On the Effect of Scale-Free Structure of Network Topology on End-to-End Performance

Hiroyuki Ohsaki, Kotaro Yagi, and Makoto Imase
Graduate School of Information Science and Technology
Osaka University, Japan
E-mail: {oosaki,kou-yagi,imase}@ist.osaka-u.ac.jp

Abstract

In recent years, it has been reported that several existing networks including the Internet have scale-free structure. In this paper, through a simple numerical analysis, we investigate effect of the scale-free structure of communication networks on their end-to-end performance. As network topologies, a random network and a scale-free network with the equal number of nodes and the equal number of links are used. We compare end-to-end performance of flows (i.e., throughput) in both random and scale-free networks. Consequently, we show that when the average degree of a network is small (i.e., whah the number of links is small), a scale-free network shows better end-to-end performance. On the contrary, when the average degree of a network is large (i.e., when the number of links is large), we show that a random network shows better end-to-end performance.

keywords: Random Network, Scale-free Network, BA (Barabasi-Albert) Model, End-to-End Performance

1. Introduction

In recent years, scale expansion and complication of a network has been rapidly progressing [10]. Among various artificial systems that exist actually, the Internet is one of the most complicated systems. Since the scale of the Internet has been continuously expanding, it is difficult to intuitively understand its behavior. For clarifying characteristics of such a large-scale system, research on large-scale and complex systems has been actively performed [12, 22].

Among those researches, in particular, attention to a topology of large-scale communication networks has been increasing [1]. Studies on a topology of communication networks have long history [17, 2], but those conventional researches have focused on comparatively small-scale communication networks. In the late 1990s, it was discovered that several real networks such as the topology of Internet ASs and the hyperlink structure of Web pages exhibit scale-free structure [12].

Such finding causes increasing concern on the optimal topology of, in particular, large-scale communication networks. Most conventional researches on a topological of communication networks have focused on regular networks such as star, ring, and mesh, and random networks where the probability that a link exists between arbitrary node pair is given by uniform distribution. However, as several interesting characteristics of a scale-free network become clear, the relation between a communication network and its scale-free structure was been attracting much attention.

Notable characteristics of a scale-free network include, for instance, that the average distance of a network (i.e., the average of shortest path lengths between arbitrary node pair) is much smaller than that of a random network, and that a scale-free network is more robust to random node failures (i.e., connectivity among nodes is more likely to be preserved). Hence, researches on a topology of large-scale communication networks for improving reliability [21] and for improving packet transfer efficiency [20, 13] have been performed.

However, conventional researches have focused only on network-level performance (e.g., reliability, resilience, and per-link transfer efficiency); end-to-end performance, which should be the most important metric to users, has not been fully investigated.

The characteristic of a scale-free network that the average distance of a network is small is advantageous for performing information retrieval [15] or maintaining reachability [19]. However, considering packet transfer over a communication network, such a small average distance of a scale-free network implies traffic concentration at hub nodes. If traffic is concentrated at hub nodes, those nodes would become the bottleneck of the network, and limit the performance of the entire network. Namely, from a viewpoint of the end-to-end performance, a small average distance of a scale-free network and traffic concentration at hub nodes should have opposite effects on the end-to-end performance.

In this paper, through simple numerical analyses, we investigate the effect of the scale-free structure of a communication network on its end-to-end performance. As topologies of a network, a random network and a scale-free network with the equal number of nodes and the equal number of links are used. We compare end-to-end throughput of flows in both random and scale-free networks. Consequently, we show that when the average degree of a network is small (i.e., the number of links is small), a scale-free network shows better end-to-end performance. On the contrary, when the average degree of a network is large (i.e., the number of links is large), we show that a random network shows better end-to-end performance.

The organization of this paper is as follows. First, related works are introduced in Section 2. Section 3 briefly explains fundamental characteristics of a random network and a scale-free network. In Section 4, by simple numerical computations, throughput of each flow is obtained by assuming that the bandwidth allocation to each flow satisfies Max-Min fairness [4]. In Section 5, effect of scale-free structure of a network on its end-to-end performance is clarified by some numerical examples. Finally in Section 6, we conclude this paper and discuss future works.

