Generalized theory for node disruption in finite-size complex networks

Bivas Mitra,[1] Niloy Ganguly,[1] Sujoy Ghose,[1] and Fernando Peruani[2,3,*]
[1] *Department of Computer Science and Engineering, Indian Institute of Technology, Kharagpur 721302, India*
[2] *CEA-Service de Physique de l'Etat Condensé, Centre d'Etudes de Saclay, 91191 Gif-sur-Yvette, France*
[3] *Institut des Systémes Complexes de Paris Île-de-France, 57/59, rue Lhomond F-75005 Paris, France*

(Received 20 April 2008; published 21 August 2008)

After a failure or attack the structure of a complex network changes due to node removal. Here, we show that the degree distribution of the distorted network, under any node disturbances, can be easily computed through a simple formula. Based on this expression, we derive a general condition for the stability of noncorrelated finite complex networks under any arbitrary attack. We apply this formalism to derive an expression for the percolation threshold f_c under a general attack of the form $f_k \sim k^{\gamma}$, where f_k stands for the probability of a node of degree k of being removed during the attack. We show that f_c of a finite network of size N exhibits an additive correction which scales as N^{-1} with respect to the classical result for infinite networks.

DOI: 10.1103/PhysRevE.78.026115 PACS number(s): 89.75.Hc, 02.50.Cw, 89.75.Fb

INTRODUCTION

The stability of graphs against various disrupting events is a central issue in the study of complex networks [1–10]. If information is transported across a network, as is the case of epidemics across social networks or information broadcast through the Internet, the "damage" of some nodes can dramatically affect the dynamics of the system. In the context of disease spreading, this could lead to an epidemic extinction, while in communications to a halt of information broadcast [11–15]. Both the topological structure of the network and the nature of the attack determine the resulting effect [16]. For example, it has been shown that scale-free (SF) networks display a high degree of tolerance against random failures [9], while, on the other hand, they are quite sensitive to intentional attacks [10]. Clearly, there are various strategies to perform an intentional attack, and each one of them requires a different level of knowledge on the network topology [16–18]. A rather general attack, proposed in [19,20] and which we will use in this paper, takes the form $f_k \sim k^{\gamma}$, where f_k denotes the probability of a node of degree k of being removed during the attack, while γ is associated with the degree of knowledge of the attacker. The analysis of this attack has revealed that in SF networks an increase of γ leads to a decrease of the critical fraction of nodes that must be removed to disintegrate the network—i.e., a decrease in the percolation threshold f_c [19,20].

Though many results have been derived for infinite networks, very little is known about the stability of finite networks. Typical examples of small-size finite networks are *ad hoc* networks of commercial mobile devices, frequently used for communication [21], temporary peer-to-peer networks formed by BitTorrent clients for efficient download of files [22], and networks of autonomous mobile robots [23]. The operation of these systems relies on the robustness of the highly dynamical underlying network. Thus, a good understanding of the stability of these small-size networks is imperative for these applications. Moreover, we can say that in general a comprehensive theory for the stability of arbitrary finite networks under any node disturbance is still lacking.

In this paper we attempt to shed some light on this matter by proposing an alternative derivation for the percolation threshold [24]. In our approach, instead of applying a generating function formalism to find an analytic expression for the percolation threshold as in [1,2,17], we used the fact that during an attack the degree distribution of the network changes (see Fig. 1). We show that the degree distribution of the distorted (uncorrelated) network, under any node disturbances, can be easily computed through a simple formula. Based on this expression, we derive a general condition for the stability of noncorrelated complex networks under any arbitrary attack. This condition applied to the study of network stability under the general attack proposed in [19,20] leads us to a general expression for the percolation threshold f_c. We show that f_c of a finite network of size N exhibits an additive correction which scales as N^{-1} with respect to the classical result for infinite networks [1,9,10]. Simulation results confirm all these findings.

