Limited Path Percolation in Complex Networks

Eduardo López,[1,*] Roni Parshani,[2,†] Reuven Cohen,[3] Shai Carmi,[2] and Shlomo Havlin[2]
[1]*CNLS & T-7, Theoretical Division, Los Alamos National Laboratory, Los Alamos, New Mexico 87545, USA*
[2]*Minerva Center & Department of Physics, Bar-Ilan University, Ramat Gan, Israel*
[3]*Massachusetts Institute of Technology, Cambridge, Massachusetts, USA*
(Received 14 February 2007; published 29 October 2007)

We study the stability of network communication after removal of a fraction $q = 1 - p$ of links under the assumption that communication is effective only if the shortest path between nodes i and j after removal is shorter than $a\ell_{ij} (a \geq 1)$ where ℓ_{ij} is the shortest path before removal. For a large class of networks, we find analytically and numerically a new percolation transition at $\tilde{p}_c = (\kappa_0 - 1)^{(1-a)/a}$, where $\kappa_0 \equiv \langle k^2 \rangle / \langle k \rangle$ and k is the node degree. Above $\tilde{p}_c$, order N nodes can communicate within the limited path length $a\ell_{ij}$, while below $\tilde{p}_c$, N^δ ($\delta < 1$) nodes can communicate. We expect our results to influence network design, routing algorithms, and immunization strategies, where short paths are most relevant.

DOI: 10.1103/PhysRevLett.99.188701 PACS numbers: 89.20.Hh, 02.50.Cw, 64.60.Ak, 89.75.Hc

The study of complex networks has emerged as an important tool to better understand many social, technological, and biological real-world systems ranging from communication networks like the Internet to cellular networks [1]. In many cases, networks are the medium through which information is transported, i.e., in social networks the propagation of epidemics, rumors, etc., and in the Internet the propagation of data packets [2–6].

An important question regarding networks is their stability, i.e., under what conditions the network breaks down [7–10]. In communications, a network breakdown means information cannot be transmitted to most nodes, and in epidemiology, that an epidemic has stopped.

The main approach for studying network stability is percolation theory [11]. In percolation, a fraction $q = 1 - p$ of the N network nodes (or L links) is removed until a critical value p_c is reached [12]. For $p < p_c$ the network collapses into small clusters, while for $p > p_c$, a spanning cluster of order N nodes appears [8,9,11,13,14]. However, even though in the original network nodes are connected through short paths, decreasing values of p imply increasing values of the typical distance between nodes, even far above p_c. In the insets of Fig. 1, we present an illustration of such an effect for Erdős-Rényi and the EpiSimS [15] networks, where one can clearly observe that as p decreases, the length increases monotonically until reaching its maximum at p_c [16]. Such path length changes can have devastating effects on a number of real-works networks, where even a slight increase in path length leads to a loss of network function. For example, in communication, long paths are usually inefficient, and in epidemics, disease spreading often decays in time due to effects such as seasonality or pathogen mutation, so for long paths the epidemic may die out before total network infection. In these cases, the interesting question changes from "when does the network break down?" as in usual percolation, to "when does the network become inefficient?"

Therefore, to correctly consider network stability in the face of functional limits on path length, we propose a new percolation model which we call limited path percolation (LPP). In this model, failures on the links are represented by the removal of a fraction $1 - p$ of the network links. For any two nodes i and j to be considered connected after removal, we require that the new shortest path between them is shorter than $a\ell_{ij}$ ($a \geq 1$), where ℓ_{ij} is the shortest path before removal. We then ask, given our new limited path constraint, what is the value p at which a spanning communicating cluster appears? The communicating cluster is defined by the number of nodes S_a a randomly chosen node can communicate with given the limited length restriction [17], and it is considered spanning in LPP when it scales as N. We find a new phase transition, dependent on a, at $\tilde{p}_c \equiv \tilde{p}_c(a)$, where $p_c < \tilde{p}_c < 1$. For $p_c < p < \tilde{p}_c$, the LPP communicating cluster is only a zero fraction (fractal) of the network, which scales as N^δ ($\delta < 1$). For $p > \tilde{p}_c$, LPP produces a spanning communicating cluster. Figure 1 illustrates the effect of imposing length restrictions on the connectivity, where the location of the transition for S_a moves to increasing values of p as a decreases.

