Attack Vulnerability of Complex Communication Networks

Yongxiang Xia, *Member, IEEE*, and David J. Hill, *Fellow, IEEE*

***Abstract*—The Internet has been studied as a typical example of real-world complex networks. In this brief, we study the traffic performance of the Internet when it encounters a random or intentional attack. Different from previous approaches, the congestion control protocols are considered so that the bandwidth can be reallocated among flows. In this way, cascading breakdown is less likely to happen. The flow rates are adjusted when a node is attacked and out of function. Consequently, the traffic utility and the utilization ratio of bandwidth are affected. We compare the real Internet data with the classic random graph and scale-free network models. The simulated results also show that the "robust yet fragile" property previously observed in the study of cascading failures in the scale-free networks is still valid in this scenario.**

***Index Terms*—Bandwidth allocation, complex networks, Internet, vulnerability.**

I. Introduction

RESEARCH on complex networks has received considerable attention during the past decade [1]. Networks in a wide range of areas are studied and statistical properties, which show the characteristics of those networks, are found. Moreover, several network models are proposed [2], [3], which help us to have a further understanding of those statistical properties. Based on these developments, the research interest has been switched to the dynamic behaviors of networks in recent years [4]. Cascading failures [5]–[7], congestions in communication networks [8], [9], epidemic spreading [10], [11] and synchronization [12]–[16] are some examples of dynamic behaviors of complex networks.

Many researchers have studied random and intentional attacks to complex networks and their effects on the traffic performance in the networks (e.g., see [4] and its references). It is a natural problem which many real-world networks, such as power grids, the Internet, telephone networks and transportation networks, may face. When a node is attacked and then out of action, the flows which originally go through the node have to reconfigure their paths and go through other nodes. The redistribution of traffic may affect loads on other nodes and possibly start a sequence of overload failures. Further study indicates that a scale-free network is highly robust to random failures yet fragile under intentional attack, whereas a random graph is robust under both random and intentional attacks [5].

In these results, the flow rates are assumed to be fixed even after the reconfiguration of flow paths. In the real Internet, however, the case is quite different [17]–[19]. When a node is out of function and flows which used to go through it reconfigure their paths, the transmission control protocols (TCP) adjust the rate of each flow using congestion control algorithms, and make the aggregate flow rate under the bandwidth limitation. Thus, the cascading breakdown is less likely to happen.

In this brief, we will study the effect of random and intentional attacks on the traffic performance in the Internet. Because flow rates are adjusted by the congestion control protocols when a node is attacked, the traffic performance is then affected. We will define some indicators to measure the traffic performance and show how they are affected.

II. Optimal Bandwidth Allocation Algorithm

Consider a network with L links and S flows. Set $A_{li} = 1$ if flow i goes through link l and $A_{li} = 0$ otherwise. Then the matrix $A = (A_{li}, 1 \le l \le L, 1 \le i \le S)$ contains all the routing information. If the rate of flow i is x_i and the capacity of link l is C_l, then we have the following inequality:

$$Ax \le C \tag{1}$$

where $x = (x_i, 1 \le i \le S)$ and $C = (C_l, 1 \le l \le L)$ are the flow rate vector and capacity vector, respectively. The above inequality gives the constraint of telecommunication channels, i.e., the aggregate rate cannot exceed the channel bandwidth.

The aim of transmitting a flow of packets from their source to the destination is to get some benefit from the information transmission (e.g., downloading a file, reading news, or making an online booking, etc.). It is natural to set a utility function U_i for flow i, and assume that U_i is related to its rate x_i. Thus, we can denote it as $U_i(x_i)$. In order to get an optimal bandwidth allocation solution, the congestion control protocols try to solve the following optimization problem [18]:

$$\max_x \sum_i U_i(x_i), \text{ subject to } Ax \le C, \text{ over } x \ge 0. \tag{2}$$

We consider two kinds of attacks: random attack and intentional attack. In the random attack scenario, the attacked node is randomly chosen from the network. The intentional attack is assumed on the node with the largest degree. When a node is attacked, it is removed from the network. The flows which originate or end at this node are removed simultaneously, while

Manuscript received March 30, 2007; revised August 1, 2007. This work was supported by the Australian Research Council's Discovery Projects Scheme under Project FF0455875. This paper was recommended by Associate Editor G. M. Maggio.

