PHYSICAL REVIEW E **66**, 065102(R) (2002)

Cascade-based attacks on complex networks

Adilson E. Motter[1,*] and Ying-Cheng Lai[1,2]
[1]*Department of Mathematics, Center for Systems Science and Engineering Research, Arizona State University, Tempe, Arizona 85287*
[2]*Departments of Electrical Engineering and Physics, Arizona State University, Tempe, Arizona 85287*
(Received 6 August 2002; published 20 December 2002)

We live in a modern world supported by large, complex networks. Examples range from financial markets to communication and transportation systems. In many realistic situations the flow of physical quantities in the network, as characterized by the loads on nodes, is important. We show that for such networks where loads can redistribute among the nodes, intentional attacks can lead to a *cascade* of overload failures, which can in turn cause the entire or a substantial part of the network to collapse. This is relevant for real-world networks that possess a highly heterogeneous distribution of *loads*, such as the Internet and power grids. We demonstrate that the heterogeneity of these networks makes them particularly vulnerable to attacks in that a large-scale cascade may be triggered by disabling a *single* key node. This brings obvious concerns on the security of such systems.

DOI: 10.1103/PhysRevE.66.065102 PACS number(s): 89.75.−k, 89.20.Hh, 05.10.−a

Complex networks are an essential part of a modern society [1,2]. It has been shown that many networks, such as the World Wide Web (WWW), the Internet, and electrical power grids, present a surprisingly small average distance between nodes and a highly organized distribution of links per node [3–5]. Generally, the average distance will not be affected by the removal of a random subset of nodes, but it will increase significantly if the removed nodes are among the most connected ones [3] (see also Refs. [6–8]). The existence of a giant connected component in the network, however, does not depend on the presence of highly connected nodes. For instance, the WWW has homepages with many thousands of hyperlinks and can remain well connected after the removal of all homepages with five or more hyperlinks [9]. In addition, the giant component itself is typically a *small-world network* [10] even after the removal of all highly connected nodes [11]. These pioneering studies on network security address mainly static properties, i.e., the effect of different network architectures. They suggest that the network connectivity, and hence its functionability, is robust against random failure of nodes [3,6,7] and to some extent is even robust against intentional attacks [9,11]. Here we show that for many physical networks, the removal of nodes can have a much more devastating consequence when the intrinsic *dynamics* of flows of physical quantities in the network is taken into account. In a power transmission grid, for instance, each node (power station) deals with a load of power. The removal of nodes, either by random breakdown or intentional attacks, changes the balance of flows and leads to a global redistribution of loads over all the network. This can trigger a cascade of overload failures [12,13], as the one that happened on August 10, 1996 in the western United States power grid [14,15]. Another example is the Internet [16–18], where the load represents the amount of information a node (router) is requested to transmit per unit of time, and overloads correspond to congestion [19]. Internet collapses caused by congestion have been reported since its very beginning [20]. In this Rapid Communication, we introduce a model for cascading failure in complex networks and show that it is applicable to realistic networks such as the Internet and power grids.

For a given network, suppose that at each time step one unit of the relevant quantity, which can be information, energy, etc., is exchanged between every pair of nodes and transmitted along the shortest path connecting them. The load at a node is then the total number of shortest paths passing through the node [21–23]. The capacity of a node is the maximum load that the node can handle. In man-made networks, the capacity is severely limited by cost. Thus, it is natural to assume that the capacity C_j of node j is proportional to its initial load L_j,

$$C_j=(1+\alpha)L_j, \quad j=1,2,\ldots N, \tag{1}$$

where the constant $\alpha\geqslant 0$ is the *tolerance* parameter, and N is the initial number of nodes. When all the nodes are on, the network operates in a free-flow state in so far as $\alpha\geqslant 0$. But, the removal of nodes, in general, changes the distribution of shortest paths. The load at a particular node can then change. If it increases and becomes larger than the capacity, the corresponding node fails. Any failure leads to a new redistribution of loads and, as a result, subsequent failures can occur. This step-by-step process is what we call a *cascading failure*, or a cascade. It can stop after a few steps but it can also propagate and shutdown a considerable fraction of the whole network [24]. A fundamental question is, under what conditions can such a global cascade take place?

