Proceedings of the 5th WSEAS Int. Conf. on MULTIMEDIA, INTERNET AND VIDEO TECHNOLOGIES, Corfu, Greece, August 17-19, 2005 (pp85-90)

The Influence of Network Topological Models on the Prediction of End-to-End Loss Probabilities

PAOLO ADDESSO, ANTONIO GRAVETTI, MAURIZIO LONGO, FABIO POSTIGLIONE
Dipartimento di Ingegneria Elettrica ed Ingegneria dell'Informazione
Università degli Studi di Salerno
Via Ponte don Melillo, I-84084 Fisciano (SA)
ITALY

Abstract: - In last years new procedures to generate Internet-like topologies have been proposed, stemming from the evidence of a *power-law* behavior of real networks. Realistic networks generators are necessary for simulation and performance evaluation of data communication systems. Assuming a simple loss model for the links, the present investigation shows that the adoption of a specific topological model for the Internet graph (here, the Waxman or the Barabási-Albert model) may affect, to a certain extent, the estimates of the loss probabilities for end-to-end communications obtained by simulation (e.g. via NS-2 package).

Key-Words: - Internet topology, Packet loss, Waxman model, Barabási-Albert model, NS-2.

1 Introduction

For some while, technical and scientific research activity related to the Internet world has been mainly focused on new kinds of applications and communication protocols in order to develop new service scenarios or guarantee the interoperability among software and hardware entities of the network. In last years, researchers have also been led to investigate new mathematical models suitable for the network traffic [1], aiming to face the increasing demand of Quality of Service (QoS). More recently, new research activities have been stimulated by the exponential growth of the number of hosts connected to the Internet and the huge amount of data traffic exchanged by them. It has been realized that a comprehensive understanding of the Internet is necessarily related to the mathematical characterization of its topological structure, since it has been shown that network topology can influence protocol behaviors [2]. Such investigations are particularly relevant for the design of new services requiring high QoS level, such as Voice over IP (VoIP) and video-streaming.

Modeling the Internet topology is an important open problem both from a theoretical point of view and for practical aspects, related to the generation of realistic graphs for the simulation of the behavior of new protocols, the evaluation of communications performance or the analysis of new service platforms for the Internet [3].

The study of the statistical aspects of the Internet topology relies on the data collected by many research projects devoted to the mapping of nodes of the Internet and connections among them, such as the *Internet Mapping Project* [4], *Skitter* by Cooperative Association for Internet Data Analysis (CAIDA) [5] or *Rocketfuel* [6].

Graph-like representations of a large portion of the Internet are now available. Recent studies on such results, due to M. Faloutsos et. al in 1999 [7], have addressed some typical features of many real complex systems.

Indeed, it has been realized that Internet is a scale-free *network* whose interconnection structure is governed by *power-law* distributions (as in the case of the *degrees* of the nodes, the eigenvalues distribution, etc.). This result is in contrast with the distributions of Internet-like networks produced by traditional generators, based on the Erdős-Rényi classical *random graph* (henceforth ER) model [8-10], including the Waxman generator [11] among others.

The discovery of the scale-free nature of the Internet stimulated the introduction of new mathematical models [12-14] reproducing such a scale-free behavior. The first and most popular model was proposed by A.-L. Barabási and R. Albert (henceforth BA) in late 1999 [12]. It is ruled by two simple concepts: *(i)* the graph grows as a consequence of the continuous addition of new nodes; *(ii)* each new node connects to the existing vertices with a probability proportional to their degree (*preferential attachment mechanism*).

Since then, other models have been proposed in order to overcome some limitations of the ER-based and BA models. For further reading Ref. [15] is suggested.

In this paper the influence is investigated of different topological models on communications performance. In particular, for a given link loss model, the end-to-end loss probabilities are evaluated in Internet-like networks, and then compared. Such networks are generated according to two different network generation algorithms: the Waxman algorithm [11] and the BA model [12], both included in the Internet-like topologies generator BRITE [16].

