Attack Induced Cascading Breakdown in Complex Networks

Liang Zhao[1], Kwangho Park[2], Ying-Cheng Lai[2] & Thiago Henrique Cupertino[1]

[1]Institute of Mathematics and Computer Science
University of São Paulo
Av. Trabalhador São-Carlense, 400
Phone: +55 (16) 33739713
FAX: +55 (16) 33739751
P.O.Box: 668, Zip 13560-970 - São Carlos - SP - BRAZIL
zhao@icmc.usp.br and thiagohc@grad.icmc.usp.br

[2]Department of Electrical Engineering
Arizona State University
Tempe, Arizona 85287, USA

Abstract

The possibility that a complex network can be brought down by attack on a single *or very few nodes through the process of cascading failures is of significant concern. In this paper, we investigate cascading failures in complex networks and uncover a phase-transition phenomenon in terms of the key parameter characterizing the node capacity. For parameter value below the phase-transition point, cascading failures can cause the network to disintegrate almost entirely. Then we show how to design networks of finite capacity that are safe against cascading breakdown. Our theory yields estimates for the maximally achievable network integrity via controlled removal of a small set of low-degree nodes.*

Keywords: complex networks, scale-free networks, power grids, computer networks, degree distribution, cascading breakdown.

1. INTRODUCTION

Complex networks arise in natural systems and they are also an essential part of modern society. Many real complex networks, such as the World Wide Web (WWW), the Internet, and some electrical power grids, were found to be heterogeneous with power-law degree distribution [1, 2, 3, 4, 5, 6, 7] which means that the probability for a subset of nodes to possess a large number of links is not exponentially small, in contrast to random networks. Because of the ubiquity of scale-free networks in natural and man-made systems, the security of these networks, i.e., how failures or attacks affect the integrity and operation of the networks, has been of great interest.

In network security study, two main points should be taken into account: one is architecture of the network; another is dynamics in the network, i.e., how information or load is distributed in the network. An intuitive reasoning based on the load distribution would suggest that, for a scale-free network, the possibility of breakdown triggered by attack on even only a single node cannot be ignored. Imagine such a network that transports some physical quantities, or load. Nodes with large numbers of links receive relatively heavier load. Each node, however, has a finite capacity to process or transport load. In order for a node to function properly, its load must be less than the capacity at all time; otherwise the node fails. If a node fails, its load will be directed to other nodes, causing a redistribution of load in the network. If the failing node deals with a small amount of load, there will be little effect on the network because the amount of load that needs to be redistributed is small. This is typically the situation of random failure of nodes. However, if the failing node carries a large amount of load, the consequence could be serious because this amount of load needs to be redistributed and it is possible that for some nodes, the new load exceeds their capacities. These nodes will then fail, causing further redistributions of load, and so on. As

a consequence, a large fraction of the network can be shutdown.

There have been many studies on cascading failures in engineering networks (for a brief review, see Sect. 3). In Ref. [8], a simple mechanism was proposed to incorporate the dynamics of load in complex networks. The model, which is referred to *ML model*, generates results that are consistent with the above intuition on cascading failures. For instance, it was demonstrated that random networks are robust against cascading breakdown but they can be easily triggered by intentional attacks in scale-free networks. The existing results are, however, largely descriptive and qualitative. The purpose of this work is to address theoretically and numerically the fundamental mechanism of cascading breakdown. In this work, we study cascading failures in complex networks with focus on scale-free networks by using the ML model. Our finding is that cascading breakdown in scale-free networks can be understood in terms of a phase transition. In particular, let α be the tolerance parameter characterizing the capacity of nodes in the network. Cascading breakdown due to attack on a single node is possible only when α is below a critical value α_c. By making use of the degree distribution of scale-free networks and the concept of betweenness [9] to characterize the load distribution, we are able to derive a theoretical formula for estimating the phase-transition point α_c, which is verified by numerical experiments. In terms of practical utility, our result enables a possible implementation of predicting and preventing mechanism for cascading breakdown in scale-free networks.

The rest of this paper is organized as follows. In Sect. 2, a brief review on complex network models is given. In Sect. 3, we discuss cascading failure in engineering and computer networks. Sect. 4 is devoted to analyze cascading failure in complex networks with focus on scale-free networks. Sect. 5 is devoted to describe and analyze the protection mechanism for scale-free networks against cascading breakdown. Finally, Sect. 6 concludes the paper.

2. COMPLEX NETWORK MODELS

In recent years, many complex network models have been discovered or developed. Here, we briefly review some classical models, such as *Random Graphs*, *Small-World Networks* and *Scale-Free Networks*, and some newly developed models, such as *Layered Networks* and *Co-evolution Networks*.

