Available online at www.sciencedirect.com
SCIENCE DIRECT®
Physica A 330 (2003) 246–252
PHYSICA A
www.elsevier.com/locate/physa

Length of optimal path in random networks with strong disorder

Sergey V. Buldyrev[a], Lidia A. Braunstein[a,c], Reuven Cohen[b], Shlomo Havlin[a,b], H. Eugene Stanley[a,*]

[a] *Department of Physics, Center for Polymer Studies, Boston University, Boston, MA 02215, USA*
[b] *Minerva Center and Department of Physics, Bar-Ilan University, 52900 Ramat-Gan, Israel*
[c] *Departamento de Física, Facultad de Ciencias Exactas y Naturales, Universidad Nacional de Mar del Plata, Funes 3350, 7600 Mar del Plata, Argentina*

Abstract

We study the optimal distance ℓ_{opt} in random networks in the presence of disorder implemented by assigning random weights to the links. The optimal distance between two nodes is the length of the path for which the sum of weights along the path ("cost") is a minimum. We study the case of strong disorder for which the distribution of weights is so broad that its sum along any path is dominated by the largest link weight in the path. We find that in random graphs, ℓ_{opt} scales as $N^{1/3}$, where N is the number of nodes in the network. Thus, ℓ_{opt} increases dramatically compared to the known small-world result for the minimum distance $\ell_{\min}$, which scales as $\log N$. We also study, theoretically and by simulations, scale-free networks characterized by a power law distribution for the number of links, $P(k) \sim k^{-\lambda}$, and find that ℓ_{opt} scales as $N^{1/3}$ for $\lambda > 4$ and as $N^{(\lambda-3)/(\lambda-1)}$ for $3 < \lambda < 4$. For $2 < \lambda < 3$, our numerical results suggest that ℓ_{opt} scales logarithmically with N.

PACS: 87.10.+e; 64.60.Cn; 75.10.Hk; 02.50.−r

Keywords: Scale-free networks; Small-world networks; Percolation; Strong disorder; Optimal path

Much attention has been focused on the topic of complex networks characterizing many biological, social, and communication systems [1–3]. The networks can be visualized by nodes representing individuals, organizations, or computers and by links between them representing their interactions. The classical model for random networks

* Corresponding author. Tel.: +1-617-3532617; fax: +1-617-3533783.
E-mail address: hes@bu.edu (H.E. Stanley).

0378-4371/$ - see front matter © 2003 Published by Elsevier B.V.
doi:10.1016/j.physa.2003.08.030

is the Erdős–Rényi model where two nodes are chosen randomly from the total N nodes in the system and are connected by a link [4]. An important quantity characterizing networks is the minimum distance $\ell_{\min}$ between two nodes in the network. For the Erdős–Rényi network, $\ell_{\min}$ scales as $\log N$, consistent with the "six degrees of separation" concept (e.g., if $N = 10^6, \ell \approx 6$).

Here, we study a more realistic problem in which all links are not assumed to be equivalent. Hence, we assign to each link a weight or "cost". For example, the cost could be the time required to transit the link, e.g., there are often many traffic routes from point A to point B with a set of delay times τ_i associated with each link along the path. The fastest (optimal) path is the one for which $\sum_i \tau_i$ is a minimum, and often the optimal path has more links than the shortest path.

If the distribution of weights is such that all the links have the same weight, the average length of the optimal path between any two nodes is the minimal length $\ell_{\min}$. In that case, it is well known that $\ell_{\min} \sim \log N$ [5] (or, for some scale-free networks $\ell_{\min} \sim \log\log N$ [6]). If the distribution is narrow, the average length of the optimal path ℓ_{opt}, in general, is greater than $\ell_{\min}$ but scales the same as $\ell_{\min}$ [7,8]. If the random distribution is broad, in the limit of infinite broadness, the disorder is called "strong" and only the largest weight in the path dominates the sum. The strong disorder limit is implemented by assigning to each link a potential barrier ε_i so that τ_i is the waiting time to cross this barrier. Thus $\tau_i = e^{\beta\varepsilon_i}$, and the optimal path corresponds to the minimum $(\sum_i \tau_i)$ over all possible paths. When $\beta = 1/kT \to \infty$, only the largest τ_i dominates the sum. Thus, $T \to 0$ (very low temperatures) corresponds to the strong disorder limit.

