

## Regular Paper

# Revisiting the Robustness of Complex Networks against Random Node Removal

KAZUYUKI YAMASHITA<sup>1,a)</sup> YUICHI YASUDA<sup>2,b)</sup> RYO NAKAMURA<sup>2,c)</sup> HIROYUKI OHSAKI<sup>2,d)</sup>

Received: November 12, 2018, Accepted: June 11, 2019

**Abstract:** It is widely known that scale-free networks are robust against random node removal, which is one of major interesting findings in network science. This suggests that, for instance, communication networks such as the Internet is robust against random node failures caused by breakdowns and/or malicious attacks if their network topologies are scale-free networks. Generally, the ratio of failed devices (e.g., routers) to operational devices is not extremely high. In this paper, we revisit the robustness of complex networks against random node removal. Through simulations, we compare the robustness of scale-free and non-scale-free networks against random node removal as well as random edge removal. Our findings include that, contrary to common understanding, non-scale-free networks are more robust than scale-free networks except under extremely high node removal ratio. We also show that the robustness of non-scale-free networks can be further improved by bounding the minimum node degree of those networks.

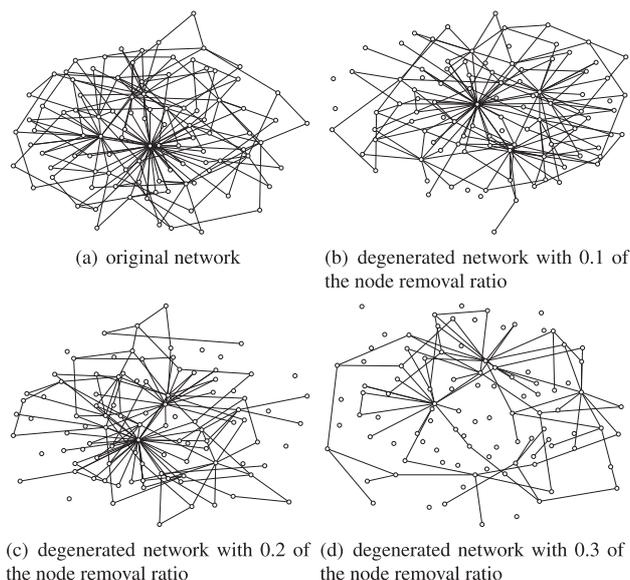
**Keywords:** robustness, complex network, scale-free network, non-scale-free network, random node removal, largest component

## 1. Introduction

In the literature, scale-free networks are widely known for their robustness against random node removal (**Fig. 1**) which is a major finding in network science and the robustness of scale-free networks against node removals (e.g., random node failures/attacks in communication networks) has been extensively studied [1], [2]. For instance, authors of Ref. [1] showed that scale-free networks are robust against random node removals since the connectivity of a network can be preserved because of the existence of *hub* nodes (i.e., a small number of high-degree nodes), even if some of the nodes were eliminated. In contrast, the authors showed that by removing high-degree nodes, the diameter (i.e., the average path length of shortest-paths between any node pair in a network) of scale-free networks rapidly increases as the node removal ratio increases.

Figure 1 illustrates how the network connectivity is degraded when a portion of the nodes (i.e., vertices) are randomly removed from the network. As the node removal ratio increases, nodes are more likely to be disconnected from the network and also clusters of nodes are isolated each other.

However, contrary to the above, several questions on the robustness of scale-free networks and its implications for communication networks have been raised [3], [4]. For instance, the authors of Ref. [3] have pointed out the confusion in Ref. [1]



**Fig. 1** An example of random node removal.

from not distinguishing between AS-level network topologies and router-level network topologies. Even if an AS-level network topology has scale-free property, it does not mean the underlying router-level topology has scale-free property. Also, authors of Ref. [3] suggest that router-level network topologies might not be scale-free since routers in the Internet have practical limitation on the number of links (i.e., the number of communication interfaces).

In this paper, we revisit the robustness of complex networks against random node removal. The above description clearly states that a scale-free network can be robust even if a significant portion of its nodes are removed. However, the finding that scale-

<sup>1</sup> Department of Informatics, School of Science and Technology, Kwansai Gakuin University, Sanda, Hyogo 669-1337, Japan

<sup>2</sup> Department of Informatics, Graduate School of Science and Technology, Kwansai Gakuin University, Sanda, Hyogo 669-1337, Japan

a) kazuyuki@kwansai.ac.jp

b) yuichi@kwansai.ac.jp

c) r-nakamura@kwansai.ac.jp

d) ohsaki@kwansai.ac.jp

free networks are robust, might not be valid under typical node removal ratios in actual computer networks. However, the finding that scale-free networks are robust might not be valid under a typical node removal ratio in actual computer networks. Generally, the ratio of failed devices (e.g., routers) and communication links which comprise computer networks such as the Internet is not extremely high [5], [6]. For this reason, it is necessary to clarify the robustness of scale-free and non-scale-free networks under typical node removal ratios in actual computer networks.

