Limited path percolation in complex networks

Eduardo López,[1,*] Roni Parshani,[2,*] Reuven Cohen,[3] Shai Carmi,[2] and Shlomo Havlin[2]

[1]*Center for Non-Linear Studies & T-13, Theoretical Division,*
Los Alamos National Laboratory, Los Alamos, NM 87545, USA

[2]*Minerva Center & Department of Physics,*
Bar-Ilan University, Ramat Gan, Israel

[3]*Massachusetts Institute of Technology, Cambridge, MA, USA*

(Dated: February 14, 2007)

Abstract

We study the stability of network communication after removal of $q = 1 - p$ 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 a new percolation transition at $\tilde{p}_c = (\kappa_o - 1)^{(1-a)/a}$, where $\kappa_o \equiv \langle k^2 \rangle / \langle k \rangle$ and k is the node degree. Below $\tilde{p}_c$, only a fraction N^δ of the network nodes can communicate, where $\delta \equiv a(1 - |\log p| / \log(\kappa_o - 1)) < 1$, while above $\tilde{p}_c$, order N nodes can communicate within the limited path length $a\ell_{ij}$. Our analytical results are supported by simulations on Erdős-Rényi and scale-free network models. We expect our results to influence the design of networks, routing algorithms, and immunization strategies, where short paths are most relevant.

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 links) are removed until a critical value p_c is reached. 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]. However, even though in the original network the nodes are connected through short paths, near p_c the paths become very long. For instance, in the original Erdős-Rényi network the typical distance between nodes is of order $\log N$ [12] compared to order $N^{1/3}$ near the percolation threshold [14]. These long distances may have a significant influence on network function. For example, in communication, long paths are usually inefficient, and in epidemics, disease spreading often decays in time due to mutations or natural immunization, so for long paths the epidemic may die out before the network collapses. In these cases the interesting question is sometimes, not when does the network break down, but when the network connectivity becomes inefficient.

To answer this question, we propose a new percolation model which we call limited path percolation (LPP). In this model, after removing a fraction $q = 1 - p$ of the network nodes, any two of these nodes, say i and j are considered connected only if the 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 constrains, what is the value p at which a spanning cluster appears. We find a new phase transition, which depends 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 spanning cluster is only a zero fraction (fractal) of the network, which scales as N^{δ} ($\delta < 1$). For $p > \tilde{p}_c$ the LPP spanning cluster is of order N.

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

Erdős-Rényi networks [12, 13] are random networks consisting of N nodes connected with probability ϕ and disconnected with probability $1-\phi$. 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 $\log N/\log\langle k\rangle$.

Next we evaluate S_a, the size of the spanning cluster under LPP. After the removal of fraction q of the links, the spanning cluster can be considered tree-like since, up to order N, loops are negligible [8]. Thus, S_a can be approximated by

$$S_a \sim c(p)[p\langle k\rangle]^{a\frac{\log N}{\log\langle k\rangle}} = c(p)N^{\delta}, \qquad \delta \equiv a\left(1 - \frac{|\log p|}{\log\langle k\rangle}\right) \leq 1 \qquad \text{(Erdős-Rényi)} \quad (1)$$

where $p\langle k\rangle$ is the average degree after removal, $c(p) \equiv c_o p\langle k\rangle/(p\langle k\rangle - 1)$ [15], 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 values of a longer paths are valid and therefore more nodes are included in the spanning cluster, leading to a higher value of δ. The exponent δ is bounded below by zero and above by 1, since N is the maximum number of nodes available. Setting $\delta = 1$ and solving for p in Eq. (1) we obtain the transition threshold

$$\tilde{p}_c(a) = \langle k\rangle^{\frac{1-a}{a}} \qquad \text{(Erdős-Rényi).} \quad (2)$$

Figure 1 presents the phase diagram for LPP. For $p_c \leq p \leq \tilde{p}_c(a)$ the spanning cluster is a fractal of size N^{δ} and δ continuously increases with p. For $p > \tilde{p}_c(a)$, a spanning cluster of order N 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 the network is no longer connected with effective paths, i.e., shorter than $a\ell_{ij}$. Note that for $a \to \infty$, when no path length restriction is imposed, we recover the usual percolation threshold $\tilde{p}_c(a \to \infty) = p_c = 1/\langle k\rangle$ [12]. Equations (1) and (2) are supported by the simulations presented in Fig. 2(a) [16, 18]. For a summary of the various equations in the article, see Table I.