The paper is organized as follows. In Sect. 2 we outline the relevant Internet topology models: ER, Waxman and BA models. In Sect. 3 we describe the loss model adopted for the links of the synthetic Internet-like graphs. In Sect. 4 numerical results via NS-2 simulations are presented. Finally, Sect. 5 yields conclusions and references to further current work.

2 Internet topological models

Internet topologies are generally represented as graphs that mime the large-scale characteristics of the maps obtained by measurements of the real network. As it is well known, Internet is a world-wide network composed of computers (or *hosts*) communicating by means of intermediate nodes (or *routers*), responsible to forward properly the information flows, and of *links* physically interconnecting each pair of nodes. It is possible to represent the network as an undirected graph, whose vertices are the routers and whose edges stand for the physical connections between pairs of them: this is often referred as the *Internet Router* (IR) *level representation*.

At an higher level, Internet can be partitioned into many autonomously administered routing domains, named *Autonomous Systems* (AS), that are groups of nodes under a common administration and sharing routing information. Information flows among them are ruled by inter-AS routing protocols, such as the Border Gateway Protocol (BGP). Then, another possible representation of Internet is an undirected graph in which vertices represent the ASs and edges are *peering relationships* between pairs of ASs: this is the so-called *AS level representation*. A peering relationship is an agreement between two ASs that exchange traffic and routing information through one or more directly connected *border* routers (gateways). Fig. 1 illustrates the representations of Internet at both levels.

The traditional approach to model data networks relied on the use of classical random graphs, introduced by P. Erdős and A. Rényi in 1959 [8,9] (ER model). Based on such a model, the computer science community developed some tools to reproduce Internet in order to test new protocols [11,17].

After the discovery of the scale-free nature of Internet [7], such models and generators became inadequate to describe or reproduce large Internet-like networks (however, they seem to give good results for small and medium networks), so the rise of new models has been stimulated in order to catch its

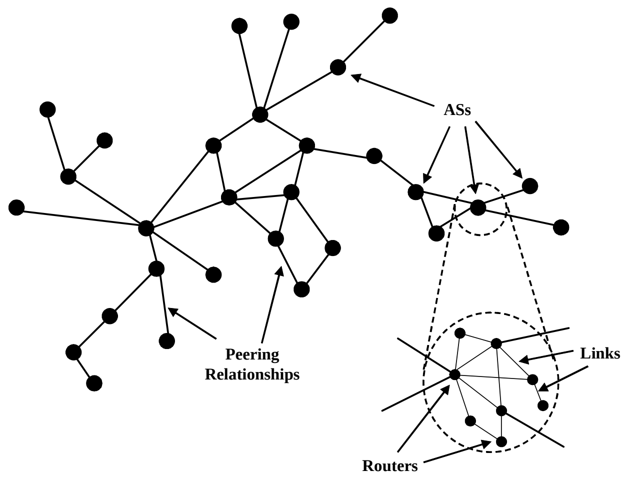

Figure 1 – A schematic representation of Internet AS-level map with a router level map of an AS.

features and new paradigms have been proposed to generate representative synthetic networks [16,18].

2.1 Static random graphs and the Waxman model

Static random graph models entail a fixed number of nodes N throughout the generation process. Typical examples of such models are the Erdős-Rényi and the Waxman models.

Technically speaking, an undirected graph G is a pair of sets $G = \{V,E\}$ where V is the set of vertices and E is the set of edges connecting two vertices of V. The size of the graph is $|V| = N$.

An ER random graph $G_{N,p}$ can be defined [19] as a graph with N nodes where each of the $N(N - 1)/2$ possible edges is present, independently of the others, with probability p, called *connection probability*, and absent with probability $1 - p$. In this ensemble, the number of edges M is a binomial random variable (r. v.), viz. $M \sim B\left(N(N-1)/2, p\right)$. Such a model, also named *binomial model*, has been widely adopted by Internet researchers.

Defining the *degree* of a node as the number of edges attached to it, the *average degree* can be easily computed as

$$\bar{k} = p(N-1) \cong pN, \qquad (1)$$

where the approximation holds for large N.

