PHYSICAL REVIEW E **67**, 015101(R) (2003)

Resilience to damage of graphs with degree correlations

Alexei Vázquez[1] and Yamir Moreno[2]
[1]*INFN and International School for Advanced Studies, Via Beirut 4, 34014 Trieste, Italy*
[2]*The Abdus Salam International Centre for Theoretical Physics, Condensed Matter Group, P.O. Box 586, Trieste I-34014, Italy*
(Received 5 September 2002; published 6 January 2003)

The existence or nonexistence of a percolation threshold on power law correlated graphs is a fundamental question for which a general criterion is lacking. In this work we investigate the problems of site and bond percolation on graphs with degree correlations and their connection with spreading phenomena. We obtain some general expressions that allow the computation of the transition thresholds or their bounds. Using these results we study the effects of assortative and disassortative correlations on the resilience to damage of networks.

DOI: 10.1103/PhysRevE.67.015101 PACS number(s): 89.75.Hc, 05.20.−y, 89.75.−k, 02.50.−r

The graphs representing many real networks are characterized by power law degree distributions [1]. The origin of these power laws can be traced back to the growing nature of real networks and to some effective preferential attachment mechanism. This later mechanism implies that when new vertices are added to the graph they are more likely linked to existing vertices with large degrees [1,2]. Recently, there has been a great interest in the study of processes running on top of these graphs due to their social, technological, and scientific relevance. Percolation processes [3,4], spreading phenomena [5–7], the Ising model [8], and searching techniques [9] are some examples for which analytical solutions have been found in random graphs with the only constraint given by the degree distribution. One of the fundamental results is that the threshold characterizing the percolation transition or an epidemic outbreak, depends on the ratio $\langle d^2\rangle/\langle d\rangle$ of the first two moments of the degree distribution [3–5,8]. Hence, if $\langle d^2\rangle$ diverges when increasing the graph size, there is no transition in the thermodynamic limit.

The topology of real networks is also characterized by degree correlations [10,11] and, therefore, the extension of previous results for uncorrelated graphs is of utmost importance. Moreover, it has been shown that growing network models with [12] and without [13] preferential attachment lead to nontrivial degree correlations. The study of models on graphs with degree correlations is quite recent [11,14–16]. Some expressions for the size of the giant component and related quantities have been obtained in Ref. [11], whereas an equation for the epidemic threshold has been provided in Ref. [14]. General statistical mechanics approaches for models on correlated graphs has also been developed in Refs. [15,16]. However, in contrast to the case of uncorrelated graphs, a general criterion for the existence or nonexistence of a transition threshold has not been proposed yet. A first step in this direction has been taken in Ref. [17] for a disease spreading model.

In this paper, we study the resilience to damage (vertex or edge removal) of random graphs with arbitrary degree distributions and correlations by addressing the problem of dilute (site or bond) percolation on these graphs. We report a general equation for the threshold and bound it. Besides, we analyze the effect of correlations considering some examples of uncorrelated, assortative, and disassortative correlated graphs or their mixture. We conclude that assortative correlations can make graphs quite robust, even with a finite $\langle d^2\rangle$. On the contrary, disassortative correlations can make graphs fragile, even with a divergent $\langle d^2\rangle$.

Let us start by considering the set of undirected graphs with N vertices and arbitrary degree distribution p_d. Following one end of a randomly chosen edge, we will find a vertex of degree d with probability $q_d = dp_d/\langle d\rangle$. We further assume correlations between adjacent vertices. The conditional probability $p(d'|d)$ that a vertex of degree d' is reached following any edge coming from a vertex of degree d explicitly depends on both d and d'. Consistency with the degree distribution requires $\Sigma_{d'} p(d'|d) = 1$. Besides, the joint probability $p(d'|d)q_d$ that the two vertices at either end of a randomly chosen edge have degrees d and d' must be symmetric. For uncorrelated networks $p(d'|d) = q_{d'}$ that is independent of d.

The problem of percolation on graphs with degree correlations has been recently studied [11] using the generating function formalism. Alternatively, one can use a more general statistical mechanics approach [16]. In this case the size of the giant component, the fraction of nodes in the largest cluster, is given by

$$S = 1 - \sum_d p_d (u_d)^d, \tag{1}$$

$$u_d = \sum_{d'} p(d'|d)(u_{d'})^{d'-1}, \tag{2}$$

where u_d is the average probability that an edge connected to a vertex of degree d leads to another vertex that does not belong to the giant component [11].

