System crash as dynamics of complex networks

SI (Supplementary Information) Appendix

Yi Yu[1], Gaoxi Xiao[1,2], Jie Zhou[3], Yubo Wang[1], Zhen Wang[4,5], Jürgen Kurths[6,7,8], Hans Joachim Schellnhuber[6]

[1] School of Electrical and Electronic Engineering, Nanyang Technological University, 50 Nanyang Avenue, Singapore 639798
[2] Complexity Institute, Nanyang Technological University, Singapore
[3] Department of Physics, East China Normal University, Shanghai, China 200241
[4] Qingdao University, Qingdao, Shandong 266071, China
[5] Interdisciplinary Graduate School of Engineering Sciences, Kyushu University, Kasuga-koen, Kasuga-shi, Fukuoka 816-8580, Japan
[6] Potsdam Institute for Climate Impact Research (PIK), 14473 Potsdam, Germany
[7] Department of Physics, Humboldt University, 12489 Berlin, Germany
[8] Institute for Complex Systems and Mathematical Biology, University of Aberdeen, Aberdeen AB24 3UE, United Kingdom
Correspondence and requests for materials should be addressed to:
`egxxiao@ntu.edu.sg`, `zhenwang0@gmail.com`, or `john@pik-potsdam.ge`

SI Appendix 1: Analytical approach

Given an infinite random network with an initial degree distribution of $p_1(k)$. In the i^{th} time step, for those nodes with degrees lower than k_s or having lost more than q proportion of their neighbors, they may be pruned following an arbitrary probability $f_i(k)$. To track the proportion of neighbors a node has lost, we introduce a *degree transition matrix* D^i where the element D^i_{jk} keeps record of the probability that a node with an original degree j has a degree k at the beginning of the i^{th} time step.

Let $p_i(k)$ be the probability that a node randomly chosen at the beginning of the i^{th} time step has a degree k at that moment, $p_i^*(k)$ be the probability that a node randomly chosen at the beginning of the i^{th} time step has a degree k at that moment and does not get pruned in this time step, and $p_i^{new}(k)$ the probability that a node randomly chosen at the beginning of the i^{th} time step has a degree k at the end of this time step. θ_i denotes the probability that a randomly followed edge at the beginning of the i^{th} time step points to a node not getting pruned in this time step. For the degree transition matrix D^i, set the initial matrix D^1 as an identity matrix.

For the pruning process in the i^{th} time step, we have

$$p_i^*(k) = p_i(k)\gamma_i(k), \tag{1}$$

where $\gamma_i(k)$ denotes the proportion of the nodes with a degree k at the beginning of the i^{th} time step which are not pruned in this time step. We have

$$\gamma_i(k) = \sum_{j=k}^{k_{max}} \varphi^i_{kj}\mu(k,j), \tag{2}$$

where φ^i_{kj} denotes the probability that a node with a degree k at the beginning of the i^{th} time step has a degree j in the original network, and

$$\mu(k,j) = \begin{cases} 1 & k \geq (1-q)j \text{ and } k \geq k_s, \\ 1 - f_i(k) & \text{otherwise.} \end{cases} \tag{3}$$

By following a randomly selected edge at the beginning of the i^{th} time step, the probability of hitting a node not pruned in this time step is:

$$\theta_i = \frac{\sum_k k p_i^*(k)}{\sum_k k p_i(k)}. \tag{4}$$

And the probability that a node randomly selected at the beginning of the i^{th} time step has a degree k at the end of this time step, before normalization, can be calculated as:

$$p_i^{new}(k) = \sum_{k'=k}^{k_{max}} p_i^*(k') \binom{k'}{k'-k} \theta_i^k (1-\theta_i)^{k'-k}. \tag{5}$$

As later will be shown in equation (12), $p_i^{new}(k)$, after normalization, shall show the degree distribution at the beginning of the $(i+1)^{th}$ time step.

The *time-step degree transition matrix* T^i, where the element $T^i_{k'k}$ denotes the probability that a node with a degree k' at the beginning of the i^{th} time step ends up with a degree k at the end of it, can be derived as:

$$T^i_{k'k} = \binom{k'}{k'-k} \theta_i^k (1-\theta_i)^{k'-k}. \tag{6}$$

Define a matrix U^i where the element U^i_{jk} denotes the probability that an original degree j node has a degree k at the beginning of the i^{th} time step and does not leave the network at the end of this time step. We have

$$U^i_{jk} = D^i_{jk}\mu(j,k). \tag{7}$$

The degree transition matrix D^{i+1} hence can be calculated by the following matrix multiplication:

$$D^{i+1} = U^i T^i. \tag{8}$$

The probability that a node with a degree k at the beginning of the $(i+1)^{th}$ time step had a degree j in the original network is

$$\varphi^{i+1}_{kj} = p_1(j) D^{i+1}_{jk} / c_1(k,j), \tag{9}$$

where $p_1(j)$ is the original degree distribution and $c_1(k,i)$ is the normalization factor that

$$c_1(k,j) = \sum_{j=1}^{k_{max}} p_1(j) D^{i+1}_{jk}. \tag{10}$$

The proportion of nodes remaining in the network at the end of the i^{th} time step compared with the network at the beginning of this time step is

$$P_i^* = \sum_k p_i^{new}(k). \tag{11}$$

Normalizing $p_i^{new}(k)$, we get the degree distribution of the network at the beginning of the $(i+1)^{th}$ time step:

$$p_{i+1}(k) = p_i^{new}(k)/P_i^*. \tag{12}$$

The proportion of nodes remaining in the network after i time steps, compared with the original network, can be calculated as:

$$P_i = \prod_{m=1}^{i} P_m^*. \tag{13}$$

The degree transition matrix and the proportion of nodes remaining in the network after each time step can be calculated iteratively using equations (1)-(13).

As mentioned in the main paper, this analysis can be easily extended to a more general case where nodes leave the network following an arbitrary criterion ϕ, including that of the classic k-core problem, as long as the new criterion can be reflected by properly changing $\mu(k, j)$ in equation (3) accordingly.

SI Appendix 2: Cascade threshold of real-life networks

As pointed out in the subsection "Measuring the resilience of some real-life systems against KQ-cascade", the real-life networks may have multiple pseudo-steady states. This makes finding their threshold values rather tricky, since there may exist some strong cores persisting to exist when most network nodes have left. To get a sense of the resilience of real-life networks, we make the simple definition that the threshold q_{th} is the maximum value of q leading to no more than 10% of nodes remaining in the final state of the network. As mentioned in the main paper, we simulate 10 rounds for each real-life network and average the results. For each round of simulation, we increase q_{th} from 0 at a step length of 0.01.

Figure S1 shows the comparison of threshold q_{th} for three real-life networks. One thing in common for these real-life networks is that with a very small value of k_s, the networks already start losing users, since they all have a significant portion of low-degree nodes. As to their behaviors under KQ-cascade, they perform more like scale-free and exponential networks than ER random networks.

Closer observations reveal that the three real-life network samples have different resilience performances under KQ cascade: with an increasing value of k_s, q_{th} increases slowly in Orkut, but much faster in YouTube; LiveJournal performs in between of these two systems. Hence the observation is that the YouTube sample appears to be the most fragile one among the three: the surviving proportion of the KQ-cascade in YouTube is highly sensitive to the value of k_s. Orkut sample, on the other hand, demonstrates opposite behaviors: even when a big mistake has been made leading to a relatively high value of k_s, the system still has a rather low threshold value of q, implying that it may still be able to stand a strong competition. LiveJournal performs in between of these two systems.

Note that the resilience of the three real-life networks is measured in their sampled subnets, which may not necessarily reflect the resilience of the whole networks. Some further discussions on the relationship between the resilience of randomly sampled subnets and their corresponding whole networks can be found in SI Appendix 6.

SI Appendix 3: Evolution of nodal degree distribution during KQ-cascade

Illustration of the evolution of nodal degree distribution during the KQ-cascade helps reveal the cascade dynamics in detail. We shall study on both cases where the network finally crashes to nonexistence or survives with a nontrivial portion of nodes remaining, respectively.

Figures S2-S7 illustrate the evolution of $p(k)$ for different values of k during the cascade process in the ER random network (Figures S2 and S3), random exponential network (Figures S4 and S5) and random scale-free network (Figures S6 and S7), respectively. Figures S2, S4 and S6 show the results where the networks finally crash; while Figures S3, S5 and S7 are for the cases where the networks survive. In addition, Figures S8-S13 illustrate the degree distribution of the remaining nodes at different time during the cascade process. All the three networks adopt the same parameters as those in the main paper. Both the analytical and simulation results are presented and the simulation results show the average in 10 different networks. As we can see, the theoretical results match very well with simulation results.

Our main observations on the evolution of nodal degree distribution during a KQ-cascade in these networks can be briefly summarized as follows:

> For the case where networks survive with a non-trivial proportion of nodes, for k slightly higher than k_s, $p(k)$ increases and then sustains, while $p(k)$ for larger values of k decreases (refer to Figures S3, S5 and S7). Such can be easily understood: many high-degree nodes lose some of their neighbors and become medium-degree nodes. Medium-degree nodes, on the other hand, will stay rather stable as long as the vulnerable nodes do not percolate into a giant cluster.

> For the case where networks finally crash, the transition of nodal degrees appears to be more complicated (refer to Figures S2, S4 and S6). Specifically, for $k \leq k_s$, $p(k)$ firstly decreases and then booms up right before the "sudden crash", whereas for k slightly larger than k_s, $p(k)$ firstly slightly increases and starts to quickly decrease at the beginning of sudden crash. Such observations help explain the pseudo-steady state and the sudden crash of the networks: at the beginning, nodes with degrees lower than k_s quickly leave, making high-degree and medium-degree nodes lose their neighbors and gradually become medium- or low-degree nodes. The vulnerable nodes thus gradually grow into a giant component, allowing the sudden crash to finally happen.

Another interesting observation, which applies to both cases whether networks finally crash or survive, is that even during the sudden crash, the degree distribution of the networks still roughly keeps their "original shapes". Specifically, an ER random network shall roughly still have a Poisson degree distribution though with a lower average degree (Figures S8 and S9), an exponential network will roughly keep its exponential nodal degree distribution (Figures S10 and S11), and a scale-free network would keep its power-law degree distribution with the power-law exponent remaining almost unchanged (Figures S12 and S13). Looking back to Figures S2 to S7, this may not be a surprise: as we can observe, regardless of whether the networks finally crash or survive, the proportion of medium-degree nodes increases and decreases following a similar pattern at different degrees. This helps keep the distribution pattern of the remaining nodes roughly unchanged.

SI Appendix 4: Degree distribution and network resilience

In the main paper, it was shown that by using the proposed analysis, the threshold values of q and k_s could be obtained adopting a simple trial-and-error approach. Here we carry out some further studies on how the resilience of the network may be affected by network degree distribution. We use three random network models with the same average degree $z = 20$ and size of $N = 10^4$: ER random network, random exponential network with a degree cut-off of 100, and random scale-free network with $\gamma = 2$, the minimum degree of 7 and degree cut-off of 100 (which makes its average degree to be about 20).

Figures S14, S15 and S16 show the cascade size in the three networks, respectively. The results are the average in 10 random networks for each case. Figure S17 shows that the three networks, though with the same average degree, demonstrate different levels of resilience in different situations: for a small value of k_s, the q required to crash the network has its lowest value in the ER random network and highest one in the scale-free network. This means that when an initial event (e.g., a mistake as that in the Friendster's crash) does not generate strong impacts, among the three networks, the most and least resilient ones are the ER random network and the scale-free network respectively. When an impactful event happens leading to a high value of k_s, however, the ER random network may crash immediately even with a high value of q: the event alone is fatal to the ER random network. The scale-free network, on the other hand, performs totally differently: it can sustain rather strong initial impacts and still keep a non-trivial proportion of its nodes. For both cases, the performance of the exponential network lies in between the other two networks. Comparing the performances of the three networks, we can see that the uniform random networks may be sensitive to an initial event leading to an increase in the value k_s. If the impacts of such an event are not fatal, the network may appear to be rather stable; otherwise it may crash almost immediately. On the contrary, networks with more heterogeneous degree distributions may lose their users easily, e.g., by making a relatively small mistake, yet they have stronger capabilities to survive from strong initial impacts. This may to a certain extent explain why some social networks, usually known to be closely resembling scale-free networks, appear to lose their popularity easily yet still manage to survive over a long period of time with a significantly reduced number of users.

As we have discussed in the main paper, the existence of communities may help further enhance the fault tolerance of the networks.

We also briefly evaluate the impacts of a few parameter values on cascade threshold in different networks. Specifically, in the ER random network with degree distribution of $p(k) = (z^k e^{-z})/k!$ and exponential network with degree distribution of $p(k) = 1/ze^{-k/z}$, we calculate the cascade threshold when the average nodal degree z changes from 5 to 50. Figures S18 and S19 show that, for a given value of q, the threshold k_s increases almost linearly with z in these two networks. Larger values of z have been tested and the conclusion basically still holds. Similar observation has been made via numerical simulation in [1].

To evaluate the relation between the cascade threshold and power-law exponent γ in the scale-free network, we still set the minimum and the maximum nodal degrees of the network to be 7 and 100 respectively and increase γ from 2 to 2.8 by a step length of 0.2. Figure S20 shows that the threshold k_{th} decreases when γ increases. This is mainly because that, in scale-free networks with larger values of γ, there are fewer high-degree nodes while most network nodes are of rather low degrees. These low-degree nodes are more "vulnerable" under the KQ-cascade, making the network easier to be crashed.

