PHYSICAL REVIEW E **66**, 065103(R) (2002)

Range-based attack on links in scale-free networks: Are long-range links responsible for the small-world phenomenon?

Adilson E. Motter and Takashi Nishikawa
Department of Mathematics, Center for Systems Science and Engineering Research, Arizona State University, Tempe, Arizona 85287

Ying-Cheng Lai
Department of Mathematics, Center for Systems Science and Engineering Research, Arizona State University, Tempe, Arizona 85287
and Departments of Electrical Engineering and Physics, Arizona State University, Tempe, Arizona 85287

(Received 31 May 2002; published 30 December 2002)

The small-world phenomenon in complex networks has been identified as being due to the presence of long-range links, i.e., links connecting nodes that would otherwise be separated by a long node-to-node distance. We find, surprisingly, that many scale-free networks are more sensitive to attacks on short-range than on long-range links. This result, besides its importance concerning network efficiency and/or security, has the striking implication that the small-world property of scale-free networks is mainly due to short-range links.

DOI: 10.1103/PhysRevE.66.065103 PACS number(s): 89.75.Hc, 87.23.Ge, 89.20.Hh, 89.75.Da

Many real networks have been identified to have an amazingly small average shortest path since Watts and Strogatz (WS) [1] introduced their model of small-world networks. This model is constructed from a sparse regular network by rewiring a small fraction of links at random. Watts [2] introduced the concept of *range* to characterize different types of links: the range of a link l_{ij} connecting nodes i and j is the length of the shortest path between nodes i and j in the absence of l_{ij} (see also Ref. [3]). In this sense, typically, local connections are short-range links but rewired connections are long-range links. A key feature in the WS model is that it clearly identifies the small shortest paths observed in locally structured, sparse networks as being due to long-range connections, while short-range links are responsible for high clustering. This remarkable observation matches very well with the known results for the Erdös-Rényi (ER) model of random graphs [4], where almost all links are long-range connections and the average shortest path increases only logarithmically with the number N of nodes [5]. In regular networks, on the other hand, all the links have small range and the average shortest path increases with a power of N.

The WS and ER models explain some important features of real networks, such as the small-world phenomenon. However, since these models are homogeneous, their connectivity distribution $P(k)$, where k is the number of links connected to a node, has an exponential tail, in contrast to the algebraic one that characterizes scale-free networks recently discovered in a variety of real-world situations [6,7],

$$P(k) \sim k^{-\gamma}, \tag{1}$$

where γ is the scaling exponent. Scale-free networks are heterogeneous as their connectivity can vary significantly from node to node and a considerable number of links can be associated with a few highly connected nodes. Barabási and Albert (BA) identified in their seminal paper [6], growth with preferential attachment as the universal mechanism generating the algebraic behavior (1). As most scale-free networks possess the small-world property, it has been *tacitly* assumed that long-range connections are responsible for the small average shortest path exhibited by these networks. In addition to the insights provided by the WS model, the main argument for this comes from the observation that the removal of a link l_{ij} of range R increases the length of the shortest path between nodes i and j by $R-1$. The length of the shortest path between nodes connected by a short-range link is then robust against the removal of the link because the second shortest path between these two nodes is still short. But this is not true for long-range links, as they connect nodes that would otherwise be separated by a long shortest path.

Scale-free networks have attracted a tremendous amount of recent interest [7]. The aim of this paper is to investigate *explicitly* the contribution of short-range links to the small-world property in scale-free networks by analyzing the impact of attacks on short-range links versus those on long-range links. Attack here is defined as the deliberate removal of a subset of selected links. The importance of studying attacks on complex networks is twofold. First, it can identify the vulnerabilities of real-world networks, which can be used either for protection (e.g., of Internet) or for destruction (e.g., of metabolic networks targeted by drugs). Second, it provides guidance in designing more robust artificial networks (e.g., power grids). Different aspects of attacks on complex networks have been analyzed recently [3,8–13]. However, to our best knowledge, almost all the previous works consider attacks on nodes rather than on links, with very few exceptions [14,15].