The authors are with the Department of Information Engineering, Research School of Information Sciences and Engineering, the Australian National University, Canberra ACT 0200, Australia (e-mail: yongxiang.xia@anu.edu.au, David.Hill@anu.edu.au).

Digital Object Identifier 10.1109/TCSII.2007.908954

the flows which go through the node reconfigure their routes to find new shortest paths. All these removals and reconfigurations modify the corresponding entries in matrix A. Based on the modified routing information, congestion control protocols compute a new optimal bandwidth allocation.

Because of the large scale of the Internet, it is necessary to consider whether using the centralized or decentralized algorithms to solve the optimization problem (2). Although the centralized algorithms have a faster convergence, they require that a link knows the information of the whole network, which is not practical in the real Internet. So the decentralized algorithm is more applicable. In this brief we use a decentralized algorithm to compute the optimal results [17]. The algorithm is as follows.

• *Link l's Algorithm:* At time $t = 1, 2, \ldots$, link l:

1) gets rates of flows which go through link l;

2) computes its price

$$p_l(t+1) = [p_l(t) + \alpha(x^l(t) - C_l)]^+$$

where $x^l(t) = \sum_{s \in S(l)} x_s(t)$ is the aggregate flow rate on link l, $S(l)$ is the set of flows which go through link l, and $[z]^+ = \max\{z, 0\}$;

3) communicates the new price $p_l(t+1)$ to all flows which go through l.

• *Flow s's Algorithm:* At time $t = 1, 2, \ldots$, flow s:

1) receives prices of links which flow s goes through;

2) chooses a new transmission rate $x_s(t+1)$ based on the sum $p^s(t) = \sum_{l \in L(s)} p_l(t)$, where $L(s)$ is the set of links which flow s goes through;

3) communicates new rate $x_s(t+1)$ to links $l \in L(s)$.

The idea of the algorithm is straightforward. At each step, each link receives the rates of those flows which go through it. Then it adjusts its price based on the flow rates and communicates the new price to all flows through it. Similarly, each flow receives the prices of links on its route, computes its new rate based on those prices, and tells its new rate to links on its route. In this way, what a link or flow needs to know is only the local information related to it. It does not need to acquire the information of the whole network.

III. Network Models

We acquire the inter-connection information of the Internet at the autonomous systems (ASs) level from the online database [20]. Then we select the largest connected part, containing 1470 nodes and 3131 links, from the raw data, as a subnetwork of the Internet. We also consider a Barabási and Albert (BA) scale-free network model proposed by Barabási and Albert [21] and a random graph model proposed by Erdős and Rényi [22] for comparison. In order to compare three network structures fairly, the number of nodes and the number of links in both random graph and BA model are also set to be 1470 and 3131, respectively. Figs. 1 and 2 illustrate the degree distributions of

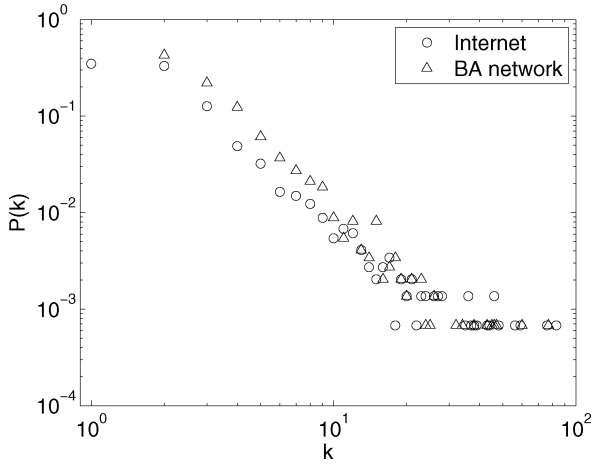

Fig. 1. Degree distribution of the subnetwork of the Internet and BA scale-free network.

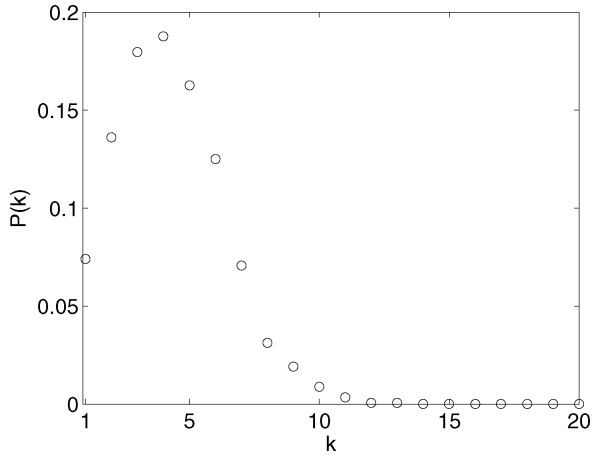

Fig. 2. Degree distribution of a random graph.

the three network models, respectively. They clearly show that both the Internet data and the BA model follow the power-law degree distribution, while the degree distribution of a random graph is Poisson.