SI Appendix 5: The speedup loss of individuals in a system crash

Intuitively, the fast decline of a system may trigger an alarm or shake the individuals' confidence, which may in turn further accelerate the crash. To better understand the system dynamics under such kind of "accelerated crash", we study a simple model where q decreases at a constant rate. Figure S21 illustrates the evolution of the cascade in a random scale-free network and the real-life Orkut online social network respectively. For the scale-free network, we assume that $k_s = 4$ and q starts with 0.28. The cascade processes with q decreasing at different rates are presented. Figure S21 shows that even a small decreasing rate of 0.002 per time step can significantly accelerate the crash of the network. The figure also demonstrates that, for cases with a dynamic value of q, our theoretical analysis results remain valid to achieve satisfactory accuracy. In all the other networks we have tested, it remains as a valid conclusion that even a slowly decreasing q significantly accelerates the crash of the systems.

SI Appendix 6: Can random sampling of a network reflect the resilience of the original network: a preliminary study

Methods for sampling real-life networks have been extensively studied [2–4]. It is shown that, though different sampling techniques may have different impacts on accuracy of the measurements of various parameters in the original networks, in many cases simple random sampling may approximately reflect the statistical characters of the original networks, e.g., the power-law exponent of the scale-free networks, and the betweenness centrality distribution, average path length, clustering coefficient and assortativity in random networks, etc. [5] It has, however, not been studied whether a sampled sub-network may properly reflect the resilience of the original network. While conducting a comprehensive study on this vitally important problem is beyond the scope of this contribution, we present some preliminary results helping to reveal whether a randomly sampled sub-network can correctly reflect the threshold of KQ-cascade of its original network.

We carry out simulations in three random network models with size of $N = 10^4$: ER random network with a average degree $z = 50$; random exponential network with a average degree of $z = 50$ and degree cut-off of 300; and random scale-free network with $\gamma = 2$, the minimum degree of 17 and degree cut-off of 300 which makes its average degree to be about 50. In addition, we also test on the Gowalla local social network [6] shown in Table I of the main paper. As to the sampling technique, we adopt the simple one evaluated in [6] which randomly chooses a proportion of nodes and keeps all the connections between the sampled nodes. In random networks, the sampled subnets shall have an average degree of αz where z is the average degree of the original network (shown in Figure S22). For the sampled networks, we evaluate the threshold value of k_s for different q (still denoted as k_{th}) for triggering the global cascade of the sampled networks. For each case, the average results of 10 networks are presented.

Figures S23-S26 show the relationship between the sampling percentage α and the cascade threshold k_{th} in the four networks respectively. We observe that there exists approximately linear relationship between α and threshold k_{th} in both synthetic and real-life networks. This may not be a surprise considering that *i)* in the sampled networks, z changes roughly linearly with α; and *ii)* as pointed out in the SI Appendix 4, k_{th} changes approximately linearly with z. Such observations reveal that, by evaluating the threshold k_{th} of a sampled network at a given value of q, we may roughly estimate the k_{th} value of the original network. However, further studies are needed to better understand how specific structures of the original networks, especially the community/clustering structures, may affect the accuracy of the sampling-based estimation results.

SI Appendix 7: Effects of community structures and degree correlations on multi-stage pseudo-steady state

To evaluate the effects of community structures and degree correlations on the emergence of multi-stage pseudo-steady state, we carried out the following randomization operations on a few real-life networks: randomly selecting links A-B and C-D and replacing them with links A-C and B-D. It is known that repeating such randomization operations for a large enough times can reduce, and finally eliminate, the community structures and degree correlations of the network without changing any node's degree. To facilitate discussions, we term the networks before and after the extensive randomization operations as original networks and randomized networks, respectively. In our extensive simulation experiences, multi-stage pseudo-steady state has never been observed in any randomized networks. Simulation results on Orkut and LiveJournal networks are presented in Figures S27 and S28. We see that multi-stage pseudo-steady state can be observed in crashes of the original networks, but not those of the randomized ones.

Another interesting observation is that the extensive randomization operations help enhance network robustness against system crash in both Orkut and LiveJournal networks. Specifically, we see that for some k_s and q values leading to crash of the original networks, their corresponding randomized networks may survive with a nontrivial proportion of network nodes. This could be understood: it is known that many real-life social networks are assortative networks [7], where high-degree nodes tend to connect high-degree nodes and low-degree nodes tend to connect to low-degree ones. Orkut and LiveJournal networks are of no exception, with assortative coefficient values 0.31 [7] and 0.5625 [8], respectively. In such networks, once some low-degree nodes leave the network as they have fewer than k_s neighbors, their neighbors, being low-degree nodes themselves, may soon find themselves satisfy the KQ-cascade conditions and hence leave as well. A global cascade may therefore be triggered. The extensive randomization operations, by weakening and finally eliminating network assortativity, make global cascade more difficult to happen.

Note that a strongly assortative network, though being more vulnerable to cascading decline, may have one or multiple small communities composed by some high-degree nodes in the original network persist to exist at the end of system crash, as we could observe in Figure S28.

References

1. Dorogovtsev, S. N., Goltsev, A. V. & Mendes, J. F. F. K-core organization of complex networks. *Phys. Rev. Lett.* **96**, 040601 (2006).

2. Leskovec, J. & Faloutsos, C. Sampling from large graphs. In *Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining*, 631-636 (ACM, Philadelphia, PA, USA, 2006).

3. Gjoka, M., Kurant, M., Butts, C. T. & Markopoulou, A. Walking in Facebook: a case study of unbiased sampling of OSNs. In *Proceedings of the 29th IEEE INFOCOM International Conference on Computer Communications*, 1-9 (2010).

4. Stutzbach, D., Rejaie, R., Duffield, N., Sen, S. & Willinger, W. Sampling techniques for large, dynamic graphs. In *Proceedings of the 25th IEEE INFOCOM International Conference on Computer Communications*, 1-6 (2006).

5. DLee, S. H., Kim, P.-J. & Jeong, H. Statistical properties of sampled networks. *Phys. Rev. E* **73**, 016102 (2006).

6. Cho, E., Myers, S. A. & Leskovec, J. Friendship and mobility: user movement in location-based social networks. In *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 1082–1090 (ACM, San Diego, California, USA, 2011).

7. Ahn, Y.-Y., Han, S., Kwak, H., Moon, S., Jeong, H. Analysis of topological Characteristics of huge online social networking services. in *Proceedings of 16th International World Wide Web Conference*, 835-844 (ACM, Banff, Alberta, Canada, 2007).

8. Sathe, S. Rumor spreading in LiveJournal. technical report EPFL-STUDENT-176326, from http://infoscience.epfl.ch/record/176326/files/project-report.pdf (2008).

Figure Legends

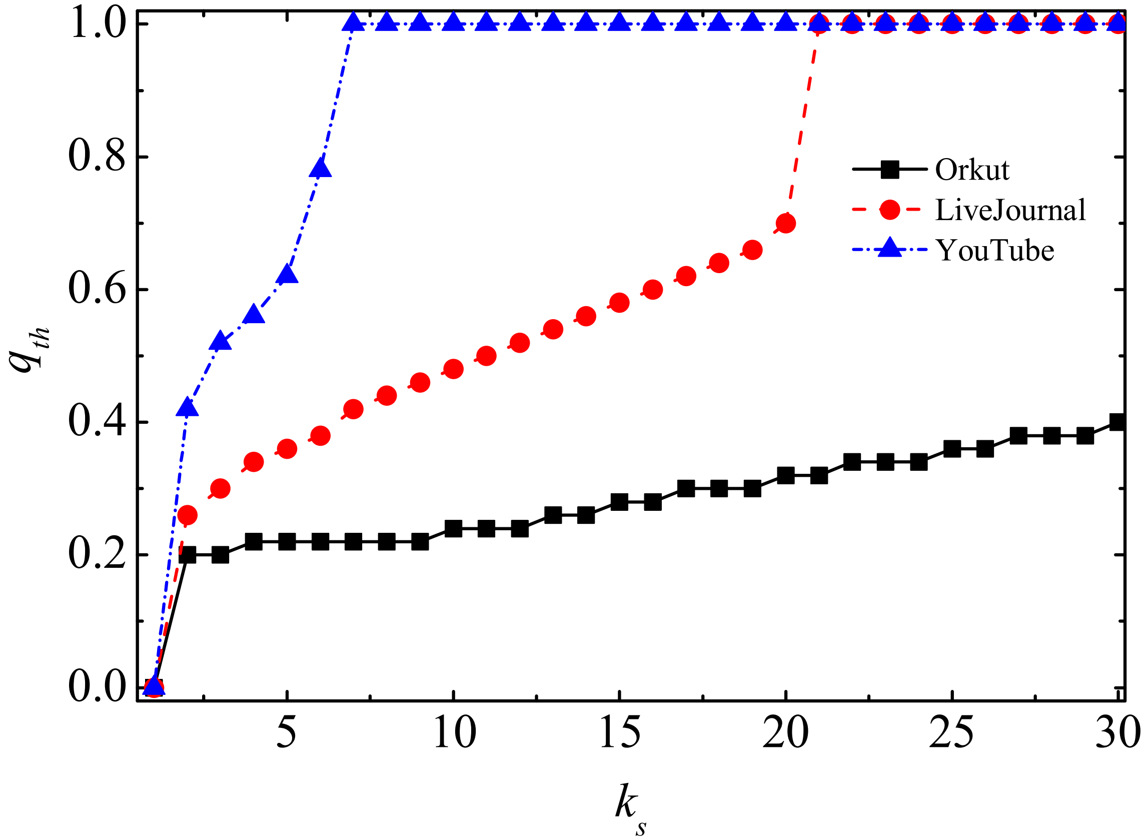

Figure S1. Comparison of the threshold value q_{th} in real-life networks.

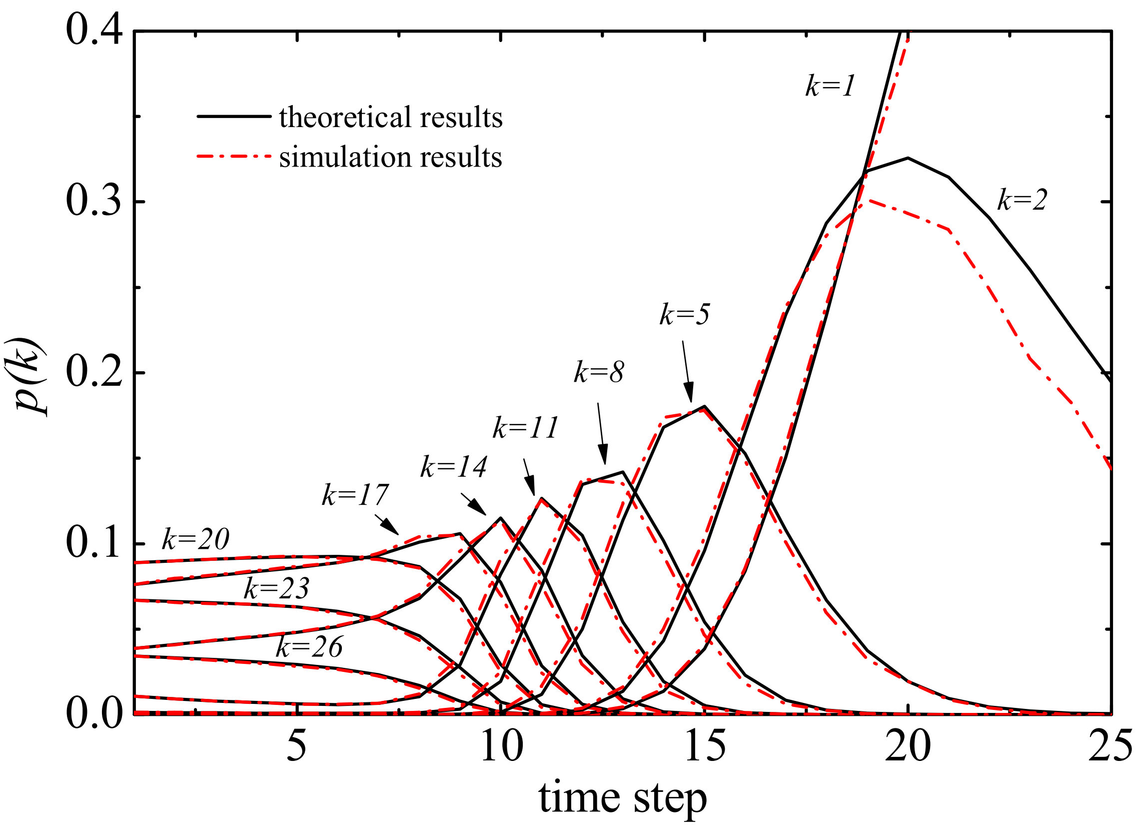

Figure S2. Evolution of $p(k)$ during the cascade process in the ER random network with $q = 0.1$ and $k_s = 14$, where the network finally crashes.

