Fractal Dimensions of Percolating Networks

Reuven Cohen [a,b] and Shlomo Havlin [b]

[a] *Department of Computer Science and Applied Mathematics, Weizmann Institute of Science, Rehovot 76100, Israel*

[b] *Minerva Center and Department of Physics, Bar-Ilan University, Ramat-Gan 52900, Israel*

Abstract

We use the generating function formalism to calculate the fractal dimensions for the percolating cluster at criticality in Erdős-Rényi (ER) and random scale free (SF) networks, with degree distribution $P(k) = ck^{-\lambda}$. We show that the chemical dimension is $d_l = 2$ for ER and SF networks with $\lambda > 4$, as in percolation in $d \geq d_c = 6$ dimensions. For $3 < \lambda < 4$ we show that $d_l = \frac{\lambda-2}{\lambda-3}$. The fractal dimension is $d_f = 4$ ($\lambda > 4$) and $d_f = 2\frac{\lambda-2}{\lambda-3}$ ($3 < \lambda < 4$), and the embedding dimension is $d_c = 6$ ($\lambda > 4$) and $d_c = 2\frac{\lambda-1}{\lambda-3}$ ($3 < \lambda < 4$). We discuss the meaning of these dimensions for networks.

Key words: Internet, scale-free, networks, fractal, networks
PACS: 89.75.Hc, 05.50.+q, 89.75.Da

Introduction

Recently, much interest has been focused on the properties of large scale networks [1–4] due to their relevance to many real-world networks. One of the main properties studied in networks has been their percolation properties [5–9]. Percolation theory is relevant to the stability of computer networks to random failures and intentional attacks by Hackers, to the stability of ecological and biological network, to the spread of viruses in computer networks and of epidemics in populations, and to the immunization of these networks.

In this paper, we study the fractal properties of scale free and Erdős-Rényi networks [10–12] near the percolation threshold. It is well known that random

* Corresponding author. E-mail: cohenr@shoshi.ph.biu.ac.il

networks, having N nodes, present very short distances between nodes, which scale as $d \sim \log N$ [13–16]. It has recently been shown that scale free networks actually present even shorter distances, which scale as $d \sim \log \log N$ [17–19]. Here we discuss the behavior of networks at the critical point of the percolation transition, and show that the distances become much larger and the network has a fractal behavior of $d \sim M^{1/d_l}$, where M is the size of the largest cluster. Our study extends known results from percolation theory, and in particular in infinite-dimensional percolation [20–22].

The generating function formalism

To study the behavior of the network near and at the percolation transition, we use the generating function formalism similar to [9,7]. We begin by building the generating function for the degree distribution:

$$G_0(x) = \sum_{k=0}^{\infty} P(k) x^k \ , \tag{1}$$

where $P(k)$ is the fraction of nodes having degree (number of connections) k. For scale free networks $P(k) = ck^{-\lambda}$ ($k > k_0$). For ER networks $P(k) = e^{-z} z^k / k!$. The probability of reaching a site by following a link is proportional to k [6,7] and therefore the generating function for the out degrees (*i.e.* the number of links except the one through which we arrived) of sites reached by following a link is:

$$G_1(x) \equiv \frac{G_0'(x)}{G_0'(0)} = \sum_{k=0}^{\infty} \frac{kP(k)}{\langle k \rangle} x^{k-1} \ . \tag{2}$$

If nodes (or alternatively, link), are removed with probability p, they exist with probability $q = 1 - p$ leading to the new generating function:

$$F_1(x) = 1 - q + q \sum_{k=0}^{\infty} \frac{kP(k)}{\langle k \rangle} x^{k-1} \ . \tag{3}$$