The systematic study of random graphs was initiated by Erdös and Rényi [10]. The term random graph refers to the disordered nature of the arrangement of links between different nodes. In their article, Erdös and Rényi proposed a model to generate random graphs with N nodes and K links, that we will henceforth call Erdös and Rényi (ER) *Random Graphs*. The model for ER random graphs consists in connecting each couple of nodes with a probability $0 < p < 1$. ER random graphs are one of the best studied among graph models, although they do not reproduce most of the properties of real networks. The great discovery of Erdös and Rényi was that many important properties of random graphs appear quite suddenly. That is, at a given probability either almost every graph has some property Q or, conversely, almost no graph has it. For many such properties there is a critical probability $p_c(N)$. If $p(N)$ grows more slowly than $p_c(N)$ as $N \to \infty$, then almost every graph with connection probability $p(N)$ fails to have Q. If $p(N)$ grows somewhat faster than $p_c(N)$, then almost every graph has the property Q, i.e.,

$$\lim_{N\to\infty} P_{N,p}(Q) = \begin{cases} 0 & if \quad \frac{p(N)}{p_c(N)} \to 0 \\ 1 & if \quad \frac{p(N)}{p_c(N)} \to \infty \end{cases}$$

A few important special cases are:

1. The critical probability of having a tree of order k if $p_c(N) = cN^{-\frac{k}{k-1}}$;
2. The critical probability of having a cycle of order k if $p_c(N) = cN^{-1}$;
3. The critical probability of having a complete subgraph of order k if $p_c(N) = cN^{-\frac{2}{k-1}}$.

Since all the nodes in a random graph are statistically equivalent, each of them has the same distribution, and the probability that a node chosen uniformly at random has degree k has the same form as $P(k_i = k)$. For large N, and fixed average degree $\langle k \rangle$, the degree distribution is well approximated by a Poisson distribution:

$$P(k) = e^{-\langle k \rangle} \frac{\langle k \rangle^k}{k!}$$

For this reason, ER graphs are sometimes called Poisson random graphs. Figure 1 shows the degree distribution of an ER graph. The network is generated by the method shown in [11].

Real networks are rarely pure random networks. The study of several dynamical processes over real networks has pointed out the existence of shortcuts, i.e. bridging links that connect different areas of the networks, thus speeding up the communication among distant nodes, which is called *Small-World* effect. The small-world property in real networks is often associated with the presence of clustering, denoted by high values of the *clustering coefficient* in a network G, defined by

$$C = \frac{3 \times number\ of\ triangles\ in\ G}{number\ of\ connected\ triples\ of\ vertices\ in\ G}$$

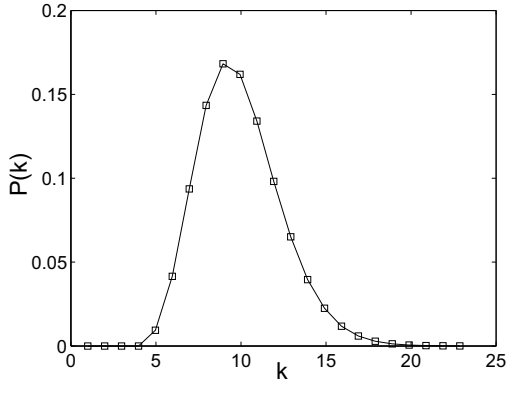

Figure 1. Degree distribution of a random network. The number of nodes $N = 20000$ and average degree $\langle k \rangle = 10$.

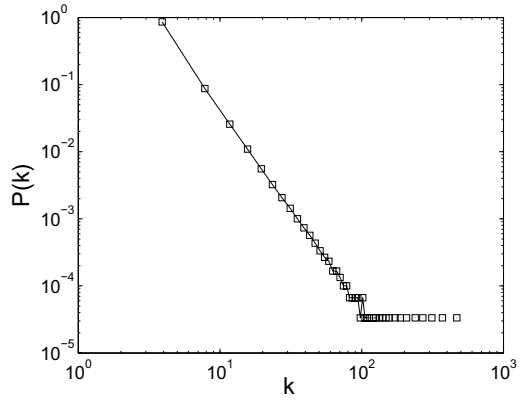

Figure 3. Degree distribution of a scale-free network. The number of nodes $N = 20000$ and average degree $\langle k \rangle = 4$.

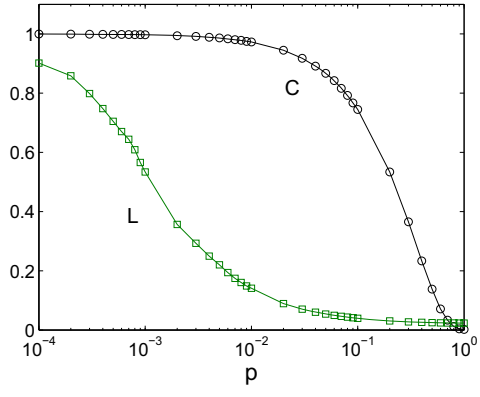

Figure 2. The average length of shortest paths L and clustering coefficient C in relation to p. The number of nodes $N = 20000$ and average degree $\langle k \rangle = 4$.

For this reason, Watts and Strogatz, in their seminal paper [11], have proposed to define small-world networks as those networks having both a small value of average shortest path length L, like random graphs, and a high clustering coefficient C, like regular lattices. Clustering, also known as transitivity, is a typical property of acquaintance networks, where two individuals with a common friend are likely to know each other. Figure 2 shows the average length of shortest paths L and the clustering coef ficient C by randomly changing a fraction p of links starting with a regular network. We see that the small-world effect occurs when p is small, for example, $p = 0.01$. Here, the networks are again generated by the model presented in [11].

