Robustness of Complex Communication Networks under Link Attacks

Yubo Wang
School of Electrical and Electronic Engineering
Nanyang Technological University
Singapore
65-6790 5363
ybwang@pmail.ntu.edu.sg

Gaoxi Xiao
School of Electrical and Electronic Engineering
Nanyang Technological University
Singapore
65-6790 4552
egxxiao@ntu.edu.sg

Tee Hiang Cheng
School of Electrical and Electronic Engineering
Nanyang Technological University
Singapore
65-6790 4534
ethcheng@ntu.edu.sg

Shi Xiao
School of Electrical and Electronic Engineering
Nanyang Technological University
Singapore
65-6790 5363
xiaoshi@ntu.edu.sg

Xiuju Fu
Institute of High Performance Computing
Singapore
65-6419 1330
fuxj@ihpc.a-star.edu.sg

ABSTRACT

Recent research results show that some most important complex communication systems, which usually can be modeled as *scale-free networks* with a power-law nodal degree distribution, may be fragile under intentional attacks that take down network hubs. We study the robustness of these networks under deliberate attacks which remove network links. Specifically, we evaluate the extreme case where an efficient graph-partitioning algorithm is applied, based on accurate network-topology information, to decide on the links to be removed. Simulation results show that even such type of calculated link-removal attack cannot easily split a complex communication network. Moreover, among the two split parts, the larger one generally remains as a scale-free network with a very small network diameter. We also consider the case where a specific set of nodes have to be split away from the major part of the network. Simulation results show that applying a graph-partitioning algorithm generally does *not* lead to a significantly more cost-effective solution than simply removing the given set of nodes together with the links connected to them.

Categories and Subject Descriptors

C.2.1.0 [**Computer Systems Organization**]: Computer-communication networks – *Network Architecture and Design.*

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

ICAIT'08, July 28-31, 2008, Da Meisha, Shenzhen, China.

Copyright 2008 ACM 1-58113-000-0/00/0004…$5.00.

General Terms

Management, Design, Economics, Reliability, Experimentation.

Keywords

Complex communication network, scale-free network, network robustness, link attack.

1. INTRODUCTION

Recent developments in computer and information technologies have enabled the in-depth studies on large-scale complex systems such as the Internet, World-Wide-Web (WWW), social connections, scientific collaborations, and airline transportation systems, etc [1, 6, 8, 20]. It has been found that when being formulated into complex network models represented by a set of nodes (entities) connected by links (relations), most of these systems share some stunning common features [1, 6, 8, 20]. Most noticeably, they usually form into scale-free networks of which the nodal degrees follow a power-law distribution. Specifically, the portion of nodes with a degree k satisfies [2, 5, 9, 23]

$$P(k) \propto k^{-r}. \tag{1}$$

In virtually all these real-world systems, the value of the scaling exponent r lies between 2 and 3 [5-6].

The wide existence of the scale-free networks has stimulated research on their robustness [3, 22]. It has been found that scale-free networks are highly robust under random node failures yet fragile under the so-called *intentional attack* which removes network nodes in decreasing order of their degrees [3, 24]. The tolerance of scale-free networks against various types of link-removal attacks has also been studied [12, 17-18]. It is found that random removals of links cannot easily break down a scale-free network [12, 17]. Deliberate link attacks which remove network links in a decreasing order of their degrees (defined as the product of the nodal degrees of its two end nodes) or betweenness (measured as the number of shortest paths passing through the link) are more effective [12]. But still, a large portion of the links has to be removed before a scale-free network can be fragmented.

We study the robustness of the scale-free networks under link removals. The main objective is to understand the network robustness against any type of link-removal attack. For that purpose, we evaluate the extreme case where link removals are pre-calculated based on accurate network-topology information aiming at splitting the network into two parts (rather than a large number of small pieces). Such type of calculated network splitting is not only among the worst attacks against a network (and therefore provides a benchmark for the less serious cases), it may also have its applications. For example, in disease control, the healthy part of a complex system may have to be split from the infected part at a minimum cost. To better resemble some real-life applications, we also study the cases where a specific set of nodes (e.g., infected nodes) have to be split away from the major part of the networks. We find that compared to the simple solution of removing the given set of nodes together with the links connected to them, applying a graph-partitioning algorithm based on accurate global network-topology information does not usually lead to a significantly more cost-effective solution.

