Robustness of Networks with Skewed Degree Distributions under Strategic Node Protection

Yui Kazawa
*Graduate School of Systems
and Information Engineering,
University of Tsukuba
1-1-1 Tennodai, Tsukuba,
Ibaraki 305-8573, Japan
Email: kzw-y@mibel.cs.tsukuba.ac.jp*

Sho Tsugawa
*Graduate School of Systems
and Information Engineering,
University of Tsukuba
1-1-1 Tennodai, Tsukuba,
Ibaraki 305-8573, Japan
Email: s-tugawa@cs.tsukuba.ac.jp*

***Abstract*—Previous studies on the robustness of networks against intentional attacks have suggested that protecting a small fraction of important nodes in a network significantly improves its robustness. In this paper, we analyze the robustness of networks under several strategic node protection schemes. Strategic node protection schemes select a small fraction of nodes as important nodes, using a network measure such as node centrality, and protect the important nodes to prevent them from being removed by intentional attacks. Our simulation results indicate that (1) strategic node protection significantly improves the robustness of networks with skewed degree distributions, (2) the efficiency of strategic node protection schemes is affected by the strength of community structure of the network being protected, and (3) strategic node protection based on betweenness centrality can effectively improve the robustness of networks regardless of the strength of community structure.**

***Keywords*-Robustness; Node protection; Network attack; Community structure**

I. Introduction

Many real-world networks have been reported to have skewed degree distributions, which means that most of the nodes have a few links and a few of the nodes have many links [1], [2]. For example, in the case of the World-Wide Web (WWW), while only a small minority of site contain millions of links, the majority of sites have only a few links [3]. In particular, networks whose degree distribution follows a power law are called scale-free networks.

Although scale-free networks are robust against random failure (i.e., uniformly random removal of nodes) they are vulnerable against intentional attack (i.e., intentional removal of important nodes) [4]–[8]. This property is caused by their skewed degree distributions. The connectivity of a scale-free network is maintained by a few important nodes. Consequently, when these nodes are removed, the network becomes fragmented into many smaller subnetworks. Hence, there is concern that real-world networks with skewed degree distributions may also be vulnerable to intentional attacks [8], [9].

In the real world, networks are required to be able to maintain overall connectivity even when some of the nodes fail [10]. The robustness of various networks against intentional attack has thus already been studied [4], [11], [12]. The robustness of a network indicates the ability to tolerate random failure and intentional attack. Studying the robustness of networks is expected to be useful for building disaster-resistant social infrastructure and resilient communication networks.

One topic that has been explored is how the robustness of a network is affected by hiding information about the topological properties of networks from an attacker who aims to fragment the network. Wu *et al.* suggested that hiding a small fraction of nodes in scale-free networks significantly decreases the efficiency of intentional attacks [11]. Shang suggested that adding error to the degree of each node improves the robustness of networks [12].

Previous studies have revealed that the robustness of networks is improved by adding random error to the degree of each node. This is because the added error prevents the attacker from being able to correctly recognize and remove nodes that are actually important. This brings us to the idea that if we protect important nodes, we can significantly improve the robustness of a network with a skewed degree distribution.

In this paper, we analyze the robustness of networks with skewed degree distributions by using strategic node protection. Strategic node protection schemes consist of the following steps: first, ranking nodes in descending order of importance; next, protecting a small fraction of the top-ranked nodes. We assume that protected nodes can be prevented from being removed by network attacks. For instance, in communication networks, strategic node protection corresponds to deploying security appliances to the important nodes (e.g., routers) for preventing malicious attacks to them.

In particular, we investigate the effectiveness of several network measures for strategic node protection schemes that improve the robustness of networks, and reveal which measure is the most effective. We address this question by performing intentional network attack simulations under strategic node protection. In the intentional network attack simulation, the nodes are sequentially attacked and removed from the network in descending order of importance, except for the small fraction of protected nodes, which are not removed even if they are attacked. Our main contributions are as follows.