We focus here on the case of strong disorder. This is believed to be the case for many computer and traffic networks, since the slowest link in communication networks determines the connection speed. We study this problem both theoretically and numerically and find that for random graphs ℓ_{opt}, the average length of the optimal path, scales as $N^{1/3}$.

To obtain the optimal path in the strong disorder limit, we present the following theoretical argument. It has been shown [9,10] that the optimal path for $\beta \to \infty$ between two nodes A and B on the network can be obtained by the following algorithm:

(1) Sort the edges by descending weight.
(2) If the removal of the highest weight edge will not disconnect A from B—remove it.
(3) Go back to step 2 until all edges have been processed.

Since the edge weights are random, so is the ordering. Therefore, in fact, one needs not even select edge weights to begin with. This "bombing" algorithm can be replaced by simply removing randomly chosen edges one at a time, where an edge is not removed if its removal will cause the connectivity between A and B to be lost. The final path left is the optimal path between A and B in the limit $\beta \to \infty$.

Since randomly removing links is a percolation process, the optimal path must be on the percolation backbone connecting A and B. Since the network is not embedded in space but has an infinite dimensionality, we expect from percolation theory that

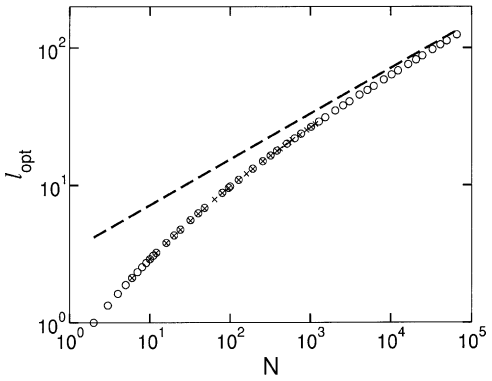

Fig. 1. Log–log plot of ℓ_{opt} as a function of N for the optimal path length in strong disorder using the two methods discussed in the text: (i) the results obtained using the "bombing" approach (○) and (ii) the results obtained using the ultrametric approach (×). The dashed line shows the slope $\frac{1}{3}$.

at criticality loops are not relevant. Thus, the shortest path must be the same as the optimal path. It is also known from percolation theory that at criticality the mass S of the incipient infinite cluster scales as ℓ_{min}^2 [11]. Since the spanning cluster S scales at criticality as $N^{2/3}$ [12], it follows that

$$\ell_{min} \sim \ell_{opt} \sim N^{\nu_{opt}} , \tag{1}$$

where $\nu_{opt} = \frac{1}{3}$.

To test Eq. (1), we apply two numerical approaches. The first approach is to find the optimal path (which minimizes the sum of weights) using the ultrametric approach described in Ref. [10]. The second approach is based on the "bombing" algorithm of Ref. [9]. In Fig. 1, we show a double logarithmic plot of ℓ_{opt} as a function of N. To evaluate the asymptotic value for ν_{opt} we use for both approaches successive slopes, defined as centered differences of the values in Fig. 1. One can see from Fig. 2 that their value approaches $\frac{1}{3}$ as $N \to \infty$, supporting Eq. (1).

In Fig. 3, using the result $\nu_{opt} = \frac{1}{3}$ obtained by extrapolation of successive slopes, we show ℓ_{opt} as a function of $N^{1/3}$. The linear behavior supports the theoretical value $\nu_{opt} = \frac{1}{3}$.

Next, we describe in more detail the two numerical methods for computing ℓ_{opt} between any two nodes in strong disorder. We can assume that the energy spectra ε_i is discrete. We can make β so large that, even for the closest values of energy spectra, the waiting times $\tau_i = \exp[\beta\varepsilon_i]$ differ by at least a factor of 2. In this limit, the sum is dominated by the maximum value $\exp[\beta\varepsilon_{max}]$. When all the links on the paths have different weights, the optimal path is the one that has the smallest maximal link weight among all the paths. In general, as a consequence of the existence of loops, there are links in common between different paths. Such a link might provide the maximum ε_i of both paths. In this case, we compare the second highest weight and take the path with the lower value and so forth until the optimal path is determined. This procedure is