In this paper we use simulations to compare the robustness of scale-free and non-scale-free networks by way of scale-free networks generated with Barabási Albert (BA) model, randomized BA model and Li-Maini model [7] and non-scale-free networks generated with Erdős-Rényi (ER) model and Degree-Bounded (DB) model against different levels of random node removal. Specifically, we compare the largest component sizes (i.e., the number of nodes contained in the largest cluster [1]) in five classes of networks (networks generated with BA, randomized BA, ER, DB, and Li-Maini models) after random node removal.

The main contributions of this paper are as follows.

- We reveal that, contrary to common understanding, random networks are more robust than scale-free networks except under extremely high node removal ratios.
- We show that the robustness of non-scale-free networks can be further improved by bounding the minimum node degree of those networks.

Our findings imply that under extremely severe failures (e.g., 90% of routers were destroyed or malfunctioning due to some reason), scale-free communication networks would actually be more robust than non-scale-free communication networks. If communication networks are scale-free, most of the remaining 10% of nodes would likely still be connected to other nodes. However, on the contrary, if communication networks are non-scale-free, those 10% nodes would be likely to be isolated with others. However, if the failure ratio is not so exceptional (e.g., if the failure ratio is around 1–20%) [5], [6], then non-scale-free communication networks are *more* robust than scale-free communication networks.

Phase transition at the critical threshold in a complex network or namely the giant cluster in the network will disappear as the node removal ratio exceeds the critical threshold is an interesting phenomenon. Hence, a vast number of studies in the literature investigate the robustness of complex networks near the critical threshold. However, we should take account the likelihood of those failures. For instance, what are the chances that 90% of routers in the Internet were destroyed? Or that a 1% node failure is likely to happen. The occurrence probability of 5% node failure is much smaller than that of 1% node failure. Our findings indicate that communication networks should be designed by taking into account the probability of different levels of network failures occurring.

This paper is organized as follows. Section 2 discusses related studies on the robustness of complex networks. Section 3 explains the method to investigate the robustness of scale-free and non-scale-free networks against random node removal. Section 4 presents our simulation results under different network sizes and

densities. Section 5 investigates the case of random *link* removal to investigate whether our findings in Section 4 are valid in regard to other types of errors. Finally, Section 6 provides the summary of this paper and discusses future works.

## 2. Related Works

Robustness of complex networks is a well-studied topics in network science and significant research has been devoted to both mathematical and experimental studies of non-scale-free and scale-free networks [1], [2], [3], [4], [8], [9], [10].

Types of failures in complex networks are classified by their *randomness* (either random or deterministic e.g. adversary) and their *location* (either node or link).

One of major findings is that scale-free networks are robust against random node removal whereas those networks are fragile against adversary node removal [1], [2], [8], [9]. In the literature, there have been many studies on the evaluation of robustness of scale-free networks against random node failure/attack in terms of the diameter and the giant component size [1], and the efficiency [8] (a measure of how efficiently information is exchanged over a network which is defined in Ref. [11]). As a result, it is shown that scale-free networks are more robust against random node failure than non-scale-free networks.

Both random node removals and random edge removals are similar, but it has been reported in Ref. [12] that, different from random node removal, random *edge* removals will not preserve scale-free property of the original network; i.e., the degree distribution of the network after edge removals does not follow a power-law. Specifically, in Ref. [12], the authors analytically derived a series of the number of nodes with an arbitrary degree (referred to as the degree sequence in Ref. [12]) in a graph after edges are randomly removed. Also, through numerical computations, they showed that the scale-free property does not remain in a network in which edges are randomly removed.

Robustness of communication networks such as router-level Internet topologies against random node/edge removal have been studied in Refs. [3], [4], [10]. In Ref. [10], the authors focused on the Internet topology at router-level and investigated the robustness of the Internet against random node/edge removal through experiments. Consequently, the authors showed that the Internet is robust against random node/link failures, on the other hand, the Internet is not robust against focused node failures (i.e., adversary attacks; that is, adversaries target and attack a few nodes playing an import role in maintaining the connectivity of a network).

Based on the above observations, several approaches have been taken to design (or re-design) of a network topology to improve its robustness against random failures [13], [14], [15]. In particular, several studies have been devoted for improving the robustness of scale-free network against adversary attacks by optimizing the topological structure of a network. The authors of Ref. [14] revealed that an optimal network which is robust against both of random node failures and attacks is comprised of one node with a high degree and other nodes with the same degree (almost the average degree of the network).

### 3. Method

Using synthetic networks, we generated scale-free and non-scale-free networks, and we compared the robustness of scale-free networks and non-scale-free networks against random node removal when changing the node removal ratio (i.e., the ratio of the number of removed nodes to the initial network size).

To generate scale-free networks, we used the BA (Barabási Albert) model [16] and the randomized BA model (an extension of the BA model).

