phys. Rep.* **578** 1–32
[8] Barabási A L, Ravasz E and Vicsek T 2001 Deterministic scale-free networks *Physica* A **299** 559–64
[9] Comellas F and Sampels M 2002 Deterministic small-world networks *Physica* A **309** 231–5
[10] Du W B *et al* 2016 Analysis of the Chinese airline network as multi-layer networks *Transp. Res.* E **89** 108–16
[11] Boccaletti S *et al* 2014 The structure and dynamics of multilayer networks *Phys. Rep.* **544** 1–122
[12] Böttcher L *et al* 2017 Failure and recovery in dynamical networks *Sci. Rep.* **7** 41729

[13] Li D et al 2015 Percolation transition in dynamical traffic network with evolving critical bottlenecks Proc. Natl Acad. Sci. 112 669–72
[14] Sole R V and Montoya M 2001 Complexity and fragility in ecological networks Proc. R. Soc. B 268 2039–45
[15] Motter A E et al 2008 Predicting synthetic rescues in metabolic networks Mol. Syst. Biol. 4 168
[16] Du W B et al 2015 Adequate is better: particle swarm optimization with limited-information Appl. Math. Comput. 268 832–8
[17] Haldane A G and May R M 2011 Systemic risk in banking ecosystems Nature 469 351–5
[18] Albert R, Albert I and Nakarado G L 2004 Structural vulnerability of the North American power grid Phys. Rev. E 69 025103
[19] Cohen R et al 2000 Resilience of the Internet to random breakdowns Phys. Rev. Lett. 85 4626
[20] Callaway D S et al 2000 Network robustness and fragility: percolation on random graphs Phys. Rev. Lett. 85 5468
[21] Schneider C M et al 2011 Mitigation of malicious attacks on networks Proc. Natl Acad. Sci. 108 3838–41
[22] Bunde A and Havlin S 2012 Fractals and Disordered Systems (New York: Springer)
[23] Stauffer D and Aharony A 2014 Introduction to Percolation Theory Revised 2nd edn (Boca Raton, FL: CRC Press)
[24] Buldyrev S V et al 2010 Catastrophic cascade of failures in interdependent networks Nature 464 1025–8
[25] Parshani R, Buldyrev S V and Havlin S 2010 Interdependent networks: reducing the coupling strength leads to a change from a first to second order percolation transition Phys. Rev. Lett. 105 048701
[26] Gao J et al 2012 Networks formed from interdependent networks Nat. Phys. 8 40–8
[27] Gao J et al 2011 Robustness of a network of networks Phys. Rev. Lett. 107 195701
[28] Gao J et al 2012 Robustness of a network formed by n interdependent networks with a one-to-one correspondence of dependent nodes Phys. Rev. E 85 066134
Zhou D, Gao J, Stanly H E and Havlin S 2013 Percolation of partially interdependent scale-free networks Phys. Rev. E 87 052812
[29] Albert R, Jeong H and Barabási A L 2000 Error and attack tolerance of complex networks Nature 406 378–82
[30] Cohen R et al 2001 Breakdown of the Internet under intentional attack Phys. Rev. Lett. 86 3682
[31] Motter A E and Lai Y C 2002 Cascade-based attacks on complex networks Phys. Rev. E 66 065102
[32] Carreras B A et al 2002 Critical points and transitions in an electric power transmission model for cascading failure blackouts Chaos 12 985–94
Carreras B A et al 2004 Complex dynamics of blackouts in power transmission systems Chaos 14 643–52
[33] Motter A E 2004 Cascade control and defense in complex networks Phys. Rev. Lett. 93 098701
[34] Gallos L K et al 2005 Stability and topology of scale-free networks under attack and defense strategies Phys. Rev. Lett. 94 188701
[35] Huang X et al 2011 Robustness of interdependent networks under targeted attack Phys. Rev. E 83 065101
[36] Dong G et al 2012 Percolation of partially interdependent networks under targeted attack Phys. Rev. E 85 016112
[37] Dong G et al 2013 Robustness of network of networks under targeted attack Phys. Rev. E 87 052804
[38] Shao S et al 2015 Percolation of localized attack on complex networks New J. Phys. 17 023049
[39] Berezin Y et al 2015 Localized attacks on spatially embedded networks with dependencies Sci. Rep. 5 8934
[40] Yuan X et al 2015 How breadth of degree distribution influences network robustness: comparing localized and random attacks Phys. Rev. E 92 032122
[41] Zhao J et al 2016 Spatio-temporal propagation of cascading overload failures in spatially embedded networks Nat. Commun. 7 10094
[42] Dong G et al 2016 Modified localized attack on complex network Europhys. Lett. 113 28002
[43] Wasserman S and Faust K 1994 Social Network Analysis: Methods and Applications (Cambridge: Cambridge University Press)
[44] Huang X et al 2013 The robustness of interdependent clustered networks Europhys. Lett. 101 18002
[45] Shao S et al 2014 Robustness of a partially interdependent network formed of clustered networks Phys. Rev. E 89 032812
[46] Newman M E J, Strogatz S H and Watts D J 2001 Random graphs with arbitrary degree distributions and their applications Phys. Rev. E 64 026118
[47] Hackett A, Melnik S and Gleeson J P 2011 Cascades on a class of clustered random networks Phys. Rev. E 83 056107
[48] Rossi R and Ahmed N 2015 The network data repository with interactive graph analytics and visualization AAAI 15 4292–3
[49] Cho E, Myers S A and Leskovec J 2011 Friendship and mobility: user movement in location-based social networks Proc. 17th ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining (ACM) pp 1082–90
[50] Richardson M, Agrawal R and Domingos P 2003 Trust management for the semantic web Int. Semantic Web Conf. (Berlin: Springer) pp 351–68
[51] Kunegis J K 2013 Konect, the koblenz network collection Proc. 22nd Int. Conf. on World Wide Web (ACM) pp 1343–50
[52] Golbeck J 2007 The dynamics of web-based social networks: membership, relationships, and change First Monday 12
[53] Cohen R, Havlin S and Ben-Avraham D 2003 Efficient immunization strategies for computer networks and populations Phys. Rev. Lett. 91 247901
[54] Gleeson J P, Melnik S and Hackett A 2010 How clustering affects the bond percolation threshold in complex networks Phys. Rev. E 81 066114
[55] Gao J et al 2013 Percolation of a general network of networks Phys. Rev. E 88 062816