Let us generalize this result to the site percolation problem. In this case a fraction f of the nodes is removed from the graph and the new giant component is computed. Since the node removal is independent of the node degree that is equivalent to replace the original degree distribution and correlations by: (i) the probability that a node selected at random has degree d and it has not been removed, and (ii) the probability that if we select a node at random and follow one of its edges we end in a node with degree d' that has not been removed, i.e.,

$$p_d \to (1-f)p_d, \quad p(d'|d) \to (1-f)p(d'|d). \tag{3}$$

Substituting Eqs. (3) in Eqs. (1) and (2) we get

$$S = 1 - f - (1-f)\sum_d p_d (u_d)^d, \tag{4}$$

$$u_d = f + (1-f)\sum_{d'} p(d'|d)(u_{d'})^{d'-1}, \tag{5}$$

where the term $-f$ (f) in Eq. (4) [Eq. (5)] gives the probability of hitting a removed node. One solution to these equations is $u_d = 1$ yielding $S = 0$. This solution is valid whenever the equation for the u_d is stable under successive approximations. That is, if we start with $u_d(n) = 1 - \rho_d(n)$ and compute the successive approximation $\rho_d(n+1)$ then we should obtain that $\rho_d(n) \to 0$ in the limit $t \to \infty$. For $\rho_d(n) \ll 1$ the last equation is approximated by the linear map

$$\rho_d(n+1) = \sum_{d'} L_{dd'} \rho_{d'}(n), \tag{6}$$

with

$$L_{dd'} = (1-f)C_{dd'}, \quad C_{dd'} = (d'-1)p(d'|d). \tag{7}$$

The stability of the solution $u_d = 1$ is then related to the largest eigenvalue of $L_{dd'}$. If it is smaller (larger) than 1 the solution is stable (unstable). Since $L_{dd'}$ is linear in f the stability condition can be written as

$$f > f_c, \quad (1-f_c)\Lambda_{max} = 1, \tag{8}$$

where Λ_{max} is the largest eigenvalue of $C_{dd'}$ provided that $\Lambda_{max} > 1$. If $\Lambda_{max} < 1$ the graph does not have a giant component even for $f = 0$. Moreover, since $C_{dd'}$ is a positive matrix then Λ_{max} has the lower and upper bounds $\min_d \Sigma_{d'} C_{dd'}$ and $\max_d \Sigma_{d'} C_{dd'}$, yielding

$$\min_d \langle d \rangle_{nn}(d) \leqslant 1 + \Lambda_{max} \leqslant \max_d \langle d \rangle_{nn}(d), \tag{9}$$

where

$$\langle d \rangle_{nn}(d) = \sum_{d'} p(d'|d) d' \tag{10}$$

is the average degree among the neighbors of a node with degree d [10]. Equation (9) can be used to determine, based on a simple topological measure, whether or not a given graph is robust under vertex removal.

In the bond percolation problem, a fraction f of the edges is removed from the graph and the new giant component is computed. Since the edge removal is made at random, this is equivalent to keep the original degree distribution and replace the degree correlations by the probability that if we select a node at random and follow one of its edges, given it has not been removed, we end in a node with degree d', i.e.,

$$p_d \to p_d, \quad p(d'|d) \to (1-f)p(d'|d). \tag{11}$$

Substitution of Eq. (11) in Eqs. (1) and (2) yields

$$S = 1 - \sum_d p_d (u_d)^d, \tag{12}$$

$$u_d = f + (1-f)\sum_{d'} p(d'|d)(u_{d'})^{d'-1}. \tag{13}$$

Note that the only difference between the site and bond percolation problems [see Eqs. (4) and (5)] is the equation for the giant component while that for u_d is identical. Hence, Eqs. (8) and (9) are also valid for the bond percolation problem.

In what follows we consider some particular graphs in order to analyze the effects of correlations. Depending on the monotony of $\langle d \rangle_{nn}(d)$ the degree correlations can be classified in: uncorrelated if it is independent of d, assortative or positive if it increases with increasing d, and disassortative or negative if it decreases with decreasing d. A similar definition has been introduced in Ref. [11] using a correlation coefficient.