Since the BA model generates a network by repeatedly adding vertices with a fixed number  $m$  of edges, it can only generate networks with specific average degrees. The randomized BA model relaxes this limitation; i.e., it can generate a scale-free network with an arbitrary average degree.

The difference between the BA model and the randomized BA model is in their preferential attachment stages. At the  $i$ -th cycle, the randomized BA model adds a node with a random number  $X_i$  of edges whereas the BA model adds a node with a fixed number  $m$  of edges. More specifically, in the randomized BA model, the number  $X_i$  of edges added at the  $i$ -th cycle is determined by the Bernoulli process with the probability of  $1/m$ .

Also, to generate scale-free networks with a cluster structure which is widely observed in many social and biological networks, we used the Li-Maini model [7]. The Li-Maini model is a network generation models for creating networks with cluster (i.e., community) structure, and an evolving network model based on inner-community and inter-community preferential attachment.

We used the ER (Erdős-Rényi) model [17] to generate non-scale-free networks. We also used another network generation model (i.e., a modified version of the ER model). As we will show in Section 4, the robustness of a network is significantly influenced by nodes with small degrees. We therefore introduce a network generation model called DB (Degree-Bounded) model. The DB model generates a network with  $N$  nodes and the average degree of  $\bar{k}$ . Hereafter, a network generated with the DB model is referred to as a *degree-bounded random network*. The DB model generates a degree-bounded random network as follows; (1)  $N$  nodes are initiated; and (2) for every node,  $\bar{k}/2$  links are added between the node and another randomly-chosen node.

Utilizing these five network generation models, we randomly generated multiple networks for given  $N$  and  $\bar{k}$ . Some example scale-free and non-scale-free networks generated with those network generation models are shown in **Table 1**.

We denote the node removal ratio by  $p$ . By removing randomly selected nodes from a generated network, we obtained a degenerated network. The original network and the degenerated network with the node removal ratio  $p$  are denoted by  $G$  and  $G(p)$ , respectively.

To investigate the robustness of scale-free and non-scale-free networks against random node removal, we obtained the largest component size in network  $G(p)$ . The largest component size is the number of nodes contained in the largest cluster [1].

To compare the largest component sizes in networks with the same size and the density, we conducted simulations by fixing the

number of nodes  $N$  and the average degree  $\bar{k}$ . Thus, the average number of links (i.e.,  $|E| = \bar{k}N/2$ ) is identical in all networks. In our simulations, we varied the number of nodes from 1,000 to 10,000 and the average degree from 4 to 6, respectively. Note that the number of nodes of ISP networks ranges between 10 and 10,000, and the average degree of ISP networks also ranges between 3.2 and 5.7 [3], [18].

We generated 100 network instances with each network generation model for given conditions (i.e., the number of nodes and the average degree). For each instance, we performed simulations to measure the largest component size while changing the node removal ratio  $p$ . From those 100 simulations, we calculated the mean and the 95% confidence interval of the largest component size for the node removal ratio  $p$ .

### 4. Results and Discussion

**Figure 2** shows the relation between the node removal ratio and the largest component size in five types of networks with  $N = 10,000$  and  $\bar{k} = 4$  (i.e.,  $|E| = 20,000$ ). **Figure 2** (b) shows the *normalized* largest component size. The normalized largest component size is defined as the ratio of the largest component size to the network size (i.e., the number of remaining nodes excluding removed nodes).

From this figure, it is found that when the node removal ratio is small, the largest component size in non-scale-free networks is larger than that of scale-free networks. In particular, it is also found that the degree-bounded random network shows the best robustness among networks generated with other network generation models. However, the normalized largest component size in scale-free networks is larger than that of non-scale-free networks when the node removal ratio is very high (i.e.,  $p \geq 0.7$ ), which coincides with the observation reported in Refs. [1], [2]. Furthermore, those results indicate that the difference in scale-free and non-scale-free networks highly affects the number of nodes in the main cluster. Specifically, under the small node removal ratio (i.e.,  $p = 0.1$ ), the number of nodes contained in the main cluster is varied around 8,500–9,000 due to the difference in network topologies.