In each network, we randomly generate $S = 4000$ flows. To do that, for each flow, the source and destination nodes are randomly chosen. Once a source-destination pair is determined, a shortest path between them is found. If there are more than one shortest path between the source and destination, then we randomly choose one.

IV. Simulations

A. Link Capacity

In the Internet, the link capacity means the bandwidth of the link. Intuitively, links connected to those popular nodes need larger capacities since more traffic loads go through them. Here we set

$$C_{ij} \propto B_i + B_j \tag{3}$$

where C_{ij} is the capacity of link between nodes i and j, "$\propto$" means proportional to, and B_i is the betweenness of node i,

which is defined as the number of shortest paths between any pair of nodes which go through i [23]. Comparing the definition of betweenness with the routing rule of traffic flows (i.e., to go through a shortest path from the source to the destination), we can conclude that the betweenness characterizes the average traffic load on a node [24]. Generally speaking, if a link is connected to a node with a larger betweenness, it will encounter more traffic loads. Therefore, it needs more capacity.

B. Performance Indicators

The traffic performance can be measured in two aspects: the utility and the utilization ratio. Since the task of optimal bandwidth allocation is adjusting the flow rates to find the maximum utility, we can use the total utility, $\sum_i U_i(x_i)$, as our first performance indicator. In the following simulations, we use

$$U_i(x_i) = d_i \log(x_i + 1) \tag{4}$$

as the utility function, where d_i is the path length of flow i. It is actually a simplified version of the utility function used in TCP Vegas [25].

The second performance indicator is the utilization ratio of bandwidth. When the total utility reaches the maximum, there is still some idle bandwidth. That is to say, a proportion of the bandwidth is wasted. From an economical point of view, we want this kind of waste to be as little as possible, or the utilization ratio of bandwidth to be as high as possible. The utilization ratio of bandwidth is defined as

$$u = \frac{\sum_l \sum_i A_{li} x_i}{\sum_l C_l} \tag{5}$$

where $\sum_l \sum_i A_{li} x_i$ is the total used bandwidth and $\sum_l C_l$ is the total provided bandwidth.

C. Simulation Results

Fig. 3 depicts the simulated results of total traffic utility in a subnetwork of the Internet. In the figure the total utility is a function of total capacity $\sum_l C_l$. When the total capacity increases, there are more bandwidth resources for traffic flows. As a result, the total utility increases. This figure also compares the performance of the original network with the cases under random and intentional attacks. The "random attack" curve overlaps with the original one, whereas the "intentional attack" curve is about 15% lower. These results show the important role of the hubs in the Internet traffic. In the network, a hub has so many connections that it is on the routes of lots of flows. Consequently, the traffic load on a hub is very heavy. If the hub is attacked, a large proportion of flows are affected, which results in the dramatic degradation of traffic performance. On the other hand, according to the power-law degree distribution, the number of hubs is quite small. In the random attack scenario, since the target is randomly selected, it is much more likely that the attacked node has only few connections. This node has less effect on the traffic performance of the whole network. That is the reason why the "random attack" curve overlaps with the original one.

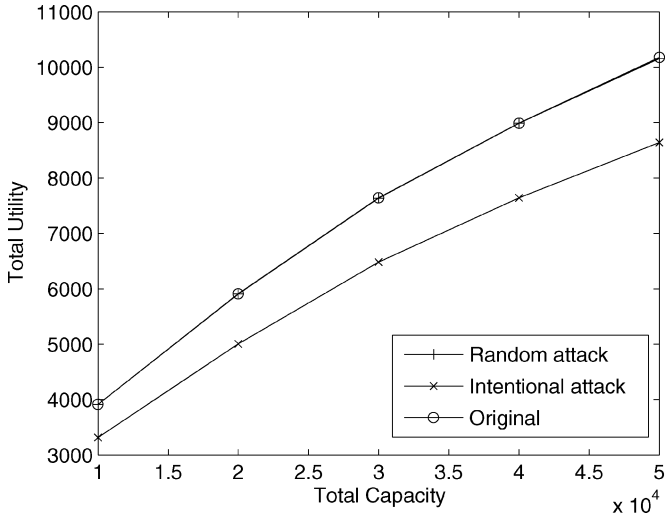

Fig. 3. Total traffic utility in a subnetwork of the Internet.

