RAPID COMMUNICATIONS

PHYSICAL REVIEW E VOLUME 60, NUMBER 2 AUGUST 1999

# Characterization and control of small-world networks

S. A. Pandit* and R. E. Amritkar†
*Physical Research Laboratory, Navarangpura, Ahmedabad 380 009, India*
(Received 29 April 1999)

Recently, Watts and Strogatz [Nature (London) **393**, 440 (1998)] offered an interesting model of small-world networks. Here we concretize the concept of a ''faraway'' connection in a network by defining a far edge. Our definition is algorithmic and independent of any external parameters such as topology of the underlying space of the network. We show that it is possible to control the spread of an epidemic by using the knowledge of far edges. We also suggest a model for better product advertisement using the far edges. Our findings indicate that the number of far edges can be a good intrinsic parameter to characterize small-world phenomena. [S1063-651X(99)50908-2]

PACS number(s): 84.35.+i, 02.40.−k, 05.45.Gg, 87.23.Ge

The properties of very large networks are mainly determined by the way the connections between the vertices are made. At one extreme are the regular networks where only the ''local'' vertices are interconnected and the ''far away'' vertices are not connected, while at the other extreme are the random networks where the vertices are connected at random. The regular networks display a high degree of local clustering and the average distance between vertices is quite large. On the other hand, the random networks show negligible local clustering and the average distance between vertices is quite small. The small-world networks [1,2] have intermediate connectivity properties but exhibit a high degree of clustering as in regular networks and a small average distance between vertices as in random networks. A very interesting model for small-world networks was recently proposed by Watts and Strogatz [3]. They found that a regular network acquires the properties of a small-world network with only a very small fraction of connections or edges (about 1%) rewired to faraway vertices. They demonstrated that several diverse phenomena such as neural networks [4], power grids, and collaboration graphs of film actors [5] can be modeled using small-world networks. Also, the spread of an epidemic is much faster in small-world networks than in the regular networks and almost close to that of random networks.

In this paper we suggest a possible way of characterizing small-world networks. The basic ingredients of small-world networks are the faraway connections. We introduce a notion of far edges in a network to identify these faraway connections. Our definition of a far edge is independent of any *external* parameters such as the topology of the underlying space of a network and depends only on the way connections or edges are made. We claim that the rapid spread of an epidemic in a small-world network as found by Watts and Strogatz [3] is due to these far edges. This allows us to propose a mechanism to control the epidemic using the same far edges that are responsible for the rapid spread. We further demonstrate the utility of our notion of far edges by offering a better method of advertising.

Consider a graph (network) with $n$ vertices and $E$ edges. Let $\mathcal{N}_{ij}^{\nu}$ denote the number of distinct paths of length $\nu$ between the vertices $i$ and $j$. For a simple graph, $\mathcal{N}_{ij}^{1}$ is one if there is an edge between vertices $i$ and $j$; otherwise it is zero. We now concretize the idea of faraway connections by defining a far edge. Let an edge $e_{ij}$ between vertices $i$ and $j$ be a far edge of order $\mu$ if it is an edge for which $\mathcal{N}_{ij}^{\mu+1}=0$ [6]. Let $\mu_{min}$ denote the minimal order of the far edge. We note that, $\mu_{min}$ satisfies the property that, for all $l \leq \mu_{min}, \mathcal{N}_{ij}^{l} \neq 0$.

Figure 1(a) shows an example of a far edge of minimal order one and Fig. 1(b) shows a typical graph with far edges and their orders. We note that none of the edges in a completely connected graph are far edges of any order, while all edges in a tree are far edges of all the orders. Henceforth we will assume that a far edge has minimal order one unless stated otherwise.

To generate small-world networks as well as other types of networks, we follow the procedure given in Ref. [3]. We start with a regular network consisting of a ring of $n$ vertices with edges connecting each vertex to its $k$ nearest neighbors. Each edge is rewired with probability $p$ avoiding multiple edges. The $p=1$ case corresponds to a random network. The networks obtained with $p \approx 0.01$ correspond to small-world networks [3].

