Evolution of the Internet topology and traffic dynamics of data packets

K.-I. Goh, B. Kahng*, and D. Kim

School of Physics and Center for Theoretical Physics, Seoul National University NS50, Seoul 151-747, Korea

Abstract. We investigate how the connection profile of the Internet at the autonomous systems (ASes) level evolves, finding that it is extremely dynamic and can be described in the framework of the multiplicative stochastic process. The fluctuations arising in the process of diversifying connections of each node play an essential role in forming the *status quo* of the Internet. Extracting relevant parameters for the growth dynamics of the Internet topology, we are able to predict the connectivity (degree) exponent $\gamma \approx 2.1$ of the Internet AS map successfully. However, the degree-degree correlation between the two nodes located at the ends of a given link is nontrivial. We introduce a toy model to illustrate the correlation, which contains the mechanism of adaptive rewiring of links during the evolution. While the adaptation model is successful for improving the stochastic model towards the real Internet map, it is not complete to describe the degree-degree correlation yet. We also introduce a quantity called the load as the capacity of node needed for handling the communication traffic and study its distribution over the Internet across years. The load distribution follows a power law with the exponent $\delta \approx 2.0$ and the load at the hub scales with the network size as $\sim N^{1.8}$.

1 Introduction

In recent years, the Internet has become one of the most influential media in our daily life, going beyond in its role as the basic infrastructure in the technological world. Explosive growth in the number of users and hence the amount of traffic poses a number of problems which are not only important in practice for, e.g., maintaining it free from any undesired congestion and malfunctioning, but also of theoretical interests as an interdisciplinary topic [1]. Such interests, also stimulated by other disciplines like biology, sociology, and statistical physics, have blossomed into a broader framework of network science [2–4]. The Internet is a primary example of complex networks. It consists of a large number of very heterogeneous units interconnected with various connection bandwidths, however, it is neither regular nor completely random. In their landmark paper, Faloutsos et al. [5] showed that the Internet at the autonomous systems (ASes) level is a scale-free (SF) network [6], meaning that the node degree k, the number of connections a node has, follows a power-law distribution,

$$p_d(k) \sim k^{-\gamma}, \tag{1}$$

*Corresponding author (E-mail: `kahng@phya.snu.ac.kr`).

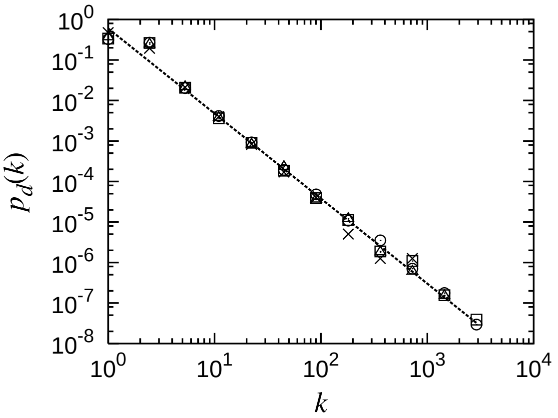

Fig. 1. The degree distribution of the Internet at the autonomous systems level. The data are sampled in January each year, 1998 ($\times$), 2000 ($\triangle$), 2002 ($\square$), and 2004 ($\circ$), from the Oregon route views project. The data are logarithmically binned, and the dashed straight line has the slope -2.1, drawn for the eye.

in node degree k, the number of connections a node has. The degree exponent γ is subsequently measured and confirmed in a number of studies to be $\gamma \approx 2.1$ (Fig.1). The power-law degree distribution implies presence of a few nodes having a hugh number of connections, called hubs, while most other nodes have a few number of connections.

Emergence of such a heterogeneous degree distribution calls for explanation and understanding of the basic mechanism underlying the growth of the Internet. Once revealed, it can be used to predict what the Internet will be like in the future, as well as how it has evolved into the present shape. In this manuscript, we will address this issue, showing that it can be described by a simple toy model based on the multiplicative stochastic process. By extracting relevant parameters for the stochastic process from the time history of the AS map archived in the Oregon route views project, we can predict the degree exponent of the Internet accurately.