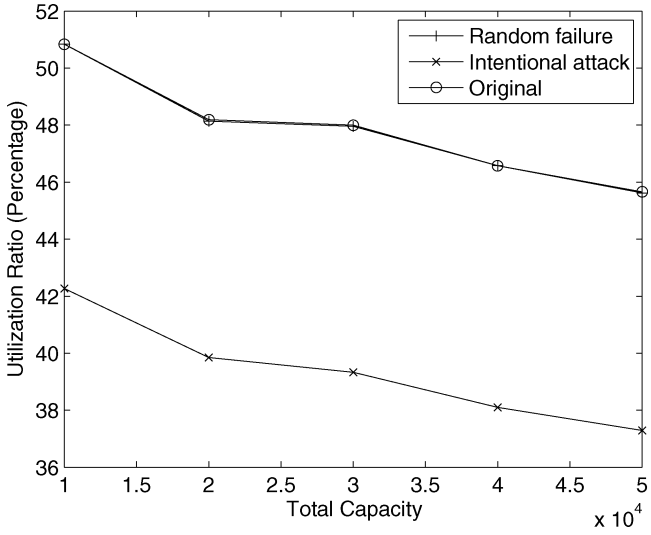

Fig. 4. Utilization ratio of the bandwidth in a subnetwork of the Internet.

Fig. 4 gives the simulated results of utilization ratio of the bandwidth in a subnetwork of the Internet. The utilization ratio of the bandwidth decreases as the total capacity $\sum_l C_l$ rises, which means that a higher percentage of bandwidth is wasted, although the total utility increases, as illustrated in Fig. 3. Comparing three cases, we find similar results as shown in 3 , i.e., the "random attack" curve overlaps with the original one, whereas the "intentional attack" curve is much lower.

Figs. 5 and 6 show the performances of the BA network. The simulated results of the BA network is similar to the Internet data, i.e., the "random attack" curve is overlapped with the original one, whereas the "intentional attack" curve is remarkably lower. The results indicate that the BA scale-free network is robust under random attack but fragile under intentional attack.

By contrast, as illustrated in Figs. 7 and 8, the random graph is robust to both random and intentional attacks. There are only slight declines when the network is attacked. The random graph is a homogeneous network, in which there is no node with an enormous number of connections. As a result, the traffic is well

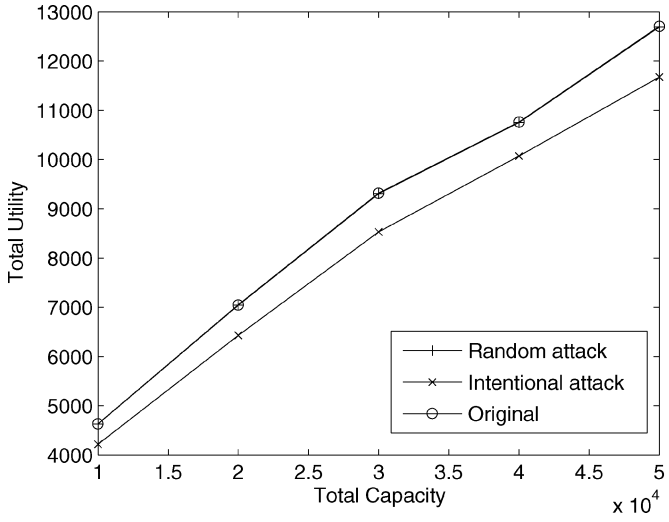

Fig. 5. Total traffic utility in a BA network.

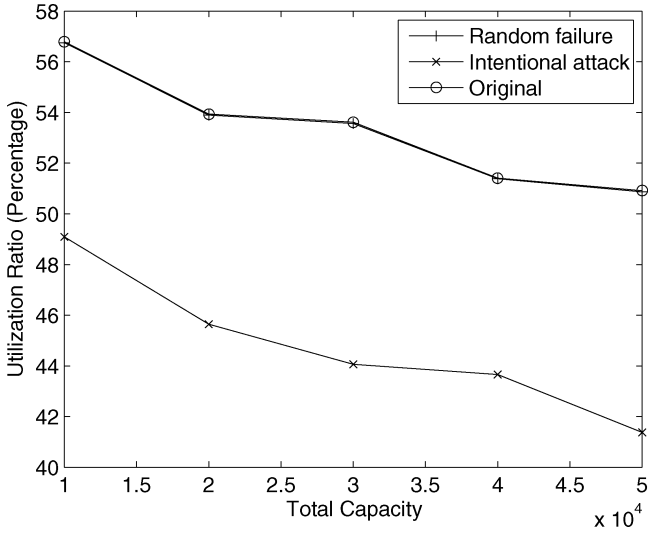

Fig. 6. Utilization ratio of the bandwidth in a BA network.