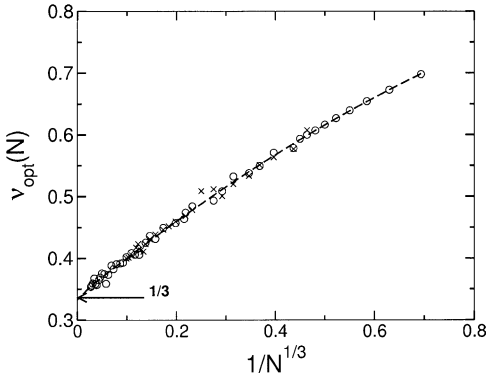

Fig. 2. Successive slopes $\nu_{opt}(N)$ as a function of $1/N^{1/3}$ for the optimal path length in strong disorder using the two methods described in the text. The symbols denote the same as in Fig. 1. The dashed line is the quadratic fitting of the results showing that the extrapolated value of the effective exponent in the limit $N \to \infty$ approaches $\frac{1}{3}$. This result coincides with our theoretical value $\frac{1}{3}$.

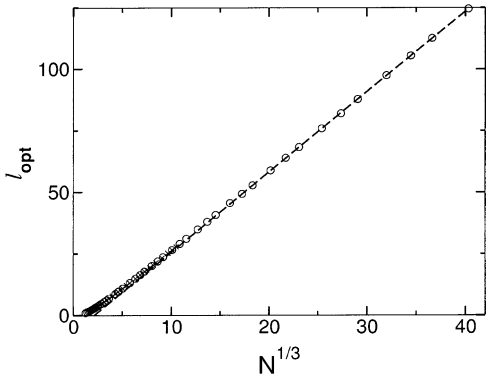

Fig. 3. ℓ_{opt} as a function of $N^{1/3}$ for the optimal path length in strong disorder using the two methods discussed in the text. The symbols denote the same as in Fig. 1. The dashed line is the linear fitting of the results showing the linear relation between ℓ_{opt} and $N^{1/3}$. This result also supports the theoretical value $\frac{1}{3}$.

equivalent to comparing integers written in binary codes and hence indeed minimizes $\sum \tau_i$ for $\beta \to \infty$.

First, we describe the ultrametric algorithm [10]. We assign weights $\tau_i = e^{\beta\varepsilon_i}$ to all the links in the graph where the order of magnitude ε_i is taken from a uniform distribution. As we see above, in the limit of strong disorder, the sum of the weights is dominated by the largest value along the path. Next, we start from one node (the origin—see Fig. 4) and visit all the other nodes connected to the origin using a burning algorithm [13]. If a node at distance ℓ_0 (from the origin) is being visited for the first time, this node will be assigned a list S_0 of weights $\tau_{0i}, i = 1,\ldots,\ell_0$ of the links by

