Research Article

Threshold for the Outbreak of Cascading Failures in Degree-Degree Uncorrelated Networks

Junbiao Liu,[1] Xinyu Jin,[1] Lurong Jiang,[2] Yongxiang Xia,[1] Bo Ouyang,[3] Fang Dong,[4] Yicong Lang,[1] and Wenping Zhang[1]

[1] *College of Information Science and Electronic Engineering, Zhejiang University, Hangzhou 310027, China*
[2] *School of Information Science and Technology, Zhejiang Sci-Tech University, Hangzhou 310018, China*
[3] *College of Electrical and Information Engineering, Hunan University, Changsha 410015, China*
[4] *School of Information and Electric Engineering, Zhejiang University City College, Hangzhou 310015, China*

Correspondence should be addressed to Xinyu Jin; jinxy@zju.edu.cn

In complex networks, the size of the giant component formed by unfailed nodes is critically important for estimating the robustness of networks against cascading failures. In order to explore the critical moment of cascading failures break-out, we provide a cascade of overload failure model with local load sharing mechanism and then deduce the threshold of node capacity when the large-scale cascading failures happen and unfailed nodes in steady state cannot connect to each other to form a large connected subnetwork. We get the theoretical derivation of this threshold in degree-degree uncorrelated networks and validate the effectiveness of this method in simulation. This threshold provides us with a guidance to improve the network robustness under the premise of limited capacity resource when creating a network and assigning load. Therefore, this threshold is useful and important to analyze the robustness of networks. We believe that our research provides us with a guidance to improve the network robustness under the premise of limited capacity resource.

1. Introduction

Cascading failures are a sort of phenomena in which a random failure or intentional attack on one or a few nodes leads to serve chain reaction in the networks, which can cause collapse of a large fraction of nodes in the network. It is widely found in many real-world networks, such as power transmission [1, 2], communication [3], economic [4], and biological [5] networks. A real example of cascading failures is the well-known Northeast Blackout in 2003 [6]. In this case, the outage of a generator led to a serve chain reaction of power blackout, which affected approximately 50 million people in North America and caused financial losses of about $6 billion.

In order to understand the essential mechanism of cascading failures, a number of cascading models have been proposed, such as betweenness-based model [7–10], sand-pile model [11, 12], and fiber-bundle model [13, 14]. In these models, a node fails when its load exceeds its capacity. The failure of this node leads to the redistribution of load in network and can cause collapse of a large fraction of the network. Therefore, the cascading failure process of these models is highly dependent on the relation between load and capacity.

When the dynamical process of cascading failures terminates in steady state, the network breaks into several connected subnetworks formed by unfailed nodes. The size of the largest connected subnetwork (i.e., the giant component) can be used to measure the severity of cascading failures, which is critically important for estimating the robustness of networks [15, 16]. It is intuitive that when the capacity increases (or the load decreases), the size of cascading failures in network reduces, which is confirmed by all of those cascading failure models above. In fiber-bundle and sand-pile model, with the capacity under a critical value (or load above a critical value), the giant component disappears, or above this critical value the size of the giant component dramatically rises up [13, 17].

This critical value is a very important feature to measure the robustness of networks. However, to our knowledge, there is little research on quantitative analysis of the relation between node capacity and this threshold for the break-out of cascading failures at present.

In this paper, we provide a cascade of overload failure model with local load sharing mechanism and then explore the threshold of node capacity when the large-scale cascading failures happen and there does not exist a large connected subnetwork formed by unfailed nodes.

2. Model

Here we provide a cascade of overload failure model with local load sharing mechanism. In this model, the statuses of nodes are divided into two categories: the *unfailed* and *failed*. We assume that all nodes in networks are unfailed at the beginning. A node fails if its load exceeds its capacity (i.e., overload). When a node fails to work, this node is considered as transferring a fixed positive load Δ to each of its unfailed neighbors and being separated from the giant component [18–20]. When the cascading failures terminate, only those nodes in the giant component are supposed to work. It is natural to assume that the capacity C_v of a node v is proportional to its initial load L_v [21, 22] as

$$C_v = (1 + \alpha) L_v, \qquad (1)$$

where the constant α is the tolerance parameter. The initial load of each node is randomly distributed following a uniform distribution on the interval $[L_{\min}, L_{\max}]$. For simplicity, we set the lower bound of initial load $L_{\min} = 0$ and the lower bound of initial load $L_{\max} = 1$ [18, 19]. Then, for arbitrary node v, the cumulative distribution function $\varphi(l)$ of L_v is defined as

$$P\{L_v < l\} \triangleq \varphi(l) = \begin{cases} 0 & l \leq 0 \\ l & 0 < l \leq 1 \\ 1 & l > 1. \end{cases} \qquad (2)$$

The numerical process of cascading failures with local load sharing mechanism is summarized as follows.