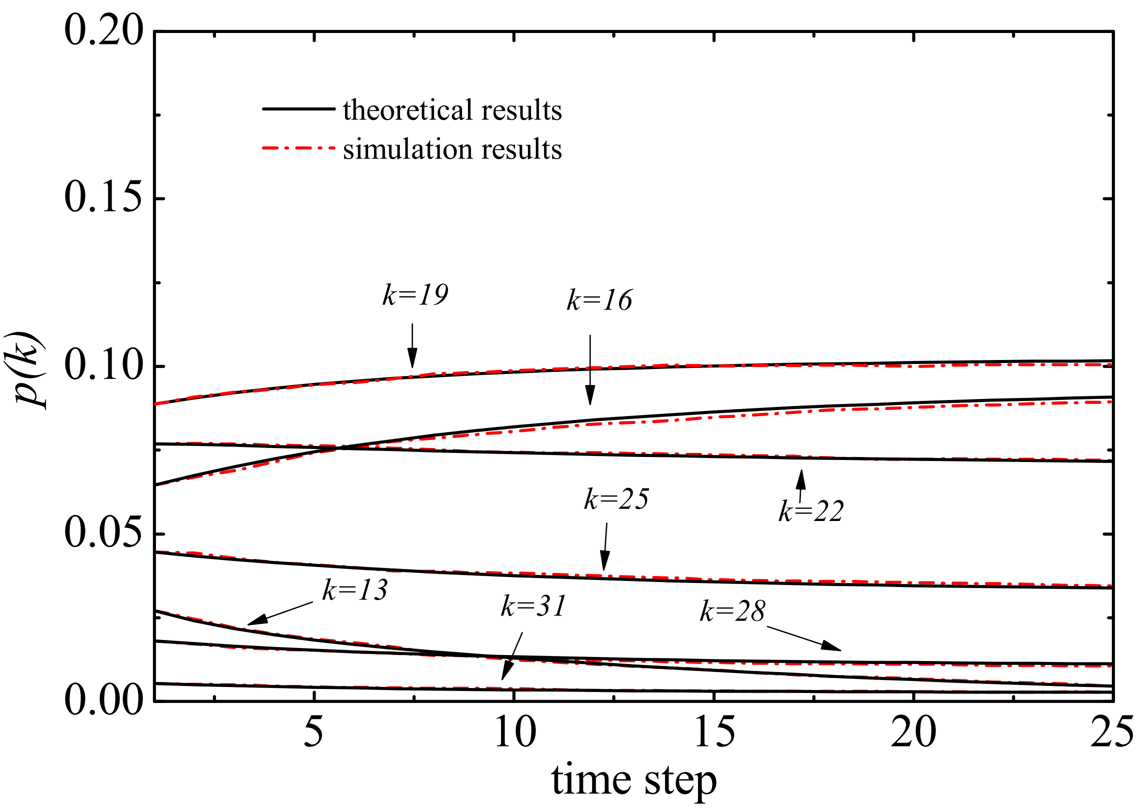

Figure S3. Evolution of $p(k)$ during the cascade process in the ER random network with $q = 0.3$ and $k_s = 14$, where the network survives. Note that for $k < 13$, $p(k)$ quickly drops to 0.

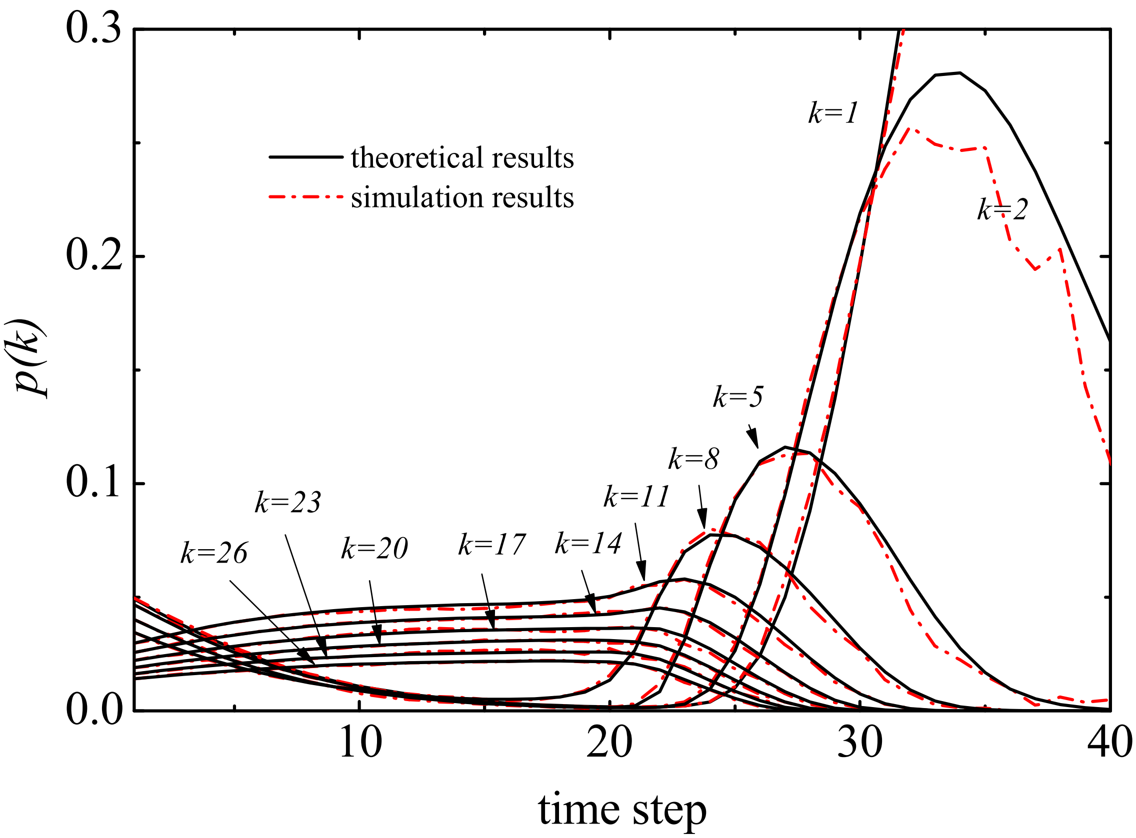

Figure S4. Evolution of $p(k)$ during the cascade process in the random exponential network with $q = 0.2$ and $k_s = 10$, where the network finally crashes.

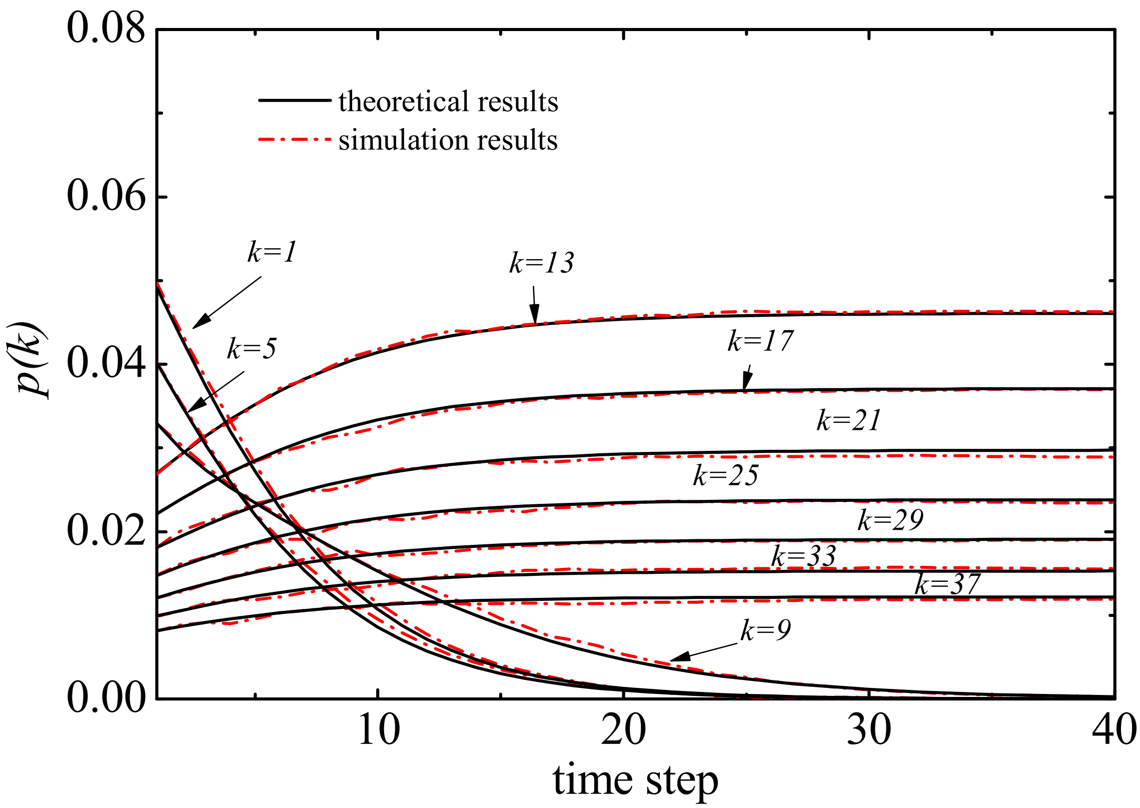

Figure S5. Evolution of $p(k)$ during the cascade process in the random exponential network with $q = 0.3$ and $k_s = 10$, where the network survives.

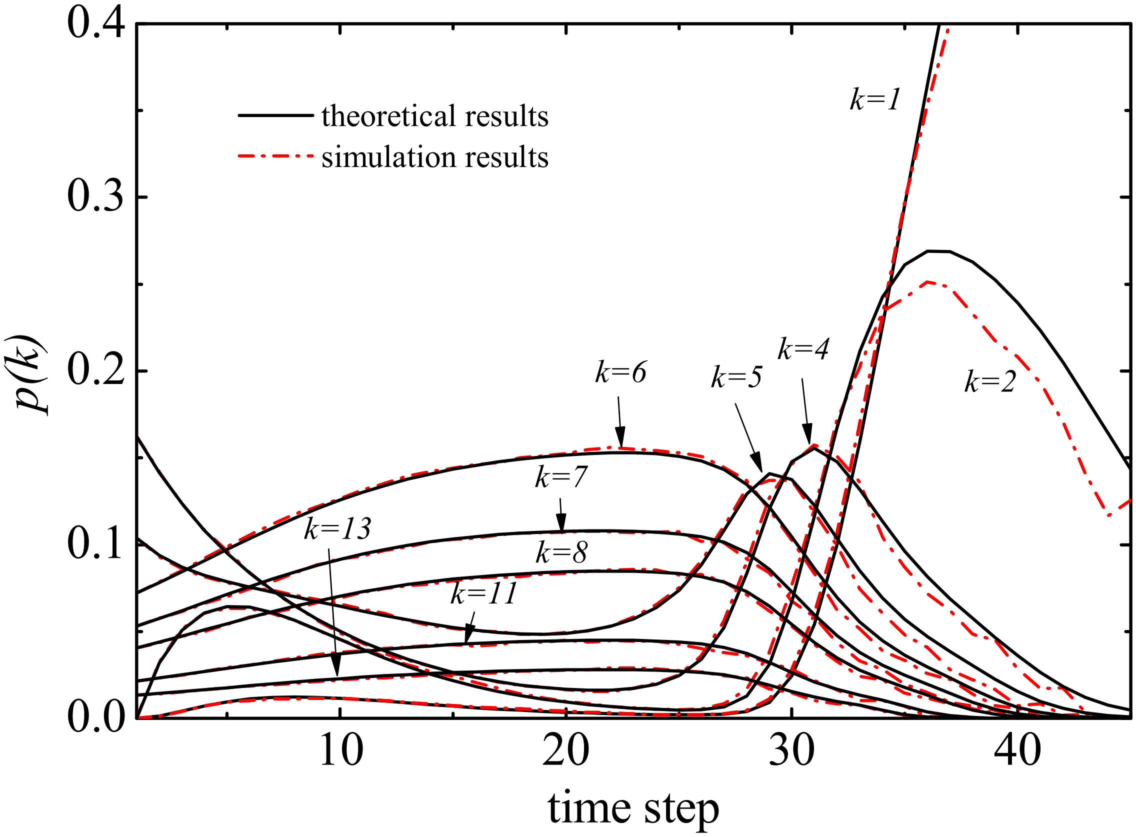

Figure S6. Evolution of $p(k)$ during the cascade process in the random scale-free network with $q = 0.4$ and $k_s = 6$, where the network finally crashes.

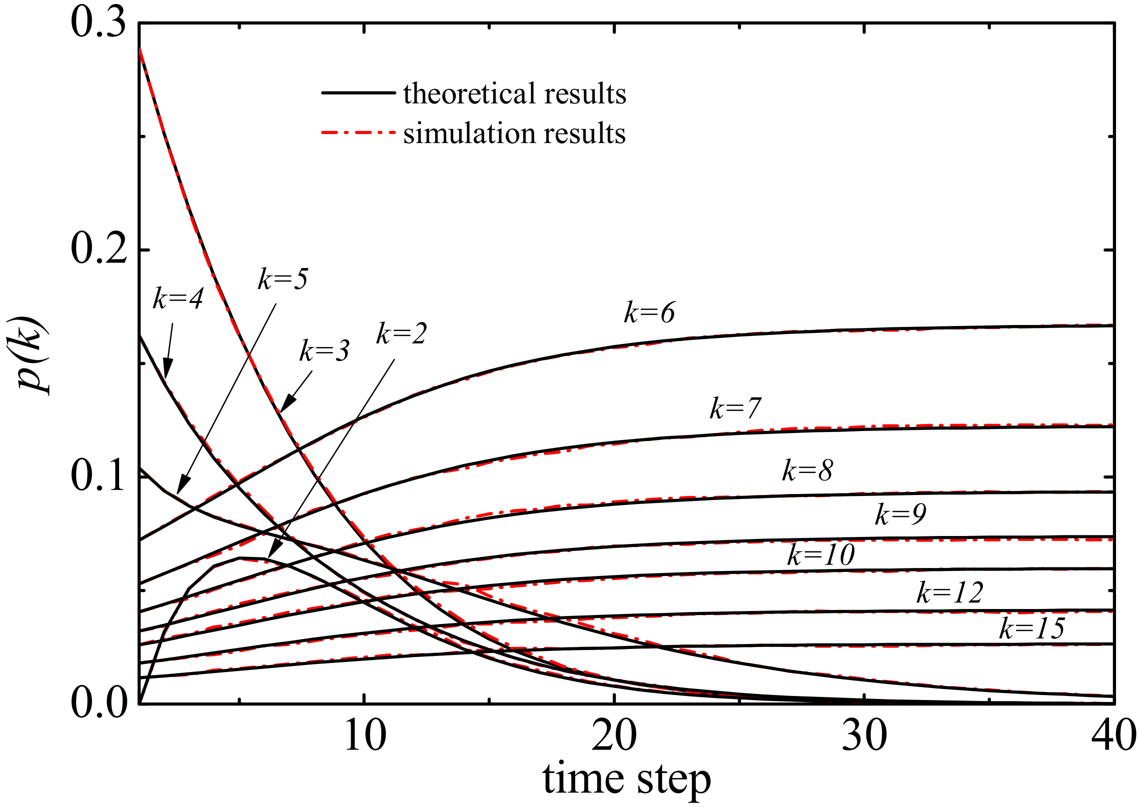

Figure S7. Evolution of $p(k)$ during the cascade process in the random scale-free network with $q = 0.5$ and $k_s = 6$, where the network survives.