The theory of random graphs studies the properties of the probability space of graphs as $N \to \infty$.

Since real networks (even evolving ones, such as the Internet) are characterized by an almost constant average degree [15], it is convenient to consider $p(N) = \bar{k}/N$, derived from eq. (1). If $\bar{k} < 1$, the network is composed by isolated subgraphs and hence it can be represented as a collection of *clusters*. On the other

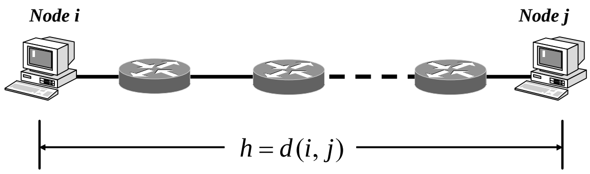

Figure 2 – Communication path between two generic nodes.

hand, when $\bar{k} > 1$ a giant cluster (aka the *giant component*) emerges which incorporates almost all nodes, as $N \to \infty$. As the average degree approaches the critical value $\bar{k} = 1$ an abrupt change in the cluster structure occurs. At the corresponding critical probability $p_c(N) = 1/N$ the random graph changes its topology abruptly from a collection of isolated components to a single giant cluster, as it also happens at the *percolation transition* in the Percolation Theory.

One of the main features of a random graph is its *degree distribution* $P(k)$. In an ER graph with N nodes and connection probability p it is

$$P(k) = \binom{N-1}{k} p^k (1-p)^{N-1-k} \cong e^{-\bar{k}} \frac{\bar{k}^k}{k!}, \qquad (2)$$

where the Poisson approximation holds for large N and for constant $\bar{k}$.

Another fundamental subject of investigation in graph theory for networks applications is the distribution of the distances among nodes, or *hop count distribution* for short, expressed, in terms of probability mass function (*pmf*) as $f(n) = \Pr\{h = n\}$, where $h = d(i, j)$ is a r. v. representing the length of the shortest path connecting a pair of randomly selected vertices i and j or, equivalently the number, plus 1, of hops to be traversed to reach node j starting from node i, as represented in Fig. 2. For ER graphs the average distance is [20]

$$\bar{h} \cong \frac{\log N}{\log \bar{k}}. \qquad (3)$$

It is evident that $\bar{h}$ is much smaller than the size of the graph N, as a consequence of the *small-world effect* [21], exhibited by many real networks.

Despite reproducing the small-world behavior, the ER model fails to predict some features of Internet topology: for instance, it yields a binomial degree distribution which decreases exponentially for large N and, then, deviates from the heavy-tailed distribution observed in measurement on the Internet.

The ER model has inspired the first Internet topology generator used for protocol testing, proposed by Waxman [11]. According to the Waxman algorithm, nodes are randomly distributed on a rectangular coordinate grid and the probability of an edge between two vertices i and j is

$$p(i, j) = \alpha\, e^{-\frac{d_E(i,j)}{\beta D}} \qquad (4)$$

where $d_E(i, j)$ is the Euclidean distance from node i to j, α and β are parameters in the range (0,1] and D is the maximum distance between two vertices. While the topological structure of the graph is not influenced by the value of D [17,22], it is highly dependent on the values of α and β: α controls directly the number of edges, while β rules the influence of the Euclidean distance between pairs of nodes.

Like ER graphs, Waxman graphs yield values of $\bar{h}$ small with respect to the size of the network and hence consistent with the small-world effect, but fail to yield the heavy-tailed degree distributions observed in Internet.

2.2 Barabási-Albert model

Many complex systems, such as Internet, show degree distributions that are not peaked around a *typical* value, the average degree $\bar{k}$, but instead highly skewed (*scale-free* behavior). The first model for computer networks producing graphs with power-law degree distributions was proposed by A.-L. Barabási and R. Albert in 1999 [12], who claimed that the network is an open system growing along the time, and that the probability that two nodes are connected depends on the degree of the nodes. It means that new edges are not placed at random but tend to connect to vertices that already have a large degree, respecting the paradigm *rich-get-richer*, introduced by H. A. Simon in 1955 [23].