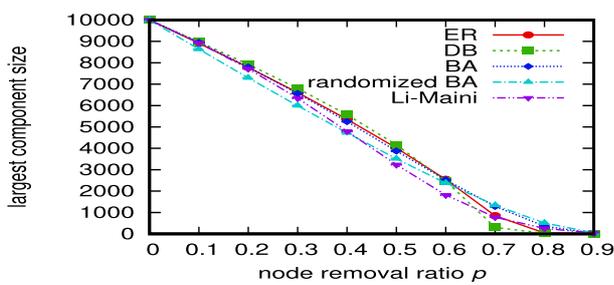
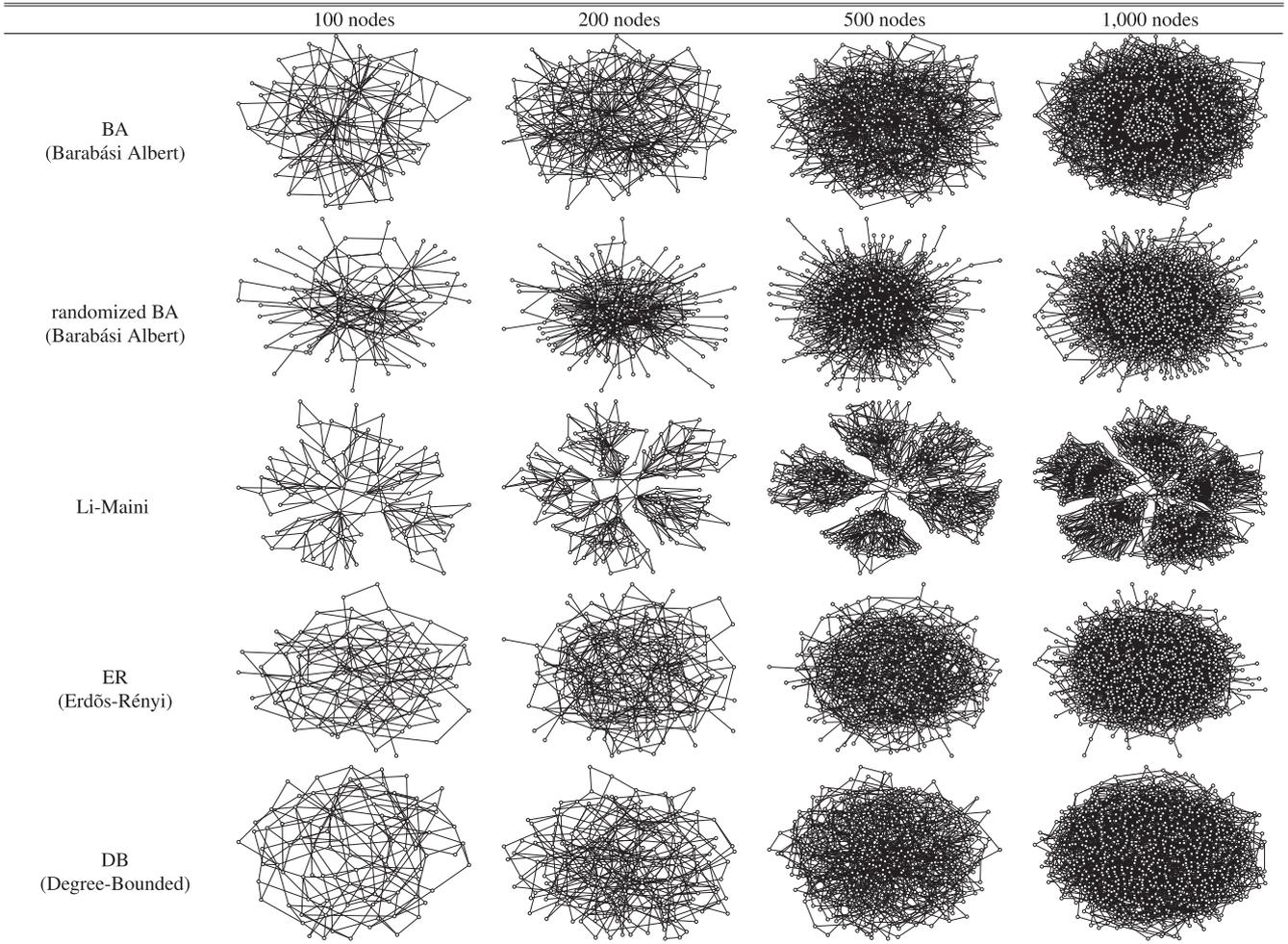
Next, we investigate whether our observations hold (or are still valid) in smaller networks. **Figure 3** shows the relation between the node removal ratio and the largest component size in five types of networks with  $N = 1,000$  and  $\bar{k} = 4$ .

Comparing Figs. 2 through 3 indicates that these results show a similar tendency. Namely, the robustness of scale-free networks and non-scale-free networks is not significantly affected by the network size. Namely our observations regarding Fig. 2 ( $N = 10,000$ ) are still valid in smaller networks.

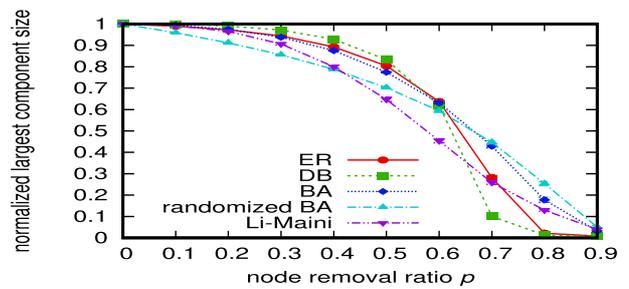
Also, we investigate how the robustness of scale-free networks and non-scale-free networks are affected by the network density (i.e., the number of links). **Figure 4** shows the relation between the node removal ratio and the largest component size in denser networks of the same size with  $N = 10,000$  and  $\bar{k} = 6$  (i.e.,  $|E| = 30,000$ ).