(1) We show that strategic node protection significantly improves the robustness of networks with skewed degree distributions.
(2) We show that the efficiency of a strategic node protection scheme is related to the strength of the community structure [13] of the network being protected. For instance, we show that the protection scheme based on degree centrality, which is effective for protecting networks with a moderate community structure, is not as effective for networks with a strong community structure.
(3) We show that strategic node protection based on betweenness centrality [14] can effectively improve the robustness of networks regardless of the strength of community structure.

The remainder of this paper is organized as follows. Section 2 introduces work related to the robustness of networks. Section 3 explains the experimental methodology of the network attack simulation. Section 4 describes the results and discussion of the simulation. Section 5 contains our conclusions and a discussion of future work.

II. Related Work

Albert *et al.* showed that, while scale-free networks are tolerant to random failure, they are vulnerable to intentional attack [4]. If we randomly and uniformly select some fraction of nodes to remove from a scale-free network, the probability of selecting nodes with small degree is high since most of the nodes in a scale-free network have small degree. Since removal of these nodes has almost no influence on the connectivity of the whole network, the network is able to maintain function even when random failures occur. However, once a few important nodes that serve as hubs are removed, the network becomes fragmented and loses function [8].

If an attacker does not have any information about the network topology, the attacker can only remove nodes randomly. This is equivalent to random failure. For this reason, it is thought that incompleteness of the information possessed by the attacker can degrade attack effectiveness and improve the robustness of networks. There have been some studies investigating the effects of incomplete information of network topology on the robustness of networks [11], [12].

Wu *et al.* [11] and Shang [12] analyzed the robustness of scale-free networks against intentional attack with incomplete information. Wu *et al.* [11] performed network attack simulations in which a fraction of nodes are hidden from the attacker. The simulation results show that the robustness of scale-free networks is significantly improved when only a small fraction of randomly selected nodes are hidden. Shang [12] showed that attacks on nodes with high degree can be interfered with by adding noise to the degree information, and therefore, the effectiveness of the network attack can be degraded.

Our work is based on these prior works, and contributes to developing an efficient strategy for improving the robustness of networks. Whereas these previous studies have investigated the robustness of networks when important nodes are indirectly protected from attackers via noise in the network or hidden nodes, we select important nodes using several measures of node importance and investigate the robustness of networks when the important nodes are directly protected.

III. Methodology

In this study, we use the following graphs for the network attack simulation. Note that the degree distributions of the following graphs are skew, and some of them follow a power law.

(a) Power grid [15][1]
This graph represents the power transmission network of the Western United States. Transmission towers correspond to nodes and transmission lines correspond to links.
(b) Internet[2]
This graph represents a snapshot of the partial structure of the Internet, reconstructed from BGP tables posted by the University of Oregon Route Views Project[3]. Autonomous systems correspond to nodes and paths connecting autonomous systems corresponds to links.
(c) Political blogs [16][4]
This graph represents the blog and link relations maintained by American politicians. Blogs correspond to nodes and hyperlinks correspond to links. Although this is a directed graph, direction is ignored in this paper.
(d) BA graph [2]
This is a scale-free graph generated by the Barabási-Albert model.
(e) CE graph [17] ($\delta = 0.2$)
This graph has a moderate community structure generated by the model proposed by Kumpula. The parameter value $\delta = 0.2$ was used. Although the model is originally intended for generation of weighted undirected graphs, the weight of the generated graph is ignored and the graph is treated as an unweighted undirected graph in this paper.
(f) CE graph ($\delta = 2.0$)
This graph has a strong community structure. The parameter value $\delta = 2.0$ was used. It can be understood from the modularity values shown in Table I that the community structure is enhanced as the value of δ increases. This graph is also treated as an unweighted undirected graph in this paper.