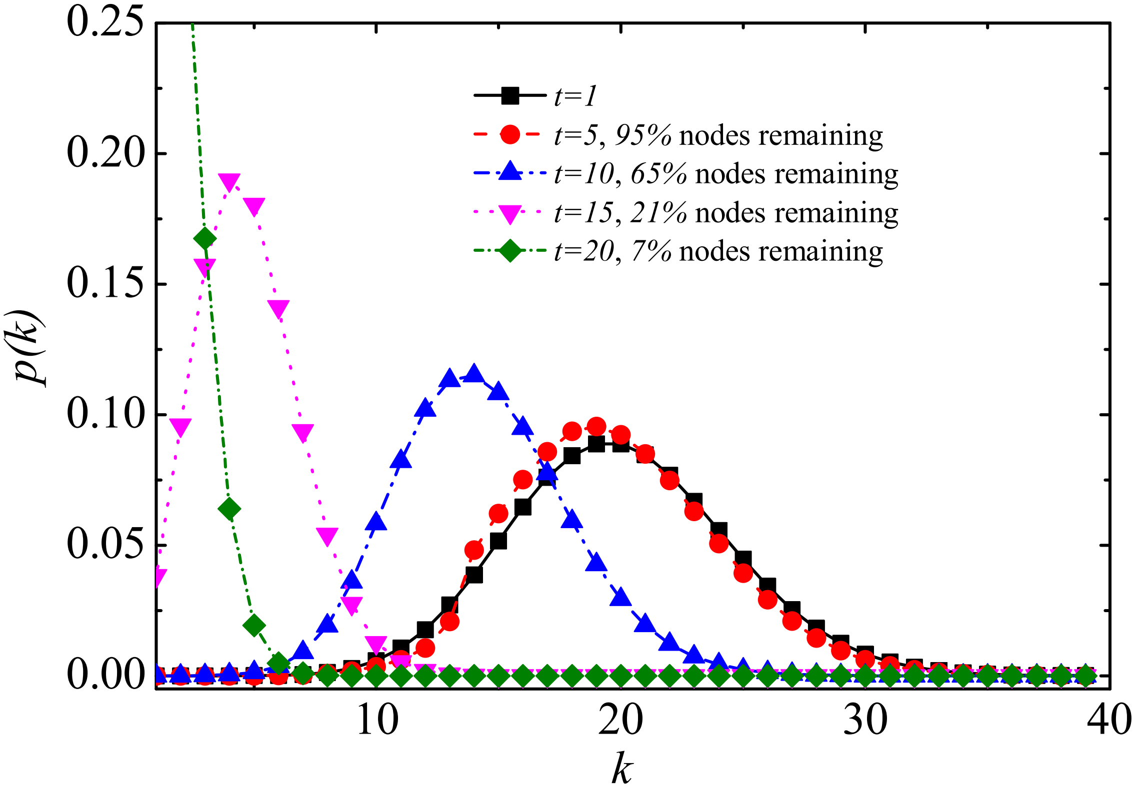

Figure S8. Degree distribution of the remaining nodes at different time during cascade process in the ER random network with $q = 0.1$ and $k_s = 14$, where the network finally crashes.

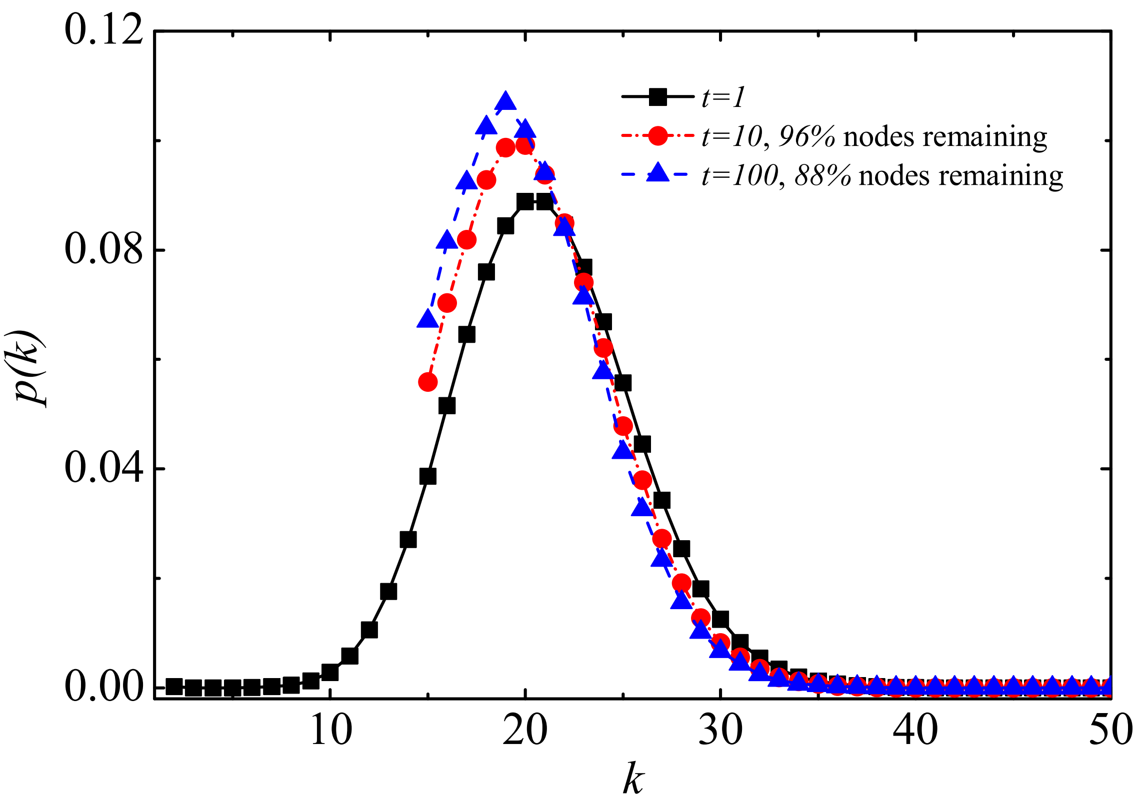

Figure S9. Degree distribution of the remaining nodes at different time during cascade process in the ER random network with $q = 0.3$ and $k_s = 14$, where the network survives.

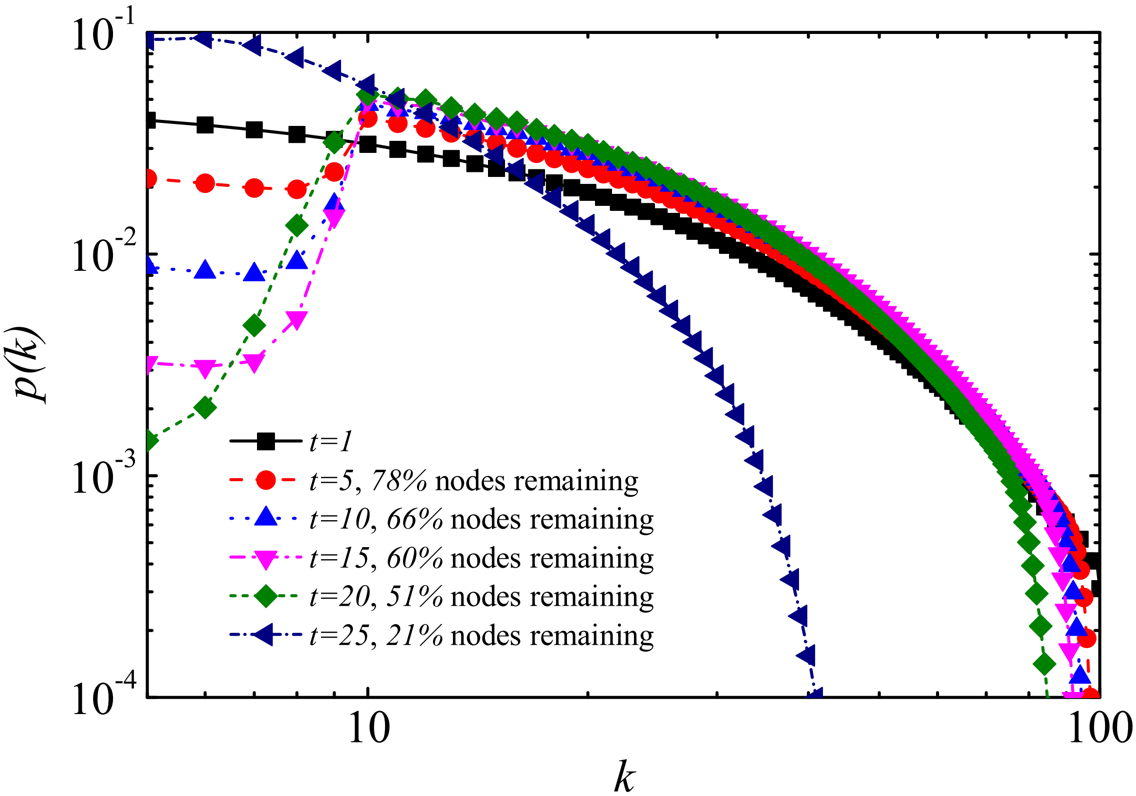

Figure S10. Degree distribution of the remaining nodes at different time during cascade process in the random exponential network with $q = 0.2$ and $k_s = 10$, where the network finally crashes.

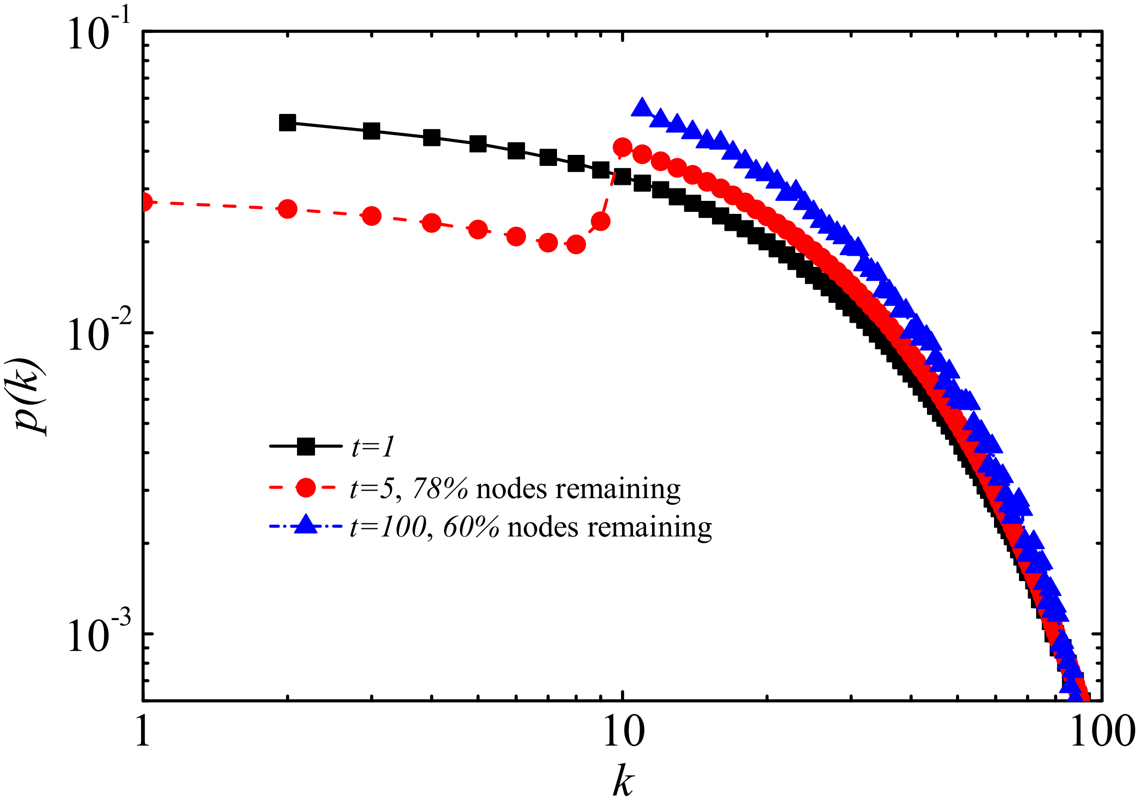

Figure S11. Degree distribution of the remaining nodes at different time during cascade process in the random exponential network with $q = 0.3$ and $k_s = 10$, where the network survives.

