Vulnerability of networks: Fractional percolation on random graphs

Yilun Shang*

Singapore University of Technology and Design, 20 Dover Drive, Singapore 138682, Singapore

(Received 29 August 2013; published 28 January 2014)

We present a theoretical framework for understanding nonbinary, nonindependent percolation on networks with general degree distributions. The model incorporates a partially functional (PF) state of nodes so that both intensity and extensity of error are characterized. Two connected nodes in a PF state cannot sustain the load and therefore break their link. We give exact solutions for the percolation threshold, the fraction of giant cluster, and the mean size of small clusters. The robustness-fragility transition point for scale-free networks with a degree distribution $p_k \propto k^{-\alpha}$ is identified to be $\alpha = 3$. The analysis reveals that scale-free networks are vulnerable to targeted attack at hubs: a more complete picture of their Achilles' heel turns out to be not only the hubs themselves but also the edges linking them together.

DOI: 10.1103/PhysRevE.89.012813 PACS number(s): 89.75.Hc, 64.60.ah, 84.35.+i

Systems in many branches of science, such as the internet backbone, airline routes, gene regulation, and biochemical pathways, can be modeled as complex networks, whose function relies crucially on the connectivity between the components [1–3]. The resilience to node and edge breakdowns is an important property of such systems. Perhaps the most widely used approach in the theoretical study of network robustness is percolation on the configuration model. The standard configuration model of network theory is an ensemble of random graphs [4,5], in which only degrees of nodes are specified and connections between nodes are random. The classical site and bond percolation process is defined through deletion of nodes and edges on the graph independently and uniformly [6]. Recently, to understand network resilience in realistic systems, some researchers have investigated nonindependent percolation on the configuration model, including degree dependent [7,8], multihop [9], core [10], and k-core percolation [11]. Others have explored the percolation process on more sophisticated configuration-like models, such as networks with degree-degree correlations [12,13], clustering [14], blocks [15], and interdependency [16–18].

Many real networks obey a power-law form in their degree distribution. Albert *et al.* [1] suggested that such networks (called scale-free networks) exhibit a drastically different character from homogeneous Erdős-Rényi (ER) networks. The existence of hubs (highly connected nodes) in a scale-free network causes remarkable robustness against random failures, but suffers from vulnerability to malicious attacks at the hubs. These discoveries have been confirmed analytically by using the percolation theory on a configuration model [2,6,19], and hubs are commonly referred to as the Achilles' heel of a scale-free network [20,21].

Previous studies of network robustness, including these mentioned above, have mostly focused on the binary node state. In other words, a node is either occupied (intact) or unoccupied (deleted). However, it is natural to think of a node as being in a state that is functional but not at full power, that is, cannot fulfill its full task. Examples include a sensor at an insufficient battery level and a substrate with reduced ability to synthesize compounds due to errors and mutations [22]. Despite the relevance of this issue, it is still unknown how the introduction of an in-between state will affect, among other things, the network robustness. In Fig. 1 is shown an example of a network structure containing fully functional (FF), partially functional (PF), and nonfunctional (NF) nodes. To accommodate this situation, we develop a theoretical framework, called fractional percolation, and find that the scale-free networks are vulnerable to attacks barely causing partial functionality of the hubs. This implies that the links between hubs play a more important role regarding the network resilience than random links. It highlights the need to consider a PF state in designing robust networks.

We start by introducing the fractional percolation on graphs with arbitrary degree distributions. Let p_k be the probability that a randomly chosen node has degree k. Initially all nodes are FF. Nodes in an FF state with degree k stay in the same state with probability $1-q_k$, become PF with probability $q_k(1-r_k)$, and become NF with probability $q_k r_k$ independently, where r_k is the probability that a PF node becomes NF. Nodes in FF and PF states are occupied, and thus they can contribute to the sizes of clusters. NF nodes are removed from the system. However, two nodes in a PF state that are connected break their edge (see Fig. 1).