For simplicity, we start our analysis with Erdős-Rényi (ER) networks and argue that the theory is valid in general for random networks. We begin with random removal and later extend our considerations to targeted removal on highly connected nodes and find that similar phenomena appear. We support our theory with simulations.

An Erdős-Rényi network [13,14] of N nodes is a random network with pairs of nodes connected with probability ϕ. The degree distribution $\Phi(k)$ is Poisson with the form $\Phi(k) = \langle k \rangle^k e^{-\langle k \rangle}/k!$, where k, the degree, is the number of links attached to a node, and $\langle k \rangle \equiv \sum_{k=1}^{\infty} k\Phi(k)$ is the average degree of the network. The typical distance between nodes is $\log N / \log\langle k \rangle$.

To evaluate S_a, we note that after the removal of fraction q of the links, the communicating cluster can be considered treelike since, up to order N, loops are negligible [8]. Thus, S_a, for $N \gg 1$, can be approximated by

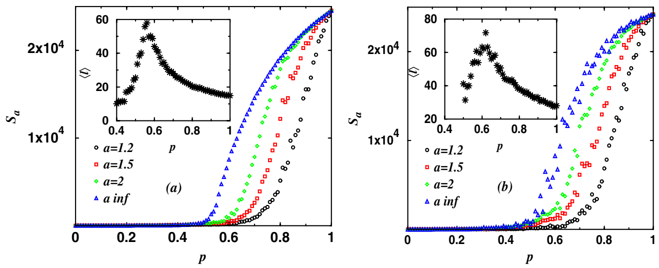

FIG. 1 (color online). The size of the communicating cluster S_a as a function of p when the length restriction for connectivity is imposed, with $a = 1.2, 1.5, 2, \infty$. As a decreases, the value of p for which S_a begins to grow substantially increases. The insets correspond to $\langle \ell \rangle$ vs p, where $\langle \ell \rangle$ is the average distance between network nodes. $\langle \ell \rangle$ reaches a maximum at the usual percolation threshold p_c, and we limit the considerations in this Letter to the range $p \geq p_c$. (a) Corresponds to Erdős-Rényi networks and (b) to the EpiSimS network [15].

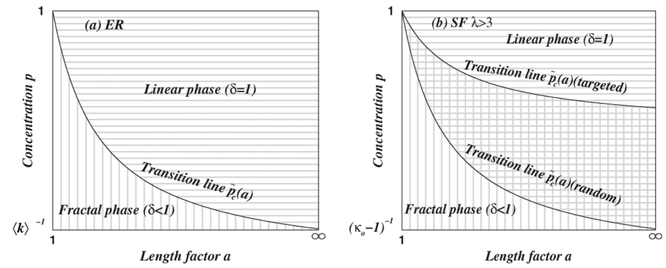

FIG. 2. (a) Phase diagram for Erdős-Rényi networks of LPP with respect to a and p, demonstrating the linear and power-law (fractal) phases for $S_a \sim N^\delta$. (b) Similar phase diagram for scale-free networks with $\lambda > 3$. The two transition lines represent networks with the same κ. Note the slow decrease of the transition line for targeted removal compared to the transition line for random removal. The region between the two lines has a power-law (fractal) phase for targeted removal and a linear phase for random removal. In both (a), (b) the regular percolation threshold is given by the limit $a \to \infty$, i.e., $p_c = \langle k \rangle^{-1}$ for ER and $p_c = (\kappa_0 - 1)^{-1}$ for SF with $\lambda > 3$.