To study *range*-based attacks on *links*, we consider the following models of scale-free networks: (1) semirandom model [16]; (2) BA model [6] and its generalization with aging [17]. In each case, we generate scale-free networks with the small-world property and a tunable scaling exponent. Because of the small-world property, one might intuitively think that these networks are much more sensitive to attacks on long-range than on short-range links. Surprisingly, our analysis and numerical computation show exactly the opposite for many scale-free networks. This result has an unexpected implication: short-range links are the vital ones for efficient communication between nodes in these net-

works. Our findings are based on the observation that the average shortest path is a global quantity that is mainly determined by links with large load, where the load of a link is defined as the number of shortest paths passing through the link [18,19]. For scale-free networks, with exponent γ in a finite interval around 3, due to heterogeneity, the load is on average larger for links with shorter range, making the short-range attack more destructive. For very large values of γ, the corresponding networks become homogeneous and, as a result, the opposite occurs.

For a given network, our attack strategy is as follows. We first compute the range for all the links. We then measure the *efficiency* of the network as links are successively removed according to their ranges: (i) for short-range attacks, links with shorter ranges are removed first; (ii) for long-range attacks, links with longer ranges are removed first [20]. In both cases, the choice among links with the same range is made at random. The efficiency is measured by the shortest paths between pairs of nodes. The shortest path between two given nodes i and j is defined as the minimal number d_{ij} of links necessary to follow from one node to the other. A convenient quantity to characterize the efficiency is then

$$E=\frac{2}{N(N-1)}\sum\frac{1}{d_{ij}}, \tag{2}$$

where the sum is over all $N(N-1)/2$ pairs of nodes. The network is more efficient when it has small shortest paths, which according to our definition corresponds to large E. Definition (2) was introduced in Ref. [21] to generalize the concept of small world, as it applies to any network regardless of its connectedness.

We first consider the semirandom model as follows. We start with N nodes $\{1,2,\ldots,N\}$ and a list of N integers representing their connectivities, i.e., the number of half-links at each node: $\{k_1,k_2,\ldots,k_N\}$, where $k_i \leqslant N-1$ and $\Sigma_{i=1}^{N}k_i$ is even. In the case of scale-free networks, this connectivity sequence is generated according to the algebraic distribution (1). Next, we pick up pairs of half-links at random and connect them to form a link and repeat this process until the last pair is connected, prohibiting self- and repeated links. In order to have nontrivial networks in the limits of small and large γ, we bound the connectivity so that $k_{min} \leqslant k_i \leqslant k_{max}$ for $i=1,2,\ldots,N$, where k_{min} and k_{max} are constant integers. For $\gamma\to\infty$, the network becomes a regular random graph, which is homogeneous with all the nodes having the same connectivity k_{min}. For $\gamma\to 0$, most of the links are associated with nodes with connectivity of the order of k_{max}, and the network becomes densely connected. The most interesting regime corresponds to intermediate values of γ because in this case, the network is highly heterogeneous but still sparse, having the number of links much smaller than $N(N-1)/2$. Consider then this case.

Employing the generating function formalism of Ref. [16], we have derived an approximate expression for the expected value of the shortest path between nodes with connectivity k_i and k_j,

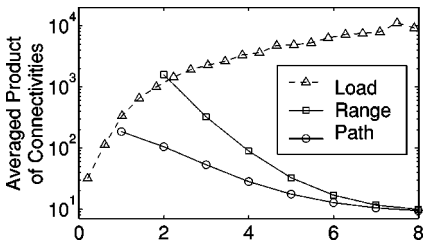

FIG. 1. Averaged product of connectivities as a function of the shortest path, range, and load for $\gamma=3$, where the load is binned and normalized by 10^4. Each curve corresponds to the average over ten realizations for $N=5000$, $k_{min}=3$, and $k_{max}=500$.

$$d_{ij}=\frac{\ln(Nz_1/k_ik_j)}{\ln(z_2/z_1)}+1, \tag{3}$$

where z_1 and z_2 are the average numbers of first and second neighbors, respectively. Accordingly, nodes with larger connectivity are on average closer to each other than those with smaller connectivity. The remarkable property of Eq. (3) is that d_{ij} depends only on the product of the connectivities k_i and k_j. This relation suggests that the range is also correlated with the product of the connectivities [22] so that short-range links tend to link together highly connected nodes, while long-range links tend to connect nodes with very few links. Moreover, links between nodes with large connectivities are expected to be passed through by a large number of shortest paths. That is, on average, these links should possess a higher load [15,23] than those connected to nodes with fewer links. These have been confirmed numerically, as shown in Fig. 1 for $\gamma=3$, where we plot the product of connectivities averaged over all pairs of nodes separated by a given shortest path length, or connected by a link with a given range or load.