(1) Initialization. Generate a degree-degree uncorrelated network with N nodes. Assume all nodes are unfailed. The load of each node is uniformly distributed in $[L_{\min}, L_{\max}]$, where we let $L_{\min} = 0$ and $L_{\max} = 1$. The capacities of nodes are determined by (1).

(2) Beginning. Choose very few nodes randomly in the network, and set them as failed.

(3) Load Redistribution. In each round, each node which is failed in the last round transfers a fixed positive load Δ to each of its unfailed neighbors. An unfailed node turns out to be failed if it overloads in this round.

(4) Halt. Repeat step (3) if there exist overloaded nodes in the network; otherwise, the process halts. Finally, only the unfailed nodes in the giant component are supposed to work.

3. Analysis

3.1. Description of Critical Conditions. When the cascading failure process ends, the unfailed nodes in the network form several connected subnetworks. The fraction of giant component, which is the relative size of the largest connected subnetwork, can be used as a measure of the network performance against cascading failures. Figure 1 shows the fraction G of the giant component as a function of tolerance parameter α in Erdös-Rényi (ER) random networks [23] and Barabási-Albert (BA) scale-free networks [24]. For each G with different average degree $\langle k \rangle$, there exists a critical tolerance parameter α_c, with α under which the giant component disappears and over which G increases dramatically and approaches 1 finally. When $\alpha < \alpha_c$, G approximately equals 0, which indicates that large-scale cascading failures occur and network breaks into extremely small clusters. There are two conditions to ensure that G approximately equals 0 when $\alpha < \alpha_c$:

(I) Large-scale cascading failures occur in the network; that is, the failed nodes connect to each other to form a large connected subnetwork.

(II) When the dynamical process of cascading failures terminates in steady state, there does not exist a large connected subnetwork formed by unfailed nodes.

In the rest of this section, we give a theoretical derivation of α_c with these two conditions above.

3.2. Large-Scale Cascading Failures Occur in the Network. First, let us consider condition (I) that large-scale cascading failures occur in the network. This condition equals the situation that the failed nodes connect to each other to form a large connected subnetwork. On average, each node fails and causes more than one of its neighbors to fail when large-scale cascading failures happen. Consider a node v of degree k_v and any of its neighbors w. The probability that node v fails and causes node w to fail is

$$P\{L_w + \Delta > (1 + \alpha) L_w\} = P\left\{L_w < \frac{\Delta}{\alpha}\right\} = \varphi\left(\frac{\Delta}{\alpha}\right), \quad (3)$$

where L_w is the initial load of node w.

The cascading failures of nodes are a sort of site percolation process. Because of the locally tree-like approximation in percolation of degree-degree uncorrelated networks, the probability that node v fails and causes m of its $k_v - 1$ neighbors (subtract the node which causes node v to fail) to fail is

$$\binom{k_v - 1}{m} \varphi\left(\frac{\Delta}{\alpha}\right)^m \left(1 - \varphi\left(\frac{\Delta}{\alpha}\right)\right)^{k_v - 1 - m}. \qquad (4)$$

Thus, on average, the node v fails and causes its

$$\sum_{m=1}^{k_v-1} m \binom{k_v - 1}{m} \varphi\left(\frac{\Delta}{\alpha}\right)^m \left(1 - \varphi\left(\frac{\Delta}{\alpha}\right)\right)^{k_v-1-m}$$