While the power-law degree distribution is intriguing, it gives only a global snapshot of the Internet structure. The degree-degree correlation function is a quantity which gives more detailed local information of the Internet structure. In this manuscript, we also study the mean degree of the nearest neighbors of nodes with degree k, denoted as $\langle k_{\rm nn} \rangle(k)$, associated with the degree-degree correlation. We show that the behavior of the mean degree function of the Internet is nontrivial. The behavior of the mean-degree function cannot be explained by the toy model with the multiplicative stochastic process only. We improve the toy model by introducing the rule of adaptive rewirings of links. The adaptation model can also generate the clustering coefficient as large as the one obtained from the real-world Internet. However, the adaptation model is not complete yet to reproduce the joint probability $P(k, k')$ that two nodes with degree k and k' are connected.

The Internet is not a quiet object. Data packets are sent and received over it constantly, causing momentary local congestion from time to time. To avoid such undesired congestion, the capacity, or the bandwidth, of the routers should be as large as it can handle the traffic. We will introduce a rough measure of such capacity, called the load and denoted as ℓ. The distribution of the load reflects the high level of heterogeneity of the Internet: It also follows a power law,

$$p_l(\ell) \sim \ell^{-\delta}, \tag{2}$$

with the load exponent $\delta \approx 2.0$. We will discuss the implication of the power-law behavior of the load distribution.

2 Internet evolution as a multiplicative stochastic process

The mechanism of the emergence of SF network is mostly captured by the Barabási-Albert (BA) model [7] which assumes the linear growth in numbers of nodes and links in time and the preferential attachment (PA) in establishing links from a new node to other previously existing ones. The PA means that the more links a node already has, the more likely it is to get another from a new node. Precisely, the probability $\Pi_i(t)$ that a node i will receive a link from the new node created at time t is linearly proportional to its present degree $k_i(t)$, i.e.,

$$\Pi_i(t) = \frac{k_i(t)}{\sum_j k_j(t)}. \tag{3}$$

The degree exponent $\gamma = 3$ for the BA model. In the generalized version which can tune γ continuously, k_i in Eq.(3) is replaced by $k_i - a$, with a a constant related to γ by $a = \langle k \rangle (3 - \gamma)/2$, $\langle k \rangle$ being the mean degree. The empirical evidence of the PA in the Internet has been reported [8, 9]. As we will see, however, the assumption that the numbers of nodes and links increase linearly in time does not apply to the real situation of the Internet. Rather, the numbers of nodes and links increase exponentially but with different rates. Furthermore, the interconnections between nodes are being updated continually in the Internet, which was not incorporated in the original BA model.

Huberman and Adamic (HA) [10] proposed another scenario for SF networks. They argued that the fluctuation effect arising in the process of connecting and disconnecting links between nodes is an essential feature to describe the dynamics of the Internet topology correctly. In the HA model, the total number of nodes $N(t)$ increases exponentially with time as

$$N(t) = N(0)\exp(\alpha t). \tag{4}$$

Next, they assumed that the degree k_i at a node i evolves through the multiplicative stochastic process,

$$k_i(t+1) = k_i(t)[1 + \zeta_i(t+1)], \tag{5}$$

where $\zeta_i(t)$ is the growth rate of the degree k_i at time t, which fluctuates from time to time. Thus, one may divide the growth rate $\zeta_i(t)$ into two parts,

$$\zeta_i(t) = g_{0,i} + \xi_i(t), \tag{6}$$

where $g_{0,i}$ is the mean value over time, and $\xi_i(t)$ the rest part, representing fluctuations over time. $\xi_i(t)$ is assumed to be a white noise satisfying $\langle \xi_i(t) \rangle = 0$ and $\langle \xi_i(t)\xi_j(t') \rangle = \sigma_{0,i}^2 \delta_{t,t'} \delta_{i,j}$, where $\sigma_{0,i}^2$ is the variance. Here $\langle \cdots \rangle$ is the sample average and $\delta_{a,b}$ is the Kronecker delta symbol. For later convenience, we denote the logarithm of the growth factor as $G_i(t) \equiv \ln[1 + \zeta_i(t)]$. Then a simple application of the central limit theorem ensures that the probability distribution of $k_i(t)/k_i(t_0)$, t_0 being a reference time, follows the log-normal distribution for sufficiently large t. To get the degree distribution, one needs to collect all contributions from different ages τ_i, growth rates $g_{0,i}$, standard deviations $\sigma_{0,i}$ and initial degree $k_i(t_0)$. HA further assumed that ζ_i are identically distributed so that $g_{0,i} = g_0$ and $\sigma_{0,i} = \sigma_0$ for all i. Then the conditional probability for degree, $P_d(k, \tau | \, k_0)$, that $k_i(t_0 + \tau) = k$, given $k_i(t_0) = k_0$ is given by