Combining the above analyses for range and load, we observe that high load should be associated mainly with short-range links. With the understanding that links with higher load should contribute more to the shortness of the paths between nodes, this correlation between load and range implies that attacks on short-range links are more destructive than those on long-range links, in contrast to what one might naively think.

Now we present numerical verification of our main result concerning the effect of attacks on links. In Fig. 2, we show the efficiency (normalized by its initial value) for both short- and long-range attacks, for different values of γ. Notably, short-range attacks are more destructive than long-range ones for intermediate values of γ, as shown in Figs. 2(a) and 2(b) for $\gamma=3$ and $\gamma=5$, respectively. The corresponding relation between the average load and range, plotted in Fig. 3 for $\gamma=3$ (open circles), confirms that higher load on links with shorter range is the mechanism underlying this phenomenon. Long-range attacks become more destructive only for networks with sufficiently small or large values of γ. In Figs. 2(c) and 2(d), we show the results for $\gamma=2.5$ and $\gamma=\infty$, respectively. The exchange of the roles of attacks on short- and long-range links for networks with small values of γ is due to the appearance of a densely connected subnetwork of

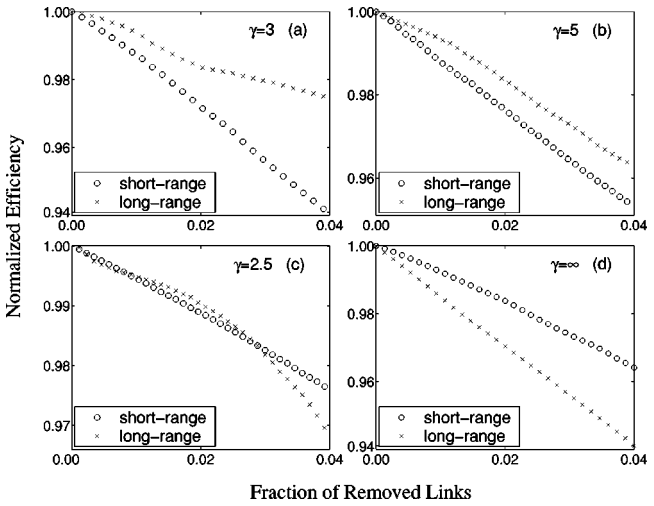

FIG. 2. Normalized efficiency for short- and long-range attacks as a function of the fraction of removed links. All the parameters other than γ are the same as in Fig. 1.

nodes with large connectivity. In this case, there are so many redundant short-range connections that the removal of one will not increase the average shortest path by much because, for a given pair of nodes, there are, in general, more than one path of minimal length which pass through *different* short-range links. For networks with large values of γ, switching of the roles of short- and long-range attacks is caused by the homogenization of the network. In a homogeneous network, all the nodes have approximately the same connectivity. Therefore, links with higher load are precisely those between distant nodes, i.e., those with larger range, as shown in Fig. 3 for $\gamma=\infty$ (open squares).

To demonstrate the generality of our results, we turn next to dynamic models of scale-free networks, where the algebraic scaling results from growth with preferential attachment, as observed in many realistic networks [6,7]. For concreteness, we consider the BA model [6] and its generalization with aging of nodes due to Dorogovtsev and Mendes [17]. The model is constructed as follows. We start at $t=0$ with N_0 nodes and zero links. At each successive time step, we add a new node with $m \leqslant N_0$ links so that each new link is connected to some old node i with probability $\Pi_i \sim \tau_i^{-\alpha}(k_i+1)$, where τ_i is the age of the node i and k_i is its connectivity. The standard BA model with scaling exponent $\gamma=3$ is recovered by taking $\alpha=0$. In general, scale-free networks with $\gamma>2$ are generated by choosing values of α in the interval $(-\infty,1]$ [17], where γ approaches the value of 2 as $\alpha \to -\infty$ and becomes infinite as $\alpha \to 1$.