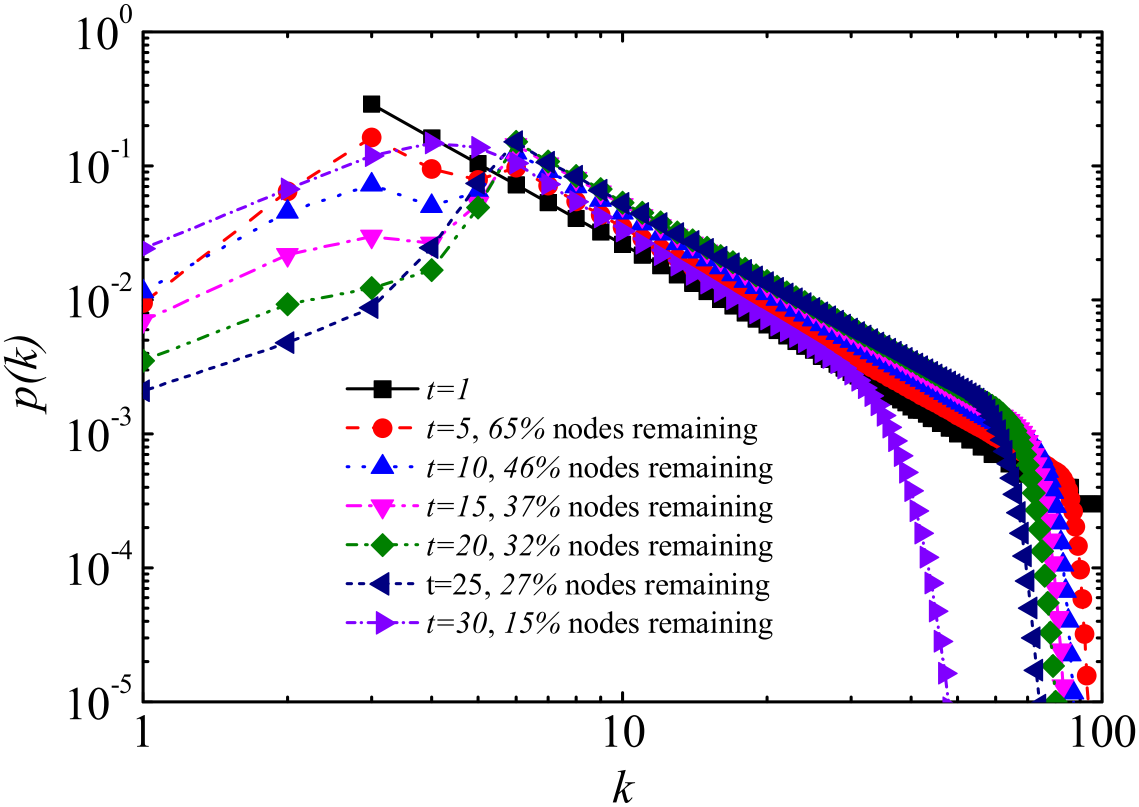

Figure S12. The degree distribution of the remaining nodes at different time in the random scale-free network with $q = 0.4$ and $k_s = 6$, where the network finally crashes.

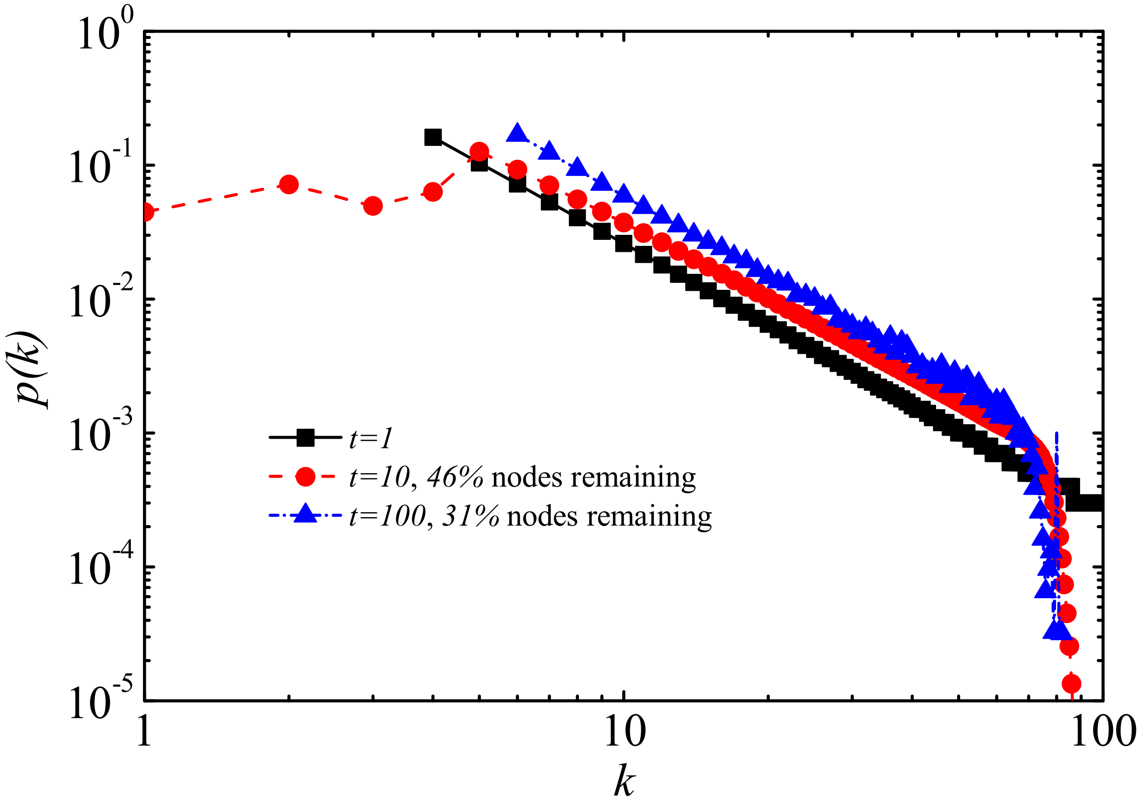

Figure S13. The degree distribution of the remaining nodes at different time in the random scale-free network with $q = 0.5$ and $k_s = 6$, where the network survives.

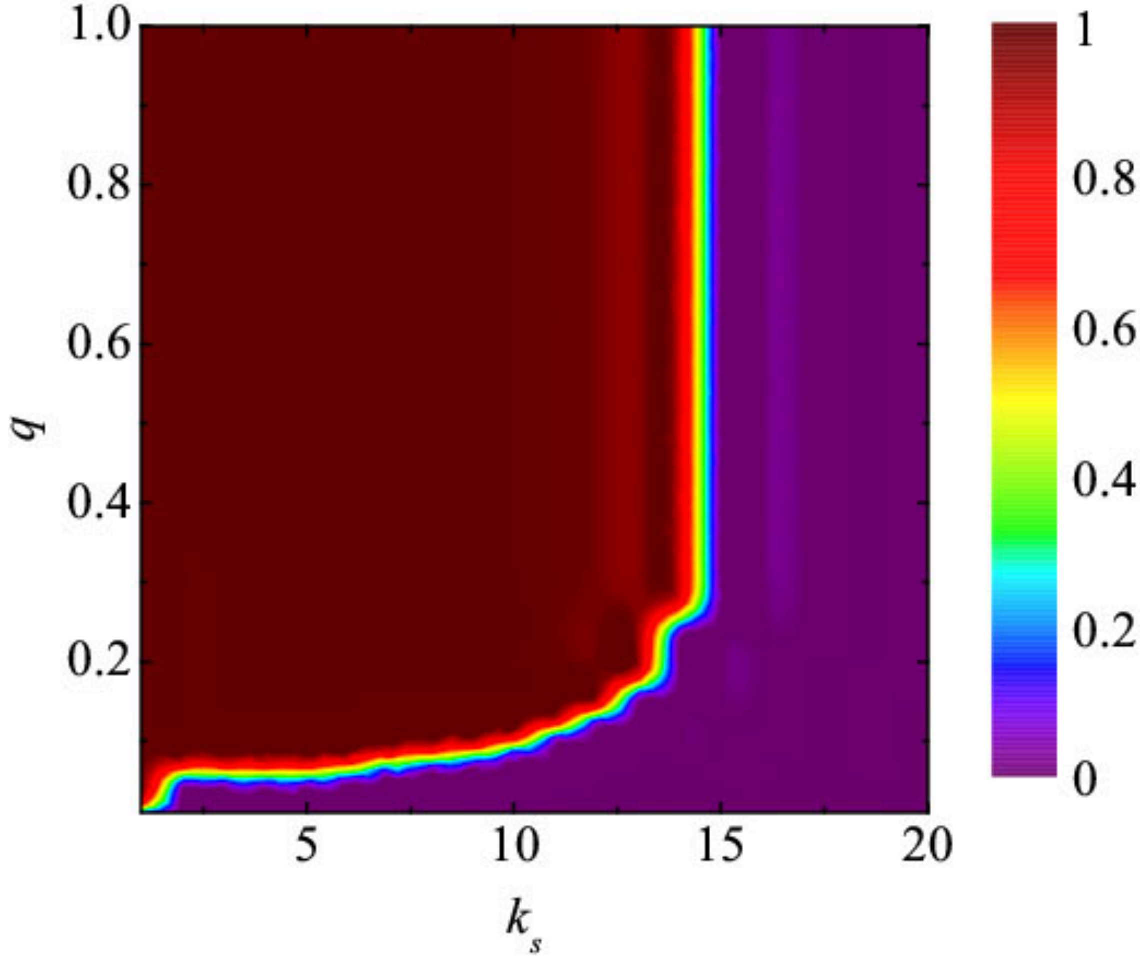

Figure S14. Cascade size of ER random network with z=20 for different values of q and k_s. The cascade size is shown in color scale.

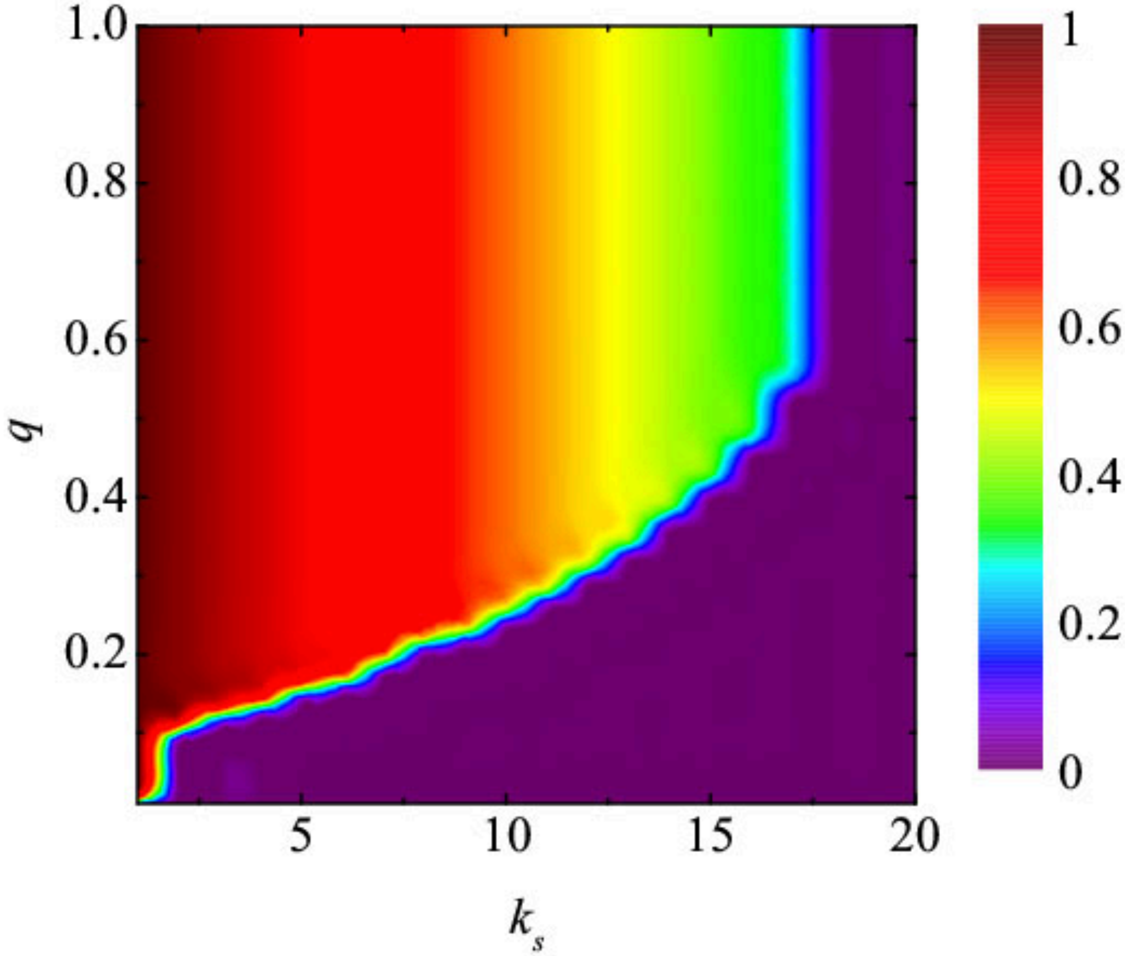

Figure S15. Cascade size of random exponential network with $z = 20$ for different values of q and k_s. The cascade size is shown in color scale.