$$S_a \sim c(p)[p\langle k \rangle]^{a(\log N/\log\langle k \rangle)} = c(p)N^\delta \quad \text{(ER)}, \quad (1)$$

where $\delta \equiv a(1 - |\log p|/\log\langle k \rangle) \leq 1$, $p\langle k \rangle$ is the average degree after removal, $c(p) \equiv c_o[p\langle k \rangle + 1]/[p\langle k \rangle - 1]$ [18], and $a \log N/\log\langle k \rangle$ is the new tree depth imposed by the limited path length restriction. The exponent $\delta = \delta(a, p, \langle k \rangle)$ is an increasing function of a; i.e., for larger a values longer paths are valid and therefore more nodes are included in the communicating cluster, leading to larger δ values. The exponent δ is bounded below by 0 and above by 1, since the number of nodes available cannot exceed N. Setting $\delta = 1$ and solving for p in Eq. (1) we obtain the transition threshold

$$\tilde{p}_c(a) = \langle k \rangle^{(1-a)/a} \quad \text{(ER)}. \quad (2)$$

Figure 2 shows the LPP phase diagram. For $p_c \leq p \leq \tilde{p}_c(a)$, the communicating cluster is a fractal of size N^δ, where δ continuously increases with p. The threshold $\tilde{p}_c(a)$ decreases from 1 at $a = 1$ to p_c at $a \to \infty$. Note that for $a \to \infty$, when no path length restriction is imposed, we recover the usual percolation threshold $\tilde{p}_c(a \to \infty) = p_c = 1/\langle k \rangle$ [13]. For $p > \tilde{p}_c(a)$, a spanning communicating cluster exists with path lengths $\ell'_{ij} \leq a\ell_{ij}$. Using the function $1 - \tilde{p}_c(a)$ we are able to calculate, for a given value of a, the percentage of links that can be removed before nodes i and j cannot communicate through paths shorter than $a\ell_{ij}$. Equations (1) and (2) are supported by the simulations presented in Fig. 3(a) [17,19]. For a summary of the results, see Table I.

Our results for the different regimes of S_a can be summarized by the scaling relation for $p > p_c$

$$S_a \sim c(p)N^\delta f\left(\frac{P_\infty N}{c(p)N^\delta}\right) \quad \text{(ER)}, \quad (3)$$

where P_∞ is the probability of an arbitrary node to belong to the usual percolation spanning cluster [11]. The function $f(x)$ scales as x when $x \ll 1$ and approaches a constant as $x \gg 1$. In Fig. 4(a), we present simulation results for several a and p values for ER networks, supporting the scaling form of Eq. (3).

The theory for LPP can be extended to all random networks with typical distance between nodes of order $\log N$ by substituting $\langle k \rangle$ with the generalized form $(\kappa_0 - 1)$, known as the branching factor, defined by $\kappa_0 - 1 \equiv \frac{\langle k^2 \rangle}{\langle k \rangle} - 1$ [8]. Replacing $\langle k \rangle$ with $(\kappa_0 - 1)$ in Eq. (1) we obtain the general equation for the communicating cluster size

$$S_a \sim c(p)(\kappa - 1)^{a[\log N/\log(\kappa_0 - 1)]} = c(p)N^\delta, \qquad \delta \equiv a\frac{\log(\kappa - 1)}{\log(\kappa_0 - 1)}, \quad (4)$$

where $\kappa_0 - 1$ is the branching factor of the original network and $\kappa - 1$ the branching factor after removal, which depends on p. When a random fraction of the network is removed, $\kappa - 1 = p(\kappa_0 - 1)$ [8]. Specifically for ER networks, $\kappa - 1 = p\langle k \rangle$ and $\kappa_0 - 1 = \langle k \rangle$, reducing Eq. (4) to Eq. (1). In the general case of random networks, the LPP transition is found by imposing $\delta = 1$, which yields

$$\tilde{p}_c(a) = (\kappa_0 - 1)^{(1-a)/a}. \quad (5)$$

The scaling of S_a is the same as Eq. (3) with δ from Eq. (4).