Until a few years ago, network study was focalized on homogeneous networks. Homogeneity means that almost all nodes are topologically equivalent, like in regular lattices or in random graphs. In contrast with all the expectancies, when the scientists approached the study of real networks from the available databases, it was found that most of the real networks were found to be heterogeneous with power-law degree distribution [1]: $P(k) \sim k^{-\gamma}$, where k is the number of links of a randomly chosen node in the network and γ is the scaling exponent (see Fig. 3). The scale-free network is generated by using the standard Barabási-Albert model [1]. This power-law distribution means that the probability for a subset of nodes to possess a large number of links is not exponentially small, in contrast to random networks. Mathematically, the power-law distribution means that statistical moments of the degree variable are generally not defined, hence the name of *Scale-Free Networks*. Examples range from the Internet, WWW, protein-protein interaction networks, some power grids, telecommunication networks, traffic flow in social networks [2, 3, 4, 5, 6, 7])

In the study of scientific collaboration networks, Newman [9] found that the degree distribution of such networks do not follow a power-law form perfectly. However, the data are well fitted by a power-law form with an exponential cutoff:

$$P(k) \sim k^{-\tau} e^{-\frac{k}{\kappa}}$$

where τ and κ are constants. This form is commonly seen in physical systems, and implies an underlying degree distribution which follows a power-law, but with some imposed constraints that place a limit on the maximum value of *k*. This kind of networks is called *Scale-Free with Exponential Cutoff*.

Many complex networks are only a part of larger systems, where a number of coexisting topologies interact and depend on each other. In Ref. [12], the authors studied the load distribution in three transportation systems, where the lower layer is the physical infrastructure and the upper layer represents the traffic flows, which is called *Layered Complex Networks*. The layered view allows us to capture different features of the same system. For in-

stance, the topologies of the Internet at the IP layer, of the World Wide Web (WWW), or of the networks formed by peer to peer (P2P) applications. Each WWW or P2P link virtually connects two IP nodes. These two IP nodes are usually distant in the underlying IP topology, and the virtual connection is realized as a path found by IP routers. In other words, the graph formed by an application is mapped on the underlying IP network.

In Ref. [13], the authors introduced a co-evolution networks, called *TARL Model* (for topics, aging, and recursive linking) that simultaneously grows coauthor and paper citation networks. The statistical and dynamic properties of the networks generated by this model are validated against a 20-year data set of articles published in PNAS. Systematic deviations from a power law distribution of citations to papers are well fitted by a model that incorporates a partitioning of authors and papers into topics, a bias for authors to cite recent papers, and a tendency for authors to cite papers cited by papers that they have read.

3. CASCADING FAILURE IN ENGINEERING AND COMPUTER NETWORKS

Cascading failures can occur in many physical systems. In a power transmission grid, for instance, each node (a generator) deals with a load of power. Removal of nodes in general can cause redistribution of loads over all the network, which can trigger a cascade of overloading failures. The recent massive power blackout caused by a series of seemingly unrelated events on August 14, 2003 in the northeastern United States and Canada seemed to have the characteristics of cascading breakdown.

Many studies have been performed on cascading failure in power transmission systems. Some recent results are briefly reviewed here. DeMarco studied cascading failure due to dynamic transients in bistable systems represented by nonlinear differential equations [14]. Carreras et. al. [15] represented cascading failure in a power system model using the DC load flow approximation and standard linear programming optimization of the generation dispatch. The model shows critical point behavior as load is increased and can show power tails similar to those observed in blackout data. Rios et. al. [16] proposed a method based on Monte Carlo simulation and taking into account time-dependent phenomena such a cascade tripping of elements due to overloads, malfunction of the protection system, and potential power system instabilities. In [17], the authors proposed risk indices for power systems, referred to online risk-based security assessment, which provide the ability to compute online probabilistic risk associated with conditions up to several hours in the future, as well as monitoring overload and cascading overload of networks. In [18], the authors studied a 15-year time series in the North American power transmission grid. They calculated long time correlations and probability distribution functions for several measurements of blackout size which show the sandpile phenomenon, known as Self-Organized Criticality (SOC). Dobson et. al. [19] proposed an analytically tractable model of loading-dependent cascading failure that captures some of the salient features of large blackouts of electric power transmission systems. A critical loading is revealed at which there is a power-law region in the distribution of number of components failed and a sharp increase in the gradient of the mean number of components failed.

Another example of cascading failure is the Internet, where the load represents data packets that a node (router) is requested to transmit and overloading corresponds to congestion [20]. The rerouting of data packets from a congested router to another may spread the congestion to a large fraction of the network. With the possibility of cascading failures, a realistic concern is attacks on complex networks. In particular, for a scale-free network, majority of the nodes deal with small amount of load, so the probability for a node with a large amount of load to fail randomly is small. This, of course, will not be the case of intentional attacks that usually target one or a few of the most heavily linked nodes. The work by Albert et al. [21] demonstrated that scale-free networks possess the robust-yet-fragile property, in the sense that they are robust against random failures of nodes but fragile to intentional attacks. Cohen et. al. [22, 23] studied internet breakdown by random failure and intentional attack. In their works, a scale-free network can become disintegrated under attacks on a small but still appreciable set of nodes that include a substantial fraction of links in the network. Attack on a single or very few nodes will in general not bring down the network. This result was actually obtained based purely on the scale-free architecture of the network. In other words, dynamics in the network, i.e., how information or load is distributed in the network, was not taken into account.