$$= \sum_{m=1}^{k_v-1} m \binom{k_v - 1}{m} \varphi\left(\frac{\Delta}{\alpha}\right)^m \left(1 - \varphi\left(\frac{\Delta}{\alpha}\right)\right)^{k_v-1-m}$$
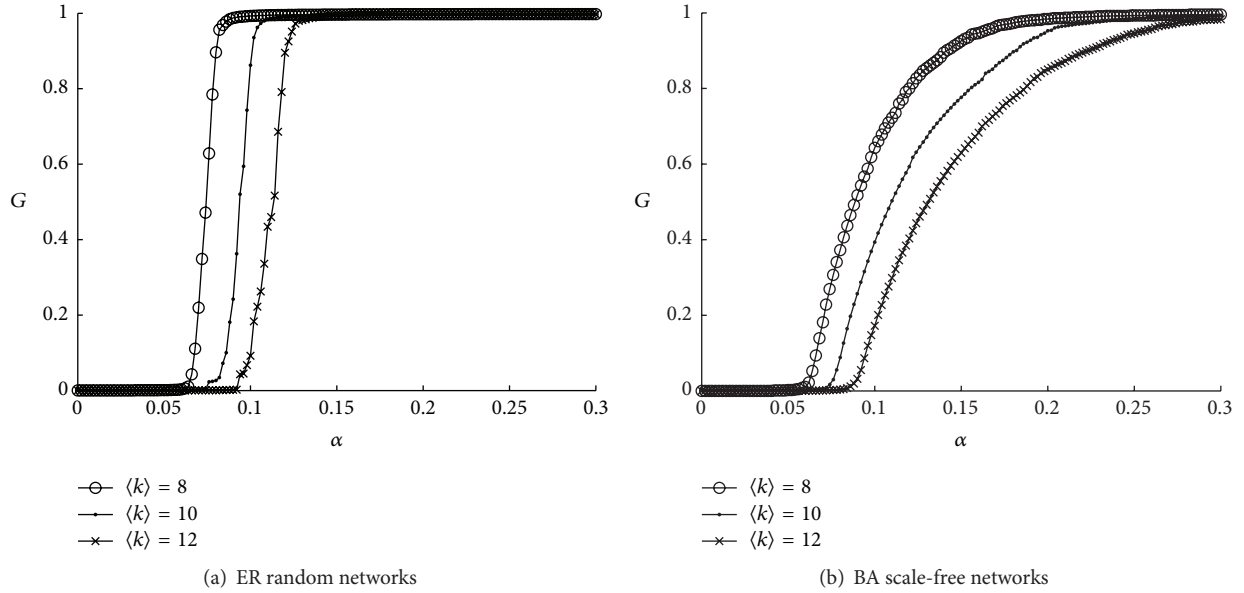
FIGURE 1: The fraction G of giant component changes with increase of tolerance parameter α in ER random and BA scale-free networks.

$$= \sum_{m=1}^{k_v-1} (k_v - 1) \binom{k_v - 2}{m - 1} \varphi\left(\frac{\Delta}{\alpha}\right)^m \left(1 - \varphi\left(\frac{\Delta}{\alpha}\right)\right)^{k_v-1-m}$$
$$= (k_v - 1) \varphi\left(\frac{\Delta}{\alpha}\right) \tag{5}$$

neighbors to fail.

Denote p_k as the probability that a randomly picked node has degree k. According to (5), the failure of arbitrary node in the network causes its $\sum_{k=0}^{\infty} p_k(k-1)\varphi(\Delta/\alpha) = (\langle k \rangle - 1)\varphi(\Delta/\alpha)$ neighbors to fail. When large-scale cascading failures occur in the network, the failure of arbitrary node causes more than one of its neighbors to fail; that is,

$$(\langle k \rangle - 1) \varphi\left(\frac{\Delta}{\alpha}\right) \geq 1. \tag{6}$$

Thus, for condition (I) that large-scale cascading failures occur in the network, we have

$$\alpha \leq (\langle k \rangle - 1) \Delta. \tag{7}$$

It is worth noting that there is not any specific assumption on degree distribution for the result above.

3.3. There Does Not Exist a Large Connected Subnetwork Formed by Unfailed Nodes. According to the local load sharing mechanism in our model, there may exist some nodes that do not propagate cascading failures. These nodes never fail with large tolerance parameter α, or their failure will not affect other nodes as all of their neighbors are already failed nodes. We call these nodes *absorbing nodes*. Before handling condition (II), we pay close attention to these absorbing nodes. Let us consider the situation that a node v of degree k_v does not fail after $k_v - 1$ of its neighbors fail and share loads $(k_v-1)\Delta$ with it. Consequently, no matter whether node v fails

or not, it can not affect other nodes any more. In this situation, node v has the ability to absorb the loads of all its neighbors by itself; thus we call this node the *independent absorbing node*. The capacity and initial load of an independent absorbing node v meet the following relation:

$$L_v + (k_v - 1) \Delta < (1 + \alpha) L_v. \tag{8}$$

But this is not the only reason that a node happens to be an absorbing node. There is another situation for a node to be an absorbing node whose load does not satisfy (8). In this situation, considering a node v, m of its neighbors are absorbing nodes; thus we only need $n_v > k_v - m - 1$ to make node v an absorbing node. We call this kind of nodes *dependent absorbing nodes* because their ability to absorb loads depends on the other absorbing nodes of their neighbors. Dependent absorbing nodes can also prevent cascading failures locally.