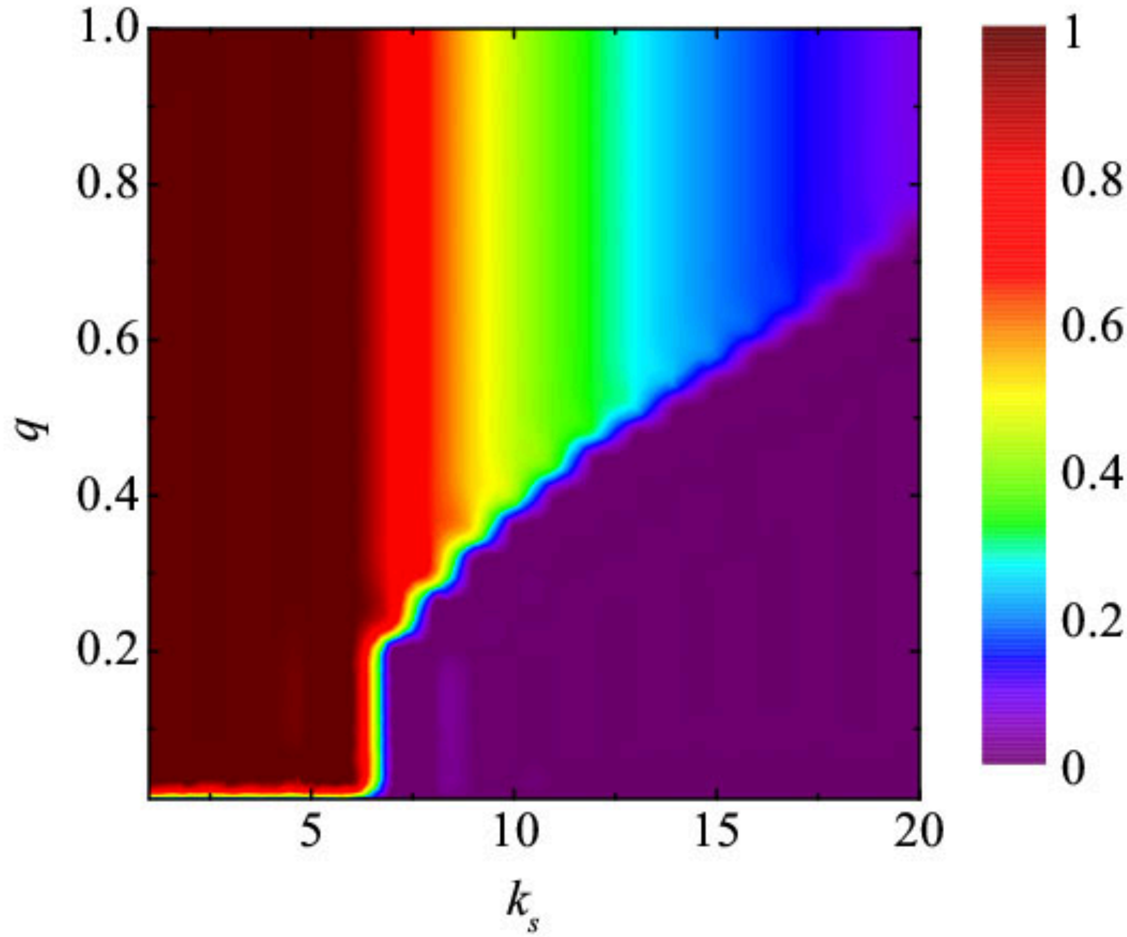

Figure S16. Cascade size of random scale-free network with $\gamma = 2$ for different values of q and k_s. The cascade size is shown in color scale.

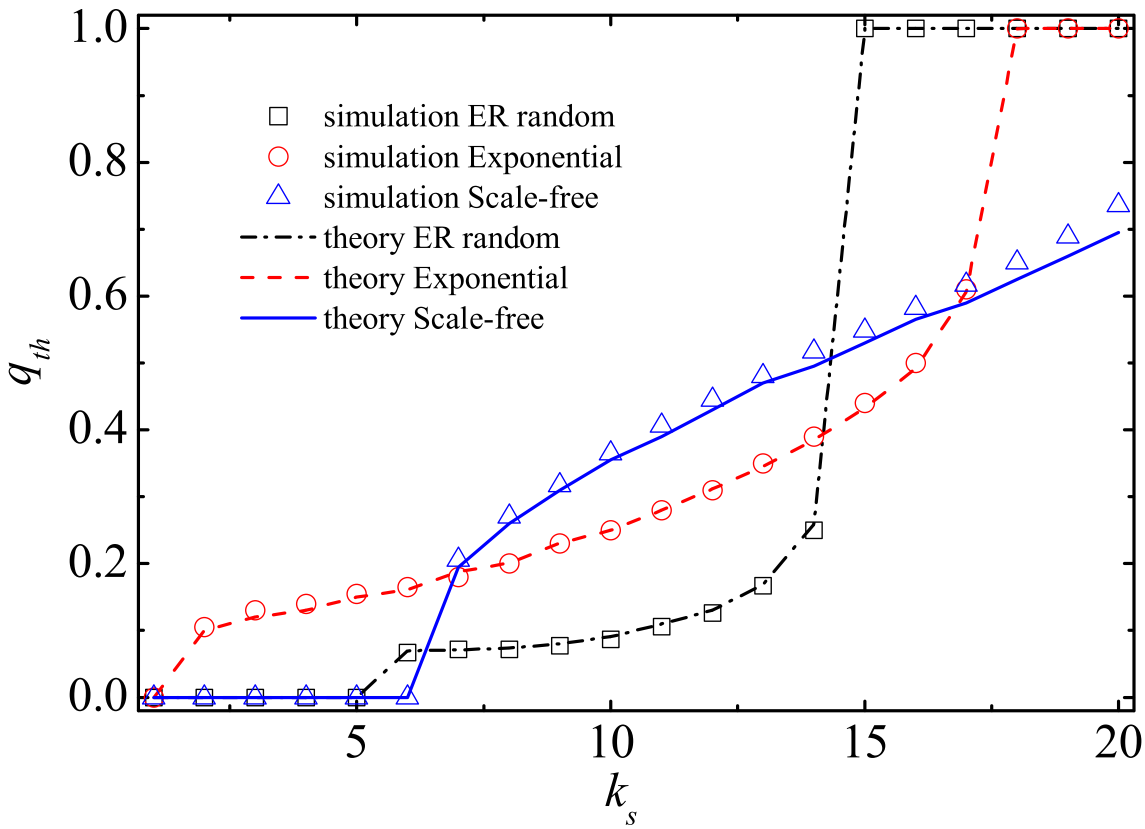

Figure S17. Comparison between the analytical (circles) and simulation results (lines) of cascade threshold in ER random network (red), exponential network (black) and scale-free network (blue) with the same average degree of $z = 20$. Both the analytical and simulation results of threshold are obtained by trial and error. Specifically, for each k_s, we test by increasing the value q by a step length of 0.01 until the threshold value is obtained. The simulation results show the average in 10 randomly generated networks with size $N = 10^4$.

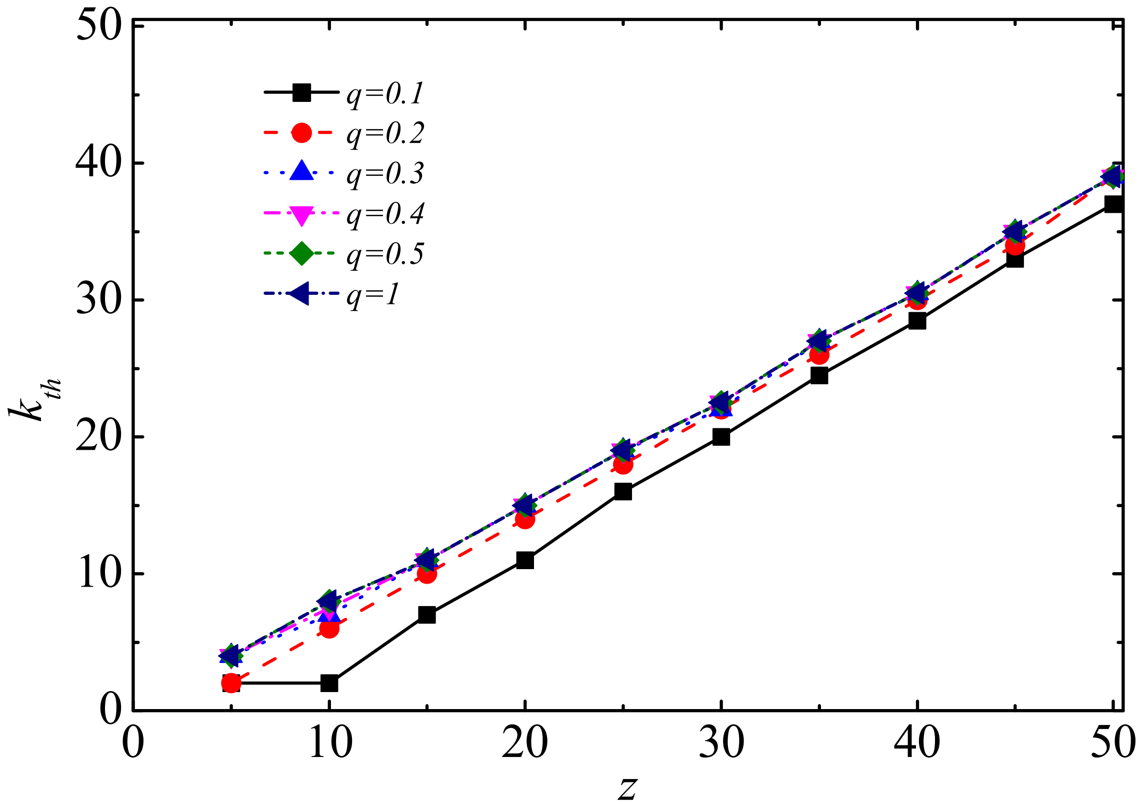

Figure S18. Relation between the average degree z and threshold k_{th} in the ER random network with different values of q.

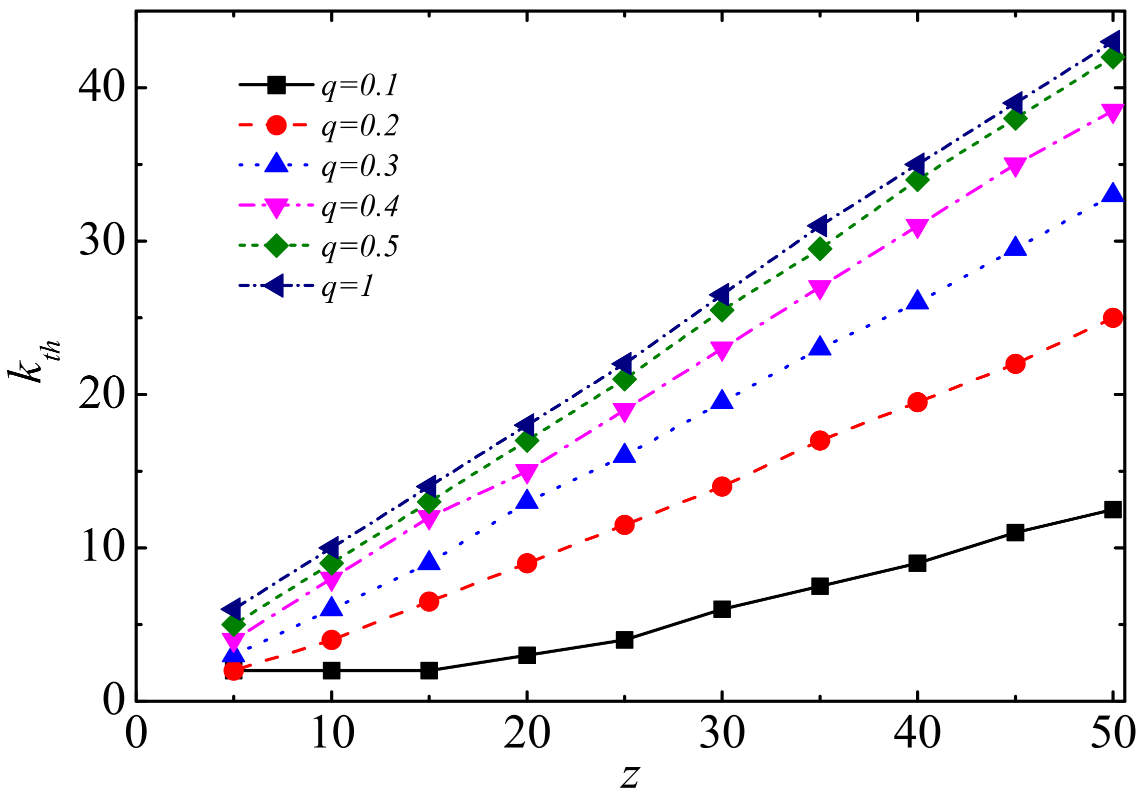

Figure S19. Relation between the average degree z and threshold k_{th} in the random exponential network with different values of q.