2. Related Works

In [14], the effect of scale-free structure of a network on router load is investigated, where *router load* is defined as the number of packets processed at a router. The authors of [14] show that the distribution of router loads follows power-law under the following conditions: (1) all nodes generate packets uniformly, and (2) routing is determined simply by the number of hops (i.e., packets traverse the shortest path among possible paths).

However, [14] investigates only scale-free networks, and does not investigate non-scale-free networks. Also, it focuses only on router load, which is one of network-level performance metrics, and does not take account of the end-to-end performance.

In [18], the optimal network topology of a packet switching network is investigated by assuming that routing is determined by a neural network. The authors of [18] show that a scale-free network is optimal as a network topology under the following conditions: (1) the distribution of buffer sizes of routers follows power-law, and (2) routing of a packet is determined by a probability proportional to the buffer size of a down-stream router. However, in [18], since the authors assume proprietary routing utilizing a neural network, results there would not be applicable to general packet switching networks. Moreover, they focus only on per-link performance, and do not take account of the end-to-end performance.

In [9], characteristic of a packet switching network is investigated by proposing a simple model for packet switching networks. The authors consider the following simple models: a fixed number of packets exist in a network, each of which moves randomly among nodes, and every router can accommodate only a single packet.

The authors of [9] show that characteristic of this network model is described by Fermi-Dirac distribution, and the probability of congestion occurrence in a network (i.e., probability that a packet cannot be accommodated in a router) is derived. However, since the authors assume quite simplified network model, it is difficult to generalize this result to other packet switching networks. Moreover, it focuses only on the per-link performance, and does not take account of the end-to-end performance.

Unlike these researches, in this paper, we conduct investigation focusing not only on the per-link performances but on the end-to-end performance, which must be one of the most important performance metrics to users.

3. Random Network and Scale-Free Network

In this section, fundamental characteristics of a random network and a scale-free network are outlined. Moreover, models that generate either a random network or a scale-free network are also explained.

A random network is a network where the probability that a link exists between arbitrary node pair is given by uniform distribution [7]. As a representative model for generating a random network, ER (Erdos-Renyi) model [11] is widely used. Two parameters, the number N of nodes and the connection probability p between nodes, are used for generating a random network. For a given N nodes, a random network is generated by creating links among all node pairs with the probability p.

As characteristics of a random network, it is known that the degree distribution $P(k)$ follows binomial distribution and, for sufficiently large N, the average distance l satisfies $l \propto \log N$ [8].

A scale-free network is a network where its degree distribution follows the following power-law [12].

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

In the literature, several models that generate scale-free networks have been proposed [6, 16]. In this paper, we explain BA model (Barabasi Albert) [3], which is one of the most representative models for generating a scale-free network. Notable features of BA model are network growth and link preferential attachment [3]. First, a connected network with a small number of nodes is created, and nodes are added to the network one-by-one. Then, a scale-free network can be generated by creating a link between a new node and existing one, which is randomly chosen from all existing nodes with a probability proportional to the degree of an existing node. It is known that the network generated by BA model has a power-law index of $\lambda = 3$.

As characteristics of a scale-free network, it is known that the average distance l is much smaller than that of a random network. For instance, it is known that, for a sufficiently large N, the average distance satisfies $l \propto \log \log N$ for $2 < \lambda < 3$ [8].

4. Network Model and Flow Throughput Computation Algorithm

In what follows, we investigate the effect of scale-free structure of a network on the end-to-end performance by simple numerical computation. We do not take account of influence of TCP congestion control; instead, we assume that bandwidth allocation to each flow satisfies Max-Min fairness [4]. This makes it possible to investigate the characteristic of a network topology, neglecting the complex and unpredictable effect of TCP congestion control.