The algorithm inside Barabási-Albert (BA) models can be summarized as follows:

- *Growth*: the network starts at time $t_0 = 0$ with a small number of nodes m_0. At every time unit a new vertex with m edges ($m < m_0$) is added and it is connected to m different nodes already present in the system.
- *Preferential attachment*: The edges of the new vertex are connected to the i-th already existing node with a probability $\Pi(k_i(t))$ proportional to its degree $k_i(t)$ at time t, such that

$$\Pi\left(k_i\left(t\right)\right) = \frac{k_i\left(t\right)}{\sum_j k_j\left(t\right)}. \qquad (5)$$

After t time units the BA procedure provides a graph $G_m^{(N)}$ with $N = t + m_0$ nodes and mt edges, whose average degree is

$$\bar{k} \cong 2m, \qquad (6)$$

where $N \gg m_0$ has been considered. Since m is an integer, eq. (6) implies that $\bar{k}$ can have integer values.

BA graphs have a power-law degree distribution $P(k)$ [12]

$$P(k) \sim k^{-\gamma}, \qquad (7)$$

where $\gamma = 3$, similar to the degree distribution measured in Internet (however, in real networks $\gamma \cong 2.1$ [15]). The power-law distribution implies that the probability of finding vertices with a very large degree is not negligible in the BA graph. These nodes represent the hubs of the network since they connect a large number of other nodes. The hubs represent the core of the BA graph and, hence, they provide many shortcuts among nodes composing the network.

BA graphs exhibits the small-world effect as happens in ER graphs. Indeed, it has been proved rigorously [24] that the *diameter*, namely the *maximum* distance in the shortest path sense, of BA networks shows different asymptotical behaviours , in the limit of large N, depending on the value of the parameter m. The average distance is supposed to behave in similar way. In particular, in BA graphs with $m = 1$, the average distance is, as $N \to \infty$,

$$\bar{h} \sim \log N$$

like ER graphs. Instead for $m \geq 2$ it is asymptotically

$$\bar{h} \sim \frac{\log N}{\log \log N}. \qquad (8)$$

3 Loss model

The hop count distribution of the Internet reflects the interconnection structure among routers (IR level representation) and hence affects significantly the end-to-end communication performance. Thus, since every model proposed to represent Internet provides, in principle, a different distribution of the distance among nodes, the predicted performance may depend on the adopted topological model of the network.

To shed some light on this subject, we carried on an evaluation of packet loss probabilities under different Internet-like topologies, given the loss model of single links.

Every communication between two end-points involves a *path* set up by the routing protocols, aiming to minimize the number of hops packets have to traverse to reach the destination. Although some intra-ASs routing policies might in principle inflate the shortest paths, the distance, h, between source and destination can be assumed to be the length of one of the shortest paths between the two end-points [25] and can be considered fixed along time.

Further, we introduce a simple link loss model, in which every packet is lost on the n-th link, *independently* of the others, with a probability

$$\lambda_n \sim U\left[0, \lambda_{max}\right]. \qquad (9)$$

Such a model introduces some relevant simplifications with respect to reality, since correlations are neglected both in space (from link to link at a given time) and in time (at different times on one link). However, it can be considered as a first step to capture some effects of the network topology on performance analysis.

Then, the end-to-end packet loss probability L can be expressed as

$$L = 1 - \prod_{n=1}^{h}\left(1 - \lambda_n\right) \qquad (10)$$

where λ_n is the loss probability of the n-th link in the path, and where h , as above, is a r. v. representing the shortest path length and conveying the influence of network topology on the packet loss rate.

In the simulations, λ_{max} has been set to the value 10^{-2}, selected to match to the order of magnitude of the maximum loss rate typically encountered inside Internet [26,27].