We have generated several networks from the regular ($p=0$) to the random ($p=1$) case. For each network we cal-

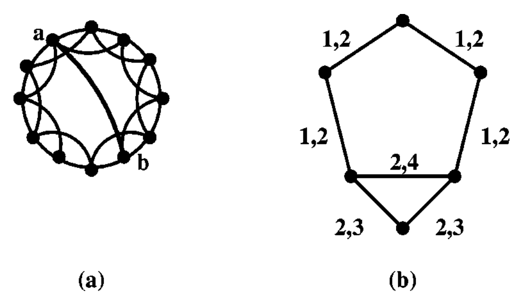


FIG. 1. Examples of networks consisting of far edges. In (a) the edge between vertices a and b is a far edge of order one, which is also its minimal order. (b) shows a general graph with far edges of different orders. The orders of the far edges are shown near each far edge.

*Electronic address: sagar@prl.ernet.in
†Electronic address: amritkar@prl.ernet.in

 

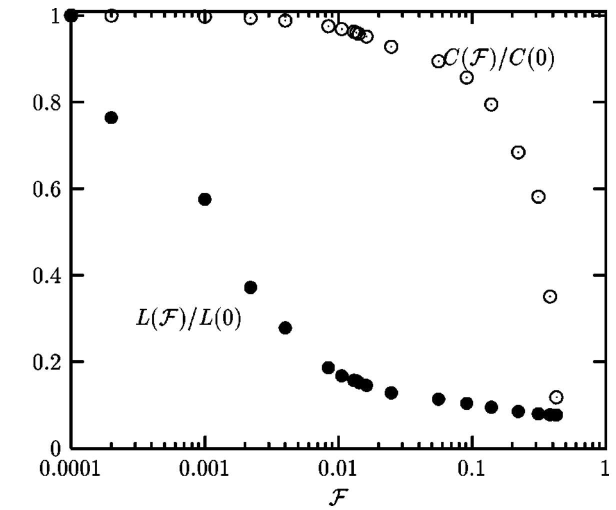


FIG. 2. Graph of $C(\mathcal{F})/C(0)$ and $L(\mathcal{F})/L(0)$ as functions of $\mathcal{F}$, where $C$ is the clustering coefficient, $L$ is the average path length, and $\mathcal{F}$ is the ratio of the number of far edges to the total number of edges. This figure is similar in nature to the plot of $C(p)/C(0)$ and $L(p)/L(0)$ as functions of $p$. The small-world networks lie around $\mathcal{F}=0.01$.

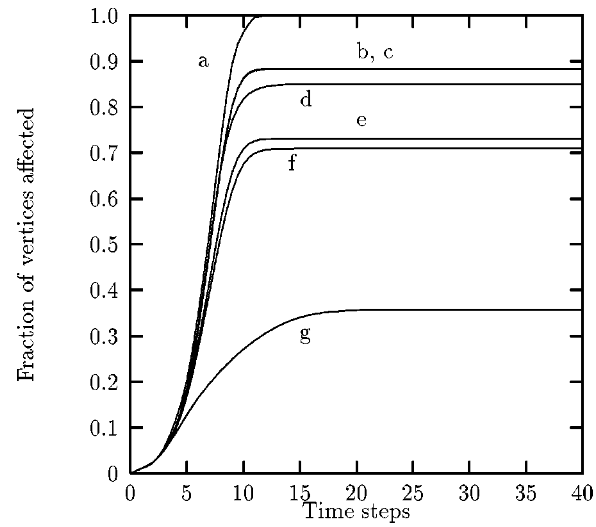


FIG. 3. Graphs of the fraction of vertices affected as a function of time steps. Curve (a) is the epidemic spread without immunization; curves (c) and (f) represent the spread when the random immunization is applied (see text) for $\tau=7$ and 2, respectively; curves (b) and (e) shows the spread if the immunization is carried out for the vertices with highest degree first and then in descending degree for $\tau=7$ and 2, respectively; curves (d) and (g) show the spread when the far edge immunization used is $\tau=7$ and 2, respectively. The simulations are carried out on a small-world network of 1000 vertices and 10 000 edges. The plotted results are averaged quantities over 500 seeds for an epidemic.