Our assumption that two PF nodes cannot sustain the "load" between them can be justified in various contexts. Imagine, for example, that each node v_i in a network possesses some "energy" $E_i \geqslant 0$. Node v_i is linked to node v_j by an edge with probability $p_{ij} = \theta(E_i + E_j - T)$, where $T > 0$ is a threshold and θ is the Heaviside step function [23]. We may think of FF, PF, and NF nodes as having energies 0, 1/2, and 1, respectively. Clearly, by choosing an appropriate T, two nodes both with lower energies cannot communicate with each other. Such realistic examples include autonomous multiagent systems [24] (E_i can be the battery level of a sensor, the sensory distance of a robot, the coverage area of a satellite, etc.) and insect colonies [25] (E_i can be the pheromone level that an ant or a bee emits). In all these examples, resilience of the signaling networks underpinning the physical systems is essential to their functionality.

By construction, when $r_k = 1$ for all k, we reproduce the ordinary site percolation: each node is deleted once it suffers from some error. If $r_k = 0$ for all k, it produces a

*shylmath@hotmail.com

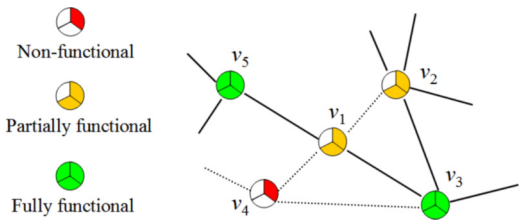

FIG. 1. (Color online) Fractional percolation on a network: the green nodes v_3 and v_5 remain intact (FF); the amber nodes v_1 and v_2 are under some nonfatal error (PF); the red node v_4 suffers from some fatal error (NF). Dotted lines mean deleted edges.

nonindependent bond percolation. The fractional percolation, in general, is a nonindependent joint site and bond percolation. When $q_k = q$ and $r_k = r$ for all k, for example, we can think of r as the intensity of error, the greater the r, the more serious the error of a node becomes, and q as the extensity of error, the greater the q, the more extensively the error occurs. The fractional percolation can be generalized to include more node states justifying its name [26]. Here we concentrate on this benchmark model and derive analytical formulas for properties of the resulting networks.

We first consider the nondegree-dependent case, where $q_k = q$ and $r_k = r$ for all k. The probability generating function for occupied node degree distribution [4] is given by

$$F_0(x) = \sum_{k=0}^{\infty}[(1-q)+q(1-r)]p_k x^k = (1-qr)G_0(x), \tag{1}$$

where the first term $1-q$ is due to the FF state and the second term $q(1-r)$ to the PF state. Here $G_0(x) = \sum_{k=0}^{\infty} p_k x^k$ is the generating function for node degree alone. If we follow a randomly chosen edge from an occupied root node, the node reached, say, v_0, may be FF with probability $1-q$, or PF with probability $q(1-r)$. In the first case, the selected edge is occupied with probability 1, while in the second case, it is occupied with average probability $1-q$ [27]. Therefore, the probability that v_0 and the edge in question are both occupied is $1-q+q(1-r)(1-q)$ [28]. The distribution of the number of edges leading out of v_0 (called the excess degree distribution [4,29]) is generated by [6]

$$\begin{aligned} F_1(x) &= \frac{\sum_{k=1}^{\infty} kp_k[(1-q)+q(1-r)(1-q)]x^{k-1}}{\langle k\rangle} \\ &= (1-q)[1+q(1-r)]G_1(x), \end{aligned} \tag{2}$$

where $\langle k\rangle = \sum_{k=0}^{\infty} kp_k$ is the average degree, and $G_1(x) = G_0'(x)/\langle k\rangle$ is the generating function for excess node degree alone. Since $F_1(x) \neq F_0'(x)/\langle k\rangle$, the fractional percolation is essentially different from the independent joint site and bond percolation [6].