Our general theory for LPP can be illustrated on scale-free (SF) networks. Scale-free networks have generated much interest due to their relation to many real-world networks, such as the Internet, WWW, social networks, cellular networks, and world-airline network [1,22–25]. Scale-free networks are characterized by a power-law degree distribution $\Phi(k) \sim k^{-\lambda}$ $(m \leq k \leq K)$, where $K \equiv mN^{1/(\lambda-1)}$ [8]. The power-law distribution allows a network to have a few nodes with a large number of links ("hubs") which usually play a critical role in network function. Calculating κ_0 for SF networks one obtains $\kappa_0 = \left(\frac{2-\lambda}{3-\lambda}\right)\frac{K^{3-\lambda} - m^{3-\lambda}}{K^{2-\lambda} - m^{2-\lambda}}$ [8]. For $\lambda > 3$, Eq. (4) is valid and thus LPP

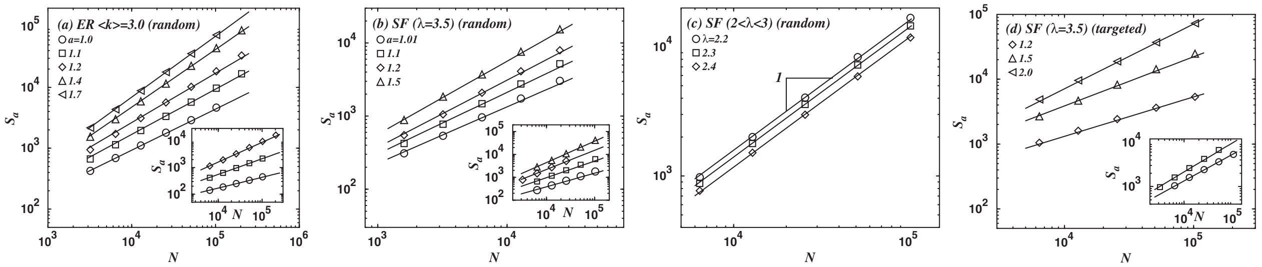

FIG. 3. Simulation results (symbols) for S_a vs N for various network types under random or targeted removal and various a and p values (indicated in plot legends), compared to the theoretically predicted power laws (solid lines), with δ calculated from Table I. Network sizes are typically between 1600 and 204 800. In all plots the simulation results agree with theoretical predictions. (a) ER networks (random) with $\langle k\rangle = 3$ for fixed $p = 0.7$ and different a values. Inset shows the same networks with fixed $a = 1.1$ and $p = 0.5$ (○), 0.6 (□), and 0.7 (◇). (b) SF networks (random) with $\lambda = 3.5$, $m = 2$, $p = 0.7$, and different a values. Inset shows the same networks with fixed $a = 1.1$ and $p = 0.5$ (○), 0.6 (□), 0. 7 (◇), and 0.8 (△). (c) SF networks (random), $\lambda = 2.2$, 2.3, and 2.4, $m = 3$, $p = 0.4$, and $a = 1$. (d) SF networks (targeted) with $\lambda = 3.5$, $m = 3$, fixed $p = 0.92$, and different a values. Inset shows the same networks with fixed $a = 1.2$ and $p = 0.92$ (○) and 0.94 (□).

is similar to ER networks, except that it depends on $\kappa_0 - 1$ instead of $\langle k\rangle$. The phase diagram of SF networks is shown in Fig. 2(b). The results of the simulations supporting the theoretical value of δ, Eq. (4), are shown in Fig. 3(b), and for the scaling form of S_a are presented in Fig. 4(b).