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

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

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. 3(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 $(\kappa - 1)$, known as the branching factor, defined by $\kappa - 1 \equiv \frac{\langle k^2 \rangle}{\langle k \rangle} - 1$ [8]. Replacing $\langle k \rangle$ with $(\kappa - 1)$ in Eq. (1) we obtain the general equation for the spanning cluster size

$$S_a \sim c(p)(\kappa - 1)^{a \frac{\log N}{\log(\kappa_o - 1)}} = N^\delta, \qquad \delta \equiv a \frac{\log(\kappa - 1)}{\log(\kappa_o - 1)} \tag{4}$$

where $\kappa_o - 1$ 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, $\kappa - 1 = p(\kappa_o - 1)$ [8]. Note that for the specific case of ER networks, $\kappa - 1 = p\langle k \rangle$ and $\kappa_o - 1 = \langle k \rangle$, reducing Eq. (4) to Eq. (1). In the general case of random networks, the LPP transition is found by imposing $\delta = 1$, which yields

$$\tilde{p}_c(a) = (\kappa_o - 1)^{\frac{1-a}{a}}. \tag{5}$$

The scaling form for S_a is the same as Eq. (3) with δ taken 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, 19–22]. 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 for SF networks one obtains [8]

$$\kappa = \left(\frac{2-\lambda}{3-\lambda}\right) \frac{K^{3-\lambda} - m^{3-\lambda}}{K^{2-\lambda} - m^{2-\lambda}}. \tag{6}$$

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

For $2 < \lambda < 3$ the typical network length scales as $\ell = 2 \log\log N / |\log(\lambda - 2)|$ [23, 24]. 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 $\ell' = 2\log\log P_\infty N/|\log(\lambda-2)|$ [25]. Solving $a\ell_{ij} = \ell'_{ij}$ for $N \to \infty$, we obtain

$$a = \frac{\ell'_{ij}}{\ell_{ij}} = \frac{\log\log P_\infty N}{\log\log N} \to 1 \qquad \text{(Scale-free } 2 < \lambda < 3\text{)}. \tag{7}$$

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. 2(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 of q nodes with the 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 given by $K' = mq^{1/(1-\lambda)}$. For $\lambda > 3$, making the substitutions $q \to q'$ and $K \to K'$ in Eq. (4) we obtain the equation for $\tilde{p}_c$ [26] 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. 1(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 $\log P_\infty N/\log(\kappa-1)$ [23, 24]. Thus, after percolation S_a can be calculated using the tree approximation which yields

$$S_a \sim (\kappa-1)^{2a\frac{\log\log N}{|\log(\lambda-2)|}} = (\log N)^{2a\frac{\log(\kappa-1)}{|\log(\lambda-2)|}} \qquad \text{(Scale-free } 2 < \lambda < 3\text{, targeted removal)}. \tag{8}$$

In this case, the phase transition to a spanning cluster of order N cannot be achieved for any finite value of a and $p < 1$, as seen from Eq. (8). Simulation results supporting Eq. (8) are shown in Fig. 3(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 the usual percolation theory cannot correctly describe connectivity when only a limited set of path lengths are useful. In usual percolation, order N of the network nodes are connected when $p > p_c$. However, in LPP, when

$p_c < p < \tilde{p}_c$, only a zero fraction of the network is connected. 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 length between nodes to $a = 1.5$ times the original length. The theory of 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 the virus typically survives up to $1.5 \log N$ steps, our theory predicts that the immunization threshold is significantly smaller, 0.31 compared to 0.67. Due to the above considerations, we expect our results to be important for network design, routing protocols and immunization strategies.

We thank the Israel Science Foundation, Israel Internet Association, Yeshaya Horowitz Association and The Center for Complexity Science (Israel), the European NEST project DYSONET, DOE and ONR for financial support, and J. C. Miller, E. Perlsman, and S. Sreenivasan for fruitful discussions.