According to Dobson et. al. [19], cascading process in power grids consists of the following essential components: components that fail when their load exceeds a threshold, an initial disturbance loading of the system, and the additional loading of components by the failure of other components. In the ML model [8], the load dynamic is represented by betweenness of the node (total number of shortest paths passing through the node). In fact, several papers have used betweenness as a measure for the vertex load in dynamical communications systems (see, for example, [25]). When an attack or random failure occurs, some nodes are removed from the network. This effect can cause redistribution of loads over all the net-

work and it is possible that for some nodes, the new load exceeds their designed capacities. These nodes will then fail, causing further redistributions of load, and so on. As a consequence, a large fraction of the network can be shutdown, a cascade of overloading failures. One can see that the ML model captures the essential components of cascading process. Moreover, it takes into account the network architecture by means of betweenness, which allow us to study cascading failure in different network topology (such as scale-free networks) and by different origins (attack or random failure). This is an important feature since many real networks, including the Internet, WWW and some power grids are scale-free networks.

4. CASCADING FAILURE IN THE ML MODEL

In the ML model [8, 26, 27, 28], the load (or betweenness) at a node i is defined as the total number of shortest paths passing through this node. 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_i of node i is proportional to its initial load $L_i(0)$,

$$C_i = (1+\alpha)L_i(0) = \lambda L_i(0), \qquad (1)$$

where the constant $\alpha \geq 0$ (or $\lambda \geq 1$) is the tolerance parameter, i is the node index, and (0) represents the process is at time 0 (before attack).

Specifically, the cascading process can be described by the algorithm shown by Fig. 4. When all nodes are on, the network operates in a free flow state insofar as $\alpha \geq 0$. 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 node fails. Any failure leads to a new distribution of load and, as a result, subsequent failures can occur. The failures can stop without affecting too much the connectivity of the network but they can also propagate and shutdown a considerable fraction of the whole network.

Cascading failures can be conveniently quantified by the relative size of the largest connected component

$$G = \frac{N'}{N}, \qquad (2)$$

where N and N' are the numbers of nodes in the largest component before and after the cascade, respectively. Thus, we have $0 \leq G \leq 1$. The integrity of the network is maintained if $G \approx 1$, while breakdown (cascading failure) occurs if $G \approx 0$. If G is far from both 0 and 1, the phenomenon is referred to partial failure.

To gain a qualitative view of cascading failure, we first provide some computer simulation results. We generate

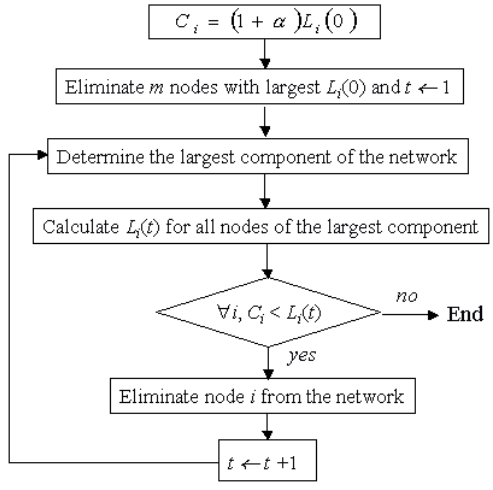

Figure 4. Algorithmic description of cascading process in the ML model.

scale-free networks by using the standard Barabási-Albert model [1]. The shortest paths and the load L_i are computed by using the algorithm developed by Newman [29]. Figure 5 shows G versus both α and $N_{trigger}$ for scale-free networks under attack, where $N_{trigger}$ represents the number of nodes that an attack targets. Here the removed nodes are those with the highest numbers of links. Note that the number of targeted nodes, while more than one, is still far small compared with the total number of nodes in the network. Practically, this means that, even if the network is designed to have a high tolerance by stipulating high capacities for its nodes, cascading failures triggered by attack on a very small subset of nodes are capable of bringing down the entire network. Partial failures can be observed too in the figure. Figure 6 shows G versus both α and $N_{trigger}$ for random networks under attack. Due to the lack of hubs (nodes with large number of links), random networks can be disintegrated only when α is small or (and) $N_{trigger}$ is large. Thus, cascading process under attack can bring much more serious consequence for scale-free networks than for random networks.

Figure 7 shows the random failure case for scale-free networks. Random failure means that m randomly selected nodes are removed from the network at the beginning of cascading process. In a scale-free network, a small portion of nodes has high degree, while a large portion of nodes has low degree. Thus, with high probability, a randomly selected node is low degree, usually having a low value of betweenness. Intuitively, there will be little effect on the network because the amount of load that needs to be redistributed is small. This point is confirmed by the simulation result shown by Fig. 7 where $G \sim 1$ even if α is small and $N_{trigger}$ is large.

To obtain an analytic estimate of the critical value of the tolerance parameter, we focus on scale-free networks

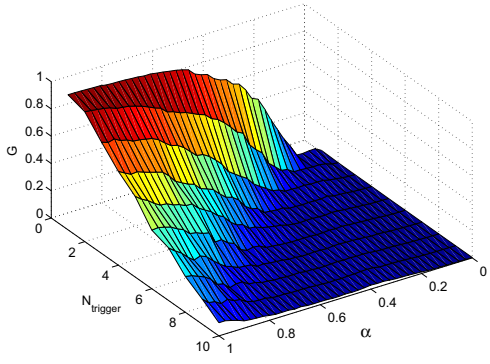

Figure 5. For a scale-free network of $N = 2000$ nodes under attack targeting multiple nodes, G versus α and $N_{trigger}$, the number of targeted nodes. For each parameter value, G is averaged over 30 realizations.