Here we focus on cascades triggered by the removal of a single node. If a node has a relatively small load, its removal will not cause major changes in the balance of loads, and subsequent overload failures are unlikely to occur. However, when the load at the node is relatively large, its removal is likely to affect significantly loads at other nodes and possibly starts a sequence of overload failures. Our result is the following: global cascades occur if (1) the network exhibits a highly heterogeneous distribution of loads; (2) the removed node is among those with higher load. Otherwise, cascades are not expected. The distribution of loads is in turn highly correlated with the distribution of links: networks with het-

*Electronic address: motter@chaos3.la.asu.edu

erogeneous distribution of links are expected to be heterogeneous with respect to load so that on average, nodes with larger number of links will have higher load [22]. This result confirms the robust-yet-fragile property of heterogeneous networks, which was first observed in Ref. [3] for the attack on *several* nodes. The cascade effect is important, however, because a large damage can be caused in this case by the attack on a *single* node. While a network with more links can be more resistant against cascading failures, in practice the number of links is limited by cost.

Now we provide evidence for our result. We study cascades triggered by random breakdown and by intentional attacks. To simulate the former, we choose a trigger at random among all the nodes of the network, as can occur in networks such as power grids [14]. In the case of attack the targeted node is selected from those with highest loads or largest *degrees* (number of links at a node). We consider heterogeneous networks with algebraic (scale-free) distribution P of links, as observed in real systems [2,5,25,26],

$$P(k) \sim k^{-\gamma}, \tag{2}$$

where k denotes the degree and γ the scaling exponent, and compare them with an equivalent homogeneous configuration. These networks are generated according to the procedure in Refs. [27,28], where the nodes are connected randomly for a given degree distribution, and self- and repeated links are forbidden. The damage caused by a cascade is quantified in terms of the relative size G of the largest connected component,

$$G = N'/N, \tag{3}$$

where N and N' are the numbers of nodes in the largest component before and after the cascade, respectively. Figure 1 shows the relative size G of the largest component after cascading, as a function of the tolerance parameter α, for a scale-free network. We can see that on average G remains close to unity in the case of random breakdowns but it is significantly reduced under intentional attacks, even for α unrealistically large. Indeed, the size of the largest component is reduced by more than 20% for $\alpha = 1$, i.e., for a capacity as large as two times the capacity required for the system to operate when all the nodes function normally. This result is in agreement with intuition, because in the case of random breakdown the trigger is probably one of the many nodes with small load, while in the case of intentional attack it is a node with very large load. The damage is larger for smaller values of α, as it is for load-based attacks when compared with degree-based attacks. For instance, in the load-based attack for $\alpha = 0.2$, more than 60% of the nodes are affected. For the 5000-node networks used in our simulations, it means that a cascade triggered by the attack on a single node shuts down and disconnects more than 3000 others!

Figure 2 shows the corresponding results for a homogeneous network with the same number of nodes and exactly three links per node. To make a meaningful comparison we display in the inset results for an algebraic network with about the same average degree (actually larger, which strengthens our conclusions). The homogeneous network does not experience cascading failures due either to random breakdown or to intentional attacks for α as small as 0.05. For the heterogeneous (scale-free) network, for the same value of α, cascades triggered by the attack on a key node can reduce the largest connected component to less than 10% of the original size, as shown in the inset. Therefore, homogeneous networks appear to be more robust against attacks than the heterogeneous ones. This conclusion does not rely on the particularities of these models, as the same was also observed for classes of networks with exponential and Poisson-like distributions of degrees (e.g., the Erdös-Rényi model [29]): their homogeneity makes them relatively resistant to cascades triggered by attacks. The networks corre-