* These authors contributed equally to this work

[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, 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, K. Faust, D. Iacobucci, M. Granovetter, *Social Network Analysis: Methods and Applications* (Cambridge, 1994).

[3] R. Pastor-Satorras and A. Vespignani Phys. Rev. Lett. **86**, 3200-3203 (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, R. Cohen, E. López, Z. Toroczkai, H. E. Stanley, Phys. Rev. E (2007) (in press); cs.NI/0604023.

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

[8] R. Cohen, K. Erez, D. ben-Avraham, and S. Havlin, Phys. Rev. Lett. **85**, 4626 (2000). *ibid.* **86**, 3682 (2001).

[9] D. S. Callaway, M. E. J. Newman, S. H. Strogatz, and D. J. Watts, 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, 2nd ed.* (Taylor and Francis, 2004); *Fractals and Disordered Systems*, 2nd. ed., edited by A. Bunde and S. Havlin (Springer, Berlin 1996).

[12] P. Erdős and A. Rényi, Publ. Math. (Debrecen), **6**, 290 (1959); P. Erdős and A. Rényi, A. Publ. Math. Inst. Hung. Acad. Sci. **5**, 1760 (1960).

[13] B. Bollobás, *Random Graphs* (Academic Press, Orlando, 1985).

[14] L. A. Braunstein et al. Phys. Rev. Lett. **91**, 168701 (2003).

[15] For a tree of degree $z+1$ and depth ℓ, the size is given by $S_a = 1 + (z+1)\Sigma_{n=0}^{\ell-1} z^n \sim (\frac{z}{z-1}) z^{\ell}$. If $z = p\langle k\rangle$ the tree size scales as $[p\langle k\rangle/(p\langle k\rangle - 1)]\,[p\langle k\rangle]^{\ell}$.

[16] To construct an Erdős-Rényi network, we begin with N nodes and connect each pair with probability ϕ. To generate a scale-free network with N nodes, we use the Molloy-Reed algorithm [17], which allows for the construction of random networks with arbitrary degree distribution. We generate k_i copies of each node i, where the probability of having k_i satisfies $P(k_i) \sim k_i^{-\lambda}$. These copies of the nodes are then randomly paired in order to construct the network, making sure that two previously-linked nodes are not connected again, and also excluding links of a node to itself.

[17] M. Molloy and B. Reed, Random Struct. Algorithms **6**, 161 (1995).

[18] The size of S_a was measured by: i) generating R realizations of network structure, ii) performing percolation on each realization, iii) choosing M nodes at random from the largest cluster, iv) for each node measuring the number of nodes connected satisfying $\ell'_{ij} \leq a\ell_{ij}$, and v) averaging over the results. Typically, $R = 1000$ and $M = 100$.

[19] A.-L. Barabási and R. Albert, Science **286**, 509 (1999). H. A. Simon, Biometrika **42**, 425 (1955).

[20] L. A. N. Amaral, A. Scala, M. Barthélémy, and H. E. Stanley, Proc. Nat. Acad. Sci. USA **97** 11149 (2000).

[21] A. Barrat et al. Proc. Nat. Acad. Sci. USA, **101**, 3747 (2004).

Quantity	ER	SF random		SF targeted	
		$2<\lambda<3$	$\lambda>3$	$2<\lambda<3$	$\lambda>3$
$\tilde{p}_c$	$\langle k\rangle^{(1-a)/a}$	0	$(\kappa_o-1)^{(1-a)/a}$	1	$\tilde{p}_c(a,\kappa,\kappa_o)$[26]
S_a	N^δ	$N\ \ (p>\tilde{p}_c)$	N^δ	$(\log N)^\delta$	N^δ
δ	$a\left(1-\frac{\lvert\log p\rvert}{\log\langle k\rangle}\right)$	1	$a\left(1-\frac{\lvert\log p\rvert}{\log(\kappa_o-1)}\right)$	$2a\frac{\log(\kappa-1)}{\lvert\log(\lambda-2)\rvert}$	$a\frac{\log(\kappa-1)}{\log(\kappa_o-1)}$

TABLE I: The functions $\tilde{p}_c$, S_a and δ for several kinds of network structures under random and targeted 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 p is always above $\tilde{p}_c$.