Suppose now that we follow some branch, by starting at a link and following it in some direction. The distribution of the links outgoing from this site is given by $F_1(x)$ and the distribution of branch sizes is given by the generating function $H_1(x)$ defined by the recursive equation:

$$H_1(x) = 1 - q + qxG_1(H_1(x)) \ , \tag{4}$$

and the distribution of cluster sizes is given by the generating function

$$H_0(x) = 1 - q + qxG_0(H_1(x)) \ . \tag{5}$$

Percolation Threshold

The total fraction of the network occupied by finite clusters is given by $H_0(1)$. This is normalized to 1 as long as the entire network consists of only finite clusters. However, for some value of q, denoted q_c the network undergoes a phase transition, where a giant component is formed and occupies a finite fraction of the network. This threshold is determined by the point at which the average number of outgoing links from a node reached by following a link is larger than 1, allowing the branching process to continue. This is determined by the criterion

$$F_1'(1) = 1 \ , \tag{6}$$

leading to the critical threshold [6,7]

$$1 - p_c \equiv q_c = \frac{\langle k \rangle}{\langle k^2 \rangle - \langle k \rangle} \ . \tag{7}$$

As can be seen from Eq. (7) the critical threshold vanishes for networks having degree distribution with a divergent second moment, such as the scale free distribution with $\lambda \leq 3$ [6].

When the network is diluted one can use Eqs. (4) and (5) to find the size of the giant component, $P_\infty(q)$. The equation for $P_\infty(q)$ can be written as

$$P_\infty(q) = q(1 - \sum_{k=0}^{\infty} P(k) u^k) \ , \tag{8}$$

where $u \equiv H_1(1)$ is the smallest positive root of

$$u = 1 - q + \frac{q}{\langle k \rangle} \sum_{k=0}^{\infty} k P(k) u^{k-1} \ . \tag{9}$$

This equation can be solved numerically and the solution may be substituted into Eq. (8), yielding the size of the spanning cluster in a network of arbitrary degree distribution, at dilution q [7].

To obtain the behavior at and near the critical threshold, Eqs. (8) and (9) should be expanded near the threshold [23]. Eq. (8) has no special behavior at $q = q_c$; the singular behavior comes from u in Eq. (9). Also, at criticality $P_\infty = 0$ and Eq. (8) imply that $u = 1$. We therefore examine Eq. (9) for $u = 1 - \epsilon$ and $q = q_c + \delta$:

$$1 - \epsilon = 1 - q_c - \delta + \frac{(q_c + \delta)}{\langle k \rangle} \sum_{k=0}^{\infty} k P(k) (1 - \epsilon)^{k-1}. \tag{10}$$

The sum in (10) has the asymptotic form

$$\sum_{k=0}^{\infty} kP(k)(1-\epsilon)^{k-1} \sim \langle k\rangle - \langle k(k-1)\rangle\epsilon + \tfrac{1}{2}\langle k(k-1)(k-2)\rangle\epsilon^2 + \cdots + c\Gamma(2-\lambda)\epsilon^{\lambda-2} \;, \tag{11}$$

where the highest-order analytic term is $O(\epsilon^n)$, $n = \lfloor \lambda - 2 \rfloor$. Substituting (11) in Eq. (10), with $q_c = 1/(\kappa - 1) = \langle k\rangle/\langle k(k-1)\rangle$, we get

$$\frac{\langle k(k-1)\rangle^2}{\langle k\rangle}\delta = \tfrac{1}{2}\langle k(k-1)(k-2)\rangle\epsilon + \cdots + c\Gamma(2-\lambda)\epsilon^{\lambda-3} \;. \tag{12}$$

The divergence of δ as $\lambda < 3$ confirms the vanishing threshold of the phase transition in that regime. Thus, in the case $\lambda > 3$, keeping only the dominant term as $\epsilon \to 0$, Eq. (12) implies

$$\epsilon \sim \begin{cases} \left(\frac{\langle k(k-1)\rangle^2}{c\langle k\rangle\Gamma(2-\lambda)}\right)^{\frac{1}{\lambda-3}} \delta^{\frac{1}{\lambda-3}} & 3 < \lambda < 4, \\ \frac{2\langle k(k-1)\rangle^2}{\langle k\rangle\langle k(k-1)(k-2)\rangle}\delta & \lambda > 4. \end{cases} \tag{13}$$