[1] <http://www-personal.umich.edu/˜mejn/netdata/power.zip>
[2] <http://www-personal.umich.edu/˜mejn/netdata/as-22july06.zip>
[3] <http://routeviews.org>
[4] <http://www-personal.umich.edu/˜mejn/netdata/polblogs.zip>

Table I
SEVERAL CHARACTERISTICS OF THE GRAPHS

	number of nodes	number of links	clustering coefficient [15]	Q-values
Power grid	4,941	6,594	0.10	0.93
Internet	22,963	48,436	0.01	0.63
Political blogs	1,490	19,090	0.23	0.42
BA graph	2,500	4,997	0.47×10^{-2}	0.53
CE graph ($\delta = 0.2$)	2,500	8,060.80	0.30	0.69
CE graph ($\delta = 2.0$)	2,500	7,044.66	0.38	0.85

Table I shows several characteristics of the graphs used in this study. For BA graph, CE graph ($\delta = 0.2$) and CE graph ($\delta = 2.0$), the average values for 50 graphs are shown. Q-value, which is also known as modularity [13], evaluates the strength of community structure of a graph. For Politica blogs, the Q-value when communities are detected using the Girvan-Newman method [18] is shown. For other graphs, those when communities are detected by the Fast Newman method [19] are shown.

In these graphs, we first estimate the importance of each node, and rank all the nodes in descending order of estimated importance. The importance of the nodes are estimated by using betweenness centrality (betweenness) [14], degree centrality (degree) [14], closeness centrality (closeness) [14], PageRank [20], and collective influence (CI) [21]. Betweenness, degree, closeness and PageRank are widely used [22] [23]. CI is a recently proposed measure, and has been shown to be effective for network attacks [21]. The CI of node v_i is defined as,

$$CI_l(v_i) = (k_i - 1) \sum_{j \in \partial Ball(i,l)} (k_j - 1) \qquad (1)$$

where $\partial Ball(i,l)$ is the set of nodes that are reachable within just l hops from node v_i along the shortest path, and k_i is the degree of node v_i. In this paper, we used $l = 3$ as the parameter for CI.

We next determine the nodes to be protected using a ranking based on their importance. Specifically, the top $\lfloor Np \rfloor$ nodes in the ranking are protected, where p is a parameter that specifies the fraction of protected nodes, and N is the number of nodes.

We then remove nodes based on the importance ranking. Specifically, we remove the top $\lfloor Nr \rfloor$ nodes in the ranking except for the $\lfloor Np \rfloor$ protected nodes, where r is a parameter representing the fraction of removed nodes.

Finally, we calculate the relative size of giant component $g(r,p)$ for each graph and for each combination of the node importance measures used for node protection and node removal. The value of $g(r,p)$ is the number of nodes belonging to the giant component normalized by the number of remaining nodes after each removal. For BA graph, CE graph ($\delta = 0.2$), and CE graph ($\delta = 2.0$), the simulation results are the average of 50 independent simulation runs. For the fixed value of r, the higher the value of $g(r,p)$, the more robust the network.

IV. RESULTS AND DISCUSSION

We first investigate the effectiveness of strategic node protection for improving the robustness of each network. For comparison, we also performed simulations using the *random* protection scheme in which the protected nodes are selected randomly and also performed simulations without strategic node protection, which are denoted *nothing*.

Fig. 1 shows that the relative size of $g(r,p)$ as a function of r from 0.1 to 0.8 in steps of 0.1 when the nodes are removed based on degree. We set $p = 0.1$ for Power grid, CE graph ($\delta = 0.2$), and CE graph ($\delta = 2.0$), and $p = 0.01$ for all other graphs. The values of $g(r,p)$ are compared among each measure used for strategic node protection.

The results show that strategic node protection can significantly improve the value of $g(r,p)$ in each of the graphs (Fig. 1). Note that the value of $g(r,p)$ does not decrease monotonically as the value of r increases since $g(r,p)$ is normalized by the number of remaining nodes but the number of nodes in the original network. As discussed below, even when the nodes are removed based on measures other than degree, the robustness of the networks also improves similarly. Moreover, we find that the value of $g(r,p)$ does not improve in any of graphs when the protected nodes are selected randomly.