NETWORK TOPOLOGY AFTER DISTURBANCE

A failure or attack can be thought of in the following way. Let p_k be the degree distribution of the network before the

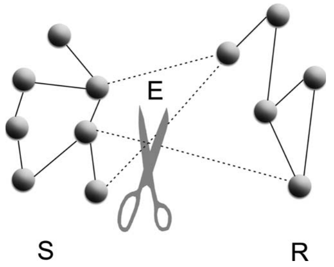

FIG. 1. The scheme illustrates an attack as consisting of two steps: (a) selection of nodes to be removed (see set R) and (b) cutting of the edges that run from the surviving nodes (represented by set S) to the set of removed nodes. As the scheme shows, the attack affects also the degree of the surviving nodes.

[*] Corresponding author: fernando.peruani@cea.fr

attack. The first step in the attack is to select the nodes that are going to be removed. Let us assume that this is performed by means of f_k, where f_k represents the probability for a node of degree k of being removed from the network. Note that the only restriction on f_k is $0 \leqslant f_k \leqslant 1$. After the node selection, we divide the network into two subsets, one subset contains the surviving nodes (S) while the other subset comprises the nodes that are going to be removed (R). Figure 1 illustrates this procedure. At the moment the nodes in R are actually removed, the degree distribution of the S nodes is changed due to the removal of the E edges that run between these two subsets. The probability ϕ of finding an edge in subset S that is connected to a node in subset R is expressed as

$$\phi = \frac{\Sigma_{i=0}^{\infty} i p_i f_i}{(\Sigma_{k=0}^{\infty} k p_k) - 1/N}. \tag{1}$$

The reasoning behind this expression is as follows. The total number of half-edges in the surviving subset, including the E links that are going to be removed, is $\Sigma_{j=0}^{\infty} j(Np_j)(1-f_j)$. The probability for a randomly chosen half-edge of being removed is simply $\Sigma_{i=0}^{\infty} i(Np_i)f_i/[\Sigma_{k=0}^{\infty} k(Np_k)-1]$. E is the number of half-edges in S times this probability, and ϕ is obtained by dividing E by the number of half-edges in the subset S. Notice that the removal of nodes can only lead to a decrease of the degree of a node. Finally, to calculate the degree distribution p'_k after the attack, we still need to estimate the probability p_q^s of finding a nodes with degree q in the surviving subset S (before cutting the E edges). This fraction takes the simple form

$$p_q^s = \frac{(1-f_q)p_q}{1-\Sigma_{i=0}^{\infty} p_i f_i}. \tag{2}$$

Now we are in condition to compute p'_k. Using Eqs. (1) and (2), we obtain the following expression for p'_k:

$$p'_k = \sum_{q=k}^{\infty} \binom{q}{k} \phi^{q-k}(1-\phi)^k p_q^s. \tag{3}$$

Equation (3) can be iteratively evaluated by replacing p_k with p'_k into Eqs. (1)–(3). It is instructive to notice that for failure—i.e., $f_k = f$—and assuming $N \gg 1$, Eq. (1) reduces to $\phi = f$, while Eq. (2) becomes $p_q^s = p_q$. In consequence, from Eq. (3) we retrieve the degree distribution p'_k after failure which reads [9] $p'_k = \Sigma_{q=k}^{\infty} \binom{q}{k} f^{q-k}(1-f)^k p_q$. A similar expression has also been used to described p'_k after an *ad hoc* attack in SF networks with $N \gg 1$ [10].

Figure 2 shows a comparison between stochastic simulations (symbols) and Eq. (3) (solid line) for two different network topologies: namely, (a) Erdős-Rényi (ER) graphs and (b) SF networks (b).[1] Removal of nodes (and edges) was performed through an attack of the form $f_k \sim k^{\gamma}$, with $\gamma = 1$. In the figure two different system sizes are shown: $N = 10^5$ and $N = 50$ (figure insets).