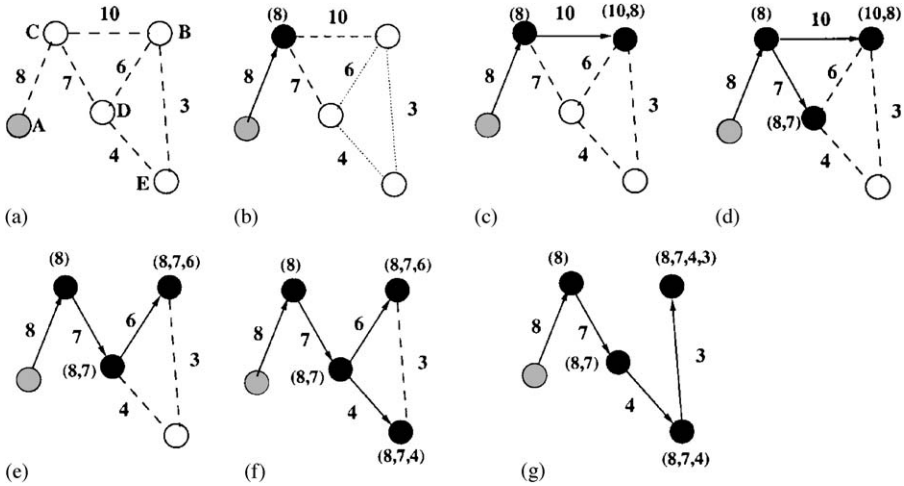

Fig. 4. In (a) we show schematically a network consisting of five nodes (A, B, C, D, and E). The links between them are shown in dashed lines. The origin (A) is marked in grey. All the links have been assigned a random weight, shown beside the links. In (b) one node (C) has been visited for the first time (marked in black) and assigned the sequence (8) of length $\ell = 1$. The path is marked by a solid arrow. Notice that there is no other path going from the origin (A) to this node (C) so $\ell_{\text{opt}} = 1$ for that path. In (c) another node (B) is visited for the first time (marked in black) and assigned the sequence (10, 8) of length 2. The sequence has the information of all the weights of that path arranged in decreasing order. In (d) another node (D) is visited for the first time and assigned the sequence (8, 7) of length 2. In (e), node (B) visited in (c) with sequence (10, 8) is visited for the second time with sequence (8, 7, 6). The last sequence is smaller than the previous sequence (10, 8) so that node (B) is reassigned the sequence (8, 7, 6) of length 3. The new path is shown as a solid line. In (f) a new node (E) is assigned with sequence (8, 7, 4). In (g) node (B) is reached for the third time and reassigned the sequence (8, 7, 4, 3) of length 4. The optimal path for this configuration from A to B is denoted by the solid arrows in (g).

which we reach that node sorted in descending order,

$$S_0 = \{\tau_{01}, \tau_{02}, \tau_{03}, \ldots, \tau_{0l_0}\} \tag{2}$$

with $\tau_{0j} > \tau_{0j+1}$ for all j. If we reach a node for a second time by another path of length ℓ_1, we define for this path a new list S_1,

$$S_1 = \{\tau_{11}, \tau_{12}, \tau_{13}, \ldots, \tau_{1l_1}\} \tag{3}$$

and compare it with a S_0 previously defined for this node.

Different sequences can have weights in common because some paths have links in common, so it is not enough to identify the sequence by its maximum weight; in this case, it must also be compared with the second maximum, the third maximum, etc. We define $S_p < S_q$ if there exists a value $m, 1 \leqslant \min(\ell_p, \ell_q)$ such that

$$\tau_{pj} = \tau_{qj} \quad \text{for } 1 \leqslant j < m$$

and

$$\tau_{pj} < \tau_{qj} \quad \text{for } j = m \tag{4}$$

or if $\ell_q > \ell_p$ and $\tau_{pj} = \tau_{qj}$ for all $j \leqslant \ell_p$.

If $S_1 < S_0$, we replace S_0 by S_1. The procedure continues until all paths have been explored and compared. At this point, $S_0 = S_{\text{opt}}$, where S_{opt} is the sequence of weights for the optimal path of length ℓ_{opt}.

A schematic representation of this ultrametric algorithm is presented in Fig. 4.

Using this method, we obtain systems of sizes up to 4000 nodes, typically 10^5 realizations of disorder. We compute ℓ_{opt} by averaging the length of the optimal path for all the pairs of nodes of the configuration and over all realizations.

An alternative method of obtaining the optimal path in strong disorder is called the "bombing" algorithm [9]. We first choose a pair of nodes on the graph and begin removing links randomly, making sure that the connectivity between the two chosen nodes is not destroyed as each link is removed. The last path remaining is equivalent to the optimal path obtained by the ultrametric algorithm.

The bombing algorithm is slow, as one must test the connectivity after removal of each link. To improve the speed, we first find the minimal path in the network and then select links in random order. We remove the selected link from the graph. If the removed link belongs to the minimal path, we check if the connectivity between the two nodes is still present and recompute the new minimal path. If the connectivity between the two nodes is destroyed, we restore the link.

The advantage of this procedure is that one has to test for connectivity only if the selected link appears to belong to the minimal path. Since checking the connectivity is the most time-consuming part in the original bombing algorithm, we could reach systems of sizes up to 2^{16} nodes with 10^5 realizations of weight disorder. Figs. 1–3 demonstrate that both algorithms yield very similar results, supporting the theoretical result $\ell_{\text{opt}} \sim N^{1/3}$.