Figure 2 gives two examples of absorbing nodes. In Figure 2(a), unfailed node 1 has five neighbors and four of them (nodes 2, 3, 4, and 5) are failed nodes. Assume that if node 6 does not fail in the following procedure of cascading failure, then node 1 does not fail either. Or in another case assume that node 6 fails and transfers a fixed value of load Δ to node 1. In the latter case, no matter whether node 1 fails or not, it can not affect any other node. Thus, node 1 is an independent absorbing node. In Figure 2(b), we assume that node 7 is not an independent absorbing node. It happens to fail if more than two of its four neighbors fail and transfer loads to it. Assume node 8 is an absorbing node; thus it never fails. Consequently, no matter whether node 7 fails or not, it can not share load with any unfailed neighbor. Therefore, node 7 in Figure 2(b) comes to be a dependent absorbing node because its ability to absorb loads depends on the absorbing node 8.
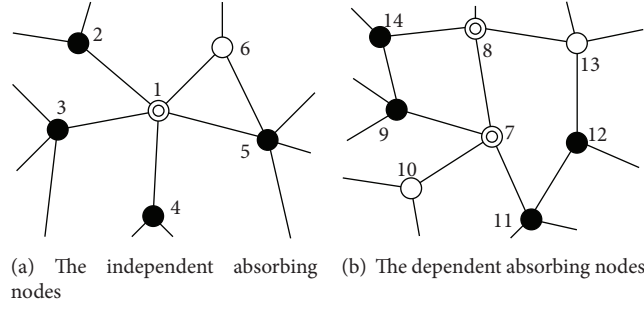
(a) The independent absorbing nodes   (b) The dependent absorbing nodes

Figure 2: Two examples of the absorbing nodes. In the procedure of cascading failures, an unfailed node ($\circ$) may turn to be a failed node ($\bullet$) when it overloads or does not fail as it is an absorbing node ($\odot$).

Then we consider condition (II) that there does not exist a large connected subnetwork formed by unfailed nodes. We have explained that absorbing nodes survive and keep to be unfailed after cascading failures. With increase of α, each node has higher probability to be assigned with more capacity and be an absorbing node on average. If there exist a large number of absorbing nodes, they may connect to each other to form a large connected subnetwork which makes G greater than 0. We have divided absorbing nodes into independent absorbing nodes and dependent absorbing nodes. Furthermore, we call an absorbing node an *m-absorbing node* ($m = 0, 1, 2, \ldots, k$), if and only if m of its neighbors are absorbing nodes. We assume that independent absorbing nodes are not equivalent to 0-absorbing node. For an interdependent node, if it has m absorbing nodes of its neighbors, we still treat this node as an m-absorbing node.

Then, we let the probability that a node of degree k happens to be an absorbing node be a_k. The probability that node v happens to be an m-absorbing node is $a_k^{(m)}$. Then we have

$$a_k = \sum_{m=0}^{k} a_k^{(m)}. \tag{9}$$

In a degree-degree uncorrelated network, consider a node v of degree k and any of its neighbors w. The probability of node w with degree j is $jp_j / \sum_i i p_i = jp_j / \langle k \rangle$. Node w happens to be an absorbing node with the probability a_j.

Then, the probability of node v's arbitrary neighbor being an absorbing node is

$$\sigma_a = \sum_j \frac{jp_j a_j}{\langle k \rangle}. \tag{10}$$

According to (10), the probability that exactly m of node v's neighbors are absorbing nodes is

$$\binom{k}{m} \sigma_a^m (1 - \sigma_a)^{k-m}. \tag{11}$$

Node v is an m-absorbing node if it stays unfailed with $k - m - 1$ neighbors fail and transfer loads to it. Therefore, the probability is that node v comes to be an m-absorbing node with probability

$$\begin{aligned}
a_k^{(m)} &= \binom{k}{m} \sigma_a^m (1 - \sigma_a)^{k-m} \\
&\quad \cdot P\left\{ L_v + (k - m - 1)\Delta < (1 + \alpha) L_v \right\} \\
&= \binom{k}{m} \sigma_a^m (1 - \sigma_a)^{k-m} \left(1 - \varphi\left(\frac{(k - m - 1)\Delta}{\alpha} \right) \right).
\end{aligned} \tag{12}$$