The rest part of this paper is organized as follows: The graph-partitioning algorithm is described in Section 2. Simulation results and discussions on different cases of splitting a few scale-free networks into two parts with pre-defined sizes are presented in Section 3. Section 4 discusses the case where a specific set of network nodes have to be split away from the major part of the network. Section 5 concludes the paper.

2. NETWORK PARTITIONING ALGORITHM

The graph-partitioning algorithm adopted in this paper is borrowed from VLSI design [3], which was developed to allocate the components of a large-scale integral circuit into multiple limited-size chips with the minimum number of inter-chip links. In this paper, by using such an algorithm, we evaluate the number of links to be removed in order to split a network into two parts with *relative sizes* of a and ($1-a$), respectively. The relative size of a split part (termed as a *partition* hereafter [3, 10, 11, 15]) denotes the number of nodes in this part versus the overall number of network nodes. Hereafter the parameter a is called as the *splitting ratio*.

The minimum graph-partitioning problem has been proved to be NP-complete [11] and a few heuristic algorithms have been developed [3]. In our study, we adopt the move-based iterative partitioning algorithms since they are simple yet efficient, and independent of the characteristics of network topology. Specifically, we apply the Fiduccia-Mattheyses (FM) algorithm [10], an improved version of the earlier Kernighan-Lin (KL) method [15].

The FM algorithm can be briefly described as follows [10].

1. Parameter setting: Set up the *balancing criterion* to ensure that the size of each partition cannot be biased from the preset splitting ratio by more than a certain number of *m* nodes. In our experiences, as long as *m* is of a small positive value (e.g., no more than 5), the FM algorithm gets virtually the same results for different values of *m*.
2. Initial separation: Separate the network nodes into two partitions according to the preset splitting ratio. Denote the inter-partition links as the network *cuts*.
3. Execute a *pass* containing the following *move* operations.
    a. Let all the nodes be unlocked. Calculate the *gain* of each unlocked node as the number of cuts that would be decreased if the node is moved to its opposite partition. Note that the gain may be negative.
    b. Move the node with the largest gain (which may be negative) to its opposite partition as long as this move does not violate the balancing criterion. Denote this node as locked. Update the gain of each node.
    c. Repeat the above procedure until either all the nodes have been locked or the balancing criterion prevents further moves. The best split (with the fewest cuts) encountered during the pass is selected as the solution. Execute those moves leading to the best split.
4. Repeat Step 3 until the cut set cannot be further reduced. □

Remark: The FM algorithm allows the moves with negative gains in order to lower the chance of being trapped in local minima. The locking of the moved nodes in Step 3 is to prevent from having an endless loop. The final result may be slightly biased from the pre-set splitting ratio, with a maximum of *m* nodes. □

3. BI-SPLITTING NETWORKS

We evaluate the robustness of a scale-free network under link attacks by measuring the size of the minimum cut set corresponding to different values of the splitting ratio a , denoted as $E_{\min}(a)$. Moreover, for each partition, we measure its (i) *largest cluster size* (LCS) [3], defined as the number of nodes in the biggest connected component versus the number of nodes in the partition; and (ii) *cluster diameter* [3], the average length of the shortest paths between all the node pairs in the largest cluster. As we will see later, these two indexes help to reveal some most important characteristics of each partition.

We present the numerical simulation results on three different scale-free network models:

- The well-known Barabási-Albert (BA) model which generates a scale-free network by *growth* and *preferential attachment* [5]. Specifically, network nodes are sequentially added, each of which being connected to a fixed number of existing nodes. The probability that a newly-added node is connected to an existing node is proportional to the degree of the existing node. In our simulations, we test the BA model with 10,000 nodes and 20,000 links.
- A real-world Internet model on the Autonomous System (AS) level as measured by the Applied Network Research (NLANR) Project on January 2, 2000 [13], which contains 6,470 nodes inter-connected by 12,566 links. It has been considered as a scale-free network [6, 9].
- A real-world router-level Internet model as measured by the CAIDA project [14]. It contains 190,914 nodes and 607,610 links, forming into a scale-free network [6, 9].