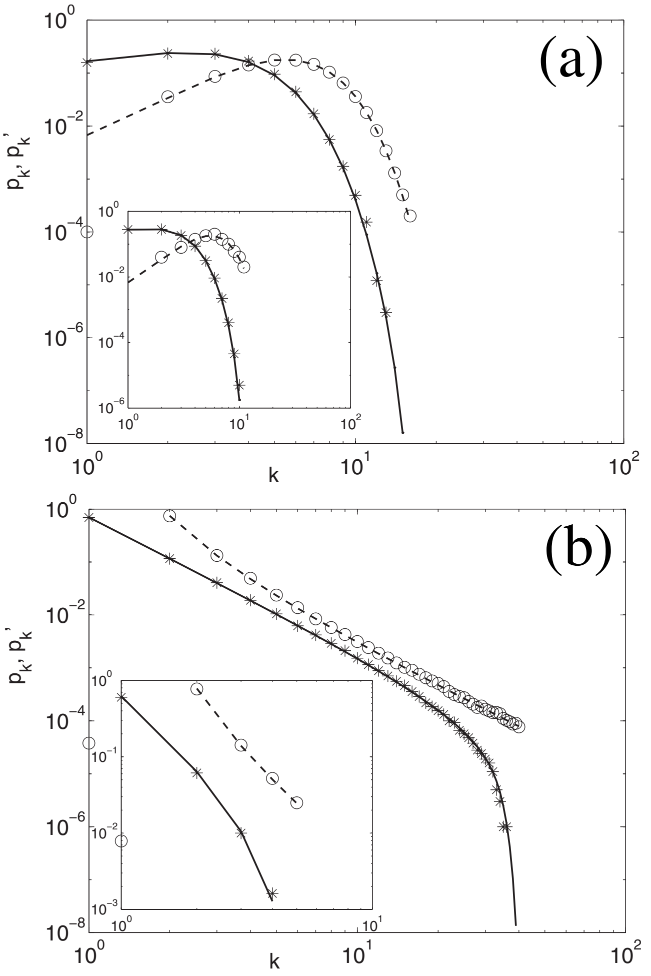

FIG. 2. Degree distribution before (circles) and after (stars) the attack. In (a) the network topology corresponds to Erdős-Rényi graphs, with $\langle k \rangle = 5$, while in (b) to SF networks, $p_k \sim k^{-\alpha}$, with $\alpha = 2.5$.[1] Symbols correspond to simulations, solid lines to the theoretical p'_k, given by Eq. (3), and dashed curves to the (theoretical) initial p_k. In (a) and (b) the main figure corresponds to networks with $N = 10^5$ nodes, while the inset shows the result for small networks with just $N = 50$ nodes.

CRITICAL CONDITION FOR FINITE NETWORKS

The following expression tells us whether an infinite network percolates after an attack [9]:

$$\kappa' = \frac{\langle k^2 \rangle'}{\langle k \rangle'} > 2, \tag{4}$$

where $\langle k \rangle'$ and $\langle k^2 \rangle'$ refer to the first and second moments of the degree distribution after the attack. We borrow the critical condition for infinite networks given by Eq. (4) to define a "percolation" criterion for finite networks. Thus, by definition we assume that the condition $\kappa' = 2$ determines the point at which the network breaks down [19,20]. To compute $\langle k \rangle'$ and $\langle k^2 \rangle'$, we utilize the generating function

[1] The cutoff degree k_M was calculated according to $Np_k \leqslant 1$, where $k > k_M$ and p_k is an initial power-law distribution which is later normalized in the interval $[1, k_M]$. The assignment of edges was performed using the *matching algorithm* and *switching method.*

$$G_0(x) = \sum_{k=0}^{\infty}\sum_{q=k}^{\infty}\binom{q}{k}\phi^{q-k}(1-\phi)^k p_q^s x^k. \qquad (5)$$

After exchanging the order of the sum, the binomial theorem can be applied, and we obtain

$$G_0(x) = \sum_{k=0}^{\infty} p_q^s[(x-1)(1-\phi)+1]^q. \qquad (6)$$

From Eq. (6), the first two moments can be easily computed as $\langle k\rangle' = dG_0(1)/dx$ and $\langle k^2\rangle' = d^2G_0(1)/dx^2 + dG_0(1)/dx$. After some algebra we obtain that the critical condition given by Eq. (4) takes the form

$$\left(\sum_k p_k(1-f_k)k\right)\left(\sum_k p_k(1-f_k)k^2 + \sum_k p_k(f_k-2)k\right) + \frac{1}{N}\left(\sum_k p_k(1-f_k)(2-k)k\right) = 0. \qquad (7)$$