4 Numerical results

In order to assess the influence of network topology on performance prediction, representative numerical experi ments have been carried out by means of *Network Simulator version 2* (*NS-2*) [28] on synthetic networks generated by BRITE [16]. The latter have been produced according to two algorithms outlined in Sect. 2: the BA model [12] and the Waxman model [11].The link loss model outlined in previous section has been implemented via the NS-2 *Error Model* class. A Constant Bit Rate (CBR) traffic flow (representative of many real-time applications, like VoIP) between two randomly selected nodes (connected by a communication path) is added to the network scenario, in order to compute the end-to-end packet loss probability L as the fraction of packets being lost among those transmitted. The duration of each NS-2 simulation has been established so that the CBR source send 10^5 packets to the destination.

For each topological model, 5000 experiments have been performed via Monte Carlo techniques in graphs with the same size (N = 1000 vertices) and average degree ($\bar{k} = 4$), in reasonable agreement to the value observed in the Internet [15]). In order to obtain the desired average degree, the m parameter of BA graphs, eq. (6), is set to the value $m = 2$.

For Waxman graphs we have selected the value of the α and β parameters of eq. (4) after thorough numerical experiments on the whole parameter space. Figure 3 represents the contour plots in (α,β) for different values of $\bar{k}$ and $\bar{h}$ obtained for Waxman graphs with $N = 1000$ nodes. Note the values of (α,β) providing graphs with the given average degree

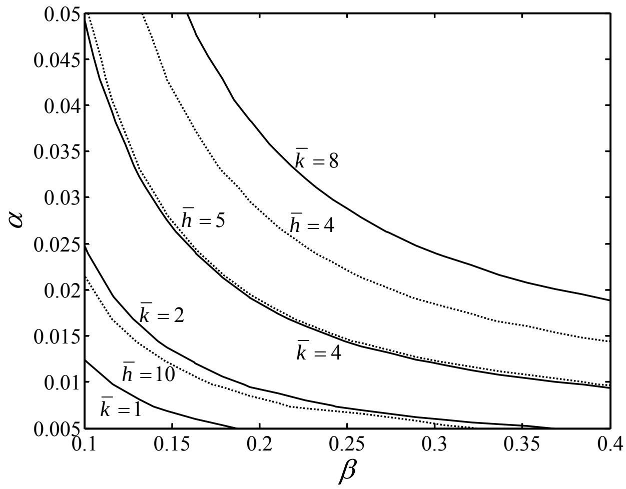

Figure 3 – Contour plots of different values of $\bar{k}$ (solid lines) and $\bar{h}$ (dotted lines) in Waxman graphs with $N = 1000$ nodes and $D = 1000\sqrt{2}$ in a fraction of the parameter space (α,β). $\bar{k} = 1, 2, 4, 8$, $\bar{h} = 4, 5, 10$ are reported.

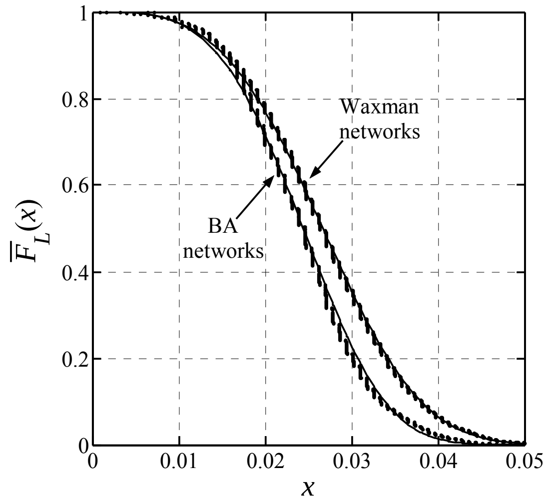

Figure 4 – Complementary Cumulative Distribution function $\bar{F}_L(x)$ of the end-to-end loss probability in BA and Waxman networks. The empirical distributions (dots) are compared with their Weibull fits (solid lines).

$\bar{k} = 4$ also provide an (almost) constant value of $\bar{h}$. Here, we have chosen α = 0.01 and β = 0.37 ($D = 1000\sqrt{2}$), but other choices are essentially equivalent.