As a network topology, a random network and a scale-free network generated by BA model are used. A random network is denoted by $G_R = (V_R, E_R)$, and a scale-free network is denoted by $G_S = (V_S, E_S)$. Moreover, the number of nodes (i.e., routers or hosts) in a network is denoted by $N (= |V_R| = |V_S|)$, and the average degree (i.e., the average of the number of links connected to a node) is denoted by k. We clarify effect of a network topology on the end-to-end performance by comparing a random network and a scale-free network with the same number of nodes N and the average degree k. For simplicity, all link bandwidths and propagation delays are identical, and are denoted by B and L, respectively.

On the random network G_R and the scale-free network G_S, flows are generated randomly. Flows on each network is given by another random network T_R with N nodes and the average degree ρk (see Fig. 1). ρ is called *load factor*. Since the number of links of G_R and G_S is $|E_R| = |E_S| = kN/2$, load factor ρ is a parameter that determines the ratio of the number of flows to the number of links in a network. In addition, routing of each flow is determined by the shortest path algorithm [5].

We assume that the bandwidth allocation to each flow satisfies Max-Min fairness. Throughput of each flow is calculated by the following algorithm (see Fig. 2).

1. Throughput allocation to each flow is initialized to 0. Namely, by letting r_i be the throughput of flow i,

$$\forall i \quad r_i \quad \leftarrow \quad 0. \tag{2}$$

2. Throughput of all flows not traversing the bottleneck link is increased by Δ.

$$\forall i \in \mathcal{F} \quad r_i \quad \leftarrow \quad r_i + \Delta,$$

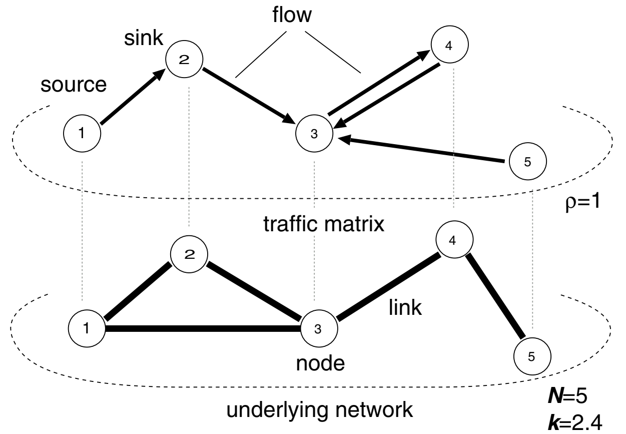

Figure 1. Flow model and underlying network model

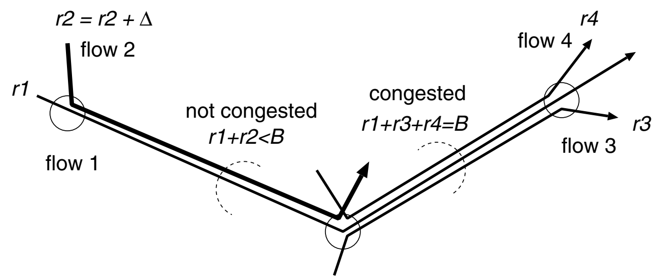

Figure 2. Computation of each flow's throughput satisfying Max-Min fairness

where $\mathcal{F}$ is a set of flows not traversing the bottleneck link.

3. If one or more of non-bottleneck links becomes bottleneck, all flows traversing those links are removed from $\mathcal{F}$. Return to step 2.

By performing the above algorithm for a sufficiently small Δ, throughput of each flow, in the case that the bandwidth allocation satisfies Max-Min fairness, can be numerically obtained.