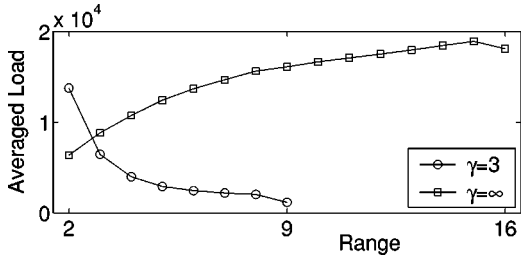

FIG. 3. Comparison between heterogeneous and homogeneous networks: averaged load as a function of the range for $\gamma=3$ and $\gamma=\infty$. All the parameters other than γ are the same as in Fig. 1.

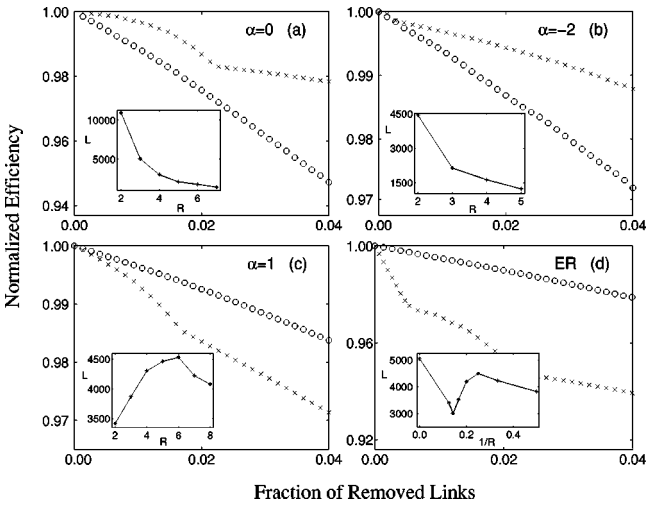

FIG. 4. Normalized efficiency for short-range attacks ($\bigcirc$) and long-range attacks ($\times$) as a function of the fraction of removed links. Each graph corresponds to the average over ten realizations for (a)–(c) $N=5000$, $N_0=3$, and $m=3$; (d) $N=5000$ and $z_1=6$. The corresponding relations between averaged loads L and ranges R are plotted in the insets. Observe that in inset (d), the horizontal axis is R^{-1}.

Most of the arguments and conclusions presented for the semirandom model are also valid for the growth model. In particular, the short-range attack is still expected to be more destructive than the long-range one at intermediate values of γ, while the opposite is expected for sufficiently large γ. However, there is an important difference for $2<\gamma<3$. Since new links come with new nodes, the subnetwork of highly connected nodes must be sparse. Accordingly, for this model, there will be no switching concerning the effect of short- versus long-range attacks at a small value of γ.

Our predictions are confirmed by numerical simulations, as shown in Fig. 4 for different values of $\alpha(\gamma)$. Indeed, short-range attacks are more destructive for $\alpha=0$ ($\gamma=3$) and also for $\alpha=-2$ ($\gamma \approx 2.3$), while long-range attacks are more destructive for $\alpha=1$ ($\gamma=\infty$). In all cases, the best strategy of attack is consistent with the correlation between load and range, as shown in the insets of Figs. 4(a)–4(c).

It is instructive to compare the results for scale-free networks with those for homogeneous networks with Poisson-like distribution of connectivities. In Fig. 4(d), we show the efficiency for the ER random model [24]. This network is more sensitive to attacks on long-range links because of the strong concentration of load on links with range infinity (see the inset). Incidentally, the long-range attack is also more destructive in the WS model [1], where the rewired connections tend to have higher load [25].

In summary, we have shown that for a wide interval of the scaling exponent γ, scale-free networks are more vulnerable to short- than long-range attacks, which results from a higher concentration of load on short-range links. In contrast to the load-based strategies of attacks considered in Ref. [15],

which are based on global information, short-range attacks are *quasilocal* in that, for a given range R, they require information only up to the $(R-1)^{th}$ neighbors [26]. Our findings have important implications that go beyond the issue of attack itself, as they provide insights into the structure and dynamics of scale-free networks. In particular, they show that short-range links are more important than long-range links for efficient communication between nodes, which is the opposite to what one might expect from other classes of small-world networks. For instance, in the network of sexual contacts, which is known to be scale-free [27], this means that the rapid spread of a disease may be mainly due to short-range contacts between people with large number of partners, in sharp contrast to its *homogeneous* counterpart [3].