In [9] it was shown that for a random graph of arbitrary degree distribution the finite clusters follow the usual scaling form:

$$n_s \sim s^{-\tau} e^{-s/s^*} \;. \tag{14}$$

At criticality $s^* \sim |q - q_c|^{-\sigma}$ diverges and the tail of the distribution behaves as a power law. We now derive the exponent τ. The probability that a site belongs to an s-cluster is $p_s = sn_s \sim s^{1-\tau}$, and is generated by H_0:

$$H_0(x) = \sum p_s x^s \;. \tag{15}$$

The singular behavior of $H_0(x)$ stems from $H_1(x)$, as can be seen from Eq. (5). $H_1(x)$ itself can be expanded from Eq. (4), by using the asymptotic form (11) of G_1. We let $x = 1 - \epsilon$, as before, but analyze at the critical point, $q = q_c$. With the notation $\phi(\epsilon) = 1 - H_1(1-\epsilon)$, we finally get (note that at criticality $H_1(1) = 1$):

$$-\phi = -q_c + (1-\epsilon)q_c\left[1 - \frac{\phi}{q_c} + \frac{\langle k(k-1)(k-2)\rangle}{2\langle k\rangle}\phi^2 + \cdots + c\frac{\Gamma(2-\lambda)}{\langle k\rangle}\phi^{\lambda-2}\right] \;. \tag{16}$$

From this relation we extract the singular behavior of H_0: $\phi \sim \epsilon^y$, for some y calculated below. Then, using Tauberian theorems [24] it follows that $p_s \sim s^{-1-y}$, hence $\tau = 2 + y$.

For $\lambda > 4$ the term proportional to $\phi^{\lambda-2}$ in (16) may be neglected. The linear term $\epsilon\phi$ may be neglected as well, due to the factor ϵ. This leads to $\phi \sim \epsilon^{1/2}$ and to the usual mean-field result [20,21]

$$\tau = \frac{5}{2} \;, \qquad \lambda > 4 \;. \tag{17}$$

For $\lambda < 4$, the terms proportional to $\epsilon\phi$, ϕ^2 may be neglected, leading to $\phi \sim \epsilon^{1/(\lambda-2)}$ and therefore [23]

$$\tau = 2 + \frac{1}{\lambda - 2} = \frac{2\lambda - 3}{\lambda - 2}\,, \qquad 2 < \lambda < 4\,. \tag{18}$$

Note that for $2 < \lambda < 3$ the percolation threshold is strictly $q_c = 0$. In that case we analyze at $q = \delta$ small but fixed, taking the limit $\delta \to 0$ at the very end. For the case $2 < \lambda < 3$, τ in Eq. (18) represents the singularity of the distribution of branch sizes. For the distribution of cluster sizes in this range one has to consider the singularity of x in Eq. (5) leading to $\tau = 3$ for this range.

For growing networks of the Albert-Barabási model with $\lambda = 3$, it has been shown that $sn_s \propto (s \ln s)^{-2}$ [25]. This is consistent with $\tau = 3$ plus a logarithmic correction. Related results for scale free *trees* have been presented in [26].

At the transition point the largest cluster, S can be obtained from the finite cluster distribution by taking the integral over the tail of the distribution to be equal $1/N$. This results in

$$S \propto N^{\tau-1} = N^{(\lambda-2)/(\lambda-1)}. \tag{19}$$

For $\lambda = 4$ this reduces to the known result $N^{2/3}$, termed by Erdös the "double jump", due to the transition of the largest cluster from order $\ln N$ for $q < q_c$, to $N^{2/3}$ at $q = q_c$, to order N at $q > q_c$ [13]. Similarly, in mean field percolation, since $d_c = 6$, it follows that for $d = 6$ dimensions, $N = L^6$ (where L is the linear size of the lattice) the size of the largest cluster is known to be $S \sim L^{d_f}$, where $d_f = 4$ is the fractal dimension. Thus $S \sim N^{2/3}$. For $\lambda \to 3$, $S \propto N^{1/2}$. It is not yet clear whether Eq. (19) has a meaningful interpretation for $\lambda < 3$.