The *Complementary Cumulative Distribution Function* (CCDF) $\bar{F}_L(x) = \Pr\{L > x\}$ of the end-to-end loss probability in BA and Waxman graphs are reported in Fig. 4. Both empirical distributions are fitted by a Weibull CCDF

$$\bar{F}(x; a, b) = e^{-ax^b}, \quad x \ge 0. \qquad (11)$$

The parameters in eq. (11) are $a = 5.7 \cdot 10^5$, $b = 3.7$ for BA networks and $a = 1.4 \cdot 10^5$, $b = 3.4$ for Waxman networks respectively.

However, as. Fig. 4 suggests by inspection, the considered graphs, although with the same size N and the same average degree $\bar{k}$, yield different packet loss probabilities. In particular, performances predicted on Waxman networks are poorer than those predicted on BA networks (in terms of average loss rates: $\bar{L}_W = 0.0267 > \bar{L}_{BA} = 0.0243$).

This results is in agreement with eq. (10) which implies that the end-to-end loss probability is related to the behavior of the h r.v., and increases as its pmf shifts to the right. Indeed, the different behavior of h under the two topological models can be appreciated from Fig. 5, which shows the empirical hop count pmfs computed by means of 5000 Monte Carlo iterations for the two models (for both models $N = 1000$ and $\bar{k} = 4$, but the average hop count for Waxman is $\bar{h}_W = 5.04$ whereas, for BA, $\bar{h}_{BA} = 4.61$).

5 Conclusion

The investigation was motivated by the fact that the various topological models of Internet yield different degree distributions, (e.g. $P(k)$ decays exponentially for Waxman model while decays as a power-law for BA model).

Under the assumed Bernoulli link loss model, our numerical experiments (performed by means of NS-2 network simulator on graphs generated by BRITE package with N = 1000 and $\bar{k} = 4$) support the following conclusions: (i) despite the said discrepancy, both end-to-end loss probabilities are fitted by a Weibull distribution; (ii) however the performance predicted in terms of average loss rates is poorer for Waxman networks than for BA networks. The latter fact has been related to the different behavior of the h r.v. for the two models.

Consequently, further investigations could compare end-to-end loss probabilities in different Internet-like graph models showing the same average distance $\bar{h}$.

The methodology presented in this paper could be applied under more realistic loss models for the links taking into account the burstiness and the (long-range) correlations of the loss process measured during real operations in the Internet.

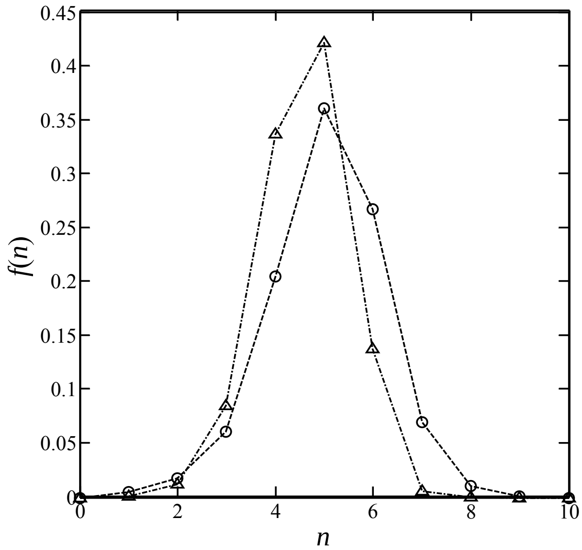

Figure 5 – Empirical pmfs of the hop count $f(n)$ in Waxman networks with α = 0.01, β = 0.37 and $D = 1000\sqrt{2}$ (circles and dashed lines) and BA networks with $m = 2$ (triangles and dot-dashed lines). All experiments are carried out in graphs with N = 1000 vertices.

References:

[1] K. Park and W. Willinger, editors. *Self-similar network traffic and performance evaluation*, J. Wiley and Sons, New York, 2000.

[2] D. Magoni and J.-J. Pansiot, Influence of network topology on protocol simulation, *IEEE ICN'01*, 9-13 July 2001, pp. 762-770.