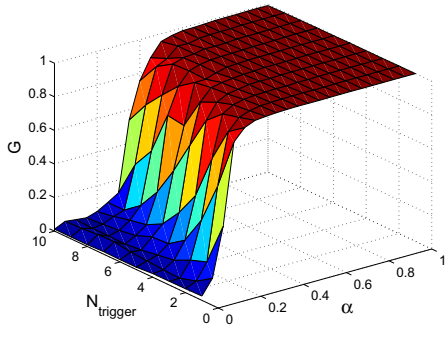

Figure 6. For a random network of $N = 2000$ nodes under attack targeting multiple nodes, G versus α and $N_{trigger}$. For each parameter value, G is averaged over 30 realizations.

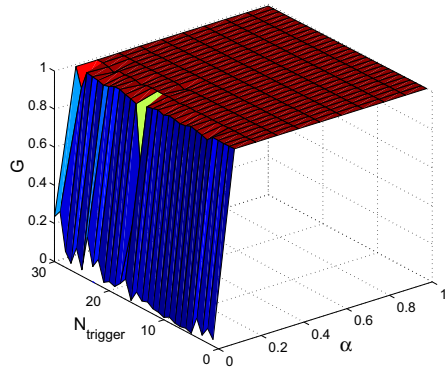

Figure 7. For a scale-free network of $N = 2000$ nodes under random failure, G versus α and $N_{trigger}$. For each parameter value, G is averaged over 20 realizations.

and the situation where cascading failures are caused by attack on the node with the largest number of links and the failures lead to immediate breakdown of the network. That is, G becomes close to zero after one redistribution of the load. For a node in the network, its load is a function of the degree variable k. For scale-free networks, we have [26],

$$L(k) \sim k^{\eta}, \tag{3}$$

where $\eta > 0$ is a scaling exponent. To proceed, we write the degree distribution as $P(k) = ak^{-\gamma}$ and the load distribution as $L(k) = bk^{\eta}$, where a and b are positive constants. Let k_{max} be the largest degree in the network. Before the attack, we have

$$\int_1^{k_{max}} P(k)dk = N \quad \text{and} \tag{4}$$
$$\int_1^{k_{max}} P(k)L(k)dk = S,$$

where S is the total load of the network. These two equations give

$$\begin{aligned} a &= \frac{(1-\gamma)N}{[k_{max}^{1-\gamma}-1]} \quad \text{and} \\ b &= \frac{\beta S}{a(1-k_{max})^{-\beta}}, \end{aligned} \tag{5}$$

where $\beta \equiv \gamma - \eta - 1$. After the removal of the highest degree node (it is only the first step of the whole cascading process), the degree and load distributions become $P'(k) = a'k^{-\gamma'}$ and $L'(k) = b'k^{\eta'}$, respectively. Since only a small fraction of nodes are removed from the network, we expect the changes in the algebraic scaling exponents of these distributions to be negligible. We thus write $P'(k) \approx a'k^{-\gamma}$ and $L'(k) \approx b'k^{\eta}$, where the proportional constants a' and b' can be calculated in the same way as for a and b. We obtain $a' = (1-\gamma)(N-1)/[k_{max}^{1-\gamma}-1]$ and $b' = \beta S'/a'(1-k_{max})^{-\beta}$, where S' is the total load of the network after the attack. For nodes with k links, the difference in load before and after the attack can be written as $\Delta L(k) \approx (b'-b)k^{\eta} = (\frac{b'}{b} - 1)L(k)$. Given the capacity $C(k)$, the maximum load increase that the nodes can handle is $C(k) - L(k) = \alpha L(k)$. The nodes still function if $\alpha > (\frac{b'}{b} - 1)$ but they fail if $\alpha < (\frac{b'}{b} - 1)$. The critical value α_c of the tolerance parameter is then

$$\begin{aligned} \alpha_c &= \frac{b'}{b} - 1 \\ &\approx \left\{1 - {k_{max'}}^{-\beta}\left(-1 + \left(\frac{k_{max}}{k_{max'}}\right)^{-\beta}\right)\right\} \\ &\quad \times \left(\frac{S'}{S}\right) - 1, \end{aligned} \tag{6}$$

where the final result is obtained by using the fact $({k_{\max'}}^{1-\gamma} - 1)/({k_{\max}}^{1-\gamma} - 1) \approx 1$. This is so because both ${k_{max'}}^{1-\gamma}$ and ${k_{max'}}^{1-\gamma}$ approach zero when $N \to \infty$ and $\gamma > 1$. In the limit $N \to \infty$, we have $k_{\max'}^{-\beta} \sim 0$, $k_{max}/k_{max'} \sim$ constant, and $S'/S \to 1$, so $\alpha_c \approx 0$, indicating that an infinite scale-free network cannot be brought down by a single attack if $\alpha > 0$. On the other hand, for finite size network, since $k_{\max'}^{-\beta} > 0$, we have $\alpha_c > 0$, suggesting that breakdown can occur for $\alpha < \alpha_c$. The practical usage of Eq. (6) is that it provides a way to monitor the state of (finite) network to assess the risk of cascading breakdown. In particular, the critical value α_c can be computed in time and comparison with the pre-designed tolerance parameter value α can be made. If α_c shows a tendency of increase and approaches α, early warning can be issued to signal an immediate danger of network breakdown.