5. Numerical Examples

Next, we discuss the effect of scale-free structure of a network on the end-to-end performance by showing several numerical examples. In what follows, unless explic-

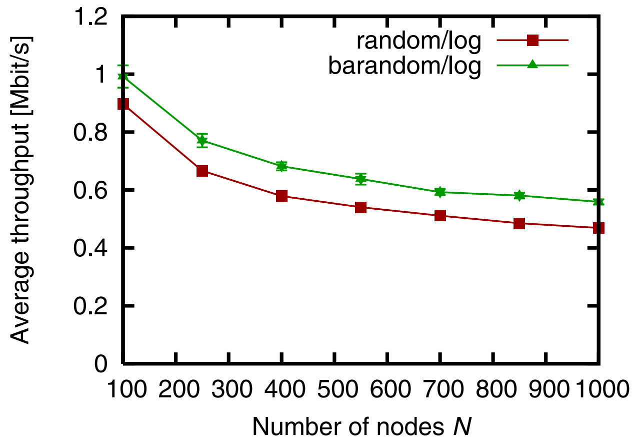

Figure 3. Average flow throughput for the different number of nodes N ($k = 3$, $B = 10$ [Mbit/s], and $\rho = 5.0$)

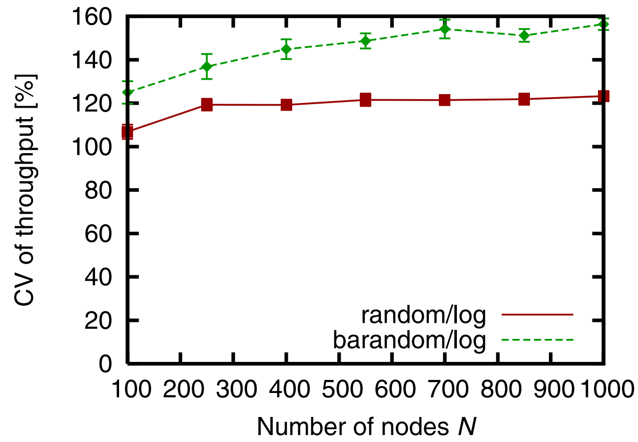

Figure 4. CV of flow throughputs for the different number of nodes N ($k = 3$, $B = 10$ [Mbit/s], and $\rho = 5.0$)

itly stated, the following parameters are used: the number of nodes $N = 1,000$, the average degree $k = 3$, the link bandwidth $B = 10$ [Mbit/s], and load factor $\rho = 5.0$.

30 random networks and scale-free networks were generated, and the throughput of each flow and the utilization of each link were numerically calculated for all networks. Note that utilization of a link is the sum of throughput of all flows passing the link normalized by the link bandwidth.

The average throughput of flows when changing the number N of nodes from 50 to 500from 100 to 1,000 is shown in Fig. 3. In the figure, 95 % confidence interval is plotted. This figure shows that the scale-free network generated by BA model shows approximately 15–20 % higher throughput regardless of the number N of nodes.

It deserves attention that Average throughputs of flows in the random network and the scale-free network differ largely, in spite of the fact that these networks have the same number N of nodes and the average degree k. In our experiment, $\rho|E_R|(= \rho|E_S|)$ flows are randomly generated on either the random network or the scale-free network. Hence, our result indicates that the scale-free network can carry 15–20% more packets than the random network.

Now, we focus on the variation in flow throughputs. The coefficient of variation of flow throughputs is shown in Fig. 4 when changing the number N of nodes from 50 to 500from 100 to 1,000. This figure shows that the scale-free network shows larger variation in flow throughputs than the random network, regardless of the number N of nodes. For instance, the cumulative probability of flow throughput for $N = 1,000$ is shown in Fig. 5. This figure shows that, as compared with the random network, the scale-free network has many flows with low throughput , but has a few flows with extremely high throughput.

As seen so far, when we focus on the end-to-end performance, the scale-free network shows higher throughput than the random network. However, focusing on per-link performance, we will see contradictory results.