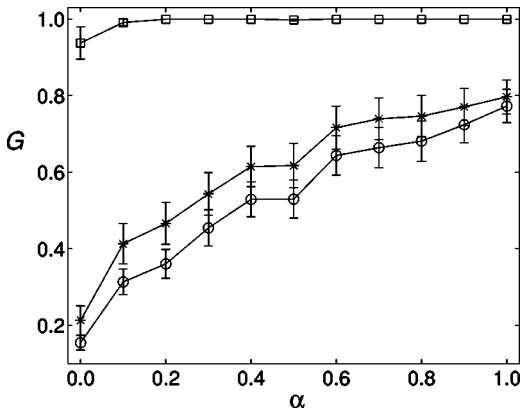

FIG. 1. Cascading failure in scale-free networks, as triggered by the removal of a single node chosen at random (squares), or among those with largest degrees (asterisks) or highest loads (circles), where α is the tolerance parameter and G is the relative size of the largest connected component. Each curve corresponds to the average over five triggers and ten realizations of the network. The error bars represent the standard deviation. The networks are generated according to the algebraic distribution (2). For the computations shown we set $\gamma = 3$ and $5000 \leqslant N \leqslant 5100$. The average degree in the largest component is $\langle k \rangle \approx 2.0$.

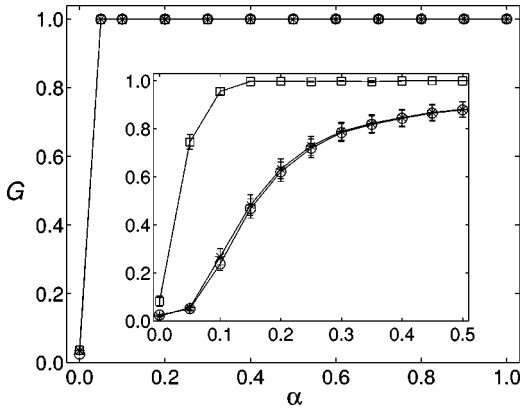

FIG. 2. Cascading failure in homogeneous networks. All nodes are set to have the same degree $k = 3$ and $N = 5000$. In the inset, the networks are generated according to the algebraic distribution (2) for $k \geqslant 2$, $\gamma = 3$, and $N = 5000$. The resulting average degree is $\langle k \rangle \approx 3.1$. The legends and other parameters are the same as in Fig. 1.

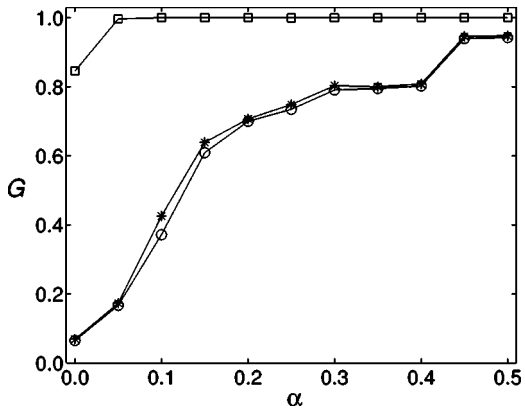

FIG. 3. Cascading failure in the Internet at autonomous system level [30]. The network has $N=6474$ nodes and $\langle k\rangle \approx 3.88$ links per node, on average. Each curve corresponds to the average over 5 triggers for attacks and 50 for random breakdown. The legends are as defined in Fig. 1.

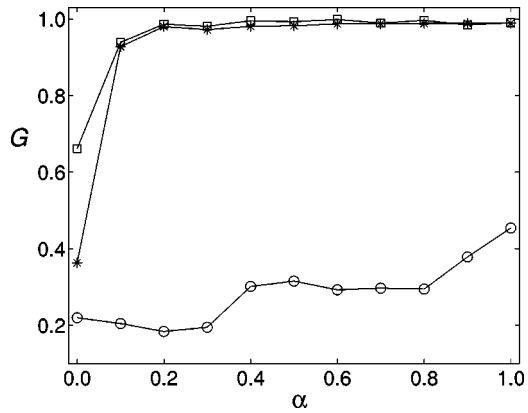

FIG. 4. Cascading failure in the western U.S. power transmission grid [31], which has $N=4941$ and $\langle k\rangle \approx 2.67$. The average is obtained via 5 triggers for attacks and 50 for random breakdown. The legends are the same as in Fig. 1.