For instance, we focus attention on the value of $g(r,p)$ when 30% of nodes are removed in descending order of degree from the graphs. For Power grid, Internet and BA graph, the value of $g(r,p)$ is greatly improved by strategic node protection. In contrast, when the nodes are removed without any node protection, the value of $g(r,p)$ falls to almost zero. In the case of the Internet where the value of $g(r,p)$ is most improved by strategic node protection among the graphs, the value of $g(r,p)$ increases from approximately 0 to 0.7. For other graphs, we can also see the improvement in the value of $g(r,p)$. From Fig. 1, we conclude that the robustness of networks with skewed degree distributions is improved by strategic node protection.

Focusing on the differences between measures used for strategic node protection, we find that strategic node protection based on betweenness is the most effective for most of the graphs (Fig. 1). Strategic node protection based on betweenness significantly improves the values of $g(r,p)$ in particular for Power grid and CE graph ($\delta = 2.0$). In contrast, strategic node protection based on degree is less effective than that based on betweenness for Power grid and CE ($\delta = 2.0$), while the degree is comparable with

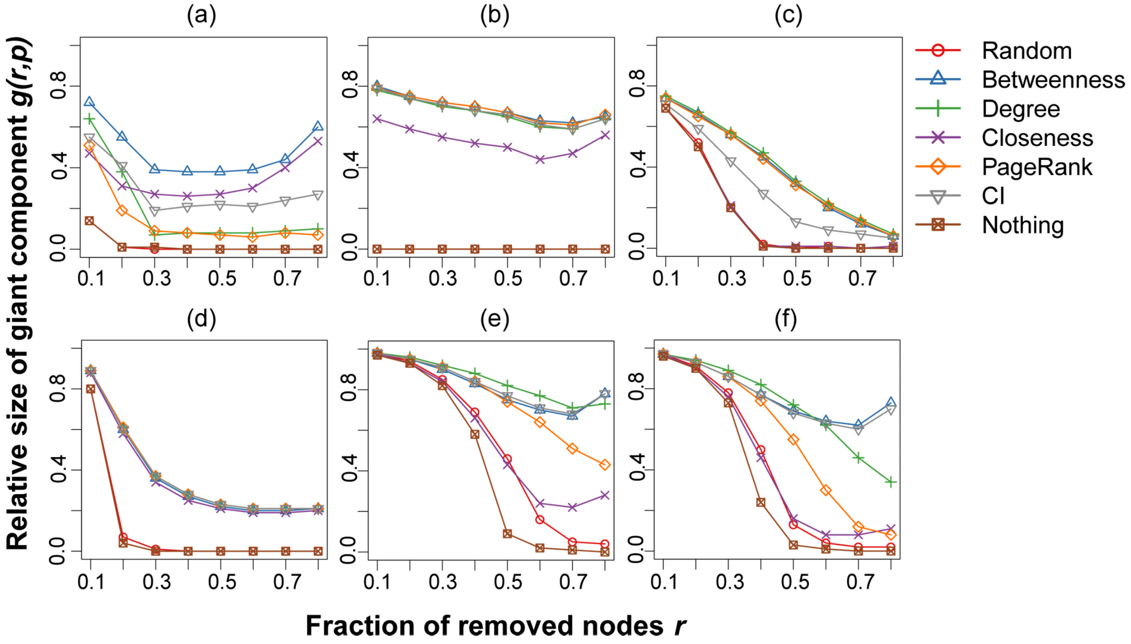

Figure 1. Fraction of removed nodes r versus relative size of giant component $g(r, p)$ in each graph when nodes are removed based on degree: (a) Power grid (b) Internet (c) Political blogs (d) BA graph (e) CE graph ($\delta = 0.2$) (f) CE graph ($\delta = 2.0$). $p = 0.1$ for (a), (e) and (f). $p = 0.01$ for (b), (c) and (d).