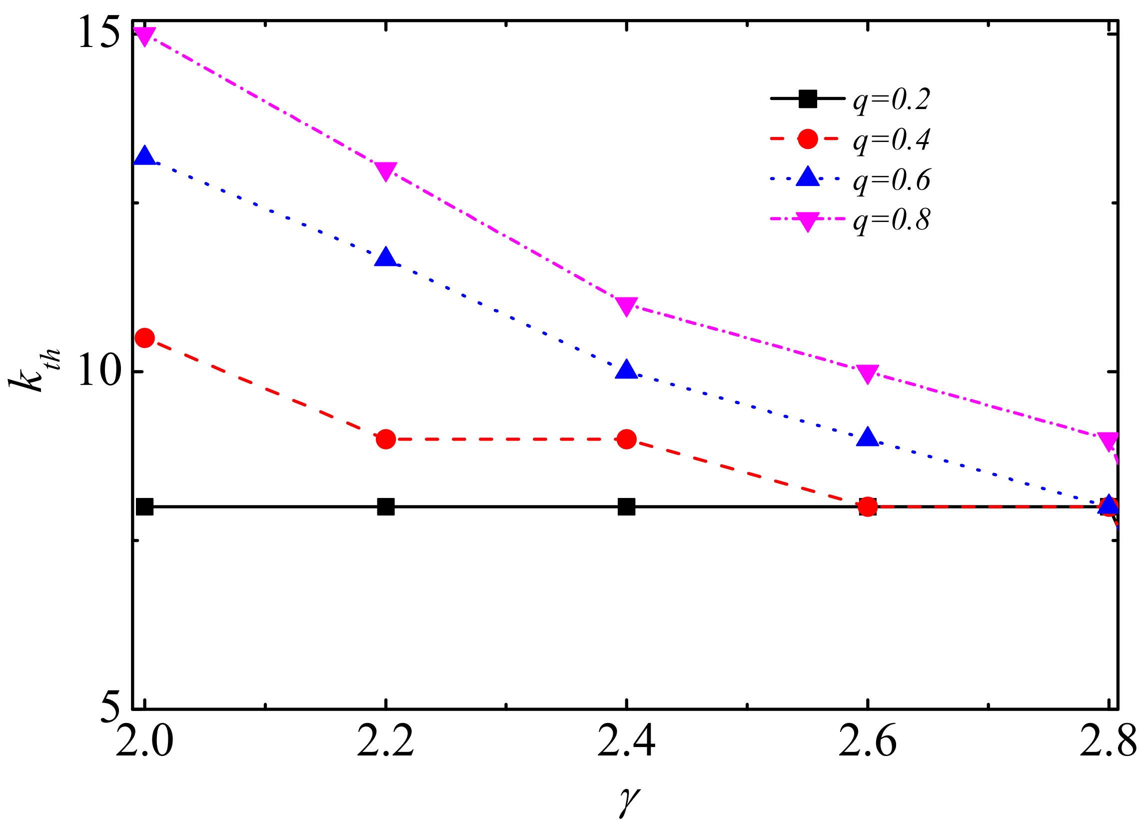

Figure S20. Relation between the power-law exponent γ and threshold k_{th} in the random exponential network with different values of q.

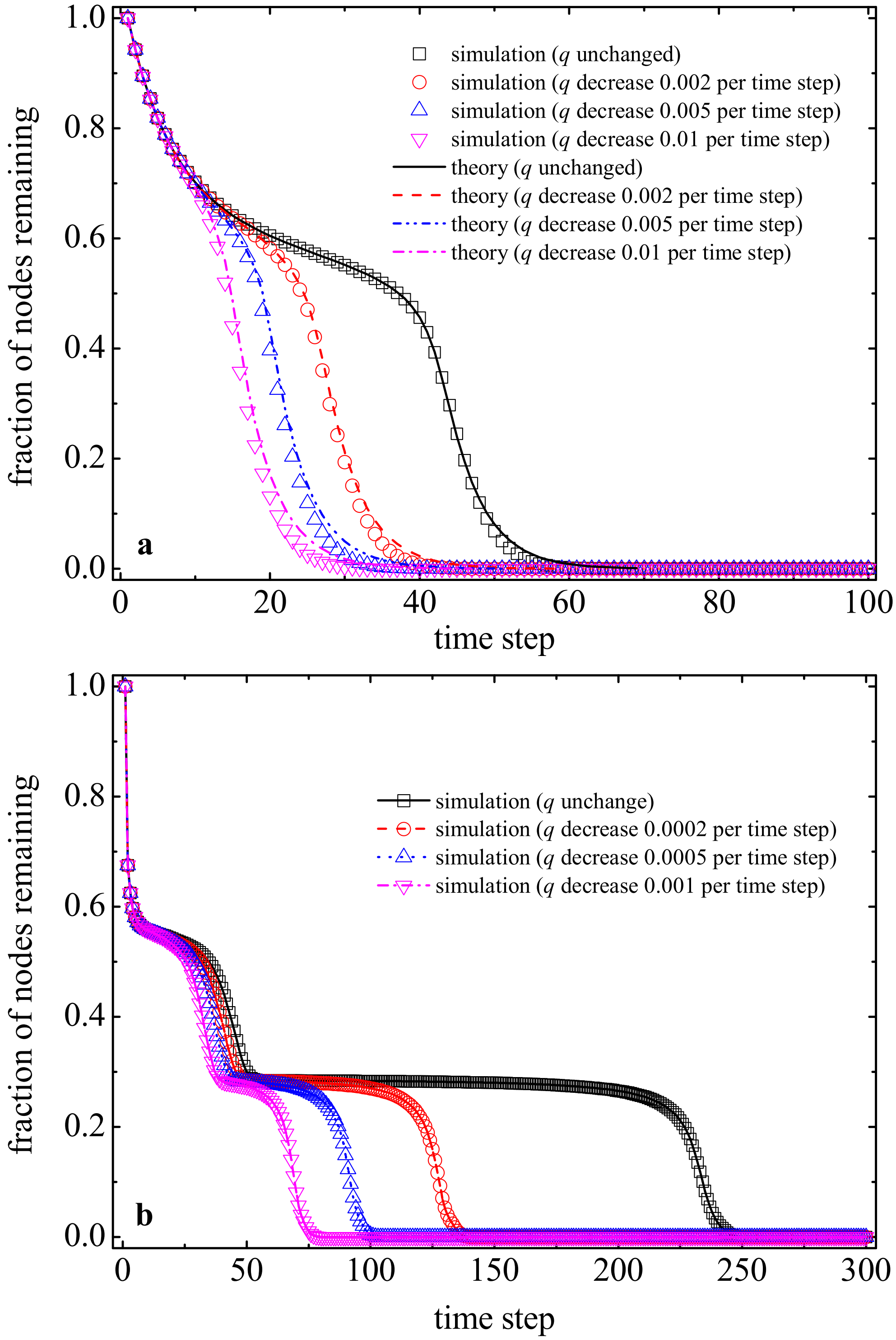

Figure S21. Decline of networks with speedup loss of nodes. Simulation results in **(a)** random scale-free network with $k_s = 4$, $q = 0.28$, $f = 0.2$, averaged over 100 independent realizations; and **(b)** the Orkut online social network with $k_s = 20$, $q = 0.3$, $f = 0.5$, averaged over 10 independent cascades.

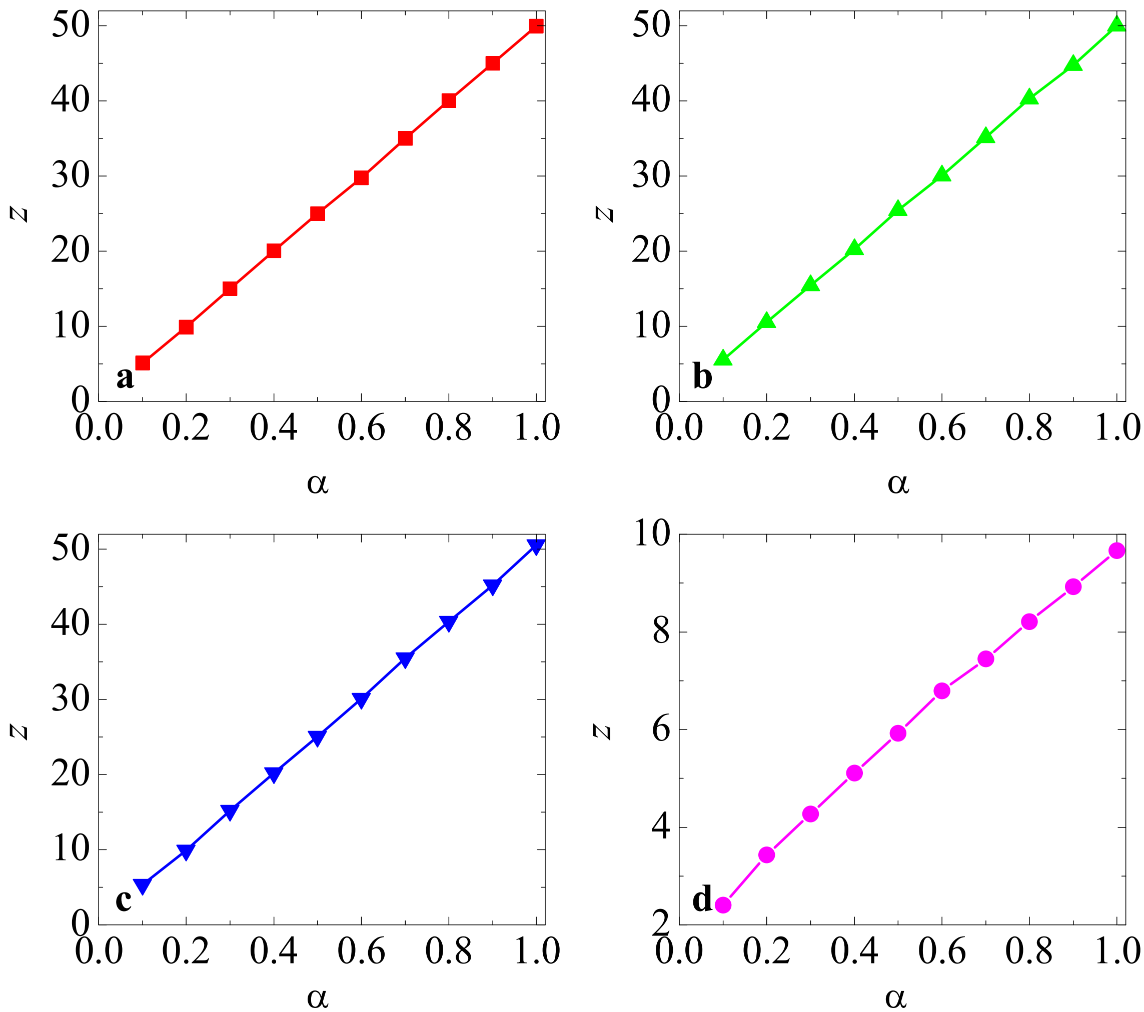

Figure S22. Sample percentage α vs. average degree z. (a) ER random network, (b) random exponential network, (c) random scale-free network; and (d) Gowalla local social network.

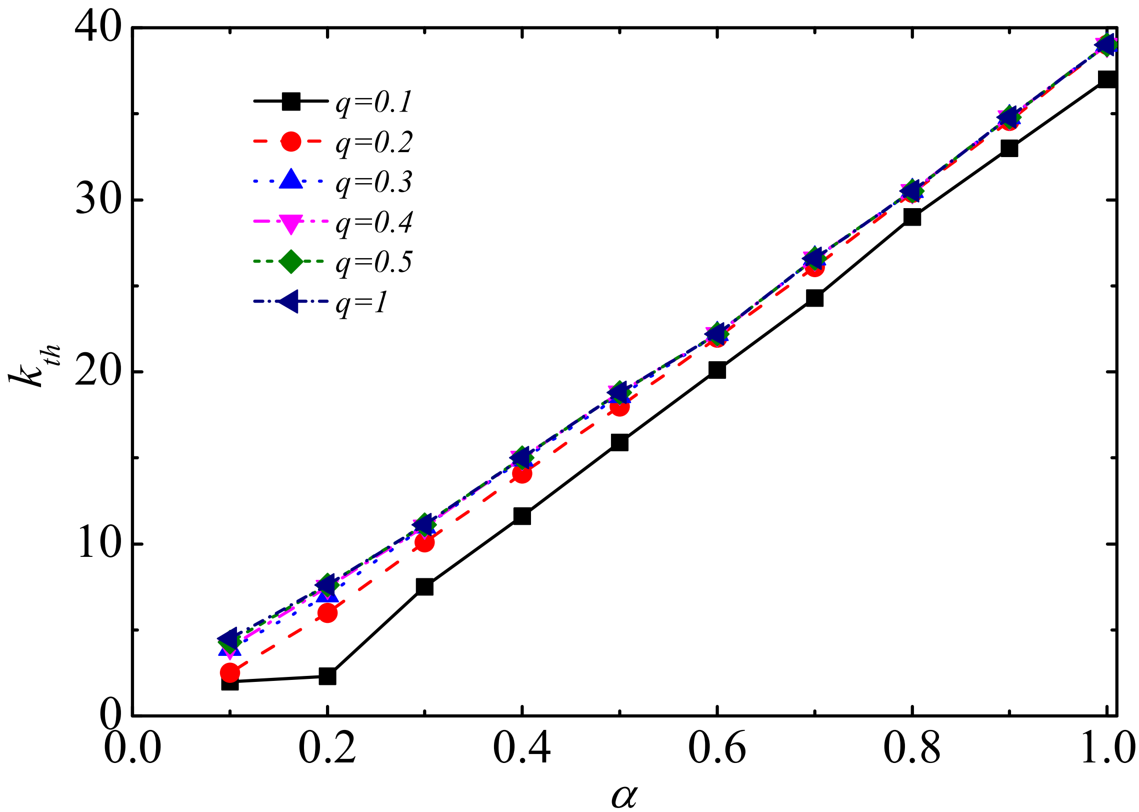

Figure S23. Relation between the sample percentage α and threshold k_{th} of the sampled subnet of the ER random network with different values of q.