We substitute (12) into (9) and get the probability that node v (i.e., arbitrary node of degree k) happens to be an absorbing node as

$$
\begin{aligned}
a_k &= \sigma_a^k + \sum_{m=0}^{k-1} \binom{k}{m} \sigma_a^m (1 - \sigma_a)^{k-m} \left(1 - \varphi\left(\frac{(k - m - 1)\Delta}{\alpha} \right) \right) \\
&= \begin{cases} \sigma_a^k + \sum_{m=0}^{k-1} \binom{k}{m} \sigma_a^m (1 - \sigma_a)^{k-m} \left(1 - \dfrac{(k - m - 1)\Delta}{\alpha} \right) & \text{if } k \leq \left[\dfrac{\alpha}{\Delta}\right] + 1 \\[2ex] \sigma_a^k + \sum_{m=k-1-[\alpha/\Delta]}^{k-1} \binom{k}{m} \sigma_a^m (1 - \sigma_a)^{k-m} \left(1 - \dfrac{(k - m - 1)\Delta}{\alpha} \right) & \text{elsewhere} \end{cases} \\
&= \begin{cases} 1 - \dfrac{(k - 1)\Delta}{\alpha} + \dfrac{k\sigma_a \Delta}{\alpha} - \dfrac{\sigma_a^k \Delta}{\alpha} & \text{if } k \leq \left[\dfrac{\alpha}{\Delta}\right] + 1 \\[2ex] \sigma_a^k + \sum_{m=k-1-[\alpha/\Delta]}^{k-1} \binom{k}{m} \sigma_a^m (1 - \sigma_a)^{k-m} \left(1 - \dfrac{(k - m - 1)\Delta}{\alpha} \right) & \text{elsewhere.} \end{cases}
\end{aligned} \tag{13}
$$