To get the suboptimal solutions of the problem, we repeat the FM algorithm for one thousand times (each time with a different initial random splitting of the network) in the first two network models. The ultra-large size of the router-level model makes extensive repetitive simulations on it prohibitively time-consuming. Thus the calculations are repeated only for five times. We present the best results among these repetitions.

Fig. 1 presents the sizes of the minimum cut sets in different network topologies, i.e., $E_{\min}(a)$, where a varies from 0 to 0.5. For comparison purpose, the link-removal sizes of a random splitting, denoted as $E_{rand}(a)$, where

$$E_{rand}(a) = 2 \times a \times (1 - a) \qquad (2)$$

have also been plotted. We observe that $E_{\min}(a)$ increases slowly with an increasing value of a. To achieve a 50% splitting ratio, in both the AS-level and router-level models, slightly less than 6.5% of all the links have to be removed. Consider the large number of links in the networks, link removals at this low percentage may still request tremendous efforts. Network splitting becomes even more difficult in the BA model, where $E_{\min}(0.5) = 18.0\%$.

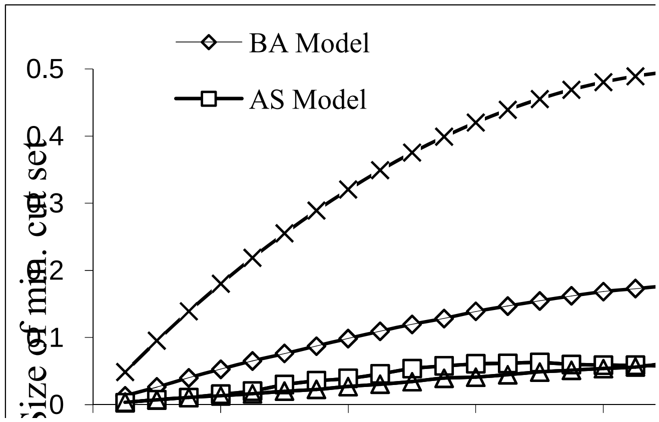
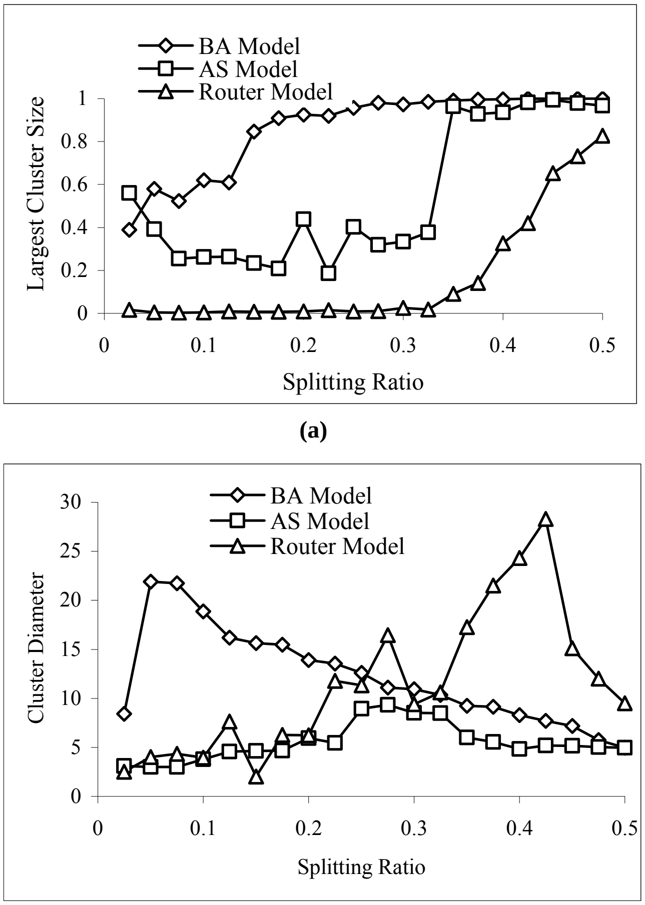


Figure 1. The minimum sizes of cut sets vs. the splitting ratios in different networks.