$$P_d(k, \tau | \, k_0) = \frac{1}{k\sqrt{2\pi\sigma_{\text{eff}}^2 \tau}} \exp\left\{ -\frac{\left(\ln\left(k/k_0\right) - g_{\text{eff}}\tau\right)^2}{2\sigma_{\text{eff}}^2 \tau} \right\}, \tag{7}$$

where $g_{\text{eff}} \equiv \langle G_i(t) \rangle$ and $\sigma_{\text{eff}}^2 \equiv \left\langle (G_i(t) - \langle G_i(t) \rangle)^2 \right\rangle$. g_{eff} and σ_{eff}^2 are related to g_0 and σ_0^2 as $g_{\text{eff}} \approx g_0 - \sigma_0^2/2$ and $\sigma_{\text{eff}}^2 \approx \sigma_0^2$, respectively [11]. Since the density of nodes with age τ is proportional to $\rho(\tau) \sim \exp(-\alpha\tau)$, the degree distribution collected over all ages becomes

$$p_d(k) = \int d\tau \rho(\tau) P_d(k, \tau | \, k_0). \tag{8}$$

The integration is evaluated by the steepest descent method in the limit $k/k_0 \to \infty$ to give $p_d(k) \sim k^{-\gamma}$ with

$$\gamma = 1 - \frac{g_{\text{eff}}}{\sigma_{\text{eff}}^2} + \frac{\sqrt{g_{\text{eff}}^2 + 2\alpha\sigma_{\text{eff}}^2}}{\sigma_{\text{eff}}^2}. \tag{9}$$

Therefore, it is instructive to know the effective values of g_{eff} and σ_{eff} to determine the degree exponent. In the next section, we will measure such parameters from the real evolutionary history of the Internet AS map and check if the HA scenario holds.

3 Growth dynamics of the Internet

A number of projects exists aiming to map the world-wide topology of the Internet. One such is the Route Views project initiated at the University of Oregon [12], the data of which were also archived at the National Laboratory of Applied Network Research (NLANR) [13]. Among the daily data from November 1997 to January 2000, we sample one AS map a month, with the total period of 26 months, and analyze them for various quantities. First we measure the growth rate of the number of ASes α. We also measure directly the growth rate of the number of links β, which can be crosschecked for consistency later.

In Fig. 2(a) and (b), we show the total number of ASes $N(t)$ and the total number of links $L(t)$ as a function of time t. The straight line in log-linear plot means $N(t)$ and $L(t)$ indeed grow exponentially. The growth rates are determined to be $\alpha \approx 0.029$ and $\beta \approx 0.034$. We also find that the newly appeared AS would connect to only one