sponding to the inset of Fig. 2 are generated according to the same scaling distribution of those in Fig. 1, except that in this case the minimal number of links at a node is set to be 2. Therefore, this inset shows that the fragility of scale-free networks is due to their heterogeneity and does not rely on the presence of nodes with degree one, which are easily disconnectable. Naturally, the increase of the average degree reduces the damage of the cascade, as can be seen from a comparison between Fig. 1 and the inset of Fig. 2.

Many real-world networks are heterogeneous and as such are expected to undergo large-scale cascades if some vital nodes are attacked, but rarely in the case of random breakdown. As an example we consider the Internet at autonomous system level [30], which displays an algebraic distribution of links [3]. The damage caused by triggers of higher load or degree is much larger than that by random breakdown, as shown in Fig. 3. The cascading failures are rarely triggered by random breakdown for $\alpha > 0.05$, but more than 20% of the nodes can be disconnected with the intentional attack on only one node for $\alpha \leqslant 0.4$. We have also considered the electrical power grid of the western United States [31]. The degree distribution in this network is consistent with an exponential [32] and is thus relatively homogeneous. The distribution of loads, however, is more skewed than that displayed by semirandom networks [27,28] with the same distribution of links, indicating that the power grid has structures that are not captured by these models. As a result, global cascades can be triggered by load-based intentional attacks but not by random or degree-based removal of nodes, as shown in Fig. 4. We see that the attack on a single node with large load reduces the largest connected component to less than a half of its initial size, even when the network is highly tolerant (e.g., $\alpha=1$).

Our result is thus that real networks are naturally evolved to be quite resistant to random failure of nodes, but the presence of a few nodes with exceptionally *large load*, which is known to be ubiquitous in natural and man-made networks, has a disturbing side effect: the attack on a single important node (one of those with high load) may trigger a cascade of overload failures capable of disabling the network almost entirely. Such an event has dramatic consequences on the network performance, because the functionability of a network relies on the ability of the nodes to communicate efficiently with each other. What is the use, say, of having a phone if you cannot call anybody?

We conclude with some thoughts on the meaning of our results for security. An effective attack relies on identifying vulnerabilities and is far from being random. Our society is geographically distributed in a way that natural hazards are by no means random [33]. An example is the crowding of people, communication, transportation, and financial centers around seismic areas, like the Pacific Rim. Natural disasters and intentional attacks can then have devastating consequences on the complex networks underlying the society. These consequences will be more severe if the damage on one or few nodes is capable of spreading over the entire network. In this sense a cascade-based attack can be much more destructive than any other strategies of attack previously considered [3,7–9,28,34–36].

The authors thank Réka Albert and Duncan J. Watts for providing the Internet and power-grid data, respectively. This work was supported by AFOSR under Grant No. F49620-98-1-0400 and by NSF under Grant No. PHY-9996454.