For $2 < \lambda < 3$, the typical distance scales as $\ell = 2\log\log N/|\log(\lambda - 2)|$ [26,27]. For this regime, our scaling approach to calculate S_a is no longer valid since the tree approximation breaks down. However, the LPP transition still exists when $a\ell_{ij} = \ell'_{ij}$, where ℓ'_{ij} is the distance after removal, with typical value $\ell' = 2\log\log P_\infty N/|\log(\lambda - 2)|$ [28]. Solving $a\ell_{ij} = \ell'_{ij}$ for $N \to \infty$ yields

$$a = \frac{\ell'_{ij}}{\ell_{ij}} = \frac{\log\log P_\infty N}{\log\log N} \to 1, \qquad [\mathrm{SF}(r),\ 2 < \lambda < 3]. \quad (6)$$

This implies that $\tilde{p}_c \to 0$ and thus, for any finite p, S_a is always of order N. The results of the simulations presented in Fig. 3(c) support our prediction.

Up to this point, we have only considered random removal of links. Another kind of removal is targeted removal where the nodes with the largest degree are removed first [8]. This kind of removal is common in many real-world scenarios such as denial of service attacks on WWW and delays in airline hubs.

In scale-free networks, targeted removal of a fraction q of the nodes with largest degree can be treated as random removal of $q' = q^{(2-\lambda)/(1-\lambda)}$ of the network links [8]. After removal, the maximum degree is $K' = mq^{1/(1-\lambda)}$. For $\lambda > 3$, making the substitutions $q \to q'$ and $K \to K'$ in Eq. (4) we obtain the equation for $\tilde{p}_c$ [21] and the scaling form for S_a (see Table I). The change to q' and K' reflects the fast collapse of the network and the rapid change in the typical network length. The transition line $\tilde{p}_c(a)$ in targeted removal decreases significantly more slowly compared to random removal as seen in Fig. 2(b).

In targeted removal for $2 < \lambda < 3$, removing even a small fraction of the hubs produces a change in the distance from $2\log\log N/|\log(\lambda - 2)|$ to $\log P_\infty N/\log(\kappa - 1)$ [26,27]. After removal, S_a can be calculated using the tree approximation yielding

$$S_a \sim (\kappa - 1)^{2a[\log\log N/|\log(\lambda-2)|]}$$
$$= (\log N)^{2a[\log(\kappa-1)/|\log(\lambda-2)|]} \qquad [\mathrm{SF}(t),\ 2 < \lambda < 3]. \quad (7)$$

In this case, the phase transition to a spanning communicating cluster cannot be achieved for any finite value of a and $p < 1$, as seen from Eq. (7). Simulation results supporting Eq. (7) are shown in Fig. 4(d). Comparing random to targeted removal for $2 < \lambda < 3$ for LPP yield entirely opposite results. In random removal, order N nodes are still connected through the original paths. On the other hand, in targeted removal for any finite a, the network collapses into logarithmically small clusters.

In summary, our results suggest that usual percolation theory cannot correctly describe connectivity when only a limited set of path lengths are useful. In usual percolation, order N network nodes are connected when $p > p_c$. However, in LPP, when $p_c < p < \tilde{p}_c$, the size of S_a corresponds to a zero fraction of the original network. Therefore, a much smaller failure of the network can lead to an effective network breakdown. As an illustration, consider an ER network with $\langle k\rangle = 3$ and limit the distance between nodes to $a = 1.5$ times the original length. Then,

TABLE I. The functions $\tilde{p}_c$, S_a, and δ for several kinds of network structures under random (r) and targeted (t) removal. The scaling of S_a is given for $p < \tilde{p}_c$ except for scale-free networks with $2 < \lambda < 3$ under random removal where $\tilde{p}_c = 0$.

Structure		$\tilde{p}_c$	S_a	δ
ER		$\langle k\rangle^{(1-a)/a}$	N^δ	$a - \frac{a\|\log p\|}{\log\langle k\rangle}$
SF (r)	$2 < \lambda < 3$	0	N^δ	1 ($p > \tilde{p}_c$)
	$\lambda > 3$	$(\kappa_0 - 1)^{(1-a)/a}$	N^δ	$a - \frac{a\|\log p\|}{\log(\kappa_0-1)}$
SF (t)	$2 < \lambda < 3$	1	$(\log N)^\delta$	$2a\frac{\log(\kappa-1)}{\|\log(\lambda-2)\|}$
	$\lambda > 3$	$\tilde{p}_c(a, \kappa, \kappa_0)$ [21]	N^δ	$a\frac{\log(\kappa-1)}{\log(\kappa_0-1)}$