Equation (7) determines the stability condition (according to the given definition) for any uncorrelated network of finite size under any arbitrary attack. In the limit of $N\to\infty$, Eq. (7) reduces to

$$\sum_{k=0}^{\infty} p_k k[k(1-f_k) + f_k - 2] = 0. \qquad (8)$$

Interestingly, Eq. (8) can be also derived through a more classical generating function formalism [25].

GENERAL EXPRESSION FOR THE PERCOLATION THRESHOLD

In the following, we model various dynamics (attacks) through a generalized equation of the form: $f_k = Ck^{\gamma}$, where γ is a real number signifying the amount of network structure information available to the attacker to breakdown the network [20] and C is a constant that we refer to as attack intensity. Clearly, $\gamma > 0$ represents a situation in which high-degree nodes are removed with higher probability, while $\gamma < 0$ models the opposite. The last case is suitable to situations in which low-degree nodes are more prone to fail. We are interested in knowing, for a given γ, the critical fraction f_c of nodes that is required to remove through such an attack in order to destroy the network—i.e., the percolation threshold. Thus, the problem reduces to compute, for a given γ, the critical attack intensity C^*. Replacing in Eq. (7) the above definition of f_k and after some algebra we obtain

$$C^* = \frac{1}{2q\langle k^{\gamma+1}\rangle}\big([2\langle k\rangle q - Q - (1/N)(q+\langle k^{\gamma+1}\rangle)] - \{Q^2 + (1/N)[2Q + (q+\langle k^{\gamma+1}\rangle)^2] - 4qp\langle k^{\gamma+1}\rangle\}^{1/2}\big), \qquad (9)$$

where $p = \langle k^2\rangle - 2\langle k\rangle$, $q = \langle k^{\gamma+1}\rangle - \langle k^{\gamma+2}\rangle$, $Q = \langle k^{\gamma+1}\rangle p + \langle k\rangle q$, and $\langle k^{\omega}\rangle$ is defined as $\langle k^{\omega}\rangle = \Sigma_k k^{\omega} p_k$. Since the fraction of removed nodes, f, after an attack is $f = \Sigma_k p_k f_k$, the expression for the percolation threshold f_c is simply

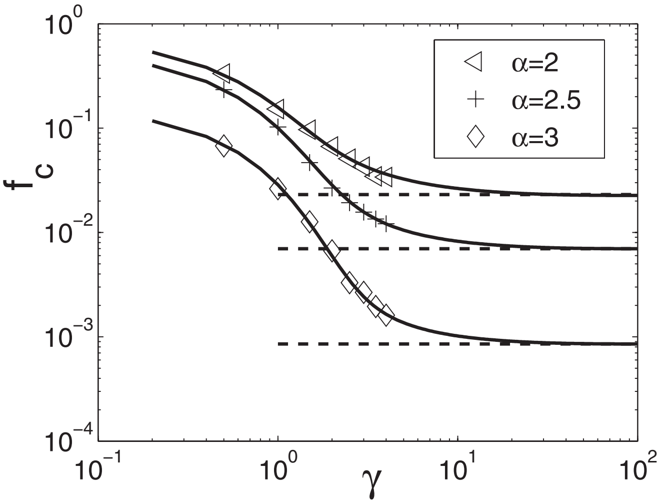

FIG. 3. Percolation threshold f_c under changes of the attack exponent γ for three different SF networks, $p_k \sim k^{-\alpha}$, with $\alpha=2$, 2.5, and 3 and $N=10^5$. Symbols correspond to stochastic simulations, while solid curves correspond to Eqs. (9) and (10). The horizontal dashed lines indicate the asymptotic value of f_c given by Eq. (11).

$$f_c = C^*\langle k^{\gamma}\rangle. \qquad (10)$$

Figure 3 illustrates the behavior of f_c on three SF networks ($\alpha=2$, 2.5, and 3) upon changes in the attack exponent γ. The symbols correspond to stochastic simulations performed on networks of size $N=10^5$, while the black curves refer to Eqs. (9) and (10). In the numerical experiments we have computed f_c following [20]: when the fraction of removed nodes is f_c, the probability F of finding the network with $\kappa' > 2$ is 1/2.