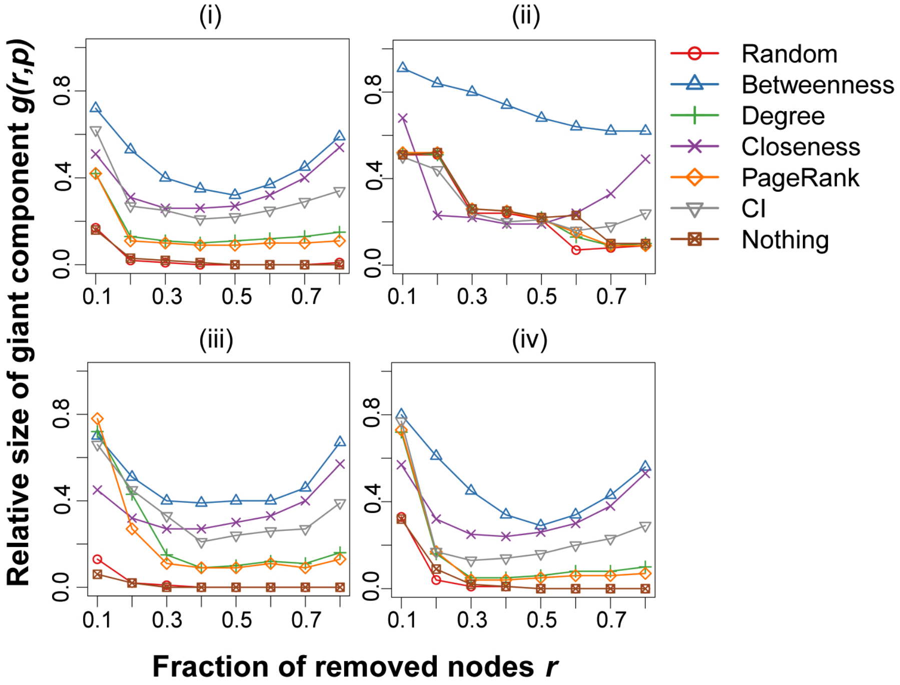

Figure 2. Fraction of removed nodes r versus relative size of giant component $g(r, p)$ in Power grid when the nodes are removed based on (i) betweenness, (ii) closeness, (iii) PageRank, and (iv) CI ($p = 0.01$)

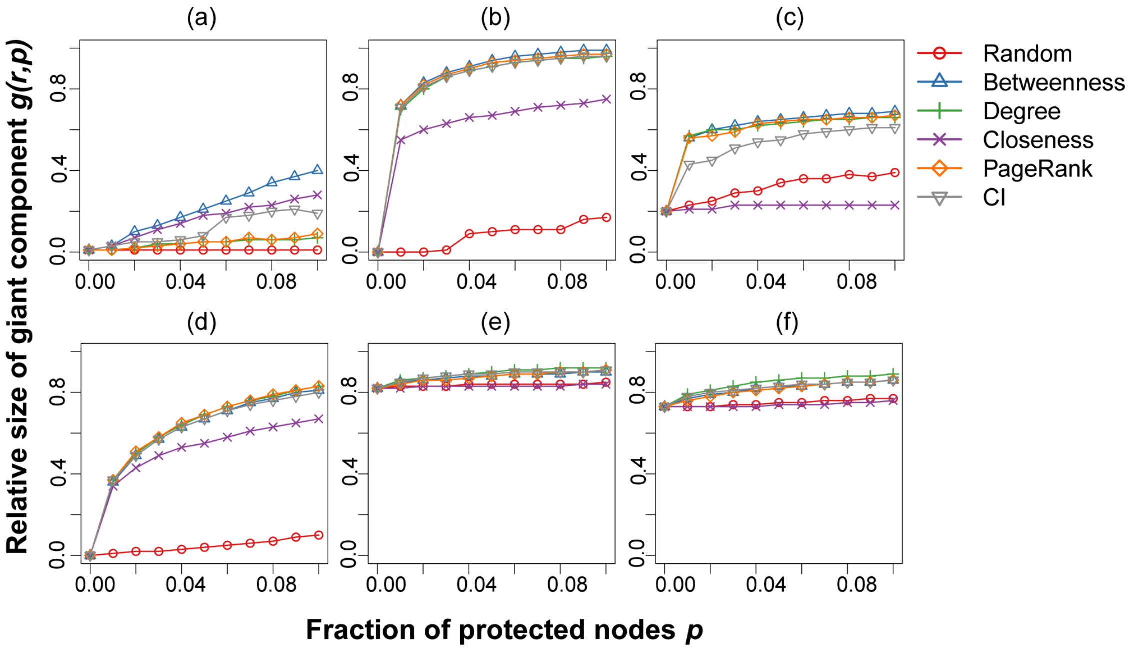

Figure 3. Fraction of protected nodes p versus relative size of giant component $g(r,p)$ in each graph when nodes are removed based on degree: (a) Power grid (b) Internet (c) Political blogs (d) BA graph (e) CE graph ($\delta = 0.2$) (f) CE graph ($\delta = 2.0$). ($r = 0.3$)