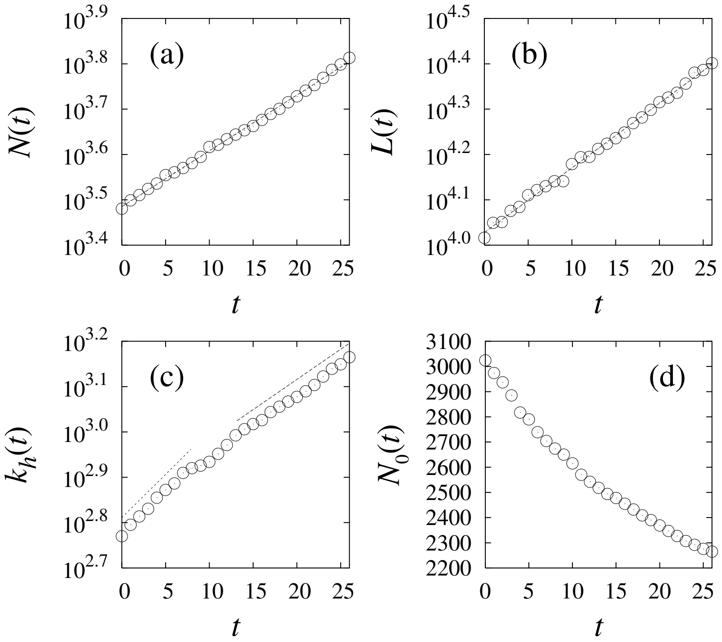

Fig. 2. The time evolution of the number of ASes $N(t)$ (a), the number of links between ASes $L(t)$ (b), the degree of the hub $k_h(t)$ (c), and the number of ASes among those existed at time $t = 0$ that have survived until t, $N_0(t)$ (d). Note that the ordinates in (a), (b) and (c) are in logarithmic scale, indicating the exponential increase of corresponding quantities. The fitted line has a slope 0.029 in (a), 0.034 in (b), and 0.043 (0.030) for the dotted (dashed) one in (c).

or two existing ASes so that the average number of links the new AS establishes is $k_{\text{new}} \approx 1.35$. Fig. 2(c) shows the growth of the degree of the hub, the node with the largest degree. It shows a change of growth rate around $t \approx 14$. $N_0(t)$, the number of ASes among those existed at time $t = 0$ that have survived until t, is also shown in Fig. 2(d).

The measurement of g_0 and σ_0 is nontrivial due to the presence of large fluctuations. To this end, we measure the degree growth rate of a node i, $G_i(t)$, defined earlier as $G_i(t) \equiv \ln[1 + \zeta_i(t)] = \ln[k_i(t)/k_i(t-1)]$. To keep $G_i(t)$ well-defined for all t, we consider only the nodes existing for the entire time range $0 \le t \le 26$, the set composed of which is denoted by S hereafter. By the existence of a node we mean that its degree is nonzero, since one cannot identify an AS with no connection. For each i ($i \in S$), let $g_i = \langle G_i(t) \rangle_t$ and $\sigma_i^2 = \left\langle (G_i(t) - g_i)^2 \right\rangle_t$, where $\langle \cdots \rangle_t$ means the temporal average over the period $16 < t \le 26$ ($T = 10$). If the HA scenario holds, the histogram of $\{g_i\}$ for all nodes would follow the Gaussian distribution with the mean g_{eff} and the variance σ_{eff}^2/T. We show such histogram in Fig. 3, the fit of which to the Gaussian gives the mean $\overline{g}$ as 0.016 and the standard deviation σ_g as 0.04. The measured values of $\{\sigma_i^2\}$ give the mean value $\overline{\sigma^2} \approx 0.017$.

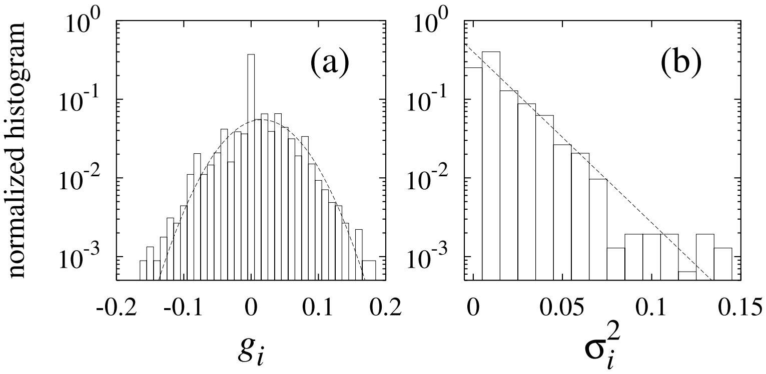

Fig. 3. The normalized histogram of g_i (a) and σ_i^2 (b). In (a), the data is fitted with a Gaussian with the mean 0.016 and the standard deviation 0.04. In (b), the data is fitted with an exponential decay $\exp(-x/x_c)$ with the characteristic scale $x_c \approx 0.02$. The measured value of the average is $\overline{\sigma^2} \approx 0.017$.