From this figure, by comparing results with the DB model and those with the BA model when  $p \leq 0.5$ , it can be found that even

**Table 1** Example networks with the average degree of 4 created with different network generation models.

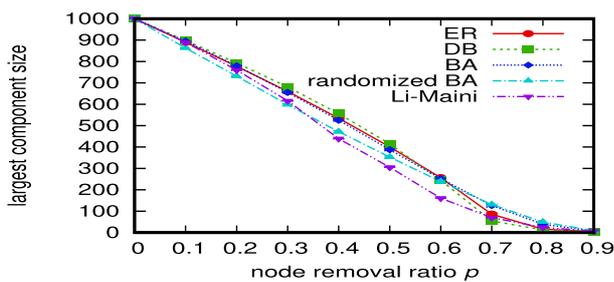


(a) largest component size

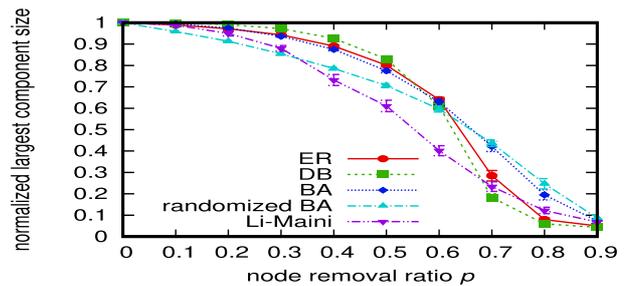


(b) normalized largest component size

**Fig. 2** Relation between node removal ratio  $\rho$  and the largest component size for  $N = 10,000$  and  $\bar{k} = 4$ .



(a) largest component size



(b) normalized largest component size

**Fig. 3** Relation between the removal ratio  $\rho$  and the largest component size for  $N = 1,000$  and  $\bar{k} = 4$ .

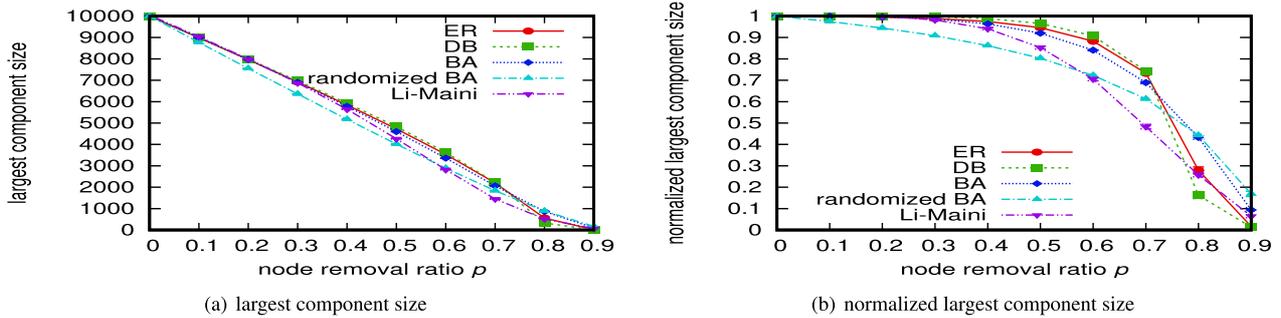


Fig. 4 Relation between node removal ratio  $p$  and the largest component size for  $N = 10,000$  and  $\bar{k} = 6$ .

though the network density is high, the largest component size in non-scale-free networks is at most 5% larger than that of scale-free networks. Hence, the difference in the normalized largest component sizes in random and scale-free networks is around 0.05. Even though the network density is high (i.e., the number of links increases), superiority of non-scale-free networks against scale-free networks is not reversed. Namely, in the case of denser networks, non-scale-free networks are more robust than scale-free networks in terms of the largest component size.

The validity of our observation — invariant superiority of non-scale-free networks to scale-free networks in terms of the robustness against random node removal — is confirmed by Figs. 2 through 4, which present how normalized largest component sizes are affected by the network size  $N$  and the average degree  $\bar{k}$ .

We then try to examine *why* non-scale-free networks are more robust than scale-free networks under modest node removal ratios. Understanding the superiority of non-scale-free networks to scale-free networks is helpful for improving the robustness of a network.

One possible explanation on the robustness of non-scale-free networks is their homogeneity or namely *unskewed* degree distributions. Except under extremely high node removal ratios, a network with homogeneous nodes can be less prone to node failures than a network with heterogeneous nodes.

The heterogeneity of five types of networks with  $N = 10,000$  and  $\bar{k} = 4$  is plotted in Fig. 5. The heterogeneity of a network is measured by three metrics: the standard deviation of node degrees (labeled as ‘stddev’), the skewness of the degree distribution (labeled as ‘skewness’), and the kurtosis of the degree distribution (labeled as ‘kurtosis’).