Let $H_1(x)$ be the generating function for the distribution of the sizes of percolation clusters that are reached by choosing a random edge and following it to its end, say, v_0. Thus, $H_1(x)$ satisfies the self-consistent condition

$$H_1(x) = 1 - F_1(1) + xF_1[H_1(x)], \tag{3}$$

where $1-F_1(1)$ represents the probability that the cluster contains zero nodes (either v_0 is NF, or the selected edge is deleted), and the term $xF_1[H_1(x)]$ takes into account an occupied (FF or PF) v_0 with k other edges leading out of it, distributed according to $F_1(x)$ [6]. The probability distribution of the size of percolation cluster to which a randomly chosen node belongs is analogously generated by $H_0(x)$, where

$$H_0(x) = 1 - F_0(1) + xF_0[H_1(x)]. \tag{4}$$

We now investigate the critical behavior associated with the fractional percolation. From (3) and (4), we obtain that, in the absence of giant (formally infinite) clusters, the mean size of cluster to which a random node belongs is given by $\langle s\rangle = H_0'(1)$. Hence,

$$\langle s\rangle = (1-qr)\left\{1 + \frac{(1-q)[1+q(1-r)]G_0'(1)}{1-(1-q)[1+q(1-r)]G_1'(1)}\right\}. \tag{5}$$

The critical probabilities (q_c, r_c) at which a giant cluster first emerges are determined by

$$1 = (1-q_c)[1+q_c(1-r_c)]G_1'(1), \tag{6}$$

where $G_1'(1)$ is the branching factor of the network. Given probabilities q and r, the giant cluster appears when $G_1'(1) > 1/\{(1-q)[1+q(1-r)]\}$. For fatal error ($r=1$), it reduces to $G_1'(1) > 1/(1-q)$, which coincides with [Ref. [6], Eq. (12)]. Note that, for a network with pure power-law distribution $p_k \sim k^{-\alpha}$, the ratio $\langle k^2\rangle/\langle k\rangle = G_1'(1)+1$ diverges when $1 \leqslant \alpha < 3$. By (6), $q_c \to 1$ for any r if $1 \leqslant \alpha < 3$. When $r=1$, we recover the result of Cohen *et al.* [2], which shows that such a network remains robust against random failure.

In Fig. 2 we show the critical probability q_c for different values of r in ER graphs with a Poisson degree distribution $p_k = e^{-\lambda}\lambda^k/k!$ ($k \geqslant 0$) and scale-free graphs with a truncated power-law distribution $p_k \sim k^{-\alpha}e^{-k/\kappa}$ ($k \geqslant 1$), featuring many real-world networks [30,31]. The agreement between the simulations and the analytical calculations is excellent. For any given λ or κ, the threshold q_c increases as r decreases, indicating that error should occur more widely (meaning higher q) in order to destroy the giant cluster when the error becomes milder (meaning smaller r). In particular, q_c becomes the highest for $r=0$, where no node is deleted.

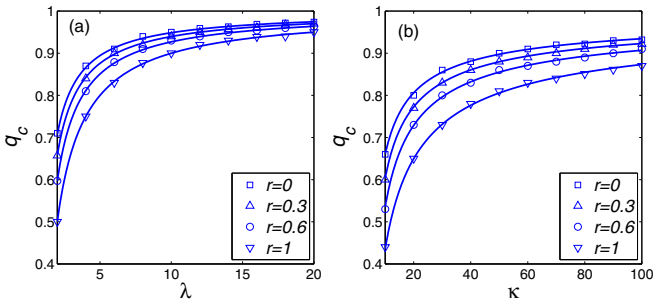

FIG. 2. (Color online) Percolation threshold q_c for networks of 10^6 nodes from numerical simulations with $r=0$ (squares), 0.3 (upper triangles), 0.6 (circles), and 1 (lower triangles), and exact solutions (solid lines): (a) for ER graphs with $\langle k\rangle = \lambda$ and (b) for scale-free networks with degree exponent $\alpha = 2.4$ and exponential cutoff κ.

We describe the algorithm used as follows. Begin with $q = 0$ and a whole list of FF nodes. Take nodes progressively from the list and do the following: (i) change its state to PF with probability q; (ii) if it becomes a "new" PF node, delete the edges between the new PF node and its neighbors in the PF state; (iii) with probability r the new PF node becomes NF; (iv) remove the NF node from the network. After checking the whole list, calculate the fraction (relative size) S of the giant cluster. Increase q and repeat the process until $S < 10^{-3}$.