It is most likely that $\overline{g}$ and $\overline{\sigma^2}$ would have a distribution over nodes. As HA assumed, however, we try to approximate the growth process by a single process whose effective mean growth rate and standard deviation are $g_{\rm eff}$ and $\sigma_{\rm eff}$, respectively. Then Eq. (7) should hold for all i and all t with a suitable choice of those parameters. For this purpose, we consider the distribution $P[k_i(t)/k_i(t_0)]$ in terms of the scaled variables x and y defined as

$$x \equiv \frac{\ln[k_i(t)/k_i(t_0)] - g_d(t-t_0)}{\sqrt{2\sigma_d^2(t-t_0)}}, \tag{10}$$

and

$$y \equiv P[k_i(t)/k_i(t_0)][k_i(t)/k_i(t_0)]\sqrt{2\pi\sigma_d^2(t-t_0)}, \tag{11}$$

where we set $t_0 = 0$ and g_d and σ_d are parameters to be chosen. From Eq. (7), with suitably chosen parameters g_d and σ_d, the distribution for different time t would collapse onto a single curve, $\ln y = -x^2$. We show such data in Fig. 4. While the data collapse breaks down for large x, the best collapse can be accomplished by choosing the parameters $g_d = 0.016$ and $\sigma_d = 0.14$, which should be identified with $g_{\rm eff}$ and $\sigma_{\rm eff}$, respectively. The effective growth parameters found in this way are in accordance with the ones estimated before as $\overline{g} = 0.016$ and $\overline{\sigma^2} = 0.017$. As noted earlier, the consistency of estimated parameters can be checked as, for example, it should satisfy $\beta = \max(\alpha, g_{\rm eff} + \sigma_{\rm eff}^2)$, for which we have $\beta = 0.034$ and $g_{\rm eff} + \sigma_{\rm eff}^2 = 0.035$, being reasonably consistent with each other. Thus we conclude the parameters $g_{\rm eff} = 0.016$ and $\sigma_{\rm eff} = 0.14$ can be regarded as the effective parameters of the degree growth dynamics of the Internet AS map as a single process. Applying those values together with

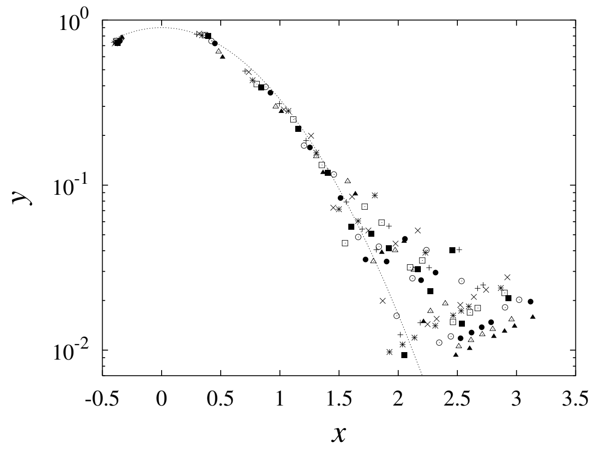

Fig. 4. Plot of $P[k_i(t)/k_i(t_0)]$ versus $k_i(t)/k_i(t_0)$ for different times $t > 16$ in terms of the scaled variables y and x defined in Eqs. (10) and (11). Larger deviations for large x indicate the drawback of the HA approach.

$\alpha = 0.029$ found earlier into Eq. (10), we found $\gamma \approx 2.1$, which is in excellent agreement with the directly measured one shown in Fig. 1.

4 Degree correlation and modular structure

It is known that the degrees of the two nodes located at the ends of a link are correlated each other. As the first step, the degree-degree correlation can be quantified in terms of the mean degree of the neighbors of a given node with degree k as a function of k, denoted by $\langle k_{\rm nn}\rangle(k)$ [8]. Then the mean degree function $\langle k_{\rm nn}\rangle(k)$ can be written in terms of the conditional probability $P(k'|k)$ that given a node with degree k is present, another node with degree k' is connected to it. That is $\langle k_{\rm nn}\rangle(k) = \sum_{k'} k'P(k'|k)$. The conditional probability is then expressed in terms of the joint probability, $P(k',k)$, as $P(k'|k) = P(k',k)/kp_d(k)$. Meanwhile, for the BA model, the joint probability that nodes of degree k' (ancestor) and k (descendent) are connected scales as [14]