Comparing Figs. 2 and 5 clearly indicates a negative correlation between the network heterogeneity and the robustness against random node removal. Also, from Fig. 5, it can be found that the standard deviation, the skewness, and the kurtosis in non-scale-free networks are smaller than those in scale-free networks.

Among the three heterogeneity metrics (i.e., standard deviation, skewness, and kurtosis), the kurtosis of the degree distribution shows the strongest correlation with the robustness in terms of the largest component size. Namely, this result indicates that a network in which the variation in degrees of all nodes is small like non-scale-free network becomes robust against random node removal.

The advantage of a low kurtosis explains why non-scale-free networks generated with our DB (Degree-Bounded) model

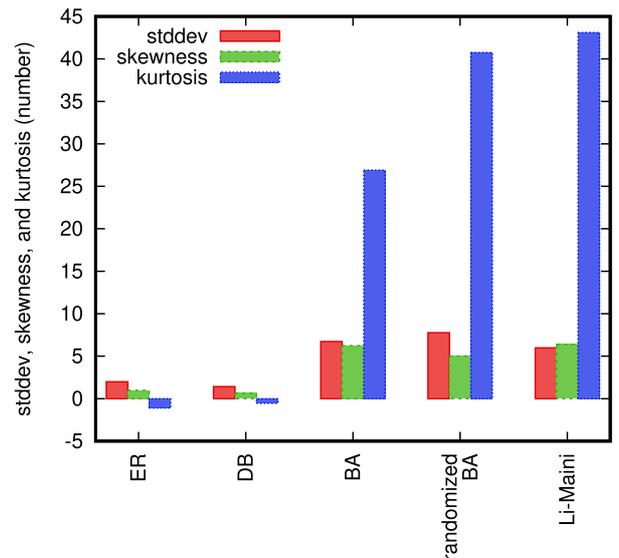


Fig. 5 Heterogeneity of a network with  $N = 10,000$  and  $\bar{k} = 4$ .

achieve the best robustness among five network generation models.

From these observations, we conclude that, contrary to common understanding, (1) non-scale-free networks are more robust than scale-free networks when the node removal ratio is not extremely high, and (2) robustness of non-scale-free networks can be further improved by bounding the minimum node degree.

## 5. Case of Random Link Removal

This section investigates the robustness of complex network against random *link* removal. In particular, we try to answer the following question: are our observations on the robustness of complex networks against random *node* removal still valid against random *link* removal?

Our methodology for investigating the robustness of complex networks against random link removal is equivalent to that against random node removal except that a fraction  $p$  of *links* are randomly removed from a generated network. We calculated the mean and the 95% confidence interval of the largest component size for the link removal ratio  $p$  while changing the link removal ratio  $p$  for 100 network instances with each network generation model.

Our simulation results with random link removal are shown in Figs. 6 through 8. Figures 6 (a), 7 (a), and 8 (a) show the relation between the link removal ratio and the largest component size in

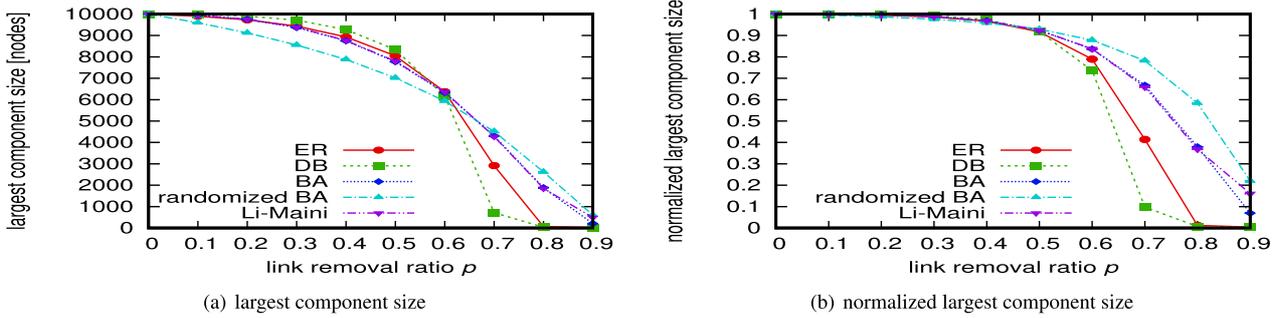


Fig. 6 Relation between link removal ratio  $p$  and the largest component size for  $N = 10,000$  and  $\bar{k} = 4$ .