We now provide numerical support for the theoretical prediction Eq. (6). Figure 8(a) shows cascading failures when a single node with different degree is removed from the network. We see that, when a node with small degree is removed, the G value remains close to one except when α is close to zero. However, when the node with the largest degree (in this case $k = 81$) is removed, nearly total breakdown of the network, as represented by values of G close to zero, occurs when $\alpha < 0.1$. The phase-transition point α_c is thus about 0.1. With numerical values of $k_{max} = 81$, $k'_{max} = 60$, $S \approx 1.86 \times 10^7$ and $S' \approx 1.91 \times 10^7$, theoretically predicted value of α_c in Eq. (6) gives $\alpha_c \approx 0.1$, which is consistent with numerics. This phase transition phenomenon seems to be robust for different sizes of network, as shown in Fig. 8(b), G versus α for $N = 1000$, $N = 2000$ and $N = 5000$, respectively.

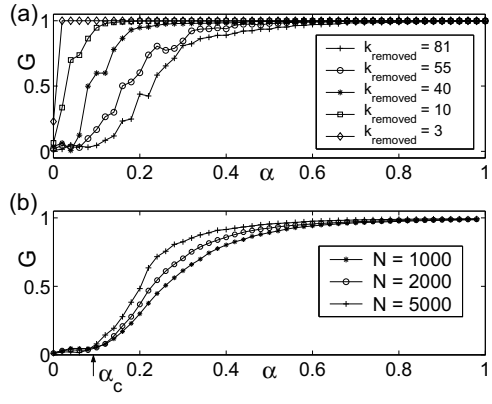

Figure 8. Cascading failure in scale-free network in relation to the tolerance parameter α. (a) Removal of the nodes with different number of links for $N = 2000$. In the case of the removal of the node with the highest degree, the phase-transition point is $\alpha_c \approx 0.1$, meaning that for $\alpha < \alpha_c$, the networks disintegrate almost entirely under intentional attack on a single node. (b) Phase transitions for networks of different sizes. The resulting data points were averaged over 30 realizations.

5. PROTECTING SCALE-FREE NETWORKS FROM CASCADING FAILURE

In this section, we study quantitatively a mechanism of protection against cascading breakdown proposed in [27]. The method consists in removing a small set of nodes that contribute to the loads in the network but they themselves otherwise process little load. Removal of these nodes and all links connected to them will not affect the functioning of the network but will help enhance the load tolerance for each remaining node, thereby helping prevent the spread of the failure or cascading.

For a scale-free network, its load distribution obeys algebraic scaling with the degree variable k [26]: $L(k) = bk^{\eta}$, where η and b are positive constants. After removing a small fraction of low-degree nodes, the average connectivity of the network changes little. Moreover, the degree distribution remains to be algebraic with approximately the same scaling exponent, which can be seen, as follows. On average, the load reduction due to the removal of a low-degree node is proportional to its original load. Let $L(k_1)$ and $L(k_2)$ be the average loads on nodes of degree k_1 and k_2, respectively. After the removal, the average loads are $L'(k'_1) = L(k_1) - c_1 L(k_1)$ and $L'(k'_2) = L(k_2) - c_2 L(k_2)$, respectively, where k'_1 and k'_2 are the new degrees. Since $c_1 \approx c_2$, we have $L'(k'_1)/L'(k'_2) \approx L(k_1)/L(k_2)$. Thus, the algebraic scaling exponent after the removal assumes approximately the same value as in the original network: $\eta' \approx \eta$, the network remains scale-free, and the load distribution can be written as $L'(k) = b'k^{\eta'} \approx b'k^{\eta}$. This property has been confirmed by our numerical simulations (to be shown below).

We can now determine the relation between the load distributions before and after removing ρ percent of low-degree nodes. For convenience, all nodes in the network are labeled by integers from 1 to N, while the removed nodes are labeled by $(1-\rho)N + 1$ to N. The total load before the removal can be written as $S = \sum_{i=1}^{(1-\rho)N} L_i + \sum_{i=N(1-\rho)+1}^{N} L_i \equiv S_0 + S_1$, where S_0 is the sum of loads of the remaining nodes before the removal and S_1 is the total load of the removed nodes. Because the removed nodes have relatively low degrees, we have $S_0 \gg S_1$ and, hence, $S \approx S_0 = \sum_{i=1}^{N(1-\rho)} L_i$. After the removal, the total load of the network is $S' = \sum_{i=1}^{N(1-\rho)} L'_i \approx \sum_{i=1}^{N(1-\rho)} \sigma L_i$, where $0 < \sigma < 1$ is a shifting constant. Since $S = N(N-1)D \approx N^2 D$, $S' = N(1-\rho)[N(1-\rho)-1]D' \approx (1-\rho)^2 N^2 D'$ and $D \approx D'$, where D and D' are the diameters of the networks before and after the removal, respectively, we have $\sigma \approx (1-\rho)^2 \approx 1 - 2\rho$. Thus, on average, the difference