Denote the two partitions as larger partition (with a relative size of $(1-a)$) and smaller partition (with a relative size of a), respectively. In all our experiences, the larger partition almost always has an LCS value close to 100%. In other words, virtually all the nodes in the larger partition are connected into a single component. The smaller partition, on the other hand, has quite different properties. As shown in Fig. 2, the smaller partition has small LCS values under low splitting ratios, denoting that under such case the smaller partition contains a number of fragmented pieces. With an increasing value of a, these fragmented pieces begin to glue together, leading to a larger value of LCS. Meanwhile, the cluster diameter also increases, revealing that the component is only sparsely connected. Finally when the splitting ratio is large enough, the cluster diameter drops down where a small-world network emerges. We observe that the BA model again exhibits relatively stronger robustness than the other two models: it achieves a large value of LCS and a small value of cluster diameter at a lower value of splitting ratio.



(a)

(b)

Figure. 2. The behaviors of the smaller partition in different networks with different splitting ratios. (a) The LCS values. (b) The cluster diameter.

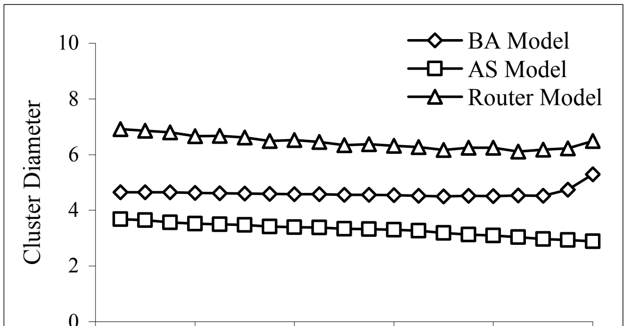
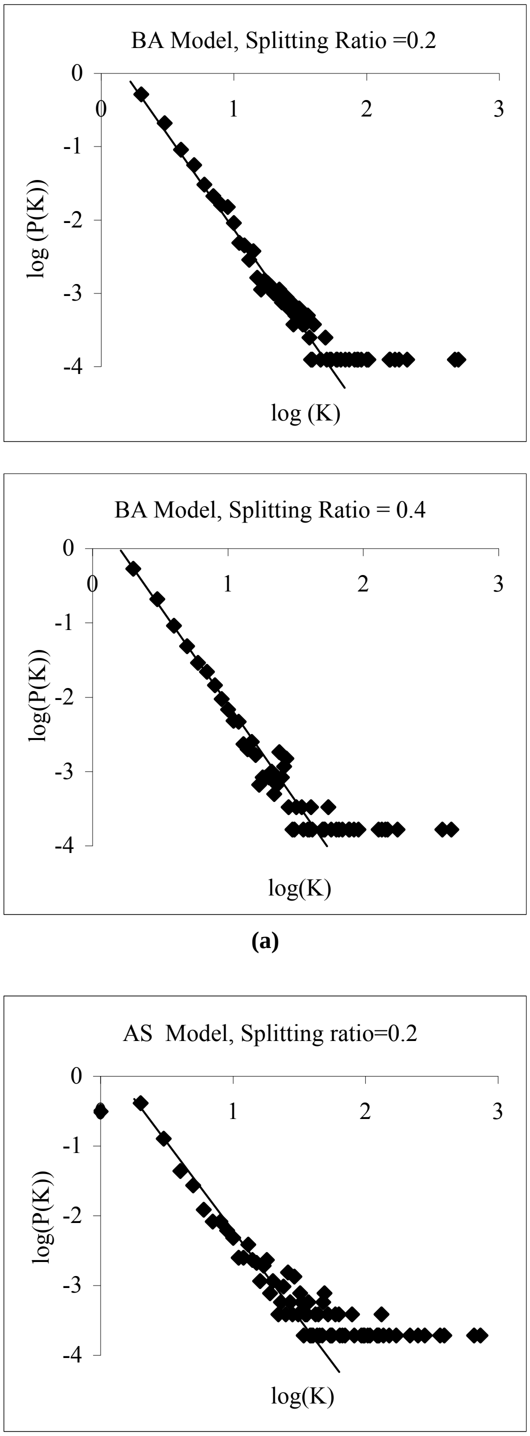
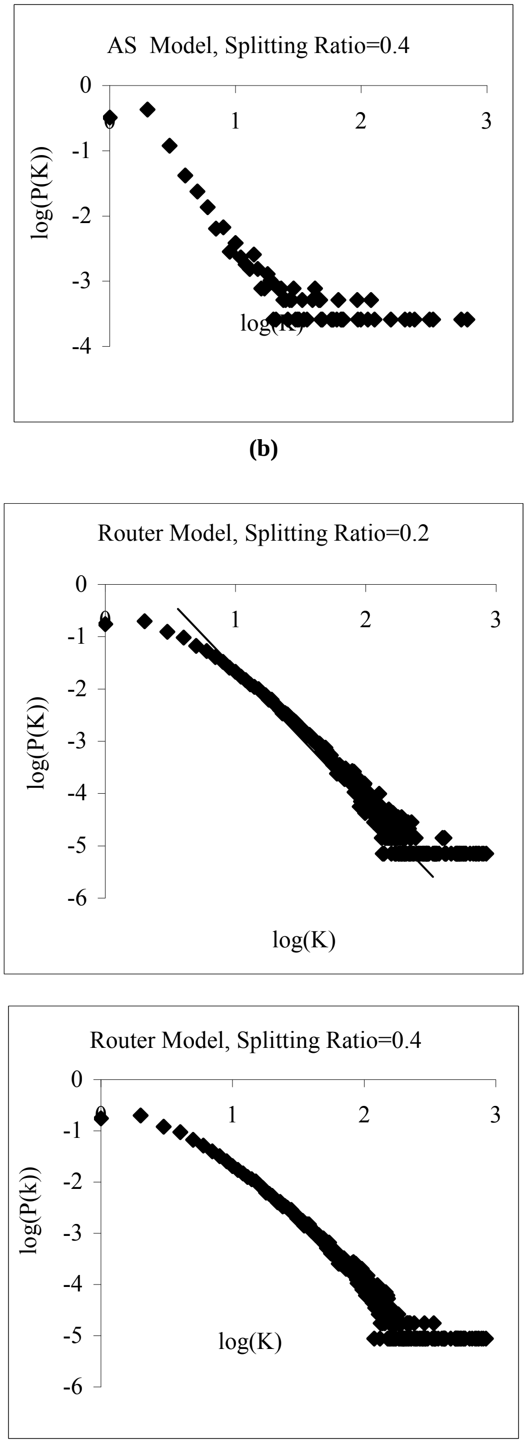