betweenness in other graphs.

The differences in the effectiveness among measures used for strategic node protection can be explained by the strength of the community structure of the graphs. From Fig. 1, we can see that the differences in the effectiveness among measures are particularly large in Power grid and CE graph ($\delta = 2.0$). Table I shows that these graphs have high Q-values, which indicates that these graphs have strong community structure. In a graph with strong community structure, both nodes connected to nodes inside the same community and node connected to nodes in different communities are important for maintaining the network connectivity. Measures such as betweenness that can successfully identify such important nodes are effective for improving the robustness of the networks whereas measures that fail to identify such important nodes may degrade their effectiveness for improving the robustness.

Next, we investigate the effectiveness of strategic node protection for improving the robustness of networks against intentional attacks based on measures other than degree. Fig. 2 shows $g(r,p)$ as a function of r from 0.1 to 0.8 in steps of 0.1 for the measures other than degree in Power grid and in the case of $p = 0.1$. The values of $g(r,p)$ are compared in terms of the measures used for strategic node protection.

From Fig. 2, we see that the value of $g(r,p)$ is remarkably increased by node protection based on betweenness in all cases of intentional attack based on any measure. Hence, we find that node protection based on betweenness is an effective way to improve the robustness of networks.

Finally, we investigate the relationship between fraction of protected nodes p and robustness of the network. Fig. 3 shows $g(r,p)$ as a function of p from 0 to 0.1 in steps of 0.01 when the nodes ranked among top 30% of degree are removed. The values of $g(r,p)$ are compared in terms of the measures used for strategic node protection.

From Fig. 3, we find that the values of $g(r,p)$ are increased by protecting only a small percent of nodes. In particular for Internet, Political blogs, and BA graph, the values of $g(r,p)$ are increased to between 0.4 to 0.7 by protecting only the top 1% of nodes based on betweenness. (Fig. 3(b), Fig. 3(c), Fig. 3(d)). These results suggest that only protecting a few percent of nodes using measures of node importance significantly improves the robustness of networks with skewed degree distributions. We should note that although we also found similar tendencies for intentional attacks using measures other than degree, the results are not shown in this paper due to space limitations.

V. Conclusion and Future Works

We investigated the robustness of real networks and networks generated by model by using strategic node protection, which uses measures identifying the most important nodes within a network and protects a small fraction of nodes in descending order of importance. Through extensive simulations, we showed the effectiveness of strategic node protection for improving the robustness of networks. Our main conclusions can be summarized as follows.

(1) Strategic node protection significantly improves the robustness of networks with skewed degree distributions.

(2) The effectiveness of strategic node protection is suggested to depend on the strength of community structure of the network being protected.

(3) Strategic node protection based on betweenness can effectively improve the robustness of networks regardless of the strength of community structure.

In future work, we are planning to investigate the relationship between the structural characteristics of a network and the effectiveness of strategic node protection in the network. Moreover, we also plan to analyze the minimum fraction of nodes that need to be protected in order to maintain a given level of network connectivity. Designing a strategic node protection scheme for dynamically changing networks such as peer-to-peer networks and mobile ad hoc networks is also important future work.

Acknowledgement

This work was supported by JSPS KAKENHI Grant Number 26280031.

References

[1] R. Albert and A.-L. Barabási, "Statistical mechanics of complex networks," *Reviews of Modern Physics*, vol. 74, no. 1, pp. 47–96, 2002.