[1] S.H. Strogatz, Nature (London) **410**, 268 (2001).
[2] R. Albert and A.-L. Barabási, Rev. Mod. Phys. **74**, 47 (2002).
[3] R. Albert, H. Jeong, and A.-L. Barabási, Nature (London) **406**, 378 (2000).
[4] D.J. Watts and S.H. Strogatz, Nature (London) **393**, 440 (1998).
[5] A.-L. Barabási and R. Albert, Science **286**, 509 (1999).
[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] A. Broder, R. Kumar, F. Maghoul, P. Raghavan, S. Rajagopalan, R. Stata, A. Tomkins, and J. Wiener, Comput. Netw. **33**, 309 (2000).
[10] D. J. Watts, *Small Worlds: The Dynamics of Networks Between Order and Randomness* (Princeton University Press, Princeton, 1999).
[11] A. P. S. de Moura, Y.-C. Lai, A. E. Motter, and P. Dasgupta (unpublished).
[12] D.J. Watts, Proc. Natl. Acad. Sci. U.S.A. **99**, 5766 (2002).
[13] Y. Moreno, J.B. Gómez, and A.F. Pacheco, Europhys. Lett. **58**, 630 (2002).
[14] B. A. Carreras, D. E. Newman, I. Dolrou, and A. B. Poole, in Proceedings of Hawaii International Conference on System Sciences, Maui, Hawaii, 2000 (unpublished).
[15] M.L. Sachtjen, B.A. Carreras, and V.E. Lynch, Phys. Rev. E **61**, 4877 (2000).
[16] R. Pastor-Satorras, A. Vázquez, and A. Vespignani, Phys. Rev. Lett. **87**, 258701 (2001).
[17] W. Willinger, R. Govindan, S. Jamin, V. Paxson, and S. Shenker, Proc. Natl. Acad. Sci. U.S.A. **99**, 2573 (2002).
[18] K.-I. Goh, B. Kahng, and D. Kim, Phys. Rev. Lett. **88**, 108701 (2002).
[19] R. Guimerà, A. Arenas, A. Días-Guilera, and F. Giralt, Phys. Rev. E **66**, 026704 (2002).
[20] V. Jacobson, Comput. Commun. Rev. **18**, 314 (1988).
[21] M.E.J. Newman, Phys. Rev. E **64**, 016132 (2001).
[22] K.-I. Goh, B. Kahng, and D. Kim, Phys. Rev. Lett. **87**, 278701 (2001).
[23] P. Holme and B.J. Kim, Phys. Rev. E **65**, 066109 (2002).
[24] A different model and mechanism for overload breakdown in growing networks has been considered by Holme and Kim in Ref. [23]. These authors focus on overloads caused by the growth of the network. Their model assigns the same capacity to every node in the network. In their analysis, when a node is overloaded, the links to that node are removed, but the node itself is not removed and can be reconnected in the future. Their conclusion is that, to avoid overloads, the capacity must grow with the size of the network. Our model is different from the model in Ref. [23] as we assume the capacity to be *node dependent* and the failed nodes to be *permanently removed* from the network. More importantly, we address the issues of *intentional attack* and *random breakdown*, and we study how the network collapses under overload failures induced by them. We assume that the time scale for these events is much smaller than the time scale in which the network grows.
[25] S. Redner, Eur. Phys. J. B **4**, 131 (1998).
[26] M. Faloutsos, P. Faloutsos, and C. Faloutsos, Comput. Commun. Rev. **29**, 251 (1999).
[27] M.E.J. Newman, S.H. Strogatz, and D.J. Watts, Phys. Rev. E **64**, 026118 (2001).
[28] A.E. Motter, T. Nishikawa, and Y.-C. Lai, Phys. Rev. E (to be published).
[29] P. Erdös and A. Rényi, Publ. Math. Inst. Hung. Acad. Sci. **5**, 17 (1960).
[30] http://moat.nlanr.net/AS/Data/ASconnlist.20000102.946809601
[31] ftp://ftp.santafe.edu/pub/duncan/power_unweighted
[32] L.A.N. Amaral, A. Scala, M. Barthélémy, and H.E. Stanley, Proc. Natl. Acad. Sci. U.S.A. **97**, 11149 (2000).
[33] D. Kennedy, Science **295**, 405 (2002).
[34] R.V. Solé and J.M. Montoya, Proc. R. Soc. London, Ser. B **268**, 2039 (2001).
[35] H. Jeong, S.P. Mason, A.-L. Barabási, and Z.N. Oltvai, Nature (London) **411**, 41 (2001).
[36] P. Holme, B.J. Kim, C.N. Yoon, and S.K. Han, Phys. Rev. E **65**, 056109 (2002).