$$P(k',k) \sim k'^{-(\gamma-1)}k^{-2}. \tag{12}$$

Then the function $\langle k_{\rm nn}\rangle(k)$ is written as [14, 3]

$$\begin{aligned}\langle k_{\rm nn}\rangle(k) &= \sum_{k'} k'P(k'|k)\\ &= \sum_{k'} \frac{k'P(k',k)}{kp_d(k)}\\ &\sim \sum_{k'} \frac{k'k'^{-(\gamma-1)}k^{-2}}{k^{1-\gamma}}\\ &\sim k^{-3+\gamma}.\end{aligned} \tag{13}$$

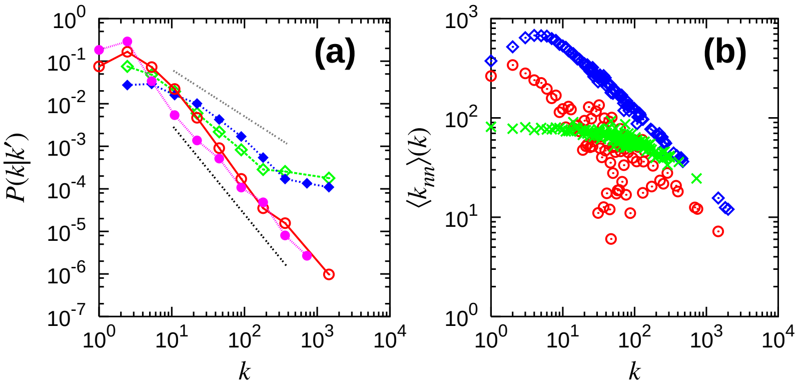

Fig. 5. (a) The conditional probability $P(k|k')$ for the dangling nodes with $k' = 1$ (diamonds) and the hub (circles). Data points with filled (open) symbols are from the Internet AS as of January, 2000 (the simulation of the adaptation model). The dashed (dotted) straight line has a slope -2.1 (-1.1), drawn for a guide to the eye. **(b)** The average degree of nearest neighbors of a node whose degree is k, $\langle k_{\mathrm{nn}}\rangle(k) = \sum_{k'} k' P(k'|k)$, as a function of k from the adaptation model (diamond), HA model (cross) and the Internet (circle).

More generally, the behavior of the average neighbor degree function is expressed by another power law as

$$\langle k_{\mathrm{nn}}\rangle(k) \sim k^{-\nu}. \tag{14}$$

For the Internet, if Eq.(12) holds, $\langle k_{\mathrm{nn}}\rangle(k)$ would decay as $\sim k^{-0.9}$, i.e., $\nu \approx 0.9$. However, it decays with $\nu \approx 0.5$ for the real-world Internet [8], suggesting that the joint probability does not behave in the way of the BA model does.

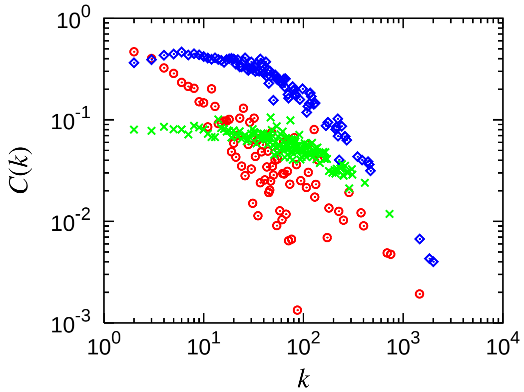

Fig. 6. The clustering function $C(k)$ for the Internet AS as of January, 2000 ($\circ$), HA model ($\times$), and the adaptation model ($\diamond$).

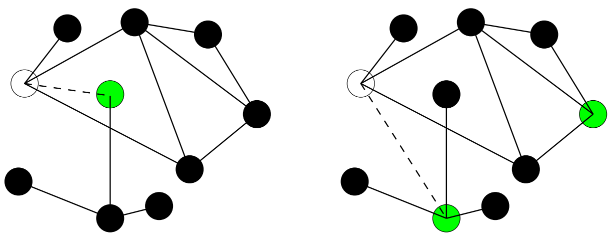

Fig. 7. Shown is the adaptive rewiring rule. A node (white) detaches one of its links from a node (green or gray) in (a), and attaches it to one of the nodes (green or gray) with degree 3, larger than 2 of the detached node, in (b).