(a) $\langle k \rangle = 8$

(b) $\langle k \rangle = 10$
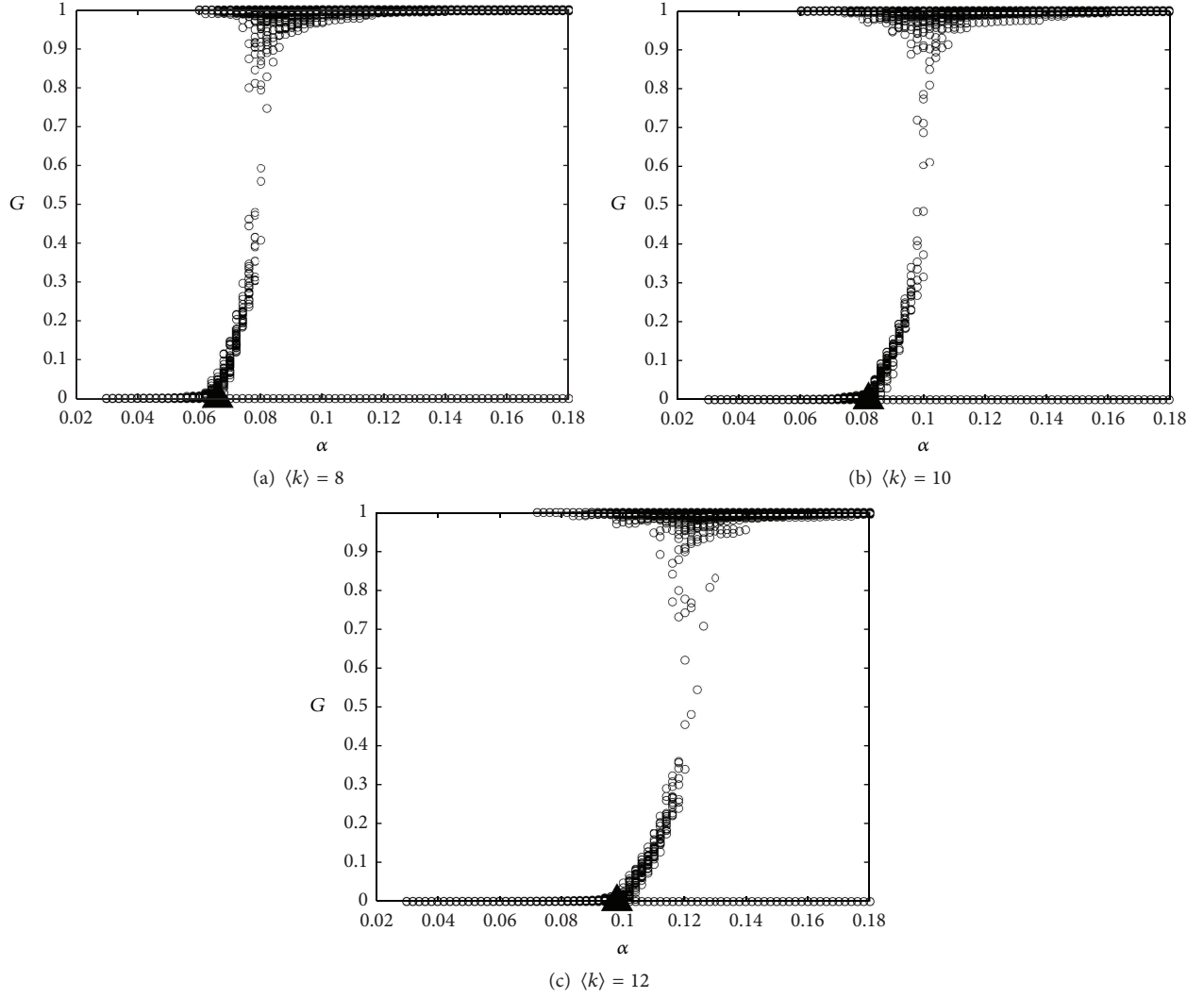

(c) $\langle k \rangle = 12$

FIGURE 3: Validation of the analytical predictions of threshold α_c where the giant component G exists in ER random networks. With increase of tolerance parameter α, we draw each simulated G of 50 realizations, respectively. The threshold α_c is marked by solid triangle (▲).

According to the *cascade condition* [25, 26], the absorbing nodes can not connect to each other to form a large connected subnetwork with the following condition:

$$\frac{1}{\langle k \rangle} \sum_{k=0}^{\infty} k(k-1) p_k a_k < 1, \tag{14}$$

where p_k is probability that a randomly picked node has degree k and a_k is the probability that a node with degree k happens to be an absorbing node.

Now we complete the derivation by (7), (10), (13), and (14). These four equations depend on each other. We can get the threshold of tolerance parameter α_c with the following iteration process:

(1) Assign $(\langle k \rangle - 1)\Delta$ as an initial value to tolerance parameter α according to (7).

(2) Get a_k via (10) and (13) with the current value of tolerance parameter α.

(3) Substitute a_k into (14). If the inequality of (14) is invalid, discount a very small value from α, and then repeat step (2); otherwise, let $\alpha_c = \alpha$ and finish the iteration process.

4. Simulation

We validate the analytical prediction of threshold α_c in ER random and BA scale-free networks with $N = 5000$ nodes and varied average degree $\langle k \rangle = 8, 10, 12$. The load of each node is uniformly distributed on the interval $[0, 1]$. We randomly set very few nodes to fail as the beginning of the dynamical procedure. When a node fails, it transfers a fixed load $\Delta = 0.01$ to each of its unfailed neighbors. Figures 3 and 4 present the giant component G with different α, and the threshold α_c for G rises from zero to nonzero. Note that different values of Δ only change the scale of abscissa. For each subfigure in Figures 3 and 4, each of 50 independent simulation results is drawn, respectively. The theoretical