Next, we study the fraction of giant cluster and mean size of nongiant clusters for the resulting network, which will help us better understand the impact of random failures and deliberate attacks on the network. Since $H_0(x)$ generates the size distribution of nongiant clusters, we have $H_0(1) = 1 - S$, where S is the fraction of giant cluster. Using (3) and (4), we obtain

$$S = 1 - H_0(1) = (1 - qr)[1 - G_0(u)], \tag{7}$$

where $u = H_1(1)$ is the smallest non-negative solution of

$$u = 1 - (1 - q)[1 + q(1 - r)][1 - G_1(u)]. \tag{8}$$

The mean size of the clusters, excluding the giant cluster to which a randomly chosen node belongs, can be expressed by

$$\langle s\rangle = \frac{H_0'(1)}{H_0(1)} = \frac{1 - qr}{1 - S}\left\{G_0(u) + \frac{(1 - q)[1 + q(1 - r)]G_0'(u)G_1(u)}{1 - (1 - q)[1 + q(1 - r)]G_1'(u)}\right\}, \tag{9}$$

which is equivalent to (5) when there is no giant cluster (i.e., $S = 0, u = 1$).

In the degree-dependent case, a node of degree k is occupied with probability $p_k(1 - q_k r_k)$. In analogy to (1) we have

$$F_0(x) = \sum_{k=0}^{\infty} p_k(1 - q_k r_k)x^k. \tag{10}$$

If we follow a randomly chosen edge from an occupied root node with degree k', the probability that the selected edge and the node reached are both occupied is $(1 - q_k) + q_k(1 - r_k)(1 - q_{k'})$ by a similar argument as above (2). Hence, the excess distribution is generated by

$$F_1(x|k') = \frac{\sum_{k=1}^{\infty} kp_k[(1 - q_k) + q_k(1 - r_k)(1 - q_{k'})]x^{k-1}}{\langle k\rangle}. \tag{11}$$

Likewise, the counterparts of (3) and (4) become, respectively,

$$H_1(x|k') = 1 - F_1(1) + \frac{x\sum_{k=1}^{\infty} kp_k(1 - q_k)H_1(x|k)^{k-1}}{\langle k\rangle} + \frac{x\sum_{k=1}^{\infty} kp_k q_k(1 - r_k)(1 - q_{k'})H_1(x|k)^{k-1}}{\langle k\rangle}, \tag{12}$$

and

$$H_0(x) = 1 - F_0(1) + x\sum_{k=0}^{\infty} p_k(1 - q_k r_k)H_1(x|k)^k. \tag{13}$$

Figure 3 shows the fraction of giant cluster and the mean size of nongiant clusters for the same systems as in Fig. 2. The main panels display the behaviors under random failures,

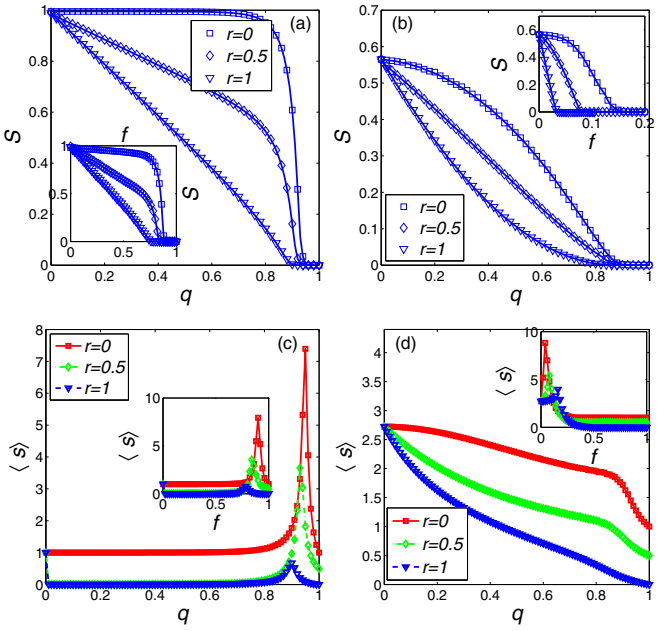

FIG. 3. (Color online) The fraction of giant cluster S and the mean size of nongiant clusters $\langle s\rangle$ as a function of q under random failures and as a function of f under targeted attacks. The same systems are used as in Fig. 2: Left column [(a) and (c)] is for ER graphs with $\lambda = 10$; right column [(b) and (d)] is for scale-free networks with $\alpha = 2.4$ and $\kappa = 60$. In (a) and (b), points are the simulation results and solid lines are the exact solutions. In (c) and (d), only exact solutions are displayed.