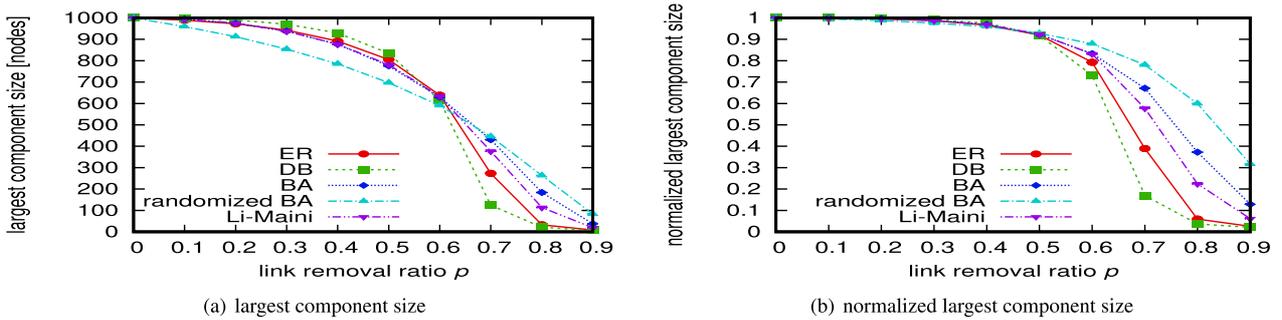


Fig. 7 Relation between link removal ratio  $p$  and the largest component size for  $N = 1,000$  and  $\bar{k} = 4$ .

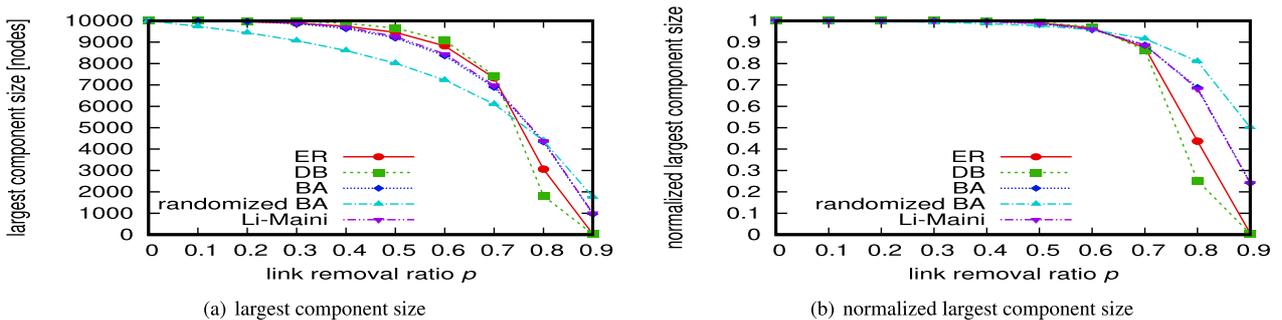


Fig. 8 Relation between link removal ratio  $p$  and the largest component size for  $N = 10,000$  and  $\bar{k} = 6$ .

five types of networks. Also, Figs. 6(b), 7(b), and 8(b) show the normalized largest component size. Hence, different from the case of the random node removal, the normalized largest component size is defined as the ratio of the largest component size to the number of remaining nodes excluding nodes whose degree is zero.

Comparison of the random link removal case (Figs. 6 through 8) with the random node removal case (Figs. 2 through 4) reveals that our observations discussed in Section 4 are still valid for random link removal; i.e., under modest link removal ratios, non-scale-free networks are more robust than scale-free networks and non-scale-free networks generated with the degree-bounded model shows the best robustness among others.

Notable differences in node removals and link removals are that the largest component size with random link removal decreases quadratically whereas that with random node removal decreases almost linearly, and that the *normalized* largest component size with random link removal are almost identical under modest link removal ratios. The reason for the former is that a network is more robust against link removal than node removal since the

network is not isolated as long as any of links connecting sub-networks remains.

## 6. Conclusion

In this paper, we compare the robustness of scale-free and non-scale-free networks against random node and link removal through simulations. Specifically, we generate multiple scale-free and non-scale-free networks using five network generation models, and compare the largest component sizes after random node removals. Our findings include that, when the node removal ratio is not extremely high, non-scale-free networks are more robust than scale-free networks. In particular, the degree-bounded random network with bounding of the minimum node degree shows the best robustness against random node removal among five types of networks. Furthermore, we have investigated the robustness of complex network against random link removal and revealed observations in the our findings for random node removal are still valid for random link removal. Our findings and common understanding on scale-free networks (e.g., Ref. [1]) are not contradictory. Namely, our findings indicate that the robustness of a network is quite dependent on the *degree* of node/link

failures. Under massive failures (e.g., 90% of node removal), scale-free network are more robust than non-scale-free networks. However, on the contrary, under moderate failures (e.g., 5% of node removal), non-scale-free networks exhibit more robustness.

As future works, we are planning to investigate the robustness of scale-free and non-scale-free networks under adversary attacks as well as non-adversary attacks.