(a) $\langle k \rangle = 8$

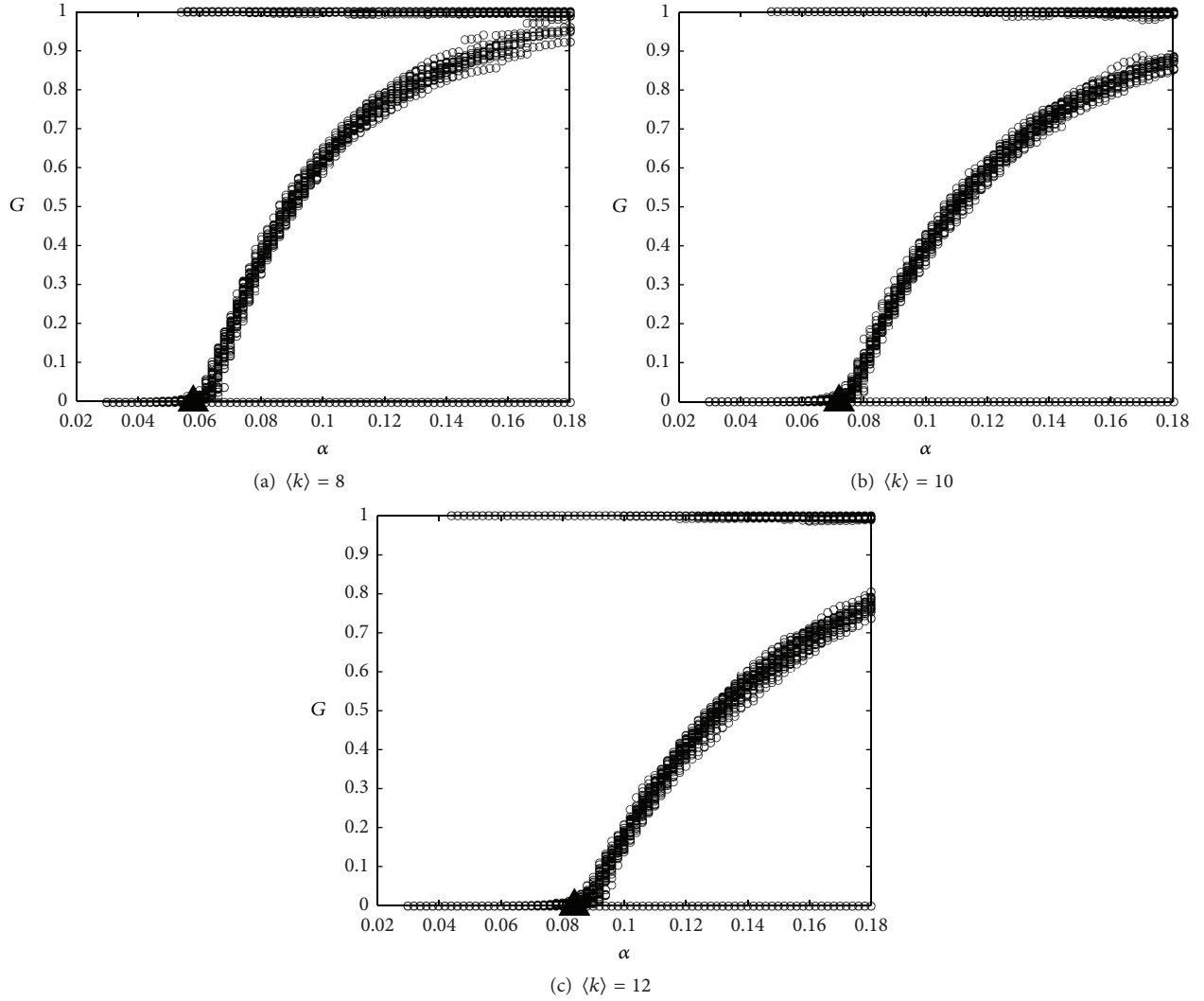(b) $\langle k \rangle = 10$

(c) $\langle k \rangle = 12$

Figure 4: Validation of the analytical predictions of threshold α_c where the giant component G exists in BA scale-free networks. With increase of tolerance parameter α, we draw each simulated G of 50 realizations, respectively. The threshold α_c is marked by solid triangle (▲).

threshold α_c is marked by the solid triangle. α_c is exactly at the moment when G rises from zero to nonzero. Therefore, our theoretical predictions are in good agreement with the simulation results. It also can be seen in both ER and BA networks that, with higher value of average degree $\langle k \rangle$, the value of threshold α_c increases and the networks turn out to be robuster against cascading failure. This appearance is consistent with previous studies on cascading failure.

5. Conclusion

In this paper, we provide a cascade of overload failure model with local load sharing mechanism and then explore the threshold α_c of tolerance parameter for capacity, when the giant component G formed by unfailed nodes rises from zero to nonzero. We provide two conditions to ensure that G approximately equals zero when $\alpha < \alpha_c$, which are as follows: (I) the large-scale cascading failures occur; (II) unfailed nodes in steady state cannot connect to each other to form a large connected subnetwork. With these two conditions, we get the theoretical derivation of α_c in degree-degree uncorrelated networks and validate the effectiveness of this theoretical derivation in simulations. We believe that when creating a network and assigning load with local load sharing mechanism, threshold α_c provides us with a guidance to improve the network robustness under the premise of limited capacity resource, especially for telecommunication networks and power grid networks.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

References

[1] Y. Koç, M. Warnier, P. Van Mieghem, R. E. Kooij, and F. M. Brazier, "The impact of the topology on cascading failures in a power grid model," *Physica A: Statistical Mechanics and Its Applications*, vol. 402, pp. 169–179, 2014.

[2] P. Henneaux, "Probability of failure of overloaded lines in cascading failures," *International Journal of Electrical Power & Energy Systems*, vol. 73, pp. 141–148, 2015.

[3] X.-L. Xu, W.-B. Du, and C. Hong, "Cascade defense via routing in complex networks," *International Journal of Modern Physics C*, vol. 26, no. 12, Article ID 1550141, 2015.

[4] X. Huang, I. Vodenska, S. Havlin, and H. Eugene Stanley, "Cascading failures in bipartite graphs: Model for systemic risk propagation," *Scientific Reports*, vol. 3, article 1219, 2013.