Fractal Dimension

It is well known that on a random network in the well connected regime, the average distance between sites is of the order $\log_{\langle k \rangle} N$ [13,14,9], and becomes even smaller smaller in scale-free networks [17–19]. However, the diluted case is essentially analogous to infinite-dimensional percolation. In this case, there is no notion of geometrical distance (since the graph is not embedded in an Euclidean space), but only of a distance along the graph (which is the shortest distance along bonds, or chemical distance). It is known from infinite-dimensional percolation theory that the chemical fractal dimension at criticality is $d_l = 2$ [21]. Therefore the average (chemical) distance d between pairs of sites on the spanning cluster for random graphs and scale free networks with

$\lambda > 4$ at criticality behaves as

$$d \sim \sqrt{M} \ , \tag{20}$$

where M is the number of sites in the spanning cluster. This is analogous to percolation theory, where in length-scales smaller than the correlation length the cluster is a fractal with dimension d_l and above the correlation length the cluster is homogeneous and has the dimension of the embedding space. In our infinite-dimensional case, the crossover between these two behaviors occurs around the correlation chemical length $\xi_l \sim |p_c - p|^{-\nu_l}$.

Next, we calculate ν_l for scale free networks with $3 < \lambda < 4$. Below the transition all clusters are finite and almost all finite clusters are trees. The correlation length can be defined using the formula [21]:

$$\xi_l^2 = \frac{\sum l^2 g(l)}{\sum g(l)}. \tag{21}$$

Where $g(l)$, the correlation function, is the mean number of sites on the same cluster at distance l from an arbitrary occupied site. The number of sites in the l shell can be seen to be approximately $\langle k \rangle (\kappa - 1)^{l-1}$ [9]. Since $\kappa - 1 = (\kappa_0 - 1)q$ and $q_c = 1/(\kappa_0 - 1)$ we get $g(l) = c(1-\delta)^l$, where $\delta = q - q_c$. This leads to $\xi_l \sim (q - q_c)^{-1}$, *i.e.* $\nu_l = 1$. Above the threshold, the finite clusters can be seen as a random graph with the residual degree distribution of sites not included in the infinite cluster [27]. That is, the degree distribution for sites in the finite clusters is

$$P_r(k) = P(k) u^k, \tag{22}$$

where u is the solution of Eq. (9). Using this distribution we can define κ_r for the finite clusters. This adds a term proportional to $\epsilon^{\lambda-3}$ to the expansion of ξ_l. But, since $\delta \propto \epsilon^{\lambda-3}$ (13), this leads again to $\nu_l = 1$.

The fractal dimensions, d_l, of scale free networks have already been derived using scaling relations in [17]. A direct method for calculating the chemical dimension is also possible. Denoting the generating function of the number of sites on the lth layer of some branch, as $N_l(x)$, we get

$$N_{l+1}(x) = G_1(N_l(x)) \ . \tag{23}$$

We are interested in the behavior of the average number of sites at a chemical distance l for those branches that have at least l layers. Since we expand exactly at criticality, the average branching factor is exactly 1, and therefore $N_l(1) = 1$ for any l. Therefore, A_l, the average number of sites for surviving branches is

$$A_l = \frac{1}{1 - N_l(0)} \ , \tag{24}$$

since $N_l(0)$ is the probability of the branching process to die out before the lth layer. At criticality the branching process will die out with probability