Figure. 3. The cluster diameter in the larger partition in different networks with different splitting ratios.

We then further investigate the properties of the larger partition. Fig. 3 shows that this partition always has a small diameter which remains largely unchanged under different splitting ratios. More significantly, as we can see in Fig. 4 and Table 1, the larger partition steadily remains as a scale-free network, with the exponent of its power-law degree distribution remains virtually un-changed under different splitting ratios. Note that we adopt the maximum likelihood estimator (MLE) in the estimations of exponents [7, 21].



(a)



(b)

(c)

Figure. 4. The log-log plots of nodal-degree distribution of the largest cluster in the larger partition in different networks with different splitting ratios. For each figure, the solid line has the slop of the estimated exponent: (a) BA model; (b) AS-level model; and (c) Router-level model.

Table 1. Estimated exponents of the largest cluster in the larger partition in different networks with different splitting ratios, assuming the largest cluster indeed follows a power-law distribution. The cases where the splitting ratio equals to "0" correspond to the original networks without splitting.

Splitting Ratio	Estimated Exponents		
	BA Model	AS-Level Model	Router-Level Model
0	2.85	2.22	2.42
0.2	2.82	2.26	2.43
0.4	2.83	2.28	2.42

4. SPLITTING PARTICULAR PARTS FROM NETWORKS

In this section, we study the splitting of networks where a specific set of nodes have to be separated away from the major part of the network. A simplest method is to remove these nodes together with all the links connected to them. It is of obvious research interest to see whether a significantly smaller cut set can be achieved at the cost of removing slightly more nodes. In some applications, e.g., disease/virus control, it may be a favorable option to sacrifice some nodes if the major part of healthy nodes can be protected at a much lower cost or a much faster speed.