[5] C. Borrvall, B. Ebenman, and T. Jonsson, "Biodiversity lessens the risk of cascading extinction in model food webs," *Ecology Letters*, vol. 3, no. 2, pp. 131–136, 2000.

[6] G. Andersson, P. Donalek, R. Farmer et al., "Causes of the 2003 major grid blackouts in North America and Europe, and recommended means to improve system dynamic performance," *IEEE Transactions on Power Systems*, vol. 20, no. 4, pp. 1922–1928, 2005.

[7] G. Barach, M. Tuchman, G. Cwilich, and S. Buldyrev, "Distributions of betweenness in cascades of overload failure in random regular networks," in *Proceedings of the APS March Meeting*, vol. 1 of *Abstract 16012*, Denver, Colo, USA, March 2014.

[8] J. Wang, L. Rong, L. Zhang, and Z. Zhang, "Attack vulnerability of scale-free networks due to cascading failures," *Physica A*, vol. 387, no. 26, pp. 6671–6678, 2008.

[9] F. Tan, Y. Xia, W. Zhang, and X. Jin, "Cascading failures of loads in interconnected networks under intentional attack," *EPL*, vol. 102, no. 2, Article ID 28009, 2013.

[10] Y. Xia, J. Fan, and D. Hill, "Cascading failure in Watts-Strogatz small-world networks," *Physica A*, vol. 389, no. 6, pp. 1281–1285, 2010.

[11] K.-I. Goh, D.-S. Lee, B. Kahng, and D. Kim, "Sandpile on scale-free networks," *Physical Review Letters*, vol. 91, no. 14, Article ID 148701, 2003.

[12] M. Hoore and S. Moghimi-Araghi, "Critical behavior of a small-world sandpile model," *Journal of Physics. A. Mathematical and Theoretical*, vol. 46, no. 19, pp. 195001–195008, 2013.

[13] Y. Moreno, J. B. Gómez, and A. F. Pacheco, "Instability of scale-free networks under node-breaking avalanches," *Europhysics Letters*, vol. 58, no. 4, pp. 630–636, 2002.

[14] D.-H. Kim, B. J. Kim, and H. Jeong, "Universality class of the fiber bundle model on complex networks," *Physical Review Letters*, vol. 94, no. 2, Article ID 025501, 2005.

[15] M. Á. Serrano, D. Krioukov, and M. Boguñá, "Percolation in self-similar networks," *Physical Review Letters*, vol. 106, no. 4, Article ID 048701, 2011.

[16] G. Ganesan and S. J. S. Sansanwal Marg, "Size of the giant component in a random geometric graph," *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, vol. 49, no. 4, pp. 1130–1140, 2013.

[17] D.-S. Lee, K.-I. Goh, B. Kahng, and D. Kim, "Sandpile avalanche dynamics on scale-free networks," *Physica A*, vol. 338, no. 1-2, pp. 84–91, 2004.

[18] I. Dobson, B. A. Carreras, and D. E. Newman, "A loading-dependent model of probabilistic cascading failure," *Probability in the Engineering and Informational Sciences*, vol. 19, no. 1, pp. 15–32, 2005.

[19] G. Sansavini, M. Hajj, I. Puri, and E. Zio, "A deterministic representation of cascade spreading in complex networks," *Europhysics Letters*, vol. 87, no. 4, 2009.

[20] L. Jiang, X. Jin, Y. Xia, B. O. Ouyang, and D. Wu, "Dynamic behavior of the interaction between epidemics and cascades on heterogeneous networks," *EPL*, vol. 108, no. 5, Article ID 58009, 2014.

[21] A. E. Motter and Y.-C. Lai, "Cascade-based attacks on complex networks," *Physical Review E*, vol. 66, no. 6, Article ID 065102, 2002.

[22] J. Lehmann and J. Bernasconi, "Stochastic load-redistribution model for cascading failure propagation," *Physical Review E*, vol. 81, no. 3, Article ID 031129, 2010.

[23] P. Erdős and A. Rényi, "On the evolution of random graphs," in *Publication of the Mathematical Institute of the Hungarian Academy of Sciences*, vol. 5, pp. 17–61, 1960.

[24] A.-L. Barabási and R. Albert, "Emergence of scaling in random networks," *Science*, vol. 286, no. 5439, pp. 509–512, 1999.

[25] M. E. J. Newman, S. H. Strogatz, and D. J. Watts, "Random graphs with arbitrary degree distributions and their applications," *Physical Review E*, vol. 64, no. 2, Article ID 026118, 2001.

[26] D. J. Watts, "A simple model of global cascades on random networks," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 99, no. 9, pp. 5766–5771, 2002.