$N_l(0) \to 1$ as $l \to \infty$, and therefore for large l we can take $N_l(0) = 1 + \epsilon_l$. Expanding G_1 at criticality one obtains (Eqs. (10) and (11), with $\delta = 0$)

$$G_1(1-\epsilon) = 1 - \epsilon + \frac{c\Gamma(2-\lambda)}{\langle k^2\rangle - \langle k\rangle}\epsilon^{\lambda-2} + \dots . \tag{25}$$

Substituting $N_l(0) = 1 - \epsilon_l$ into Eq. 23 one obtains

$$1 - \epsilon_{l+1} = 1 - \epsilon_l - \frac{c\Gamma(2-\lambda)}{\langle k^2\rangle - \langle k\rangle}\epsilon_l^{\lambda-2} + \dots . \tag{26}$$

Guessing a solution of the form $\epsilon_l \approx Bl^{-d}$ we get

$$B(l+1)^{-d} \approx B(l^{-d} - dl^{-d-1}) = Bl^{-d} - \frac{c\Gamma(2-\lambda)}{\langle k^2\rangle - \langle k\rangle}(Bl^{-d})^{\lambda-2} . \tag{27}$$

implying that $d = 1/(\lambda-3)$, and $N_l(0) \sim 1 - Bl^{-d}$. Noting that the mass of the branch is the sum of the layers up to the lth one, we get $d_l = d + 1$, and thus,

$$d_l = \frac{\lambda-2}{\lambda-3}, \qquad 3 < \lambda < 4 . \tag{28}$$

Similar results have been obtained by Burda et al.[26] for scale-free trees. Since every path when embedded in a space above the critical dimension can be seen as a random walk it is known that $\nu = \nu_l/2$ [21]. Therefore, the fractal dimension is,

$$d_f = 2d_l = 2\frac{\lambda-2}{\lambda-3}, \qquad 3 < \lambda < 4 . \tag{29}$$

The upper critical dimension of the embedding space is,

$$d_c = \frac{1}{\nu\sigma(\tau-1)} = 2\frac{\lambda-1}{\lambda-3}, \qquad 3 < \lambda < 4 . \tag{30}$$

Those dimensions reduce to the known 2, 4, and 6, respectively, for $\lambda \geq 4$. It is due to the different topology of the network (for $\lambda < 4$) that the phase transition has different critical exponents compared to the regular infinite dimensional lattice percolation system.

Discussion

The behavior of the fractal dimensions of random and scale free networks percolating clusters at criticality have been presented. It has been shown that the fractal dimensions of ER and SF networks with $\lambda > 4$ is similar to that of regular infinite dimensional percolation. The embedding dimension obtained here is similar to percolation's upper critical dimension $d_c = 6$, and the fractal and chemical dimensions of the spanning cluster (giant component) are consistent with those of percolation at and above the upper critical dimension, *i.e.*, $d_l = 2$ and $d_f = 4$. The exponents for the correlation length, ν and ν_l are also the same as in infinite dimensional percolation.

The meaning of the chemical dimension, d_l, is clear for networks. It describes the behavior of the spanning cluster at the percolation threshold, as seen by traversing the network by following its links. The meaning of the fractal and the upper critical dimension, however, is less apparent. It should be related to an embedding Euclidean space of the network. This can be obtained by two different processes.

One process is embedding the percolation cluster in a Euclidean space, while keeping connected nodes geometrically close to each other. The meaning assigned to the fractal dimensions then will be that the embedding space will have to be at least of dimension d_c and the spanning cluster will not fill the embedding space, but rather behave as a fractal with dimension d_f. The finiteness of d_c implies that, in contrast to the non-diluted network, which can not be embedded in any finite dimension, the diluted network at criticality can be embedded in a finite dimensional space.

The other process is creating an embedded network to begin with and then diluting it. Some properties of embedded networks on lattices have been studied in [28,29]. In this case the meaning ascribed to the fractal dimensions will be that such a network, embedded in a Euclidean lattice, with dimension at least d_c, will behave at the percolation threshold similarly to the mean field network discussed above. The value of d_c implies that for $d < d_c$ the geometry will effect the critical exponents, while for $d \geq d_c$ the percolating network will behave like in infinite dimensions, and the above results, Eqs. (28) and (29) will be satisfied.