It is interesting to observe that for any SF network, the minimum fraction Φ_c of nodes that is required to be removed to break down the network is obtained by taking the limit $\gamma\to\infty$ of Eq. (10):

$$\Phi_c = \lim_{\gamma\to\infty} f_c(\gamma,\alpha) = h(\alpha)\frac{1}{k_M(k_M-1)}, \qquad (11)$$

where $h(\alpha)$ is $h(\alpha) = \langle k^2\rangle - 2\langle k\rangle$ and k_M is the maximum degree of the original network. Notice that Eq. (11) represents an attack performed having full knowledge of the network topology. The asymptotic values corresponding to Eq. (11) are shown in Fig. 3 as horizontal dashed lines. Figure 3 indicates that typically an increase of the information about the network topology, respectively γ, helps the attacker to break down the network with the removal of a smaller number of nodes. However, information becomes redundant as the asymptotic value Φ_c is approached.

EFFECT OF FINITE NETWORK SIZE

To illustrate the effect of network size N upon the percolation threshold $f_c(N)$, we customize Eq. (9) for random attack or failure. When $\gamma=0$ and f_k is independent of k, we obtain

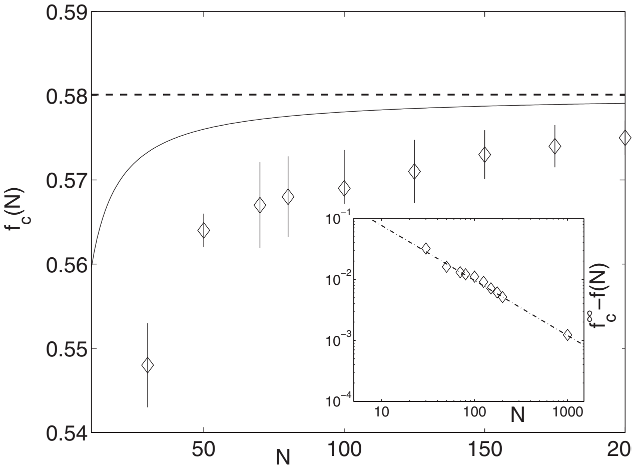

FIG. 4. Percolation threshold f_c as function of N. Symbols correspond to stochastic simulation, solid curve to Eq. (12), and dashed curve to the classical result f_c^∞. The error bars represent the confidence interval that contains $f_c^\infty(N)$. Simulations were performed with ER networks with average degree $\langle k \rangle = 3$ and maximum degree $k_M = 5$. The inset shows the scaling of $f_c^\infty - f_c(N)$ with respect to N. The best fitting of the data corresponds to a slope -0.92 ± 0.05 (dashed line).

$$f_c(N) = f_c^\infty + \frac{1}{N}\left(\frac{2 - (\langle k^2 \rangle / \langle k \rangle)}{\langle k^2 \rangle - \langle k \rangle}\right), \tag{12}$$

where f_c^∞ is the well-known percolation threshold for infinite networks under failure [1,9], which reads $f_c^\infty = 1 - [1/(\langle k^2 \rangle / \langle k \rangle - 1)]$.

Figure 4 shows a comparison between Eq. (12) (solid line), f_c^∞ (dashed line), and stochastic simulations (symbols) for ER networks of different sizes. Notice that Eq. (12) predicts the correct scaling of f_c with N—i.e., $f_c(N) - F_c^\infty \sim N^{-1}$. The observed deviation between Eq. (12) and simulations can be arguably attributed to correlations effects, which have been ignored in the current approach.