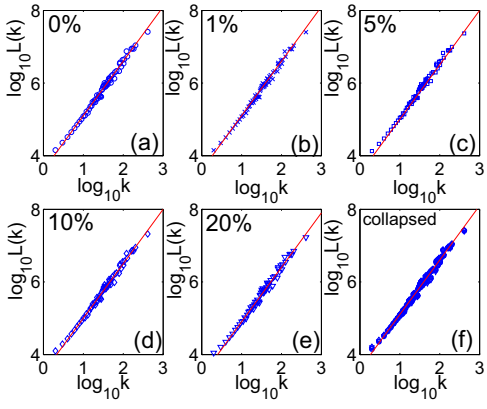

Figure 9. (a) Algebraic scaling of the load distribution $L(k)$ for a scale-free network of $N = 10000$ nodes, $\gamma = 3$, and $\langle k \rangle = 4$. (b-e) Load distributions after removing 1%, 5%, 10%, and 20% of the lowest-degree nodes. (f) Rescaled plots of all curves in (a-e). The algebraic scaling exponent is $\eta \approx 1.5$.

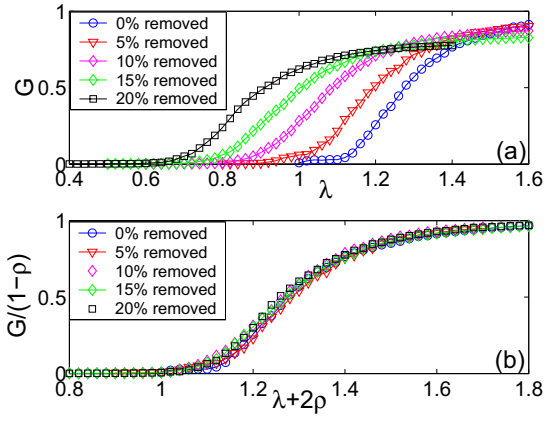

Figure 10. For a scale-free network with $N = 3000$, (a) $G(\lambda, \rho)$ versus λ for five different values of ρ and (b) properly rescaled plots that exhibit a universal relation.

between the loads of node i before and after the removal is $\Delta L_i = L_i - L'_i \approx 2\rho L_i$, which is independent of the parameter λ. Since, initially, the load tolerance of node i is $(\lambda - 1)L_i$ and the process of removal results in equivalently an extra amount of load tolerance $2\rho L_i$, the node will not fail unless the load increment due to an attack exceeds $(\lambda - 1 + 2\rho)L_i$. Controlled removal of ρ percent of low-degree nodes is thus equivalent to increasing the parameter λ to $\lambda + 2\rho$ in the original network:

$$\lambda' \approx \lambda + 2\rho. \tag{7}$$

We now present numerical support for our theoretical result Eq. (7). Again, we generate scale-free networks with degree exponent $\gamma = 3$ and average connectivity $\langle k \rangle = 4$ by using the standard Barabási-Albert model. Figure 9(a) shows the algebraic scaling of the load distribution of the network without any removal of nodes. Approximately the same scaling behavior is observed when some small fractions of nodes with the lowest degrees are removed, as shown in Figs. 9(b-e) for $\rho = 1\%, 5\%, 10\%$, and 20%, respectively. That the intentional removal of a small set of nodes does not change the algebraic load distribution can be seen more clearly in Fig. 9(f), where all plots in Figs. 9(a-e), rescaled by some proper constants, apparently collapse into a single curve. In particular, the algebraic scaling exponent η remains approximately the same, regardless of the value of ρ.

Figure 10(a) shows $G(\lambda, \rho)$ versus λ for different values of ρ, where an attack on the node with the largest degree is assumed. In all cases, there exists a critical capacity λ_c below which the network breaks down entirely as a result of the attack. This value becomes smaller as ρ is increased from zero (curves shifts toward the left), indicating that for some fixed value $\lambda > \lambda_c$, the network is more robust against cascading breakdown due to the attack. This clearly illustrates the protective role played by selectively removing a small set of low-degree nodes. Figure 10(b) shows that the relation between the rescaled quantities $G(\lambda, \rho)/(1 - \rho)$ and $\lambda + 2\rho$ is independent of the value of ρ, as predicted.

6. CONCLUSIONS

We investigated cascading failures triggered by attacks on a single or a few nodes in scale-free networks and focused on the fundamental and practically important question of whether such failures can lead to disintegration of the network. Our finding is a phase-transition like phenomenon in terms of the network tolerance parameter characterizing the node capacity, where the two distinct phases correspond to the situations where the network under attack remains largely integrated or disintegrated as a result of cascading failures.

By analyzing the dynamics of load redistribution resulted from selectively removing a small set of low-degree nodes, we obtained a criterion which allows the lower bound of the capacity parameter for cascade-free scale-free networks and the optimal fraction of intentionally removed nodes to be determined.

The model we study here not only captures the essential components of cascading process, but also incorporates network topology by means of betweenness. Thus, the understanding of the model certainly has positive meaning for engineering and computer network design and protection.

Acknowledgements

This work was supported by NSF under Grant No. ITR-0312131, by AFOSR under Grant No. F49620-01-

1-0317, by FAPESP under Grant No. 02/10707-0 and 05/55212-6 and by CNPq.