For scale free network with $\lambda < 4$ the fractal dimension increase, which is expected due to the abundance of high degree nodes, leading to a higher needed dimension. At $\lambda = 3$ the fractal dimensions diverge, with seems to indicate an exponential behavior of the spanning cluster with the distance at criticality for $\lambda \leq 3$ [30].

Acknowledgments

We thank the Israel Science Foundation for financial support.

References

[1] R. Albert and A.-L. Barabási, *Rev. of Mod. Phys.* **74**, 47 (2002).

[2] J. F. F. Mendes, S. N. Dorogovtsev and A. F. Ioffe, *Evolution of Networks: From Biological Nets to the Internet and WWW* (Oxford University Press, Oxford, 2003).

[3] M. E. J. Newman, *SIAM Review* **45**, 167 (2003).

[4] R. Pastor-Satorras and A. Vespignani, *Evolution and Structure of the Internet : A Statistical Physics Approach* (Cambridge University Press, Cambridge, to appear, 2003).

[5] R. Albert, H. Jeong, and A. L. Barabási, *Nature*, **406**, 6794, 378 (2000).

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

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

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

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

[10] P. Erdős and A. Rényi, *Publicationes Mathematicae* **6**, 290 (1959).

[11] P. Erdős and A. and Rényi, *Publications of the Mathematical Institute of the Hungarian Academy of Sciences* **5**, 17 (1960).

[12] P. Erdős and A. Rényi *Acta Mathematica Scientia Hungary* **12**,261 (1961).

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

[14] F. Chung and L. Lu, *Adv. Appl. Math.*, **26**, 257, (2001).

[15] D. J. Watts and S. H. Strogatz, *Nature* **393**, 440 (1998).

[16] D. J. Watts, *Small Worlds.* (Princeton University Press, Princeton, 1999).

[17] R. Cohen, S. Havlin and D. ben-Avraham, Chap. 4 in *Handbook of Graphs and Networks*, Eds. S. Bornholdt and H. G. Schuster, (Wiley-VCH, 2002).

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

[19] S.N. Dorogovtsev, J. F. F. Mendes, and A.N. Samukhin, *Nuclear Physics B* **653 (3)**, 307 (2003).

[20] D. Stauffer and A. Aharony, *Introduction to Percolation Theory*, 2nd edition (Taylor and Francis, London, 1991).

[21] A. Bunde, and S. Havlin (editors), *Fractals and Disordered System* (Springer, New York, 1996).

[22] D. ben-Avraham and S. Havlin, *Diffusion and Reactions in Fractals and Disordered Systems* (Cambridge University Press, 2000).

[23] R. Cohen, D. ben-Avraham and S. Havlin, *Phys. Rev. E* **66**, 036113 (2002).

[24] G. H. Weiss, *Aspects and Applications of the Random Walk* (North-Holland, Amsterdam, 1994).

[25] S. N. Dorogovtsev, J. F. F. Mendes, and A. N. Samukhin, *Phys. Rev. E* **64**, 066110 (2001).

[26] Z. Burda, J. D. Curreira, and A. Krzywicki, *Phys. Rev. E* **64**, 046118 (2001).

[27] M. Molloy and B. Reed, *Combinatorics, Probability and Computing* **7**, 295 (1998).

[28] A. F. Rozenfeld, R. Cohen, D. ben-Avraham and S. Havlin, *Phys. Rev. Lett.* **89**, 218701 (2002).

[29] C. P. Warren, L. M. Sander and I. M. Sokolov, *Phys. Rev. E* **66**, 56105 (2002).

[30] L. A. Braunstein, S. V. Buldyrev, R. Cohen, S. Havlin and H. E. Stanley, *Phys. Rev. Lett.* **91**, 168701 (2003).