This work was supported by AFOSR under Grant Nos. F49620-01-1-0317 and F49620-98-1-0400, and by NSF under Grant No. PHY-9996454.

[1] D.J. Watts and S.H. Strogatz, Nature (London) **393**, 440 (1998).
[2] D.J. Watts, *Small Worlds* (Princeton University Press, Princeton, 1999).
[3] S.A. Pandit and R.E. Amritkar, Phys. Rev. E **60**, R1119 (1999).
[4] P. Erdös and A. Rényi, Publ. Math. Inst. Hung. Acad. Sci. **5**, 17 (1960).
[5] B. Bollobás, *Random Graphs* (Academic Press, London, 1985).
[6] A.-L. Barabási and R. Albert, Science **286**, 509 (1999).
[7] R. Albert and A.-L. Barabási, Rev. Mod. Phys. **74**, 47 (2002).
[8] R. Albert, H. Jeong, and A.-L. Barabási, Nature (London) **406**, 378 (2000).
[9] D.S. Callaway, M.E.J. Newman, S.H. Strogatz, and D.J. Watts, Phys. Rev. Lett. **85**, 5468 (2000).
[10] A. Broder, R. Kumar, F. Maghoul, P. Raghavan, S. Rajagopalan, R. Stata, A. Tomkins, and J. Wiener, Comput. Netw. **33**, 309 (2000).
[11] R.V. Solé and J.M. Montoya, Proc. R. Soc. London, Ser. B **268**, 2039 (2001).
[12] R. Cohen, K. Erez, D. ben-Avraham, and S. Havlin, Phys. Rev. Lett. **86**, 3682 (2001).
[13] H. Jeong, S.P. Mason, A.-L. Barabási, and Z.N. Oltvai, Nature (London) **411**, 41 (2001).
[14] M. Girvan and M.E.J. Newman, Proc. Natl. Acad. Sci. U.S.A. **99**, 8271 (2002).
[15] P. Holme, B.J. Kim, C.N. Yoon, and S.K. Han, Phys. Rev. E **65**, 056109 (2002).
[16] M.E.J. Newman, S.H. Strogatz, and D.J. Watts, Phys. Rev. E **64**, 026118 (2001).
[17] S.N. Dorogovtsev and J.F.F. Mendes, Phys. Rev. E **62**, 1842 (2000).
[18] M.E.J. Newman, Phys. Rev. E **64**, 016132 (2001).
[19] K.-I. Goh, B. Kahng, and D. Kim, Phys. Rev. Lett. **87**, 278701 (2001).
[20] We choose to sort the links according to the initial distribution of ranges, instead of an updated distribution, because we want to address the relative importance of short-range and long-range links for the original network. In addition, in terms of attack efficiency, updating is time consuming.
[21] V. Latora and M. Marchiori, Phys. Rev. Lett. **87**, 198701 (2001).
[22] Indeed, the range of a link can be regarded as the length of the second shortest path between the nodes that are connected to the link. Since we are considering the semirandom model, for which everything other than the connectivity distribution is random, the length of the second shortest path should also be correlated with the product of connectivities.
[23] For pairs of nodes connected by $n \geqslant 1$ shortest paths, the contribution to the load due to each path is $1/n$.
[24] In this model, we start with N nodes and zero links. Then for each pair of nodes, with probability p, we add a link between them. The resulting network has on average $z_1 = p(N-1)$ links per node.
[25] The same tendency displayed in Figs. 2 and 4 was observed for larger fractions of removed links. In particular, short-range attack is still the most effective one for scale-free networks with scaling exponent around 3. We observe, however, that the removed fraction shown in these figures is already unrealistically large for many practical situations.
[26] Reference [15] also considers local strategies of attack.
[27] F. Liljeros, C.R. Edling, L.A.N. Amaral, H.E. Stanley, and Y. Aberg, Nature (London) **411**, 907 (2001).