[3] V. Paxson and S. Floyd. Why we don't know how to simulate the Internet, *WSC'97*, Dec. 1997.

[4] Internet Mapping Project. *http://research.lumeta.com/ches/index.html*.

[5] Cooperative association for Internet data analysis. The CAIDA website. *http://www.caida.org*.

[6] Rocketfuel. *http://www.cs.washington.edu/research/networking/rocketfuel/*.

[7] M. Faloutsos, P. Faloutsos, and C. Faloutsos. On power-law relationships of the Internet topology, *ACM SIGCOMM Computer Communications Review*, Vol. 29, 1999, pp. 251-263.

[8] P. Erdős and A. Rényi. On random graphs I. *Publicationes Mathematicae Debrecen*, Vol. 6, 1959, pp. 290-297.

[9] P. Erdős and A. Rényi. On the evolution of random graphs. *Public. Math. Inst. Hungar. Acad. Sci.*, Vol. 5, 1960, pp. 17-61.

[10] B. Bollobás. *Random Graphs*. Cambridge University Press, Cambridge (UK), Second edition, 2001.

[11] B. M. Waxman. Routing in multipoint connections, *IEEE Journal of Selected Areas in Communications*, Vol. 6, 1988, pp. 1617-1622.

[12] A.-L. Barabási and R. Albert. Emergence of scaling in random networks. *Science*, Vol. 286, 1999, pp. 509-511.

[13] W. Aiello, F. Chung, and L. Lu. A random graph model for power law graphs. *Experimental Mathematics*, Vol. 10, 2001, pp. 53-66.

[14] R. Albert and A.-L. Barabási. Statistical mechanics of complex networks. *Reviews of Modern Physics*, Vol. 74, 2002, pp. 47-97.

[15] R. Pastor-Satorras and A. Vespignani. *Evolution and Structure of the Internet*. Cambridge University Press, Cambridge (UK), 2004.

[16] BRITE, The Boston university Representative Internet Topology gEnerator. *http://www.cs.bu.edu/brite/*.

[17] E. W. Zegura, K. L. Calvert, and M. J. Donahoo. A quantitative comparison of graph-based models for Internet topology. *IEEE/ACM Transactions on Networking*, Vol. 5, No. 6, 1997, pp. 770-783.

[18] C. Jin, Q. Chen, and S. Jamin. INET: Internet topology generator. Technical report, University of Michigan, 2000.

[19] E. N. Gilbert. Random graphs, *Annals of Mathematical Statistics*, Vol. 30, 1959, pp. 1141-1144.

[20] B. Bollobás. Degree sequences of random graphs, *Discrete Mathematics*, Vol. 33, 1981, pp. 1-19.

[21] S. Milgram. The small world problem, *Psychology Today*, Vol. 2, 1967, pp. 60-67.

[22] P. Van Mieghem. Paths in simple random graph and the Waxman graph, *Probability in the Engineering and Informational Sciences*, Vol. 15, 2001, pp. 535-555.

[23] H. A. Simon. On a class of skew distribution functions. *Biometrika*, Vol. 42, 1955, pp. 425-440.

[24] B. Bollobás and O. M. Riordan. The diameter of a scale-free random graph. *Combinatorica*, Vol. 24, 2004, pp. 5-34.

[25] H. Tangmunarunkit, R. Govindan, S. Shenker, D. Estrin, The impact of routing policy on Internet paths. In *IEEE INFOCOM'01*, Vol. 2, April 22-26 2001, pp. 736-742.

[26] W. Jiang and H. Schulzrinne, Modeling of Packet Loss and Delay and Their Effect on Real-time Multimedia Service Quality, In *ACM NOSSDAV*, Chaper Hill (North Carolina), June 2000.

[27] Y. Zhang, N. Duffield , V. Paxson, S. Shenker. On the constancy of the Internet path properties. In *ACM SIGCOMM Internet Measurement Workshop*, 2001.

[28] NS-2, the network simulator (ver. 2), UCB/LBNL/VINT.*http://www.isi.edu/nsnam/ns*