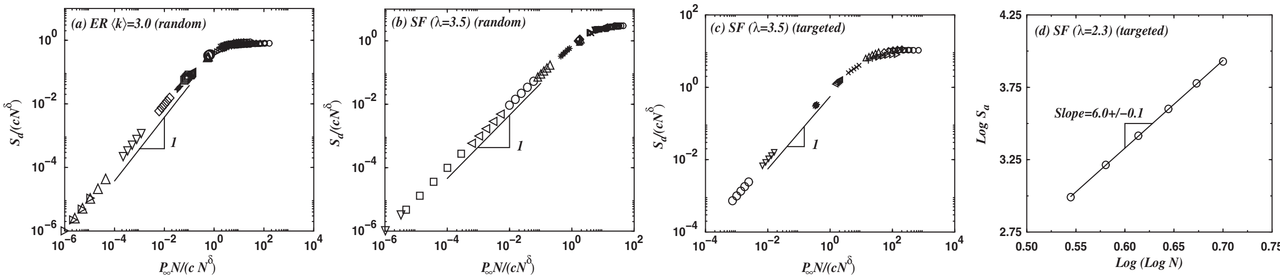

FIG. 4. Simulation results [(a)–(c)] for the scaling of $S_a/(c(p)N^{\delta})$ vs $P_{\infty}N/(c(p)N^{\delta})$ for various types of networks, for random and targeted removal for N between 1600 and 25 600, and a between 1.0 and 4. (a) ER networks (random) with $\langle k\rangle = 3$ and $p = 0.5, 0.6,$ and 0.7. (b) SF networks (random) with $\lambda = 3.5$, $m = 2$ and $p = 0.6, 0.7,$ and 0.8. (c) SF networks (targeted) for $\lambda = 3.5$, $m = 3$ and a between 1.01 and 3.0, and targeted removal with $p = 0.92$ and 0.94. (d) Simulation results of $\log S_a$ vs $\log\log N$ and fit line, close to the theoretical prediction (≈ 5.8), for δ for SF networks (targeted) for $\lambda = 2.3$, $m = 3$, $a = 1.5$, $p = 0.97$.

LPP predicts that the removal of $q = 0.31$ of the network links is enough to break down the network, compared to $q = 0.67$ in regular percolation. In the context of infectious diseases, if a virus typically survives up to $1.5 \log N$ steps, LPP predicts that the immunization threshold is significantly smaller, 0.31 compared to 0.67. The above considerations indicate that our results are important for network design, routing protocols, and immunization strategies.

We thank the Israel Science Foundation, Yeshaya Horowitz Association, and The Center for Complexity Science, NEST project DYSONET, DOE, and ONR for financial support. S. C. acknowledges support from the Israel Academy of Sciences and Humanities and S. H. from NSF-HSd. We thank J. C. Miller, E. Perlsman, and S. Sreenivasan for discussions.

*Corresponding author.
edlopez@lanl.gov
†Corresponding author.
parshani.roni@gmail.com