For random graphs with no constraint other than the one imposed by the degree distribution we have $p(d',d) = q_{d'}$. In this case, the lower and upper bounds in Eq. (9) are equal giving for the largest eigenvalue

$$\Lambda_{max}^{unco} = \frac{\langle d^2 \rangle}{\langle d \rangle} - 2. \tag{14}$$

Alternatively, one can compute Λ directly from the eigenvalue problem of $C_{dd'}$. Then from Eq. (8) we obtain $1 - f_c = 1/(\langle d^2 \rangle / \langle d \rangle - 2)$ [4]. Hence, if the second moment $\langle d^2 \rangle$ diverges the threshold equals 1, i.e., the network is robust under random vertex or edge removal. Furthermore, consider the case in which the degree correlations can be decomposed into two components

$$p(d'|d) = \alpha q_{d'} + (1-\alpha)\delta p(d'|d) \tag{15}$$

with $0 < \alpha < 1$ and $\delta p(d'|d) > 0$ for all (d,d'). Varying the parameter α, one interpolates between the uncorrelated graphs ($\alpha = 1$) and a graph with arbitrary degree correlations given by $\delta p(d'|d)$. In this case from Eq. (9) we obtain $\Lambda_{max} \geqslant \alpha \Lambda_{max}^{unco}$ and, therefore, if the network is robust for the uncorrelated case it will also be robust for any $\alpha > 0$. This immediately implies that any graph with a divergent second moment and a finite amount of random mixing of edges does not have a percolation threshold.

Assortative correlations allow us to show that the divergence of the second moment is not a necessary condition for the absence of the threshold. Let us consider a network with degree correlations

$$p(d'|d) = \alpha \delta_{dd'} + (1-\alpha)\delta p(d'|d), \tag{16}$$

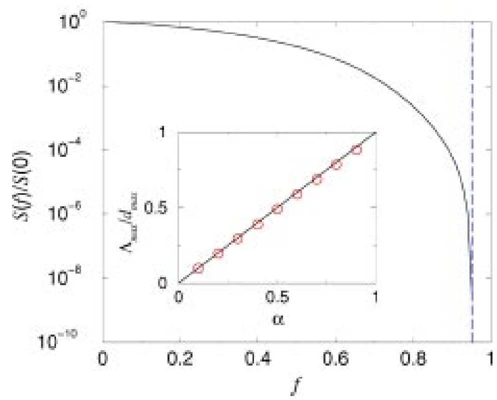

FIG. 1. Size of the giant component for a graph with $p_d = cd^{-3.5}$ ($2 \leqslant d \leqslant d_{max}$, $d_{max}=100$) and degree correlations $p(d'|d) = \alpha\delta_{dd'} + (1-\alpha)q_{d'}$ with α, as computed from Eq. (4). The dashed line marks the percolation threshold obtained using perturbation theory [Eq. (18)]. The inset shows the largest eigenvalue relative to d_{max} as a function of α. The points were computed numerically and the line is the perturbation theory dependency $\Lambda_{max}/d_{max} = \alpha$.

with $0<\alpha<1$ and $\delta p(d'|d)>0$ for all (d,d'). $\alpha=1$ corresponds to a fully assortative graph made up of subgraphs with fixed degree. In this case, $C_{dd'} = d'\delta_{dd'}$ [see Eq. (7)] is already diagonal. The largest eigenvalue is $\Lambda_{max} = d_{max}$, where d_{max} is the largest degree. If d_{max} diverges for $N \to \infty$ then $f_c = 1$. For the more general case $0<\alpha<1$ we compute the largest eigenvalue using perturbation theory [18] around $\alpha=1$, obtaining

$$\Lambda_{max}(\alpha) = \alpha d_{max} + (1-\alpha) C_{d_{max}d_{max}}. \tag{17}$$

This result is valid whenever $(1-\alpha)C_{d_{max}d_{max}} \ll \alpha d_{max}$. In general $C_{d_{max}d_{max}}$ decreases with increasing d_{max}, resulting

$$\Lambda_{max}(\alpha) \approx \alpha d_{max}, \tag{18}$$

for $d_{max} \gg 1/\alpha$. Hence, for any $\alpha>0$ and any unbounded degree distribution we have $f_c=1$, i.e., there is no percolation threshold. In Fig. 1 we show the validity of the perturbation theory for a particular perturbation $\delta p(d'|d)$. Thus, as in the fully assortative case, if $\alpha>0$ and d_{max} diverges $f_c=1$. Therefore, we can conclude that the divergence of the second moment is not a necessary condition.