On the other hand, the degree-degree correlation can also be described in terms of the assortativity coefficient introduced by Newman [15], which is defined as

$$r = \frac{\langle k_1 k_2 \rangle - \langle (k_1 + k_2)/2 \rangle^2}{\langle (k_1^2 + k_2^2)/2 \rangle - \langle (k_1 + k_2)/2 \rangle^2}, \tag{15}$$

where k_1 and k_2 are the degree of two end nodes, respectively, of a link, and $\langle \cdots \rangle$ denotes the average over all links. It is nothing but the Pearson correlation coefficient for the degrees of two end nodes over all links, normalized -1 to 1. r is negative when the function $\langle k_{\rm nn} \rangle(k)$ exhibits decreasing behavior like the case of the Internet, while it is positive when the function increases with k as occurring in social networks. In fact, the assortativity coefficient r was introduced to characterize social networks. We find the assortativity coefficient of the Internet to be $r = -0.18$ as of Jan. 2001.

The Internet has modules within it. Such modular structures arise due to regional control systems, and often form in a hierarchical way [16]. Recently, it was argued that such modular and hierarchical structures can be described in terms of the clustering coefficient. Let C_i be the local clustering coefficient of a node i, defined as $C_i = 2e_i/k_i(k_i - 1)$, where e_i is the number of links present among the neighbors of node i, out of its maximum possible number $k_i(k_i - 1)/2$. The clustering coefficient of a network, C, is the average of C_i over all nodes. $C(k)$ means the clustering function of a node with degree k, i.e., C_i averaged over nodes with degree k. When a network is modular and hierarchical, the clustering function follows a power law,

$$C(k) \sim k^{-\beta}, \tag{16}$$

for large k, and C is independent of system size N, while C depends on N as $C \sim N^{-1}$ for the BA model [17, 18].

For the Internet, the clustering coefficient is measured to be $C_{\rm AS} \approx 0.25$, which is different from the value obtained from the HA model, $C_{\rm HA} \approx 0.01(1)$. The clustering function $C_{\rm AS}(k)$ also behaves differently from what the HA model predicts as shown in Fig. 6.

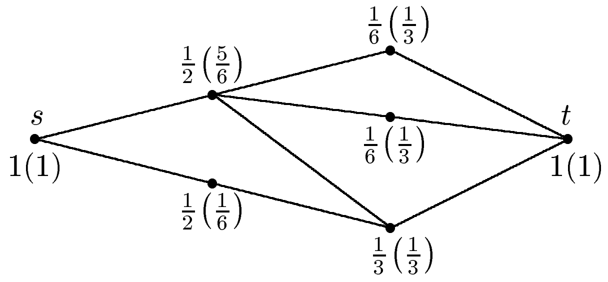

Fig. 8. The load at each node due to a unit packet transfer from the node s to the node t, $\ell_i^{s\to t}$. In this diagram, only the nodes along the shortest paths between (s,t) are shown. The quantity in parentheses is the corresponding value of the load due to the packet from t to s, $\ell_i^{t\to s}$.

5 Adaptation model

Using the measured values, α, $g_{\rm eff}$, and $\sigma_{\rm eff}$, and modifying the HA idea, we construct a stochastic model evolving through the following four rules:

- (i) **Geometrical growth**: At time t, geometrically increased number of new nodes, $\alpha N(t-1)$, are introduced in the system, and following the fact $\langle k_{\rm new}\rangle_t \approx 1.34$, each of them connects to one or two existing nodes according to the PA rule.
- (ii) **Accelerated growth**: Each existing node increases its degree by the factor $g_0 \approx g_{\rm eff} + \sigma_{\rm eff}^2/2$. These new internal links are also connected following the PA rule.
- (iii) **Fluctuations**: Each node disconnects existing links randomly (resp. connects new links following the PA rule) when the noise ($\xi_i(t)$ in Eq.(4)) is chosen to be negative (resp. positive). The variance of the noise is given as $\sigma_0^2 \approx \sigma_{\rm eff}^2$.
- (iv) **Adaptation**: When connecting in step (iii), the PA rule is applied only within the subset of the existing nodes consisting of those having more degree than the one previously disconnected. This last constraint accounts for the adaptation process in our model, which is the distinct feature from the HA model. The adaptive rewiring rule is depicted in Fig. 7.