The average link utilization when changing the number N of nodes from 50 to 500from 100 to 1,000 is shown in Fig. 6. Note that, even when utilization of bottleneck links is

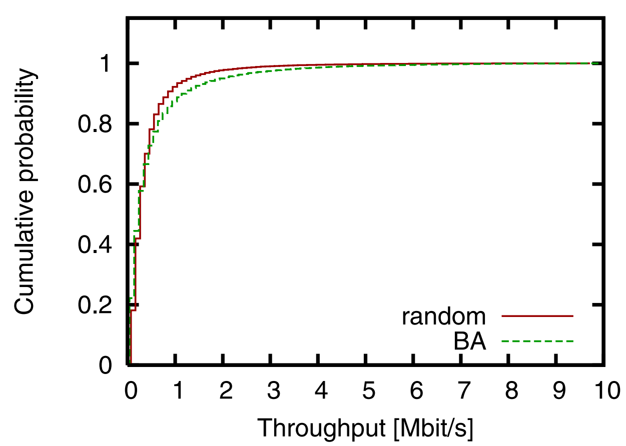

Figure 5. Cumulative probability of flow throughput for $N = 1,000$

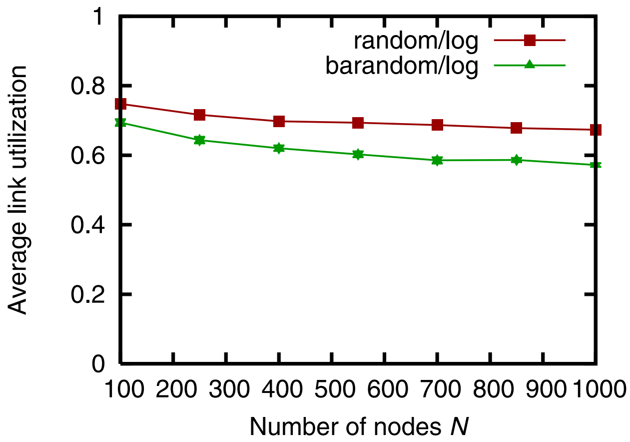

Figure 6. Average link utilization for the different number of nodes N ($k = 3$, $B = 10$ [Mbit/s], and $\rho = 5.0$)

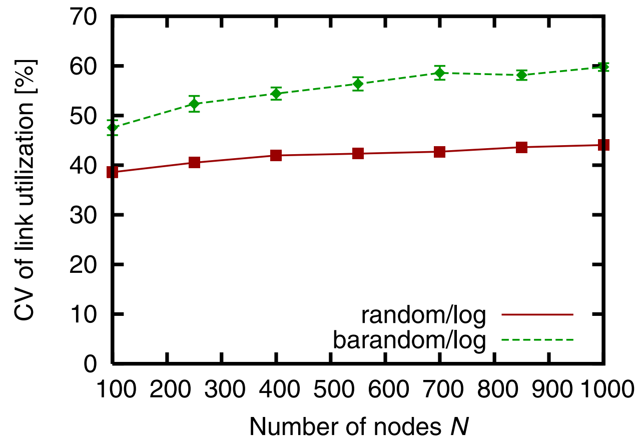

Figure 7. CV of link utilization for the different number of nodes N ($k = 3$, $B = 10$ [Mbit/s], and $\rho = 5.0$)

1.0, the average link utilization should take a value smaller than 1.0, since there exist non-bottleneck links and/or links that are not utilized by any flow. This figure shows that, although it depends on the number N of nodes in a network, the random network shows 5–15% higher throughput than the scale-free network.

This phenomenon can be explained as follows. In the scale-free network, since the average distance is smaller than that of the random network, each flow generally traverses fewer links. Consequently, in the scale-free network, each flow consumes less network resources (i.e., link bandwidth). For instance, the average path length for $N = 1,000$ was approximately 6.4 and 4.6 in the random network and the scale-free network, respectively.

In this case, each flow in the scale-free network consumes only as approximately 72% (= 4.6/6.4) of network resources as in the random network.

However, in the scale-free network, due to existence of hubs (i.e., nodes with extremely many links), link utilization of all links differs largely. In Fig. 7, the coefficient of variation of link utilization is shown when changing the number N of nodes from 50 to 500from 100 to 1,000. This figure shows that the variation in link utilization in the scale-free network is larger than in the random network. This tendency becomes noticeable as the number N of nodes becomes large.