culate the average path length $L(p)$ and clustering coefficient $C(p)$. The quantity $L(p)$ denotes the average length of the shortest path between two vertices, and $C(p)$ denotes the average of $C_v$ over all the vertices $v$, where $C_v$ is the number of edges connecting the neighbors of $v$ normalized with respect to the maximum number of possible edges between these neighbors [3]. Next we determine the far edges in these networks. Let $\mathcal{F}$ denote the ratio of the number of far edges to the total number of edges. We find that initially, to a good approximation, $\mathcal{F}$ is equal to $p$ for $p\leqslant 0.1$ and then it increases slowly until it saturates to a value of about 0.4 for $p=1$. The saturation value depends on the chosen realization and could vary from 0.2 to 0.8. It turns out that the number of far edges of minimal order higher than one are negligible.

In Fig. 2 we plot $C(\mathcal{F})/C(0)$ and $L(\mathcal{F})/L(0)$ as functions of $\mathcal{F}$. This figure is similar in nature to the plot of $C(p)/C(0)$ and $L(p)/L(0)$ as functions of $p$ (Fig. 2 of Ref. [3]). The small-world networks can be identified as those with $C(p)/C(0)\approx 1$ and $L(p)/L(0)\approx L(1)/L(0)$. From Fig. 2 we see that this corresponds to $\mathcal{F}\approx 0.01$. Thus $\mathcal{F}$ can be used as a parameter with which to characterize networks that interpolate between regular and random cases. We note that unlike $p$, $\mathcal{F}$ is an intrinsic quantity. The quantity $\mathcal{F}$ is defined for any general a network and does not depend on any specific algorithm used for generating a network. Hence $\mathcal{F}$ should prove to be a better parameter than $p$.

To further investigate the importance of far edges, we consider the problem of spread of an epidemic [7]. Consider an epidemic starting from a random vertex (seed). We assume that at each time step all the neighbors of infected vertices are affected with probability one, which is the most infectious case, and the vertices that are already affected die and play no further role in the spread of the epidemic. Here, neighbors of a given vertex means all the vertices connected to it by edges. As found by Watts and Strogatz [3], the spread of an epidemic in small-world networks is almost as fast as that in the random case. We propose that the mechanism for the rapid spread of the epidemics in small-world networks is due to the traversal of the disease along the far edges. Each such traversal opens a virgin area for the spread of the epidemic, leading to rapid growth.

Clearly, if the far edges are responsible for the rapid growth of the epidemic, then we should be able to effectively control the spread by preventing the traversal of the epidemic along the far edges. To test this hypothesis, we propose the following mechanism to control an epidemic. We assume that we have sufficient knowledge of the network and we have identified all the far edges. We note that identification of far edges requires only the knowledge of vertices and edges and hence should be possible in many practical situations. Let $\tau$ denote the time steps that have elapsed between the beginning of the epidemic and its detection. Let $m$ denote the number of vertices that can be immunized at each time step. To block a far edge, we first immunize one of the two vertices connected by this far edge. Immunization is carried out by first blocking all the far edges and then immunizing at random. If the number of far edges is greater than $m$, then blocking all the far edges will take more than one time step.

In Fig. 3 we show the fraction of vertices affected as a function of time steps for a small-world network. Curve (a) shows the uncontrolled spread of the epidemic. Curves (d) and (g) show the spread of epidemic with the control method suggested above for $\tau=7$ and 2, respectively. For comparison, we show, by curves (c) and (f), the epidemic with only random immunization for $\tau=7$ and 2, respectively. It is obvious that the far edge control mechanism proposed here is very effective. For larger $\tau$ some of the far edges are already traversed by the epidemic, decreasing the efficiency of our control mechanism. Comparing the far edge immunization