SUMMARY

We have proposed a general procedure to calculate the distorted degree distribution of uncorrelated finite-size networks under arbitrary failure and attack. Using the expression for the distorted degree distribution we have derived the critical condition for the stability of finite-size networks. The formalism has been further applied to derive an expression for the percolation threshold under a general attack. Finally, it was shown that the obtained percolation threshold predicts an additive correction which scales as N^{-1} with respect to the classical result for infinite networks, as observed in simulations.

The results derived throughout this paper are valid only for uncorrelated networks. The effect of correlations on the percolation threshold for finite (and infinite) networks remains as one of the major challenges. Further extensions of this theory will be focused in that direction.

ACKNOWLEDGMENTS

The authors thank H. Chaté, F. Ginelli, L. Brusch, and A. Deutsch for insightful comments on the manuscript. F.P. acknowledges the hospitality of IIT-Kharagpur. This work was partially financed by Indo-German collaboration project PZP/2006/BMBF.

[1] D. S. Callaway, M. E. J. Newman, S. H. Strogatz, and D. J. Watts, Phys. Rev. Lett. **85**, 5468 (2000).

[2] M. E. J. Newman, S. H. Strogatz, and D. J. Watts, Phys. Rev. E **64**, 026118 (2001).

[3] R. Albert, H. Jhong, and A. L. Barabási, Nature (London) **406**, 378 (2000).

[4] J. L. Guillaume, M. Latapy, and C. Magnien, Lect. Notes Comput. Sci. **3544**, 186 (2005).

[5] A. X. C. N. Valente, A. Sarkar, and H. A. Stone, Phys. Rev. Lett. **92**, 118702 (2004).

[6] T. Tanizawa, G. Paul, R. Cohen, S. Havlin, and H. E. Stanley, Phys. Rev. E **71**, 047101 (2005).

[7] J. G. Liu, Z. T. Wang, and Y. Z. Dang, Mod. Phys. Lett. B **19**, 785 (2005).

[8] A. E. Motter, Phys. Rev. Lett. **93**, 098701 (2004).

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

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

[11] R. Pastor-Satorras and A. Vespignani, Phys. Rev. E **65**, 035108(R) (2002); **65**, 036104 (2002).

[12] V. Colizza and A. Vespignani, Phys. Rev. Lett. **99**, 148701 (2007).

[13] R. Cohen, S. Havlin, and D. ben-Avraham, Phys. Rev. Lett. **91**, 247901 (2003).

[14] A. Trusina, M. Rosvall, and K. Sneppen, Phys. Rev. Lett. **94**, 238701 (2005).

[15] Y. Xia and D. J. Hill, IEEE Trans. Circuits Syst., II: Express Briefs **55**, 65 (2008).

[16] P. Holme, B. J. Kim, C. N. Yoon, and S. K. Han, Phys. Rev. E **65**, 056109 (2002).

[17] J. Wu *et al.*, J. Phys. A **40**, 2665 (2007).

[18] G. Paul, R. Cohen, S. Sreenivasan, S. Havlin, and H. E. Stanley, Phys. Rev. Lett. **99**, 115701 (2007).

[19] L. K. Gallos *et al.*, Physica A **344**, 504 (2004).

[20] L. K. Gallos, R. Cohen, P. Argyrakis, A. Bunde, and S. Havlin, Phys. Rev. Lett. **94**, 188701 (2005).

[21] M. Grossglauser and D. N. C. Tse, IEEE/ACM Trans. Netw. **10**, 477 (2002).

[22] D. Qiu and R. Srikant, Comput. Commun. Rev. **34**, 367 (2004).

[23] W. F. W. Othman *et al.*, in *Proceedings of EUROCON 2007*, edited by M. P. Kazmierkowski and J. Modelski (IEEE, Washington, D.C., 2007).

[24] Let us clarify that in this paper we are interested in understanding only the emergence of a giant component at percolation threshold. Hence, the terms *percolation* and *emergence of giant component* can be considered synonymous.

[25] B. Mitra *et al.*, in *Proceedings of the 14th ACM Conference on Computer and Communications Security*, edited by S. De Capitani di Vimercati, P. Syverson, and D. Evans (ACM, New York, 2007).