where a fraction q of the nodes are randomly selected (become PF), and each of them is deleted (become NF) with probability r as described in the algorithm above. The insets display the behaviors under targeted attacks, where the only difference is that we select nodes in decreasing order of their degree k [32]. In the language of fractional percolation, this is equivalent to setting $q_k = \theta(k - k_c)$, where k_c is a cutoff ranging from 0 to the maximum degree of the network [6,33]. We define the fraction $f = f(k_c)$ of the most connected nodes by the number of nodes with degree larger than k_c divided by 10^6, the total number of nodes.

First, we observe that the simulated results agree with their analytical counterparts, and the phase transition point at $S = 0$ coincides with the critical probability q_c in Fig. 2.

Second, there is a clear gap between the values of S for different r in both ER graphs [Fig. 3(a)] and scale-free networks [Fig. 3(b)]. A reduction in r, i.e., the intensity of error, systematically yields a change in the convexity of S and therefore the fraction of giant cluster increases at any given value of q. For example, consider $q = 0.5$ in Fig. 3(b), that is, half nodes in the network suffer from error. The giant cluster still consists of nearly 40% nodes for $r = 0$ while it consists of only about 10% nodes for $r = 1$.

Third, for given r, there is no big difference between the fractions of giant clusters corresponding to random failures and targeted attacks in ER graphs [Fig. 3(a)] owing to the homogeneity of the network. In contrast, a pronounced difference is seen for scale-free networks [Fig. 3(b)]: S decreases rapidly under targeted attacks on all three levels

of error intensity considered. This phenomenon for $r = 1$ has been attributed to the existence of hubs (Achilles' heel) [1,2,6]. When $r < 1$, we find that the links between two hubs also contribute substantially to the formation of a giant cluster. (The case $r = 0$ under targeted attacks is similar to the centrally biased bond percolation [7], where links between hubs are prone to attacks but all nodes are intact.) For example, even where no node is deleted when $r = 0$ [Fig. 3(b) inset], the giant cluster falls apart as 14% hubs become PF. However, the giant cluster remains almost intact when the same fraction of random nodes turn into PF [Fig. 3(b) main panel].

Fourth, there is always a peak in $\langle s \rangle$ characteristic of a phase transition at criticality (except for scale-free networks under random failures). The fact that $\langle s \rangle$ is decreasing monotonically in the main panel of Fig. 3(d) indicates that the scale-free network is deflated by nodes breaking off one by one under random failures: the increasing error scope q leads to the isolation of single nodes, not clusters of nodes. This phenomenon was observed [1] for the case of $r = 1$.

To conclude, we have developed a novel, simple framework to study nonbinary and nonrandom percolation on general random graphs so that both intensity and extensity of errors are well characterized in the model parameters. The robustness-fragility transition point is pinned down at degree exponent 3 for scale-free networks. We obtain exact solutions for the percolation threshold, the fraction of giant cluster, and the mean size of small clusters. Due to the existence of a PF state, the edges between hubs are found to play a more vital role regarding the network resilience than random edges. The agreement between the theory and the simulations is excellent. Further application of fractional percolation on other network models (e.g., interdependent networks, networks with community structures, etc.) or extension of the framework itself could deepen our understanding of network resilience and dynamical processes on networks [34].

This work has been supported by a SUTD-MIT International Design Center Grant.