From a viewpoint of the end-to-end performance, scale-free structure causes positive effect such that the number of hops between nodes is small, and also negative effect such that traffic is likely to be concentrated on hubs. The strength of these positive and negative effects should determine the end-to-end performance of scale-free networks.

For instance, in Fig. 3, the average throughput of the scale-free network is larger than that of the random network, regardless of the number N of nodes. This is probably because the positive effect, such that the number of hops between nodes is small, is significant.

However, as the average degree becomes large, the negative effect of scale-free structure becomes noticeable. In this case, the random network shows better end-to-end performance than that of scale-free network. In the numerical results so far, the average degree is fixed at $k = 3$. In Fig.Figs. 8 and 9, the average flow throughput isand the coefficient of variation of flow throughputs are shown when changing the average degree k from 2 to 8.

This figureFigure 8 shows that the throughput of the random network is higher than that of the scale-free network, when the average degree k is larger than or equal to 6. This phenomenon can be explained as follows. The number of links in a network increases as the average degree k becomes large. The average distance of a network becomes small as the number of links increases. Hence, the difference in average distances of the random network and the scale-free network becomes small. On the contrary, in our experiments, routing of each flow is simply determined by the shortest path algorithm. Hence, when a network has scale-free-network structure, flows are likely to traverse hub nodes. Consequently, hub nodes become congested, leading throughput degradation for many flows. This phenomenon can be confirmed from Fig. 9; i.e., the scale-free network shows larger variation in flow throughputs than the random network.

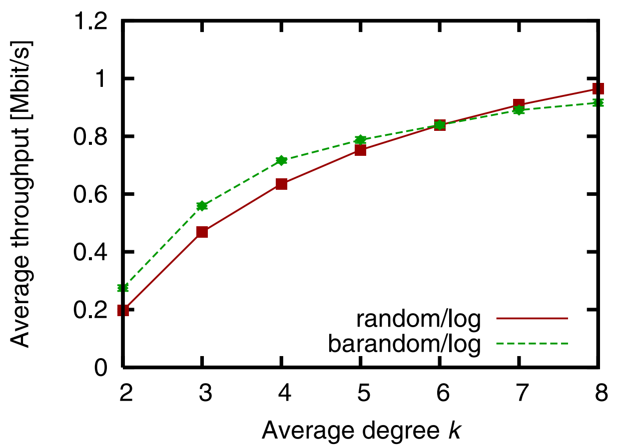

Figure 8. Average flow throughput for a different average degree k **(**$N = 1,000$**,** $B = 10$ **[Mbit/s], and** $\rho = 5.0$**)**

6. Summary and Future Works

In this paper, through simple numerical analyses, we have investigated the effect of scale-free structure of a network on its end-to-end performance. Consequently, we have found: (1) when the average degree of a network is small (i.e., there are not many links in a network), a scale-free network shows higher end-to-end performance than a random network, and (2) conversely, when the average degree of a network is large (i.e., many links in a network), a random network shows higher end-to-end performance than a scale-free network.

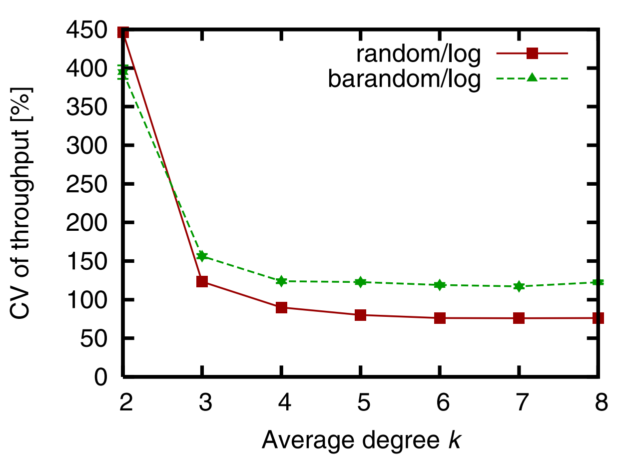

Figure 9. CV of flow throughputs for a different average degree k **(**$N = 1,000$**,** $B = 10$ **[Mbit/s], and** $\rho = 5.0$**)**