[22] P. L. Krapivsky, S. Redner, and F. Leyvraz, Phys. Rev. Lett. **85**, 4629 (2000).

[23] R. Cohen and S. Havlin Phys. Rev. Lett. **90**, 058701 (2003).

[24] R. van der Hofstad, G. Hooghiemstra, D. Znamenski, math.PR/0502581 (2005).

[25] Indeed, for $\lambda>3$, $a\ell_{ij}=\ell'_{ij}$ reduces to Eq. (4).

[26] The function $\tilde{p}_c(\kappa,\kappa_o)$ is given implicitly by solving $\delta=1$ from Eq. (4), using κ from Eq. (6) with $q=q'$ and $K=K'$ for targeted removal.

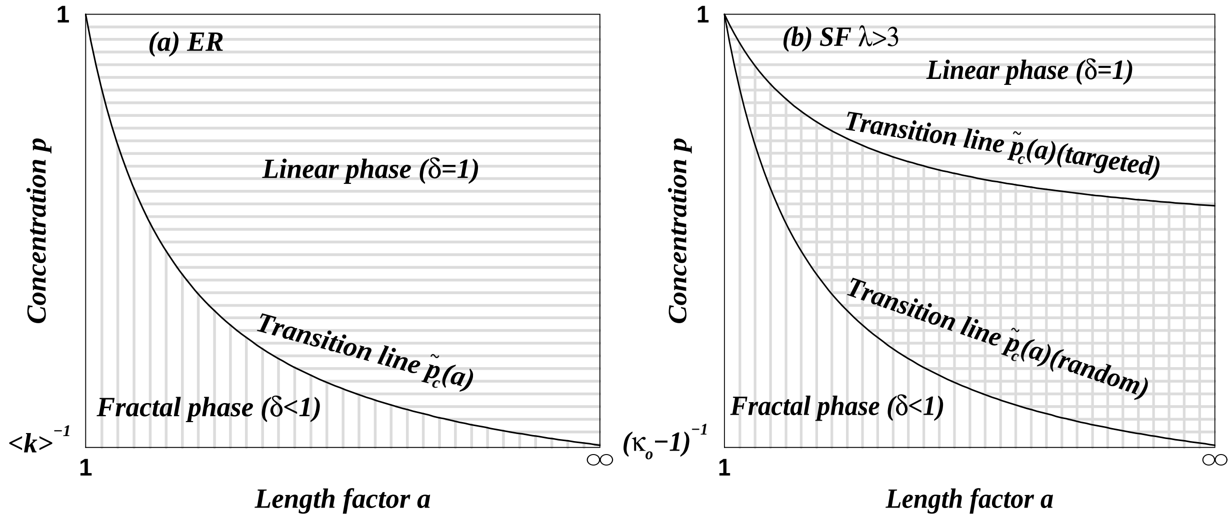

FIG. 1: (a) Phase diagram for Erdős-Rényi networks of LPP with respect to parameters a and p, demonstrating the linear and power law (fractal) phases for $S_a \sim N^{\delta}$. (b) Similar phase diagram for scale-free networks with $\lambda > 3$. The two transition lines represent networks with the same 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) and (b) the regular percolation threshold is fiven by the limit $a \to \infty$, i.e, $p_c = \langle k \rangle^{-1}$ for ER and $p_c = (\kappa_o - 1)^{-1}$ for SF with $\lambda > 3$.