[1] R. Albert, H. Jeong, and A.-L. Barabási, Nature (London) **406**, 378 (2000).
[2] R. Cohen, K. Erez, D. ben-Avraham, and S. Havlin, Phys. Rev. Lett. **85**, 4626 (2000).
[3] M. E. J. Newman, SIAM Rev. **45**, 167 (2003).
[4] M. E. J. Newman, S. H. Strogatz, and D. J. Watts, Phys. Rev. E **64**, 026118 (2001).
[5] M. Molloy and B. Reed, Random Struct. Algorithms **6**, 161 (1995).
[6] D. S. Callaway, M. E. J. Newman, S. H. Strogatz, and D. J. Watts, Phys. Rev. Lett. **85**, 5468 (2000).
[7] A. A. Moreira, J. S. Andrade Jr., H. J. Herrmann, and J. O. Indekeu, Phys. Rev. Lett. **102**, 018701 (2009).
[8] Y. Shang, J. Biol. Phys. **39**, 489 (2013).
[9] Y. Shang, W. Luo, and S. Xu, Phys. Rev. E **84**, 031113 (2011).
[10] Y.-Y. Liu, E. Csóka, H. Zhou, and M. Pósfai, Phys. Rev. Lett. **109**, 205703 (2012).
[11] S. N. Dorogovtsev, A. V. Goltsev, and J. F. F. Mendes, Phys. Rev. Lett. **96**, 040601 (2006).
[12] A. Vázquez and Y. Moreno, Phys. Rev. E **67**, 015101 (2003).
[13] A. Srivastava, B. Mitra, N. Ganguly, and F. Peruani, Phys. Rev. E **86**, 036106 (2012).
[14] M. E. J. Newman, Phys. Rev. Lett. **103**, 058701 (2009).
[15] T. P. Peixoto and S. Bornholdt, Phys. Rev. Lett. **109**, 118703 (2012).
[16] S. V. Buldyrev, R. Parshani, G. Paul, H. E. Stanley, and S. Havlin, Nature (London) **464**, 1025 (2010).
[17] J. Gao, S. V. Buldyrev, S. Havlin, and H. E. Stanley, Phys. Rev. Lett. **107**, 195701 (2011).
[18] A. Bashan, R. Parshani, and S. Havlin, Phys. Rev. E **83**, 051127 (2011).
[19] Y. Shang, Europhys. Lett. **95**, 28005 (2011).
[20] A.-L. Barabási, in *Handbook of Graphs and Networks: From the Genome to the Internet*, edited by S. Bornholdt and H. G. Schuster (Wiley-VCH, Berlin, Germany, 2003), pp. 69–84.
[21] Cover of Nature (London), **406**(6794), (2000).
[22] S. N. Dorogovtsev and J. F. F. Mendes, *Evolution of Networks: From Biological Nets to the Internet and WWW* (Oxford University Press, Oxford, 2003).
[23] Similarly, we may equally well imagine another form of connection probability as $p_{ij} = \theta(E_i \cdot E_j - T)$.
[24] Y. Shang, Int. J. Nonlin. Sci. Num. Simul. **14**, 355 (2013).
[25] T. D. Wyatt, *Pheromones and Animal Behaviour: Communication by Smell and Taste* (Cambridge University Press, Cambridge, 2003).
[26] A similar concept in spirit is the fractional chromatic number, generalizing ordinary chromatic number in graph theory.
[27] We use a mean field approximation here: When v_0 is PF, $\mathbb{P}$ (the root is connected to v_0) $= 1 \cdot \mathbb{P}$ (root is FF) $+ 0 \cdot \mathbb{P}$ (root is PF) $= 1 - q$.
[28] The site and bond percolations are decoupled by assuming that occupation of the selected edge is independent of v_0.
[29] M. E. J. Newman, Phys. Rev. E **76**, 045101 (2007).
[30] M. E. J. Newman, Proc. Natl. Acad. Sci. U.S.A. **98**, 404 (2001).
[31] L. A. N. Amaral, A. Scala, M. Barthélémy, and H. E. Stanley, Proc. Natl. Acad. Sci. U.S.A. **97**, 11149 (2000).
[32] We take a fraction f of most connected nodes from the list of FF nodes, and change their states to PF.
[33] L. D. Valdez, P. A. Macri, H. E. Stanley, and L. A. Braunstein, Phys. Rev. E **88**, 050803(R) (2013).
[34] A. Barrát, M. Barthélemy, and A. Vespignani, *Dynamical Processes on Complex Networks* (Cambridge University Press, Cambridge, 2008).