We consider two different typical cases: (i) the nodes which have to be split away are randomly distributed; and (ii) these randomly selected nodes are connected into a single cluster. The latter case may find its application again in disease or virus control where infected nodes are usually connected [20]. The procedures for generating these two different cases are defined as follows:

- **Random Node Set Case**: Denote all the nodes as uncolored. In each iteration, randomly choose an uncolored node and change it to be colored. Repeat the procedure until the requested number of nodes has been colored. The colored nodes form into the node set that has to be split away from the other nodes.
- **Connected Node Set Case**: Denote all the nodes as uncolored. Randomly choose one node and denote it as colored. In each of the following iterations, randomly choose an uncolored node which is adjacent to at least one colored node and change it to be colored. Repeat the procedure until the requested number of nodes has been selected. The colored nodes form into the node set that has to be split away from the other nodes.

Note that in the connected node set case, the probability that an uncolored degree-k node is selected to be colored is approximately proportional to $kP(k)$ [1, 6], while in the random node set case this probability is proportional to $P(k)$. Comparing these two different cases, we see that high-degree nodes have higher probabilities to be colored in the connected node set case. This to some extent better resembles the case in real-life virus spreading where higher-degree nodes have a higher chance to be infected [6].

The FM algorithm is adopted with slight modifications. Specifically, we set the initial separation to let all the colored nodes be in the smaller partition and the rest be in the larger partition. The colored nodes are never allowed to be moved to the opposite partition. In other words, they are locked throughout the calculations. The balancing criteria are relaxed: move operations would not be blocked as long as the relative size of the smaller partition is not larger than 0.5. Similarly to that in our previous simulations, we repeat the calculations for at least one thousand times and then choose among them the best solution. Due to the time limit, simulations have been conducted only on the BA model and AS-level model.

Table 2. Results for splitting the BA model with a given set of colored nodes: (a) Random node set case. (b) Connected node set case.

Before FM Operations		After FM Operations	
Colored-Node Set	Link Cuts	Smaller Partition	Link Cuts
0.01	381	0.0195	376
0.05	1886	0.0828	1850
0.1	3547	0.2092	3365

(a)

Before FM Operations		After FM Operations	
Colored-Node Set	Link Cuts	Smaller Partition	Link Cuts
0.01	2515	0.1291	2307
0.05	4954	0.3265	3946
0.1	6437	0.4657	4056

(b)

Table 3. Results for splitting the AS-level model with a given set of colored nodes. (a) Random node set case. (b) Connected node set case.

Before FM Operations		After FM Operations	
Colored-Node Set	Link Cuts	Smaller Partition	Link Cuts
0.01	161	0.0122	146
0.05	1636	0.1409	1450
0.1	1867	0.1749	1674

(a)

Before FM Operations		After FM Operations	
Colored-Node Set	Link Cuts	Smaller Partition	Link Cuts
0.01	5110	0.4997	1287
0.05	6789	0.4998	2203
0.1	7637	0.5	2630

(b)

As shown in Tables 2 and 3, for the random node set case, it is difficult to significantly reduce the link cuts for separating the colored and uncolored nodes by sacrificing some uncolored nodes: while the relative size of the smaller partition is almost doubled (or in other words, the number of sacrificed uncolored nodes almost equal to the number of colored nodes), the size of the minimum cut set is only slightly reduced. In the connected node set case, the situation does not get any better: link cuts can only be significantly reduced at an unacceptable sacrifice of a large number of uncolored nodes. To summarize, we see that in both cases, there is not a much more cost-effective solution than splitting away the colored nodes by simply cutting those links connected to them.

5. CONCLUSION

In this paper, we studied the robustness of scale-free communications networks under link-removal attacks. By adopting the FM partitioning algorithm to calculate the minimum cut set, we have evidently shown that splitting a scale-free network into two partitions, even when complete network-topology information is available to the attacker, may remain as a serious challenge. More interestingly, after such type of calculated attack, the larger partition would steadily remain a scale-free network and hence may continue to function reasonably well. Furthermore, we found that it is surprisingly difficult to separate a specific set of nodes away from the major part of scale-free networks. Even with careful computations based on accurate network-topology information, a more cost-effective solution that cutting all the links connected to this specific set of nodes cannot be easily achieved and probably, not exist at all.