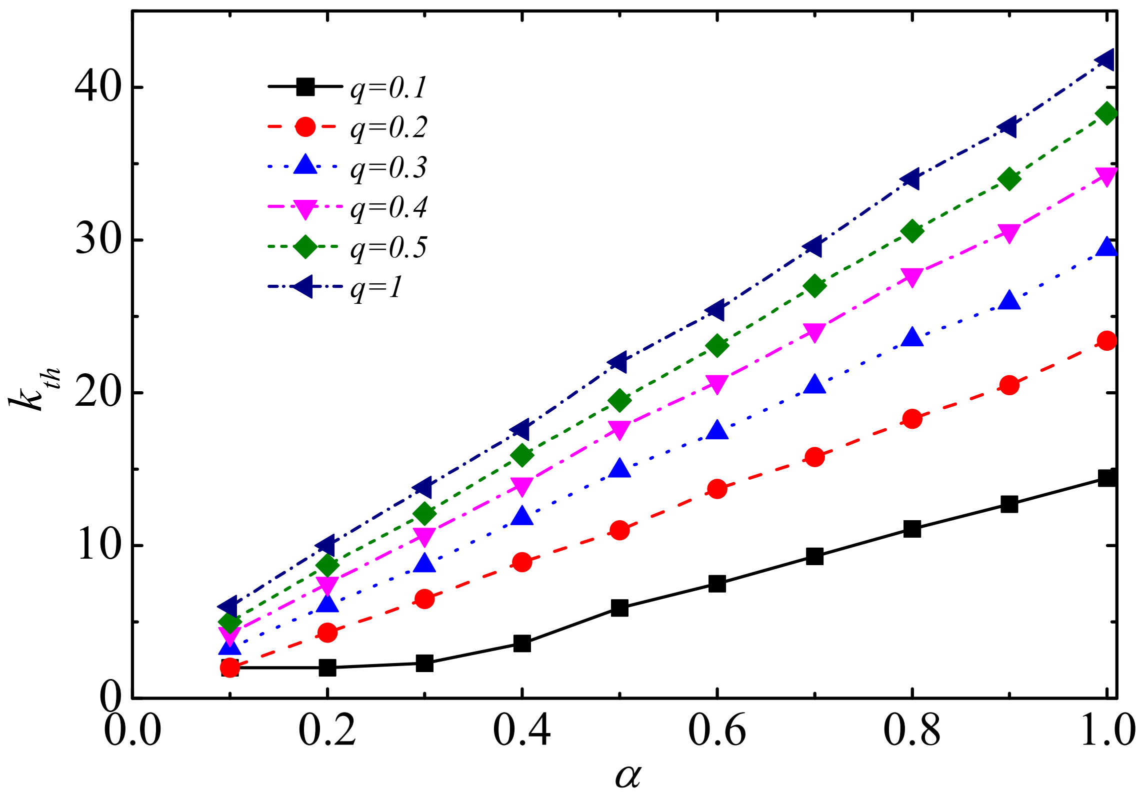

Figure S24. Relation between the sample percentage α and threshold k_{th} of the sampled subnet of the random exponential network with different values of q.

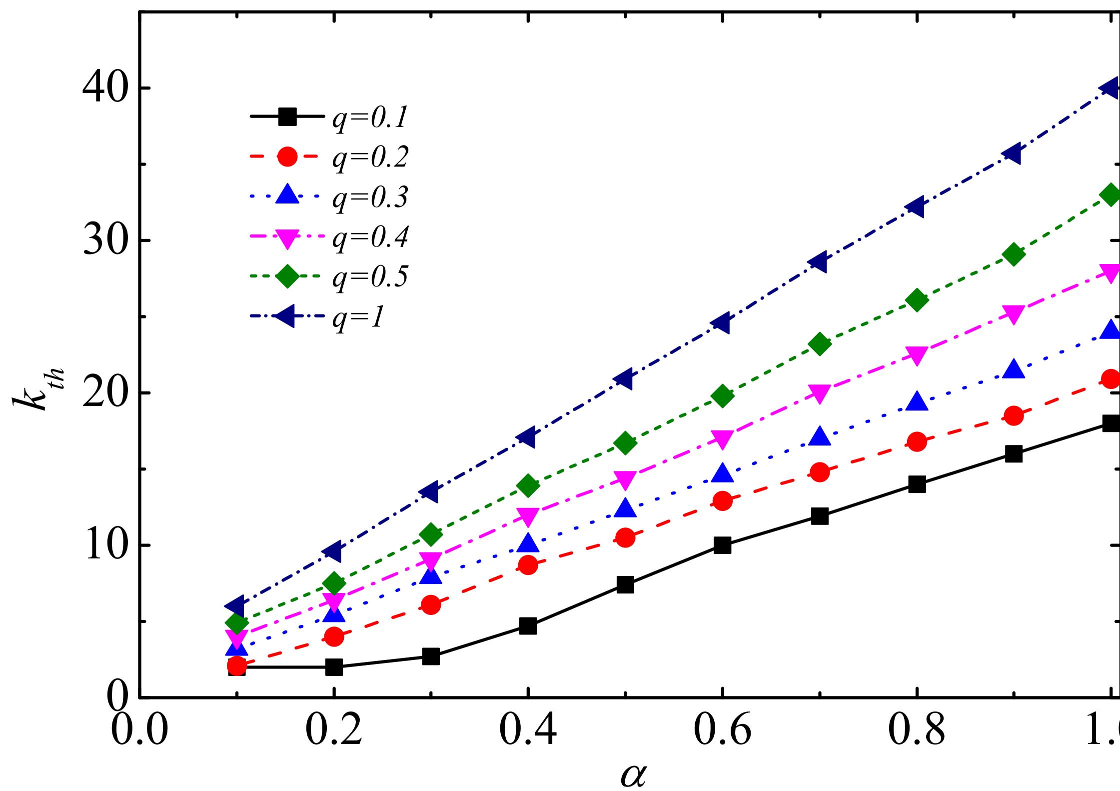

Figure S25. Relation between the sample percentage α and threshold k_{th} of the sampled subnet of the random scale-free network with different values of q.

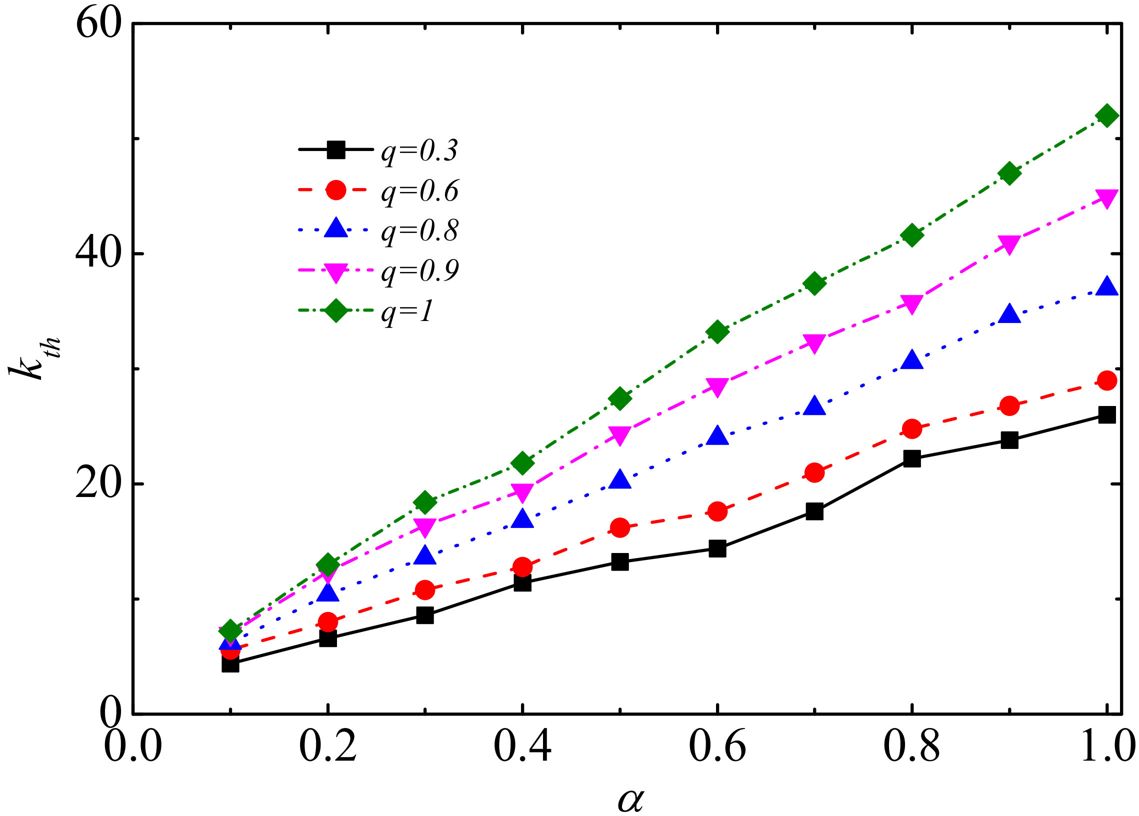

Figure S26. Comparison between the sample percentage α and threshold k_{th} of the sampled subnet of the Gowalla network with different values of q.

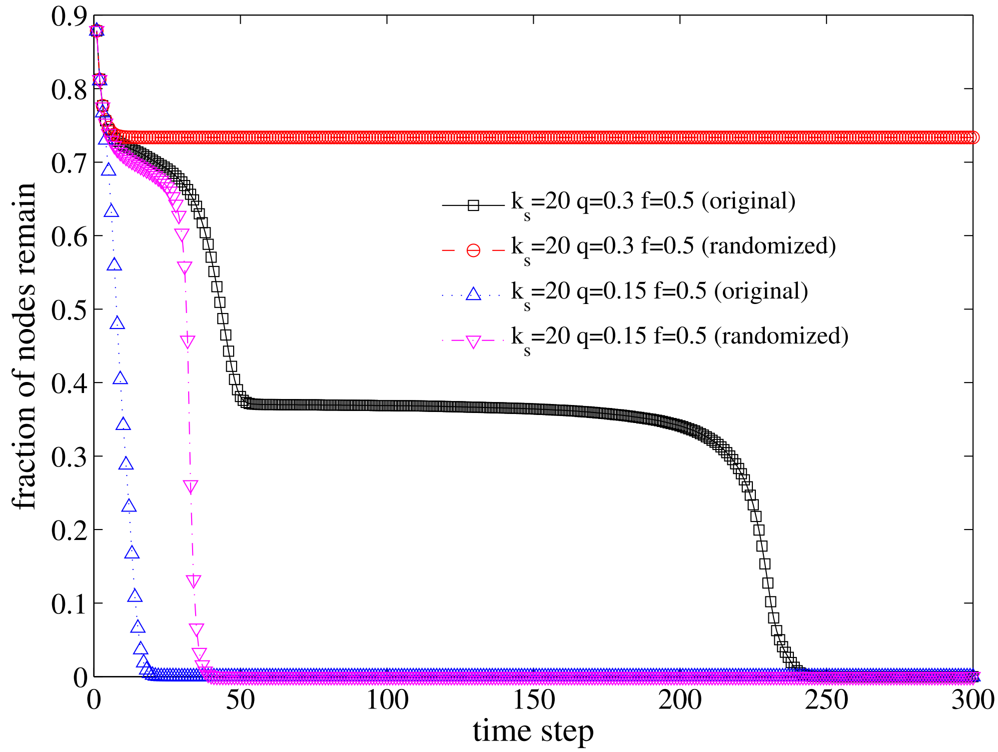

Figure S27. Simulation results on the Orkut network. The randomized network is generated by applying the randomization operation for 100-million times on the original network.

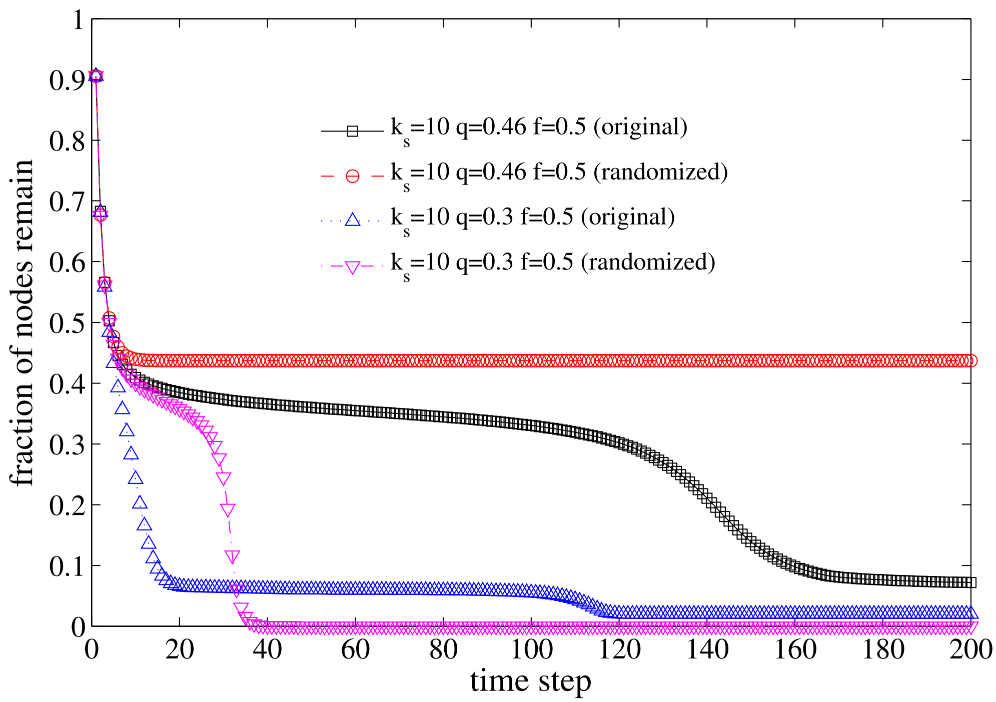

Figure S28. Simulation results on the LiveJournal network. The randomized network is generated by applying the randomization operation for 100-million times on the original network.