RAPID COMMUNICATIONS



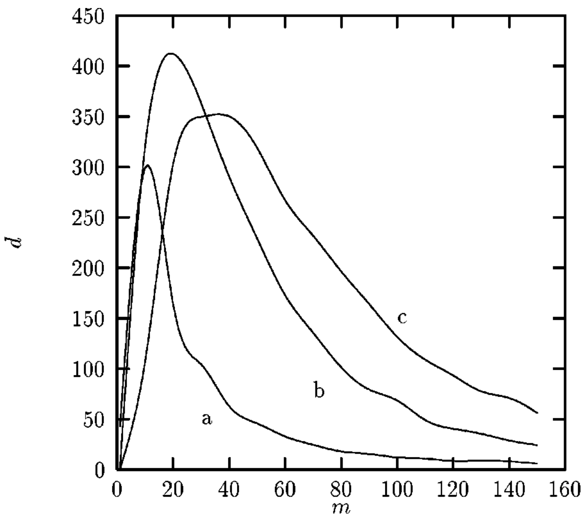


FIG. 4. Graph of the asymptotic difference between the number of affected vertices in random and far edge immunization, $d$, as function of number of vertices immunized in one time step, $m$. The three curves (a), (b), and (c) are for $\mathcal{F}=0.0022$, 0.0084, and 0.0162, respectively. Curve (b) corresponds to the small-world network. The other curves demonstrate the behavior of $\mathcal{F}$ on either side of the small-world network. The other parameters are as in Fig. 3.

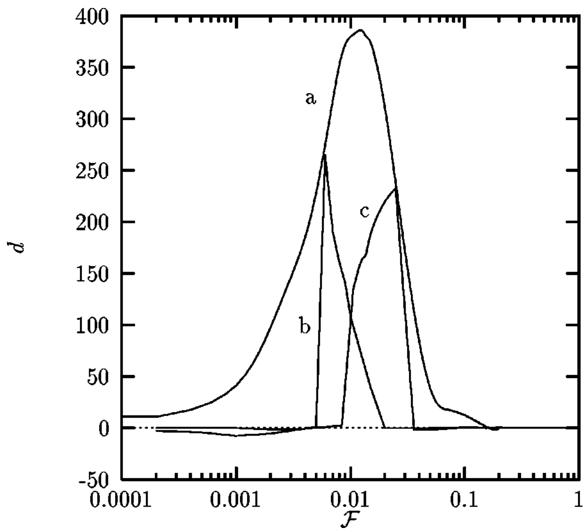


FIG. 5. Graph of $d$ as a function of $\mathcal{F}$. The three curves (a), (b), and (c) are plotted for $m=30$, 10, and 80, respectively. The figure shows that the immunization method suggested here is most effective in small-world networks.

and the random immunization, we find that the far edge immunization decreases the rate of spread of the epidemic more effectively but takes longer to completely stop the spread [see Fig. 3, curves (d) and (g)]. Further, to test the effectiveness of our method we compare the results with another method of immunization. We order the vertices by degree. Immunization is carried out by starting with the vertex with the largest degree and then going down from there. The results for $\tau=7$ and 2 are shown as curves (b) and (e) in Fig. 3, respectively. We note that results for immunization using degree are similar to that of the random immunization. This is an interesting result which shows that the degrees of the vertices do not play a significant role in the spread of the epidemic.

Let $d$ denote the asymptotic difference between the number of affected vertices in random and far edge immunization. We plot $d$ as a function of $m$ for three different values of $\mathcal{F}$ (or $p$) in Fig. 4. The plot shows that the far edge immunization is most effective when $m$ is about half the number of far edges. The reason for the decrease of $d$ for large $m$ is that the probability that random immunization blocks a far edge keeps on increasing as $m$ increases, thereby decreasing the difference between the two methods. The plot of $d$ as a function of $\mathcal{F}$ for different values of $m$ is shown in Fig. 5. The figure shows that the far edge immunization is more effective for small-world networks. Also from Figs. 4 and 5 it is clear that the far edge immunization offers a substantial benefit in terms of number of unaffected vertices in the small-world case and this number can be as large as 410, which is more than 40% of the total number of vertices.