6. ACKNOWLEDGMENTS

This work was supported in part by Biomedical Research Council (BMRC) under Grant 06/1/21/19/457.

7. REFERENCES

[1] Albert, R., Barabási, A.-L. 2002. Statistical mechanics of complex networks. Rev. Mod. Phys., vol. 74, pp. 47-97.

[2] Albert, R., Jeong , H., Barabasi , A.-L. 1999. Diameter of the World Wide Web. Nature 401, pp. 130-131.

[3] Albert, R., Jeong, H., and Barabási, A.-L. 2000. Error and attack tolerance of complex networks. Nature 406, pp. 378–382.

[4] Alpert, C. J., Kahng, A. B. 1995. Recent directions in netlist partitioning: a Survey. Integration: the VLSI Journal, 19 (1-2), pp. 1-81.

[5] Barabási, A.-L., Albert, R. 1999. Emergence of scaling in random networks, Science, vol. 286, pp. 509-512.

[6] Bornholdt, S., Schuster, H. G. (Eds.) 2003. Handbook of graphs and networks: from the genome to the Internet. Wiley-VCH.

[7] Clauset, A., Shalizi, C. R., Newman, M. E. J. 2007. Power-law distributions in empirical data. Physics/0706.1062.

[8] Dorogovtsev, S. N., Mendes, J. F. F. 2002. Evolution of Networks. Adv. Phys. 51, 1079.

[9] Faloutsos, M., Faloutsos, P., Faloutsos, C. 1999. On power-law relationships of the internet topology. In Proc. ACM / SIGCOMM, Comput. Commun. Rev. 29,251-260.

[10] Fiduccia, C., Matheyses, R. 1982. A linear time heuristic for improving network partitions. Proc. ACM/ IEEE Design Automation Conference, pp. 175-181.

[11] Garey, M. R., Johnson, D. S. 1983. Computers and intractability: a guide to the theory of NP-Completeness. New York: W. H. Freeman. ISBN 0-7167-1045-5.

[12] Holme, P., Kim, B. J., Yoon, C. N., Han, S. K. 2002. Attack vulnerability of complex networks. Phys. Rev. E 65, 056109.

[13] http://moat.nlanr.net/Routing/rawdata

[14] http://www.caida.org/tools/measurement/skitter/router_topology/

[15] Kernighan, B. W., Lin, S. 1970. An efficient heuristic procedure for partitioning graphs. Bell Systems Tech. J., 49, pp. 291-308.

[16] Magoni, D. 2003. Tearing down the Internet, Selected Areas in Communications. IEEE Journal, vol. 21, no. 6, pp. 949-960.

[17] Martin, S., Carr, R. D., Faulon, J.-L. 2006. Random removal of edges from scale free graphs. Phys. A 371, pp. 870-876.

[18] Motter, A. E., Nishikawa, T., Lai, Y.-C. 2002. Range-based attack on links in scale-free networks: are long-range links responsible for the small-world phenomenon?. Phys. Rev. E 66, 065103.

[19] Murray, J. D. 1993. Mathematical biology. Springer Verlag. Berlin.

[20] Newman M. E. J. 2003. The structure and function of complex networks. SIAM Review, 45:167–256.

[21] Newman, M. E. J. 2005. Power laws, Pareto distributions and Zipf's law. Contemporary Physics 46: 323–351.

[22] Pastor-Satorras, R., Vespignani, A. 2002. Immunization of complex networks. Phys. Rev. E 65, 036104.

[23] Watts, D. J., Strogatz, S. H. 1998. Collective dynamics of 'small-world' networks. Nature, vol. 393: pp. 440-442.

[24] Xiao, S., Xiao, G., Cheng, T. H. 2008. Tolerance of intentional attacks in complex communications networks, IEEE Commun. Mag., vol. 46, no. 1: pp. 146-152.