As future work, we are planning to investigate the effect of the scale-free structure of a network on other end-to-end performance metrics such as delay and packet loss probability as well as application-level performance metrics. Also, we are planning to investigate the effect of scale-free structure of a network on several network protocols, for instance, the congestion control mechanism of TCP.

References

[1] R. Albert and A. L. Barabasi. Statistical mechanics of complex networks. *Thesis (PhD). University of Notre Dame*, 74:47–97, July 2001.

[2] R. Albert, H. Jeong, and A. L. Barabsi. Diameter of the world wide web. *Nature*, 401:130–131, Sept. 1999.

[3] A. L. Barabsi and R. Albert. Emergence of scaling in random networks. *Sience*, 286:509–512, Oct. 1999.

[4] D. Bertsekas and R. Gallager. *Data Networks*. Prentice-Hall, Englewood Cliffs, New Jersey, 1987.

[5] D. Bertsekas and R. Gallager. *Data Networks*. Prentice Hall, May 1990.

[6] G. Bianconi and A. L. Barabasi. Competition and multiscaling in evolving networks. *Europhysics Letters*, 54:436–442, May 2001.

[7] B. Bloolbas. *Random Graphs Second Edition*. Cambridge Univ Press, Oct. 2001.

[8] R. Cohen and S. Havlin. Scale-free network are ultra small. *Physical Review Letters*, 90(5):058701, Feb. 2003.

[9] A. P. S. de Moura. Fermi-dirac statistics and traffic in complex networks. *Physical Review E*, 71.066114:1–6, June 2005.

[10] S. N. Dorogovtsev and J. F. F. Mendes. Evolution of networks. *Advances in Physics*, 51:1079–1187, Sept. 2002.

[11] P. Erdos and A. Renyi. On random graphs. *Mathematicae*, 6:290–297, 1959.

[12] M. Faloutsos, P. Faloutsos, and C. Faloutsos. On power-law relationships of the Internet topology. *ACM SIGCOMM Computer Communication Review*, 29:251–263, May 1999.

[13] C. M. Ghim, E. Oh, K. I. Goh, B. Kahng, and D. Kim. Packet transport along the shortest pathways in scale-free networks. *The European Physical Journal B*, 38:193–199, Feb. 2004.

[14] K. I. Goh, B. Kahng, and D. Kim. Packet transport and load distribution in scale-free network models. *Physica A*, 318:72–79, Feb. 2003.

[15] B. Kim, C. Yoon, S. Han, and H. Jeong. Path finding strategies in scale-free networks. *Physical Review E*, 65(2):027103, Jan. 2002.

[16] K. Klemm and V. M. Eguiluz. Highly clustered scale-free networks. *Physical Review E*, 65:036123, Feb. 2002.

[17] H. L. Topology control for multihop packet radio networks. *IEEE Transactions on Communications*, 41:1474–1481, Oct. 1993.

[18] J. Ohkubo and T. Hosiguchi. Scale-free property of optimal network for packet flow by a packet routing control. *Physica A*, 353:649–660, Aug. 2005.

[19] S. Rhea and Others. Maintenance-free global data storage. *IEEE Internet Computing*, 5:40–49, Oct. 2001.

[20] B. Tadic and G. J. Rodgers. Packet transport on scale free networks. *Advances in Complex Systems*, 5(4):445–456, July 2002.

[21] Y. Tu. How robust is the Internet? *Nature*, 406:353–354, May 2000.

[22] A. Vazquez, R. Pastor-Satorras, and A. Vespignani. Large-scale topological and dynamical properties of the Internet. *Physical Review E*, 65:066130, June 2002.