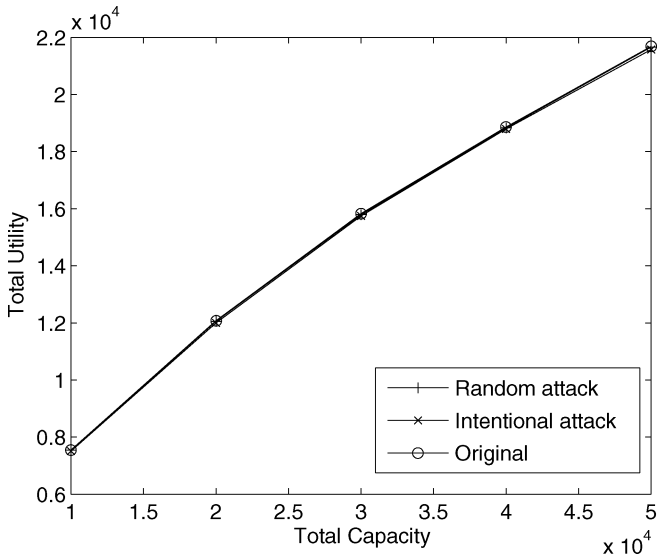

Fig. 7. Total traffic utility in a random graph.

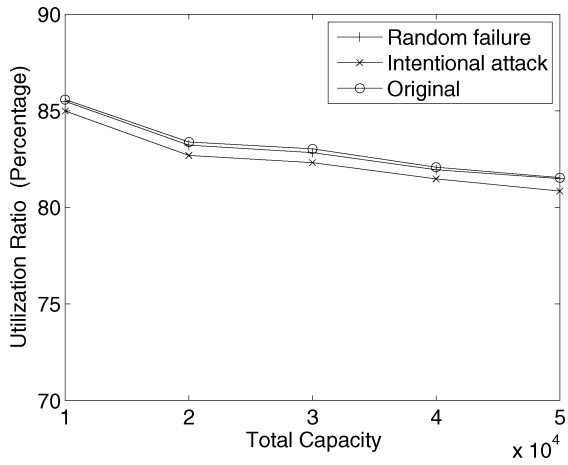

Fig. 8. Utilization ratio of the bandwidth in a random graph.

distributed among all nodes. Therefore, the attack at one node (no matter randomly or intentionally) has little effect on the traffic performance of the whole network. These results also agree with the results in the study on the cascading failures in the random graph [5].

The simulations confirm the "robust yet fragile" property of scale-free networks, whereas the random graph is robust to both random and intentional attacks. Although these results are also found in the study on the cascading failures in complex networks [5], our mechanisms and performance measurements are quite different from those in previous study. In the study on cascading failures, because the flow rates are fixed, the failure of one node may cause a sequence of overload failures. Under this condition, the size of the largest connected component is a good measurement of network performance. In our scenario, on the other hand, the optimal bandwidth allocation algorithm reallocates the flow rates to avoid cascading failures. We use the total utility and the utilization ratio of the bandwidth as the performance indicators. The comparison with cascading failure results indicates that the "robust yet fragile" property is intrinsic to scale-free networks.

V. Conclusion

This brief has studied the attack vulnerability of the Internet in consideration of the real TCP congestion control protocols. Both real Internet data and the typical BA network show a considerable decline in performance when they encounter an intentional attack. However, a random attack does not significantly affect the network performance. As to a random graph, both random and intentional attack have little effect on the network traffic.

Interesting enough, although our traffic model and performance indicators are quite distinct from the cases in previous cascading failures results, we get similar conclusion: the scale-free network has "robust yet fragile" property, whereas the random graph is robust to both random and intentional attacks.