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

[3] R. Albert, H. Jeong, and A.-L. Barabási, "Internet: Diameter of the World-Wide Web," *Nature*, vol. 401, no. 6749, pp. 130–131, 1999.

[4] R. Albert, H. Jeong, and A.-L. Barabási, "Error and attack tolerance of complex networks," *Nature*, vol. 406, no. 6794, pp. 378–382, 2000.

[5] D. S. Callaway, M. E. J. Newman, S. H. Strogatz, and D. J. Watts, "Network robustness and fragility: Percolation on random graphs," *Physical Review Letters*, vol. 85, no. 25, p. 5468, 2000.

[6] R. Cohen, K. Erez, D. Ben-Avraham, and S. Havlin, "Resilience of the Internet to random breakdowns," *Physical Review Letters*, vol. 85, no. 21, p. 4626, 2000.

[7] R. Cohen, K. Erez, D. Ben-Avraham, and S. Havlin, "Breakdown of the Internet under intentional attack," *Physical Review Letters*, vol. 86, no. 16, p. 3682, 2001.

[8] A.-L. Barabási and F. Jennifer, *Linked: The New Science of Networks Science of Networks*. Basic Books, 2002.

[9] E. Estrada, "Network robustness to targeted attacks. the interplay of expansibility and degree distribution," *The European Physical Journal B-Condensed Matter and Complex Systems*, vol. 52, no. 4, pp. 563–574, 2006.

[10] N. Katayama, T. Fujimura, H. Miwa, N. Kamiyama, H. Hasegawa, and H. Yoshino, "Design method of robust networks against performance deterioration during failures," in *Proceedings of the IEEE Global Communications Conference (GLOBECOM 2009)*. IEEE, 2009, pp. 1–6.

[11] J. Wu, H.-Z. Deng, Y.-J. Tan, and D.-Z. Zhu, "Vulnerability of complex networks under intentional attack with incomplete information," *Journal of Physics A: Mathematical and Theoretical*, vol. 40, no. 11, p. 2665, 2007.

[12] Y. Shang, "Robustness of scale-free networks under attack with tunable grey information," *Europhysics Letters*, vol. 95, no. 2, p. 28005, 2011.

[13] M. E. J. Newman, "Modularity and community structure in networks," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 103, no. 23, pp. 8577–8582, 2006.

[14] L. C. Freeman, "Centrality in social networks conceptual clarification," *Social Networks*, vol. 1, no. 3, pp. 215–239, 1979.

[15] D. J. Watts and S. H. Strogatz, "Collective dynamics of 'small-world' networks," *Nature*, vol. 393, no. 6684, pp. 440–442, 1998.

[16] L. A. Adamic and N. Glance, "The political blogosphere and the 2004 US election: divided they blog," in *Proceedings of the 3rd International Workshop on Link Discovery*. ACM, 2005, pp. 36–43.

[17] J. M. Kumpula, J.-P. Onnela, J. Saramäki, J. Kertesz, and K. Kaski, "Model of community emergence in weighted social networks," *Computer Physics Communications*, vol. 180, no. 4, pp. 517–522, 2009.

[18] M. Girvan and M. E. J. Newman, "Community structure in social and biological networks," *Proceedings of the National Academy of Sciences*, vol. 99, no. 12, pp. 7821–7826, 2002.

[19] M. E. J. Newman, "Fast algorithm for detecting community structure in networks," *Physical Review E*, vol. 69, no. 6, p. 066133, 2004.

[20] L. Page, S. Brin, R. Motwani, and T. Winograd, "The PageRank citation ranking: bringing order to the web." Stanford InfoLab, Tech. Rep., 1999.

[21] F. Morone and H. A. Makse, "Influence maximization in complex networks through optimal percolation," *Nature*, vol. 524, no. 7563, pp. 65–68, 2015.

[22] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D.-U. Hwang, "Complex networks: Structure and dynamics," *Physics Reports*, vol. 424, no. 4, pp. 175–308, 2006.

[23] M. E. J. Newman, "The structure and function of complex networks," *SIAM Review*, vol. 45, no. 2, pp. 167–256, 2003.