In addition, we studied this problem for scale-free graphs [1–3]. In this case, the number k of links is distributed as a power law $P(k) \sim k^{-\lambda}$. Presenting analogous arguments as for the random graph, we conjecture that ℓ_{opt} must scale as the shortest path $\ell_{\min}$ at the percolation threshold. The percolation problem was solved analytically for a tree-like graph using a generating function approach [14] (see also Refs. [15,16]). This is, in fact, an exact solution, since at the critical point a graph in an infinite-dimensional space can be approximated by a tree [17]. Using these results we find that at criticality, the average largest cluster size S scales with the distance ℓ as ℓ^{d_ℓ}, where $d_\ell = 2$ for $\lambda > 4$ and $d_\ell = (\lambda - 2)/(\lambda - 3)$ for $3 < \lambda < 4$. Further, S scales as $N^{2/3}$ for $\lambda > 4$ and as $N^{(\lambda-2)/(\lambda-1)}$ for $3 < \lambda < 4$. Hence we conclude that $\ell_{\min}$, which is equal to ℓ_{opt} at criticality, scales as $N^{1/3}$ for $\lambda > 4$ and as $N^{(\lambda-3)/(\lambda-1)}$ for $3 < \lambda < 4$. Our simulation results for the bombing and ultrametric approaches confirm this prediction for ℓ_{opt}. For $\lambda < 3$, we find numerically that ℓ_{opt} scales logarithmically with N.

Finally, we repeated our simulations for the case in which disorder weights are associated with the nodes of the graph, and obtained the same scaling laws as for the disordered links case.

In summary, we study the optimal distance in random networks in the presence of strong disorder. We find that in random graphs the optimal distance ℓ_{opt} scales as $N^{1/3}$. We also study scale-free networks theoretically and by simulations and find that ℓ_{opt} scales as $N^{1/3}$ for $\lambda > 4$ and as $N^{(\lambda-3)/(\lambda-1)}$ for $3 < \lambda < 4$. Also, our simulations suggest that, for $2 < \lambda < 3, \ell_{\text{opt}}$ scales logarithmically with N, which is also much faster than the ultrasmall world result $\ell \sim \log\log N$ [6]. Thus, the optimal distance in strong disorder increases dramatically when it is compared to the known small-world result for the distance $\ell \sim \log N$ for random graphs and scale-free graphs with $\lambda > 3$.

Acknowledgements

We thank A.-L. Barabási and S. Sreenivasan for helpful discussions, and ONR for financial support.

References

[1] R. Albert, A.-L. Barabási, Rev. Mod. Phys. 74 (2002) 47–97.
[2] J.F.F. Mendes, S.N. Dorogovtsev, A.F. Ioffe, Evolution of Networks: From Biological Nets to the Internet and the WWW, Oxford University Press, Oxford, 2003.
[3] R. Pastor-Satorrad, A. Vespignani, Evolution and Structure of the Internet: a Statistical Physics Approach, Cambridge University Press, Cambridge, 2004.
[4] P. Erdős, A. Rényi, Publ. Math. 6 (1959) 290–297.
[5] B. Bollobás, Random Graphs, Academic Press, London, 1985.
[6] R. Cohen, S. Havlin, Phys. Rev. Lett. 90 (2003) 058 701.
[7] I. Smailer, J. Machta, S. Redner, Phys. Rev. E 47 (1993) 262.
[8] P. van der Hofstad, G. Hooghienstra, P. van Mieghen, Probab. Eng. Inform. Sci. 15 (2001) 225.
[9] M. Cieplak, A. Maritan, J.R. Banavar, Phys. Rev. Lett. 72 (1994) 2320–2323; M. Cieplak, A. Maritan, J.R. Banavar, Phys. Rev. Lett. 76 (1996) 3754–3757.
[10] L.A. Braunstein, S.V. Buldyrev, S. Havlin, H.E. Stanley, Phys. Rev. E 65 (2001) 056 128.
[11] A. Bunde, S. Havlin (Eds.), Fractals and Disordered Systems, Springer, New York, 1996.
[12] P. Erdős, A. Rényi, Publ. Math. Inst. Hungarian Acad. Sci. 5 (1960) 17–61.
[13] H.J. Herrmann, D.C. Hong, H.E. Stanley, J. Phys. Lett. A 17 (1984) L261–L265.
[14] R. Cohen, S. Havlin, D. ben-Avraham, in: S. Bornholdt, H.G. Shuster (Eds.), Handbook of Graphs and Networks, Wiley, New York, 2002 (Chapter 4).
[15] S.V. Buldyrev, S. Havlin, J. Kertész, R. Sadr, A. Shehter, H.E. Stanley, Phys. Rev. E 52 (1995) 373–388.
[16] A.-L. Barabási, S.V. Buldyrev, H.E. Stanley, B. Suki, Phys. Rev. Lett. 76 (1996) 2192–2195.
[17] R. Cohen, K. Erez, D. ben-Avraham, S. Havlin, Phys. Rev. Lett. 85 (2000) 4626.