REFERENCES

[1] X. F. Wang and G. Chen, "Complex networks: Small-world, scale-free and beyond," *IEEE Circuits Syst. Mag.*, vol. 3, no. 1, pp. 6–20, Jan. 2003.

[2] D. J. Watts and S. H. Strogatz, "Collective dynamics of 'small-world' networks," *Nature*, vol. 393, pp. 440–442, Jun. 1998.

[3] A.-L. Barabási and R. Albert, "Emergence of scaling in random networks," *Science*, vol. 286, pp. 509–512, 1999.

[4] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D.-U. Hwang, "Complex networks: Structure and dynamics," *Phys. Rep.*, vol. 424, pp. 175–308, 2006.

[5] A. E. Motter and Y.-C. Lai, "Cascade-based attacks on complex networks," *Phys. Rev. E*, vol. 66, p. 065102(R), 2002.

[6] L. Zhao, K. Park, and Y.-C. Lai, "Attack vulnerability of scale-free networks due to cascading breakdown," *Phys. Rev. E*, vol. 70, p. 035101(R), 2004.

[7] P. Crucitti, V. Latora, and M. Marchiori, "Model for cascading failures in complex networks," *Phys. Rev. E*, vol. 69, p. 045104(R), 2004.

[8] Y. Xia, C. K. Tse., F. C. M. Lau., W. M. Tam., and X. Shan, "Traffic congestion analysis in complex networks," in *Proc. ISCAS '06*, Kos, Greece, May 2006, pp. 2625–2628.

[9] D. Arrowsmith, M. di Bernardo, and F. Sorrentino, "Communication models with distributed transmission rates and buffer sizes," in *Proc. ISCAS '06*, Kos, Greece, May 2006, pp. 5047–5050.

[10] R. M. May and A. L. Lloyd, "Infection dynamics on scale-free networks," *Phys. Rev. E*, vol. 64, p. 066112, 2001.

[11] R. Pastor-Satorras and A. Vespignani, "Epidemic spreading in scale-free networks," *Phys. Rev. Lett.*, vol. 86, pp. 3200–3203, 2001.

[12] X. F. Wang and G. Chen, "Synchronization in scale-free dynamical networks: Robustness and fragility," *IEEE Trans. Circuits Syst. I, Fundam. Theory Appl.*, vol. 49, no. 1, pp. 54–62, Jan. 2002.

[13] P. Checco, M. Biey, G. Vattay, and L. Kocarev, "Complex network topologies and synchronization," in *Proc. ISCAS '06*, Kos, Greece, May 2006, pp. 2641–2644.

[14] Z. Li and G. Chen, "Global synchronization and asymptotic stability of complex dynamical networks," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 53, no. 1, pp. 28–33, Jan. 2006.

[15] J. Zhou and T. Chen, "Synchronization in general complex delayed dynamical networks," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 53, no. 3, pp. 733–744, Mar. 2006.

[16] F. M. Atay, T. Biyikoglu, and J. Jost, "Synchronization of networks with prescribed degree distribution," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 53, no. 1, pp. 92–98, Jan. 2006.

[17] S. H. Low and D. E. Lapsley, "Optimization flow control — I: Basic algorithm and convergence," *IEEE/ACM Trans. Netw.*, vol. 7, no. 6, pp. 861–874, Dec. 1999.

[18] F. P. Kelly, A. K. Maulloo, and D. K. H. Tan, "Rate control for communication networks: Shadow prices, proportional fairness and stability," *J. Oper. Res. Soc.*, vol. 49, no. 3, pp. 237–252, Mar. 1998.

[19] S. Liu, T. Basar, and R. Srikant, "Controlling the Internet: A survey and some new results," in *Proc. 42nd IEEE Conf. Decision Control*, Maui, HI, Dec. 2003, pp. 3048–3057.

[20] Routing, MOAT, NLANR, La Jolla, CA, 2006 [Online]. Available: http://moat.nlanr.net/Routing/rawdata

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

[22] P. Erdős and A. Rényi, "On the evolution of random graphs," *Publ. Mah. Inst. Hung. Acad. Sci.*, vol. 5, pp. 17–60, 1960.

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

[24] K.-I. Goh, B. Kahng, and D. Kim, "Univeral behavior of load distribution in scale-free networks," *Phys. Rev. Lett.*, vol. 87, p. 278701, 2001.

[25] S. H. Low, F. Paganini, and J. C. Doyle, "Internet congestion control," *IEEE . Syst. Mag.*, vol. 20, pp. 2–43, Feb. 2002.