Now, we consider an interesting model of product advertisement. Let $r$ be the number of vertices or centers from which a product is advertised. The information about the product spreads by word of mouth to the neighbors with the probability $q_t$, where $t$ is the time elapsed since the initial advertisement. We compare the results of two different ways of choosing the initial centers. In one way the centers are chosen at random and in the other they are chosen as one of the vertices in a far edge. Figure 6 shows the number of people informed about the product as a function of $t$. It is clear that the choice of centers using far edges has a definite advantage over that of random choice.

To conclude, we have introduced the concept of *far edges* in networks. Our definition of a far edge is in accordance with the intuitive idea of a far away connection between two vertices. The advantage of our definition of far edge is that it is independent of the underlying topology of the network (e.g., the underlying topology of the network in [3] is the topology of a circle and our definition does not depend on it).

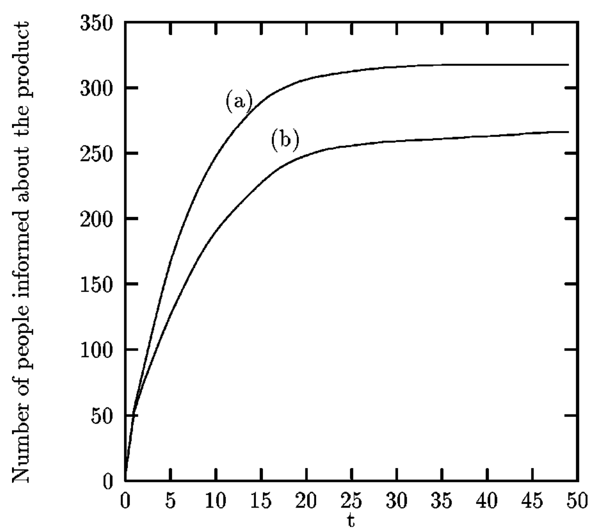


FIG. 6. Graph of the number of people informed as a function of $t$. Curves (a) and (b) show the result for far edge centers and random centers, respectively. The simulation is carried out on a small-world network with 1000 vertices and 10 000 edges. The initial advertising is done from five centers. The probability function $q_t$ is chosen as $q_1=0.8$ and $q_i=0.18$, where $i\geqslant 2$.

Also, the definition is algorithmic in nature and allows the determination of far edges only from the knowledge of vertices and edges. We have also applied the idea of far edges to the networks that are not generated by the algorithm given in Ref. [3] and arrived at similar conclusions [8].

We have demonstrated the use of far edges in the control of the spread of an epidemic and in advertisement of products [9]. Our simulations show that the far edges are indeed important in the spread of an epidemic, particularly in the small-world networks. It is also observed that the degrees of vertices do not play a significant role in the spread. We have shown that the knowledge of far edges can be fruitfully utilized to control the spread of an epidemic and to achieve better advertising. Our results strongly indicate that the far edges are the key elements responsible for the special properties of small-world phenomena.

[1] S. Milgram, Psychol. Today **2**, 60 (1967).

[2] J. Gaure, in *Six Degrees of Separation: A Play* (Vintage Books, New York, 1990).

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

[4] J. J. Hopfield and A. V. M. Herz, Proc. Natl. Acad. Sci. USA **92**, 6655 (1995).

[5] J. J. Collins and C. C. Chow, Nature (London) **393**, 409 (1998).

[6] From the definition of a far edge it is clear that for a far edge $e_{ij}$, of order $\mu$ there does not exist a path of length $\mu+1$ connecting vertices $i$ and $j$. Intuitively, the absence of a path implies farness; in our definition, it means the farness of a given order $\mu$.

[7] Although we consider the spread of an epidemic, the results are equally applicable for any quantity that spreads on a network through edges, e.g., the spread of rumors, information spread in neural networks, the spread of a virus in a computer network, the spread of a disturbance in an electrical network, etc.

[8] S. A. Pandit and R. E. Amritkar (unpublished).

[9] As pointed out in [7], the results are applicable for any situation in which a quantity spreads on a network through edges. In some cases complete information about the network may not be known. Even in these cases the definition of far edge is useful; e.g., if only a fraction of edges are known and with this information it turns out that some edge, say $e_{\{ij\}}$, is not a far edge, then after adding the information about remaining edges, $e_{\{ij\}}$ cannot become a far edge [8].