Let us now analyze if the divergence of the second moment is a sufficient condition for $f_c=1$, using an example of a disassortative graph. Consider a vertex with degree $d>1$ and an edge incident to it. Then with probability g_d a vertex at the other end is chosen at random among all vertices with degree $d'>1$, otherwise it is connected to a vertex with $d'=1$ chosen at random, i.e.,

$$p(d'|d) = \frac{(1-g_{d'})d'p_{d'}}{\Sigma_s(1-g_s)sp_s}\Theta(d'-1)\delta_{d,1} + (1-g_d)\delta_{d',1} \times \Theta(d-1) + g_d \frac{g_{d'}d'p_{d'}}{\Sigma_s g_s s p_s}\Theta(d'-1)\Theta(d-1), \tag{19}$$

where $\Theta(x)$ is the unitary step function [$\Theta(x)=0$ for $x \leqslant 0$ and $\Theta(x)=1$ for $x>0$]. Moreover, the fraction of nodes with degree 1 is obtained self-consistently from the condition $p_1 = \Sigma_{d>1}(1-g_d)dp_d$. The average degree of the neighbors of a node with $d>1$ is given by

$$\langle d \rangle_{nn} = 1 + g_d\left(\frac{\Sigma_{d'>1} g_{d'} d'^2 p_{d'}}{\Sigma_s g_s s p_s} - 1\right), \tag{20}$$

and, therefore, these graphs are disassortative for any monotonic decreasing function g_d. To analyze the percolation properties of this graph we computed exactly the largest eigenvalue of $C_{dd'} = (d'-1)p(d'|d)$, resulting

$$\Lambda_{max} = \frac{\Sigma_d g_d^2 (d-1) d p_d}{\Sigma_s g_s s p_s}. \tag{21}$$

Hence, the conditions for the existence of a giant component ($\Lambda_{max}>1$) or resilience to damage ($\Lambda_{max}=\infty$) are modulated by g_d and, therefore, the disassortative correlations given by g_d have a great impact on the percolation properties. For instance, let us consider $g_d = d^{-\alpha}$ and a power law degree distribution $p_d = cd^{-\gamma}$ with $\gamma<3$ ($\langle d^2 \rangle = \infty$). From Eq. (20) it follows that $\langle d \rangle_{nn} - 1 \sim d^{-\alpha}$, so that when increasing α the graph gets more and more disassortative. Moreover, Λ_{max} diverges for $\alpha < \alpha_c = (3-\gamma)/2$ and it is finite otherwise. Thus, for small values of α the graph is robust but for $\alpha > \alpha_c$ it becomes fragile. It is worth noting that the value of α above which the giant component disappears ($\Lambda_{max}<1$) is larger than α_c. Besides, for large degrees, the degree distribution of the vertices in the giant component is still a power law, but it decays slower than that of the whole graph. Thus, disassortative correlations compete against the formation of the giant component and the divergence of $\langle d^2 \rangle$ is not a sufficient condition to get a robust graph with $f_c=1$.

The connection between percolation theory and models of epidemic spreading is well known [19]. The two general classes of epidemiological models can be related to percolation problems, the susceptible-infected-removed (SIR) and the susceptible-infected-susceptible (SIS) classes. The SIR model assumes that individuals can exist in three classes and that once they get infected they cannot catch the infection again. This model can be mapped into a bond percolation problem taking f as the probability that the disease will be transmitted from one node to another and the size of the giant component as the size of the outbreak. Hence, all the conclusions drawn above for the bond percolation problem can be translated to the language of epidemic spreading for the SIR model on top of the graphs with degree correlations, extending in this way previous studies in Refs. [6,7] for uncorrelated graphs.

On the other hand, the SIS model allows individuals to move through the cycle of infection so that the prevalence (number of infected individuals) attains a stationary value. The SIS model on top of the graphs with degree correlations has been recently analyzed in Refs. [14,17]. They obtained the epidemic threshold (the value of λ above which the solution with zero prevalence is unstable) $\lambda = 1/\Lambda'_{max}$, where

Λ'_{max} is the largest eigenvalue of the matrix $C'_{dd'} = dp(d'|d)$. This approach is quite similar to that presented here for site percolation with the remark that $C'_{dd'}$ is different [see Eq. (7)]. In fact, if y_d is an eigenvector of $C'_{dd'} = dp(d'|d)$ corresponding to the eigenvalue Λ' then y_d/d is an eigenvector of $C''_{dd'} = d'p(d'|d)$ corresponding to the same eigenvalue. This last matrix is that of Eq. (7), but replacing d' by $d'-1$. However, this subtle difference makes the SIS and dilute percolation different. We have computed the largest eigenvalue of $C''_{dd'}$ for the disassortative graph considered above [Eq. (19)]. Taking the limit $\langle d^2\rangle \gg 1$ one gets

$$\Lambda'_{max} \approx \frac{\Sigma_d(1-g_d)d^2p_d}{\Sigma_s(1-g_s)sp_s}, \tag{22}$$

where g_d is again a decreasing function of d. In this case, independent of the form of g_d, the divergence of the second moment of the degree distribution implies the divergence of Λ'_{max}. Moreover, the same conclusion is obtained if g_d is an increasing function of d. The conditions for the existence of a finite prevalence in the SIS model have been recently addressed in Ref. [17], where the divergence of the second moment has been shown to be a sufficient condition for the absence of the phase transition in the SIS model. Nevertheless, we have shown that this conclusion does not hold for dilute percolation. This essential difference is rooted in the existence of an additional dimension in the SIS model, given by the time evolution of the density of infected sites.