**Acknowledgments** This work was partly supported by JSPS KAKENHI Grant Number 16H02815 and 18J10278.

## References

- [1] Albert, R., Jeong, H. and Barabási, A.-L.: Error and attack tolerance of complex networks, *Nature*, Vol.406, pp.378–382 (2000).
- [2] Albert, R. and Barabási, A.-L.: Statistical mechanics of complex networks, *Reviews of Modern Physics*, Vol.74, No.1, pp.47–97 (2002).
- [3] Alderson, D., Li, L., Willinger, W. and Doyle, J.C.: Understanding Internet topology: Principles, models, and validation, *IEEE/ACM Trans. Networking*, Vol.13, No.6, pp.1205–1218 (2005).
- [4] Doyle, J.C. et al.: The “robust yet fragile” nature of the Internet, *PNAS*, Vol.102, No.41, pp.14497–14502 (2005).
- [5] Paxson, V.: End-to-End Routing Behavior in the Internet, *Proc. ACM SIGCOMM '96*, pp.25–38 (1996).
- [6] Dahlin, M., Chandra, B.B.V., Gao, L. and Nayate, A.: End-To-End WAN Service Availability, *IEEE/ACM Trans. Networking*, Vol.11, No.2, pp.300–313 (2003).
- [7] Li, C. and Maini, P.K.: An evolving network model with community structure, *Journal of Physics A: Mathematical and General*, Vol.38, No.45, pp.9741–9749 (2005).
- [8] Crucitti, P., Latora, V., Marchiori, M. and Rapisarda, A.: Error and attack tolerance of complex networks, *Physica A: Statistical Mechanics and its Applications*, Vol.340, pp.388–394 (2002).
- [9] Magnien, C., Latapy, M. and Guillaume, J.-L.: Impact of Random Failures and Attacks on Poisson and Power-Law Random Networks, *ACM Computing Surveys*, Vol.43, No.3, pp.13:1–13:31 (2011).
- [10] Palmer, C.R., Siganos, G., Faloutsos, M., Faloutsos, C. and Gibbons, P.B.: The Connectivity and Fault-Tolerance of the Internet Topology, *Proc. Workshop on Network-Related Data Management (NRDM 2001)*, pp.1–9 (2001).
- [11] Latora, V. and Marchiori, M.: Efficient Behavior of Small-World Networks, *Physical Review Letters*, Vol.87, No.19, pp.198701-1–198701-4 (2001).
- [12] Martin, S., Carr, R.D. and Faulon, J.-L.: Random removal of edges from scale free graphs, *Physica A: Statistical Mechanics and its Applications*, Vol.371, pp.870–876 (2006).
- [13] Shargel, B., Sayama, H., Epstein, I.R. and Bar-Yam, Y.: Optimization of Robustness and Connectivity in Complex Networks, *Physical Review Letters*, Vol.90, No.6, pp.068701-1–068701-4 (2003).
- [14] Paul, G., Tanizawa, T., Havlin, S. and Stanley, H.E.: Optimization of the robustness of complex networks, *The European Physical Journal B*, Vol.38, pp.187–191 (2004).
- [15] Tanizawa, T., Paul, G., Havlin, S. and Stanley, H.E.: Optimization of the robustness of multimodal networks, *Physical Review E*, Vol.74, pp.016125-1–016125-14 (2006).
- [16] Barabási, A.-L. and Albert, R.: Emergence of Scaling in Random Networks, *Science*, Vol.286, No.5439, pp.509–512 (1999).
- [17] Erdős, P. and Rényi, A.: On random graphs I, *Mathematica*, Vol.6, No.26, pp.290–297 (1959).
- [18] Spring, N., Mahajan, R., Wetherall, D. and Anderson, T.: Measuring ISP Topologies with Rocketfuel, *IEEE/ACM Trans. Networking*, Vol.12, pp.2–16 (2004).



**Yuichi Yasuda** received his B.E. degree in the Informatics from Kwansai Gakuin University, Japan, in 2018. He is currently a graduate student at Department of Informatics, Graduate School of Science and Technology, Kwansai Gakuin University, Japan. His research work is in the area of design and evaluation of Information-

Centric Networking. He is a student member of IEEE and IEICE.



**Ryo Nakamura** received his M.E. degree in the Informatics from Kwansai Gakuin University, Japan, in 2017. He is currently a doctoral student at Department of Informatics, Graduate School of Science and Technology, Kwansai Gakuin University, Japan. His research work is in the area of performance analysis and evaluation of Information-Centric Networking. He is a student member of IEEE, IEICE, and IPSJ.

tion of Information-Centric Networking. He is a student member of IEEE, IEICE, and IPSJ.



**Hiroyuki Ohsaki** received his M.E. degree in the Information and Computer Sciences from Osaka University, Osaka, Japan, in 1995. He also received his Ph. D. degree from Osaka University, Osaka, Japan, in 1997. He is currently a professor at Department of Informatics, School of Science and Technology, Kwansai Gakuin

University, Japan. His research work is in the area of design, modeling, and control of large-scale communication networks. He is a member of IEEE and IEICE.



**Kazuyuki Yamashita** is currently an undergraduate student at Department of Informatics, School of Science and Technology, Kwansai Gakuin University, Japan. His research work is in the area of evaluation of complex networks in terms of robustness. He is a student member of IEEE, IEICE and IPSJ.