[1] R. Albert and A.-L. Barabási, Rev. Mod. Phys. **74**, 47 (2002); R. Pastor-Satorras and A. Vespignani, *Structure and Evolution of the Internet: A Statistical Physics Approach* (Cambridge University Press, Cambridge, England, 2004); S. N. Dorogovtsev and J. F. F. Mendes, *Evolution of Networks: From Biological Nets to the Internet and WWW* (Oxford University Press, Oxford, 2003).
[2] S. Wasserman *et al.*, *Social Network Analysis: Methods and Applications* (Cambridge University Press, Cambridge, England, 1994).
[3] R. Pastor-Satorras and A. Vespignani, Phys. Rev. Lett. **86**, 3200 (2001).
[4] M. E. J. Newman, Phys. Rev. E **66**, 016128 (2002).
[5] E. López *et al.*, Phys. Rev. Lett. **94**, 248701 (2005).
[6] S. Sreenivasan *et al.*, Phys. Rev. E **75**, 036105 (2007).
[7] R. Albert *et al.*, Nature (London) **406**, 378 (2000).
[8] R. Cohen *et al.*, Phys. Rev. Lett. **85**, 4626 (2000); **86**, 3682 (2001).
[9] D. S. Callaway *et al.*, Phys. Rev. Lett. **85**, 5468 (2000).
[10] T. Tanizawa *et al.*, Phys. Rev. E **71**, 047101 (2005).
[11] S. Kirkpatrick, Rev. Mod. Phys. **45**, 574 (1973); D. Stauffer and A. Aharony, *Introduction to Percolation Theory* (Taylor and Francis, London, 2004), 2nd ed.; *Fractals and Disordered Systems*, edited by A. Bunde and S. Havlin (Springer, Berlin, 1996), 2nd ed.
[12] Note that the original network structure, e.g., Erdős-Rényi [13], is regarded here as $p = 1$.
[13] P. Erdős and A. Rényi, Publ. Math. (Debrecen) **6**, 290 (1959); Publ. Math. Inst. Hung. Acad. Sci. **5**, 1760 (1960).
[14] B. Bollobás, *Random Graphs* (Academic, New York, 1985).
[15] The EpiSimS network [S. Eubank *et al.*, Nature (London) **429**, 180 (2004)] is a precise reproduction of a day of activity in the city of Portland, Oregon, USA, constructed to forecast epidemic propagation. From this network, we have chosen individuals with at least one interaction lasting 8.5 hours or more, such as family members.
[16] L. A. Braunstein *et al.*, Phys. Rev. Lett. **91**, 168701 (2003).
[17] The size of the communicating cluster S_a is measured by: (i) generating R network realizations, (ii) performing percolation on each realization, (iii) choosing M random nodes from the largest cluster, (iv) for each node counting the number of nodes satisfying $\ell'_{ij} \le a\ell_{ij}$, and (v) averaging the results. Typically, $R = 1000$ and $M = 100$.
[18] For a tree of degree $z + 1$ and depth ℓ, the size is given by $S_a = 1 + (z+1)\Sigma_{n=0}^{\ell-1} z^n \sim (\frac{z+1}{z-1}) z^{\ell}$. If $z = p\langle k\rangle$ the tree size scales as $[(p\langle k\rangle + 1)/(p\langle k\rangle - 1)][p\langle k\rangle]^{\ell}$.
[19] We generate networks of N nodes as follows: (i) Erdős-Rényi by connecting pairs with probability ϕ and (ii) scale-free by using the Molloy-Reed algorithm [20].
[20] M. Molloy and B. Reed, Random Struct. Algorithms **6**, 161 (1995).
[21] The function $\tilde{p}_c(a, \kappa, \kappa_0)$ is given implicitly by solving $\delta = 1$ from Eq. (4), using κ for scale-free networks with $q = q'$ and $K = K'$ for targeted removal.
[22] A.-L. Barabási and R. Albert, Science **286**, 509 (1999); H. A. Simon, Biometrika **42**, 425 (1955).
[23] L. A. N. Amaral *et al.*, Proc. Natl. Acad. Sci. U.S.A. **97**, 11 149 (2000).
[24] A. Barrat *et al.*, Proc. Natl. Acad. Sci. U.S.A. **101**, 3747 (2004).
[25] P. L. Krapivsky *et al.*, Phys. Rev. Lett. **85**, 4629 (2000).
[26] R. Cohen and S. Havlin, Phys. Rev. Lett. **90**, 058701 (2003).
[27] R. van der Hofstad *et al.*, Elect. J. Prob. **12**, 703 (2007).
[28] Indeed, for $\lambda > 3$, $a\ell_{ij} = \ell'_{ij}$ reduces to Eq. (4).