In summary, we have studied the percolation problem on top of the random networks with arbitrary degree distribution and correlations, making its generalization to site and bond percolations. The connection with the spreading phenomena was also analyzed. We provide some general expressions to obtain or bound the transition threshold. Using these results we have shown that the existence of a finite amount of random mixing of the connections between vertices is sufficient to make the graph robust under vertex or edge removal provided $\langle d^2\rangle \to \infty$. Assortative correlations makes the situation even better; they can lead to a graph robust to random damage even with a finite second moment of the degree distribution. On the contrary, disassortative correlations compete again the formation of the giant component and can make a graph fragile even with a divergent second moment.

We thank A. Vespignani, R. Pastor-Satorras, and M. Weigt for helpful comments. This work has been partially supported by the European commission FET Open Project No. COSIN IST-2001-33555.

[1] R. Albert and A.-L. Barabási, Rev. Mod. Phys. **74**, 47 (2001).
[2] S.N. Dorogovtsev and J.F.F. Mendes, Adv. Phys. **51**, 1079 (2002).
[3] R. Albert, H. Jeong, and A.-L. Barabási, Nature (London) **406**, 378 (2000); M.E.J. Newman, S.H. Strogatz, and D.J. Watts, Phys. Rev. E **64**, 026118 (2001); D.S. Callaway, M.E.J. Newman, S.H. Strogatz, and D.J. Watts, Phys. Rev. Lett. **85**, 5468 (2000); R. Cohen, K. Erez, D. ben-Avraham, and S. Havlin, *ibid.* **86**, 3682 (2001).
[4] R. Cohen, K. Erez, D. ben-Avraham, and S. Havlin, Phys. Rev. Lett. **85**, 4626 (2000).
[5] R. Pastor-Satorras and A. Vespignani, Phys. Rev. Lett. **86**, 3200 (2001); R. Pastor-Satorras and A. Vespignani, Phys. Rev. E **63**, 066117 (2001); A.L. Lloyd and R.M. May, Science **292**, 1316 (2001).
[6] Y. Moreno, R. Pastor-Satorras, and A. Vespignani, Eur. Phys. J. B **26**, 521 (2002).
[7] M.E.J. Newman, Phys. Rev. E **66**, 016128 (2002).
[8] A. Aleksiejuk, J.A. Holyst, and D. Stauffer, Physica A **310**, 260 (2002); S.N. Dorogovtsev, A.V. Goltsev, and J.F.F. Mendes, Phys. Rev. E **66**, 016104 (2002); M. Leone, A. Vázquez, A. Vespignani, and R. Zecchina, Eur. Phys. J. E **28**, 191 (2002).
[9] S.H. Strogatz, Nature (London) **410**, 268 (2001); L.A. Adamic, R.M. Lukose, A.R. Puniyani, and B. Huberman, Phys. Rev. E **64**, 046135 (2001).
[10] R. Pastor-Satorras, A. Vázquez, and A. Vespignani, Phys. Rev. Lett. **87**, 258701 (2001)
[11] M.E.J. Newman, Phys. Rev. Lett. **89**, 208701 (2002).
[12] P.L. Krapivsky and S. Redner, Phys. Rev. E **63**, 066123 (2001).
[13] D. Callaway, J.E. Hopcroft, J.M. Kleinberg, M.E.J. Newman, and S.H. Strogatz, Phys. Rev. E **64**, 041902 (2001).
[14] M. Boguñá and R. Pastor-Satorras, Phys. Rev. E **66**, 047104 (2002).
[15] J. Berg and M. Lässig, Phys. Rev. Lett. **89**, 228701 (2002).
[16] A. Vázquez and M. Weigt, Phys. Rev. E (to be published), cond-mat/0207035.
[17] M. Boguñá, R. Pastor-Satorras, and A. Vespignani, e-print cond-mat/0208163.
[18] G.W. Stewart and J.g. Sun, in *Matrix Perturbation Theory* (Academic Press, San Diego, 1990).
[19] P. Grassberger, Math. Biosci. **63**, 157 (1983).