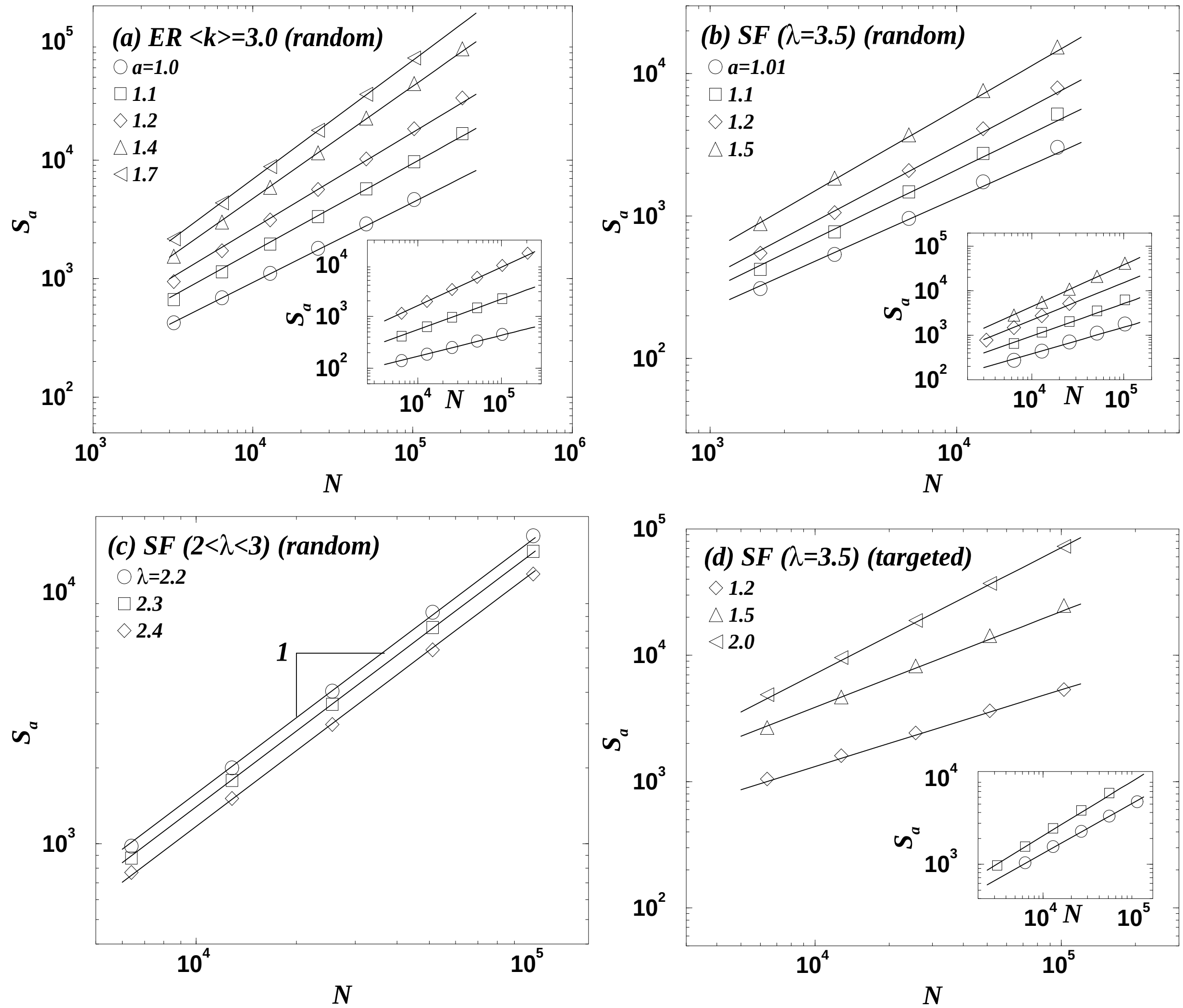

FIG. 2: Simulation results (symbols) for S_a vs. N for various network types under random or targeted removal and different values of a and p (indicated in plot legends), compared to the theoretically predicted power laws (solid lines), with δ calculated from Table I. Network sizes are typically between 1600 and 204800. In all plots the simulation results agree with the 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 $p = 0.5(\bigcirc), 0.6(\Box)$ and 0.7 ($\Diamond$). (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 $p = 0.5(\bigcirc), 0.6(\Box), 0.7(\Diamond)$ and 0.8 ($\triangle$). (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 $p = 0.92(\bigcirc)$ and 0.94 ($\Box$).