REFERENCES

[1] A.-L. Barabási e R. Albert, Emergence of scaling in random networks, *Science*, 286:509-512, 1999.

[2] R. Albert e A.-L. Barabási, Statistical mechanics of complex networks, *Rev. Mod. Phys.*, 74:47-97, 2002.

[3] Newman, M. E. J., The structure and function of complex networks, *SIAM Review*, 45(2):167-256, 2003.

[4] R. Albert, H. Jeong, and A.-L. Barabási, Diameter of the World-Wide Web, *Nature*, 401:130-131, 1999.

[5] G. Siganos, M. Faloutsos, P. Faloutsos, and C. Faloutsos, Power Laws and the AS-Level Internet Topology, *IEEE/ACM Trans. Networking*, 11:514-524, 2003.

[6] X. F., Wang and G.-R. Chen, Complex Networks: Small-World, Scale-Free and Beyond, *IEEE Circuit and System Magzine*, First Quarter:6-20, 2003.

[7] K. Sun, Complex networks theory: A new method of research in power grid, *2005 IEEE/PES Transmission and Distribution Conference and Exhibition: Asia and Pacific*, pages 1-6, 2005.

[8] A. E. Motter and Y.-C. Lai, Cascade-based attacks on complex networks, *Phys. Rev. E*, 66:065102(1-4), 2002.

[9] M. E. J. Newman, The structure of scientific collaboration networks, *Proc. Natl. Acad. Sci. U.S.A.*, 98:404-409, 2001.

[10] P. Erdös e A. Rényi, On the strength of connectedness of a random graph, *Acta Math. Acad. Sci. Hungar.*, 12:261-267, 1961.

[11] D. J. Watts e S. H. Strogatz, Collective dynamics of "small-world" networks, *Nature*, 393:440-442, 1998.

[12] M. Kurant and P. Thiran, Layered complex networks, *Phys. Rev. Letts.*, 96:138701(1-4), 2006.

[13] K. Bömer, J. T. Maru and R. L. Goldstone, The simultaneous evolution of author and paper networks, *Proc. Natl. Acad. Sci. U.S.A.*, 101:5266-5273, 2004.

[14] C. L. DeMarco, A phase transition model for cascading network failure, *IEEE Control Systems Magazine*, 21:40-51, 2001.

[15] B. A. Carreras, V. E., Lynch, I. Dobson, and D. E. Newman, Critical points and transitions in an electric power transmission model for cascading failure blackouts, *Chaos*, 12:985-994, 2002.

[16] M. A. Rios, D. S. Kirschen, D. Jayaweera, D. P. Nedic, and R. N. Allan, Value of Security: Modeling Time-Dependent Phenomena and Weather Conditions, *IEEE Trans. Power Systems*, 17:543-548, 2002.

[17] M. Ni, J. D. McCalley, V. Vittal, and T. Tayyib, Online Risk-Based Security Assessment, *IEEE Trans. Power Systems*, 18:258-265, 2003.

[18] B. A. Carreras, D. E. Newman, I. Dobson, and A. B. Poole, Evidence for Self-Organized Criticality in a Time Series of Electric Power System Blackouts, *IEEE Trans. Circuit and Systems - I*, 51:1733 - 1740, 2004.

[19] I. Dobson, B. A. Carreras, D. E. Newman, A loading-dependent model of probabilistic cascading failure, *Probability in the Engineering and Informational Sciences*, 19:15-32, 2005.

[20] A. Arenas, A. Díaz-Guilera, e R. Guimerà, Communication in networks with hierarchical branching, *Phys. Rev. Letts.*, 86(14), pp. 3196-3199, 2001.

[21] R. Albert, H. Jeong and A.-L. Barabási, Error and attack tolerance of complex networks ", *Nature* (London) 406:378-382, 2000.

[22] R. Cohen, K. Erez, D. b-Avraham, and S. Havlin, Resilience of the Internet to Random Breakdowns, *Phys. Rev. Letts.*, 85:4626-4628, 2000.

[23] R. Cohen, K. Erez, D. b-Avraham, and S. Havlin, Breakdown of the Internet under Intentional Attack, *Phys. Rev. Letts.*, 86:3682-3685, 2001.

[24] D. J., Watts, A simple model of global cascades on random networks, *Proc. Natl. Acad. Sci. U.S.A.*, 99:5766-5771, 2002.

[25] P. Holme, Congestion and centrality in traffic flow on complex networks, *Advances in Complex Systems*, 6:163-176, 2003.

[26] L. Zhao, K. Park, and Y.-C. Lai, Attack vulnerability of scale-free networks due to cascading breakdown, *Phys. Rev. E*, 70:035101(1-4), 2004.

[27] A. E. Motter, Cascade Control and Defense in Complex Networks, *Phys. Rev. Letts.*, 93:098701(1-4), 2004.

[28] L. Zhao, K.-H. Park, e Y.-C. Lai, Tolerance of scale-free networks against attack-induced cascades, *Phys. Rev. E*, 72:025104(1-4), 2005.

[29] M. E. J. Newman, Scientific collaboration networks. II. Shortest paths, weighted networks, and centrality, *Phys. Rev. E*, 64:016132(1-7), 2001.