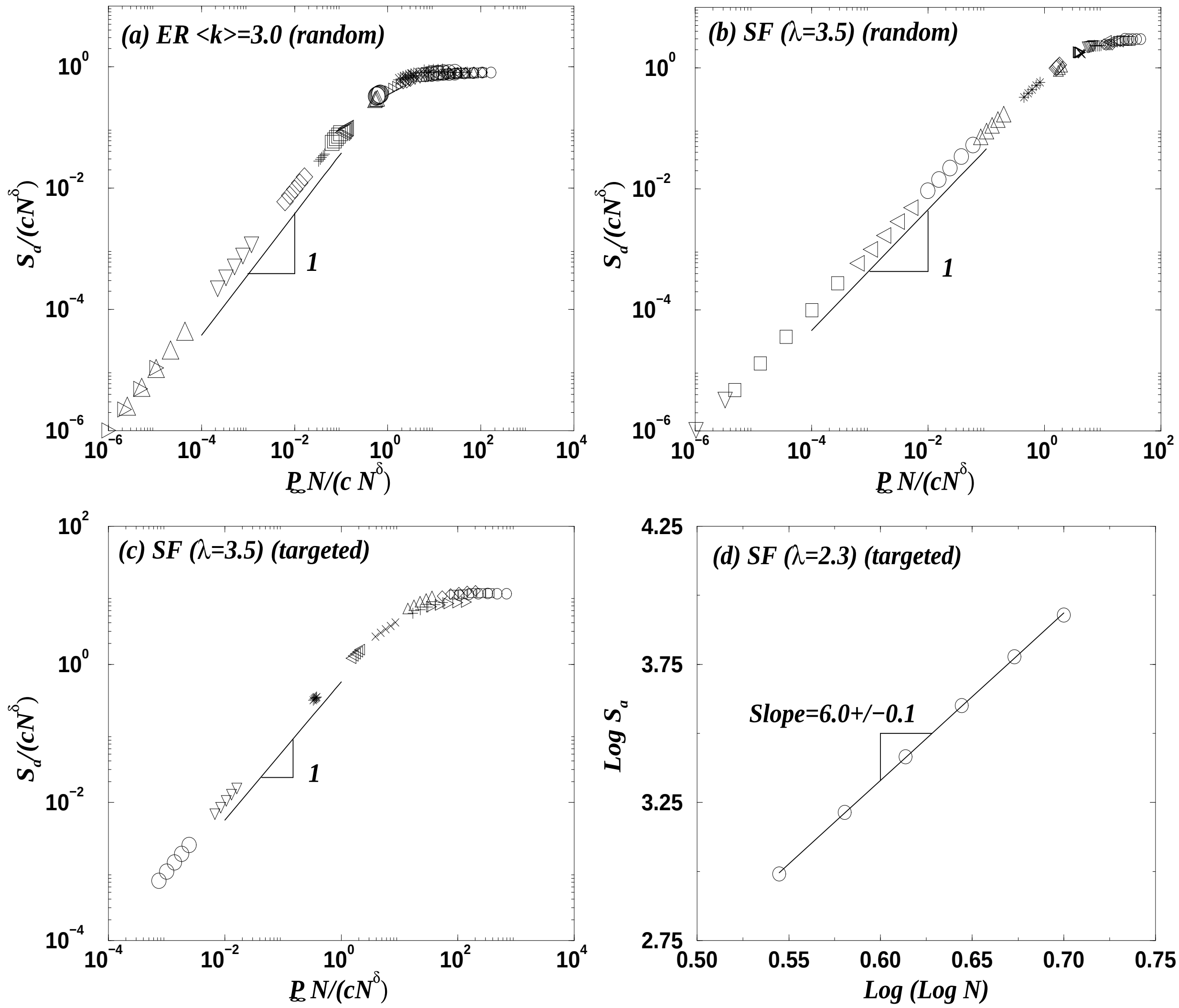

FIG. 3: Simulation results ((a) through (c)) for the scaling of $S_a/(c(p)N^\delta)$ vs. $P_\infty N/(c(p)N^\delta)$ for various types of networks, for random and targeted removal for N between 1600 and 25600, and a between 1.0 and 4. (a) Erdős-Rényi networks (random) with $\langle k\rangle = 3$ and $p = 0.5, 0.6$ and 0.7. (b) Scale-free networks (random) with $\lambda = 3.5$, $m = 2$ and $p = 0.6, 0.7$ and 0.8. (c) Scale-free 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 comparison to the theoretical prediction (line) for δ for SF networks (targeted) for $\lambda = 2.3$, $m = 3$, $a = 1.5$ and $p = 0.97$.