Through this adaptation process, the model can reproduce generic features of the Internet topologies which are as follows. First, the degree exponent is measured to be $\gamma_{\rm model} \approx 2.2$, close to the empirical result $\gamma_{\rm AS} \approx 2.2(1)$. Second, the clustering coefficient is measured to be $C_{\rm model} \approx 0.15(7)$, comparable to the empirical value $C_{\rm AS} \approx 0.25$. Note that without the adaptation rule, we only get $C \approx 0.01(1)$. The clustering function $C(k)$ also behaves in a similar pattern to that of the real-world Internet shown in Fig. 6, that is, decaying in a similar fashion for large k, but the overall curve shifts upward and the constant behavior for small k appears. Third, the mean degree function $\langle k_{\rm nn}\rangle(k)$ also behaves similarly to that of the real-world networks as shown in Fig. 5, but it also shifts upward overall. In short, the behaviors of $C(k)$ and $\langle k_{\rm nn}\rangle(k)$ of the adaptation model are much closer to those of the real Internet AS map than those of the HA model. However, the discrepancies are still significant.

6 Load distribution of the Internet

To a large extent, the Internet is the medium of communication. The continuous communication between hosts generates certain amount of data traffic. To make the best use of it, we have to avoid congestions, from which we suffer possible delays and the loss of information. What's worse, one doesn't know when and how much a host will generate the traffic. Absent is the central regulation in the Internet, hence each node should do its best for its own ends. To give a measure of such activity, we define the load ℓ_i as the amount of capacity or bandwidth that a node i can handle in unit time [19]. Not knowing the level of traffic, one assumes that every node sends a unit packet to everyone else in unit time. One further assumes that the packets are transfered from the source to the target only along the shortest paths between them, and divided evenly upon encountering any branching point. To be precise, let $\ell_i^{s\to t}$ be the amount of packet sent from s (source) to t (target) that passes through the node i (see Fig. 8). Then the load of a node i, ℓ_i, is the accumulated sum of $\ell_i^{s\to t}$ for all s and t, $\ell_i = \sum_{s\neq t} \ell_i^{s\to t}$. In other words, the load of a node i gives us the information how much the capacity of the node should be in order to maintain the whole network in a free-flow state. However, due to local fluctuation effect of the concentration of data packets, the traffic could be congested even for the capacity of each node being taken as its load. To calculate load on each node, we use the modified version of the breath-first search algorithm introduced by Newman [20] and independently by Brandes [21], which can evaluate $\{\ell_i\}$ in time of order $\mathcal{O}(N^2)$ for sparse binary graphs. The algorithm works as follows:

(1) The shortest paths to a node i from every other node are identified by using the breadth-first search and each node j is labeled by the distance d_{ij} from the source node i.
(2) A variable f_j preset to the initial value 1 except for the source node i is assigned to each node j. This will account for the contribution to load due to a unit packet fired from the node i to j. For the source node i, f_i is initially preset to 0 since we do not consider a packet transfer to itself.
(3) Going through the nodes j in descending order of their distance from i, the value of f_j is added to the corresponding variable on the predecessor node of j. Here predecessors of a node j are the nodes connected to j, being closer to the source node i than j. If j has more than one predecessor, then f_j is divided equally by the number of predecessors and each fraction is passed to each predecessor.
(4) When we have completed all nodes in this way, the variable f_j represents the number of packets (out of total $N-1$) that run through the node j. In our convention, the value f_j includes the contributions of the end points of each path. To get the load for all paths, f_j's are added to the load ℓ_j and the entire calculation is repeated for each of the N possible source node i.

For a number of SF networks in nature and society, the load distribution is also found to follow a power law, Eq. (2) [22]. The Internet AS map is of no exception and the load exponent δ of the power law is estimated to be approximately $\delta \approx 2.0$ [22, 23]. The power-law load distribution means that a few ASes should handle an extraordinarily large amount of load while most others may do only a little.

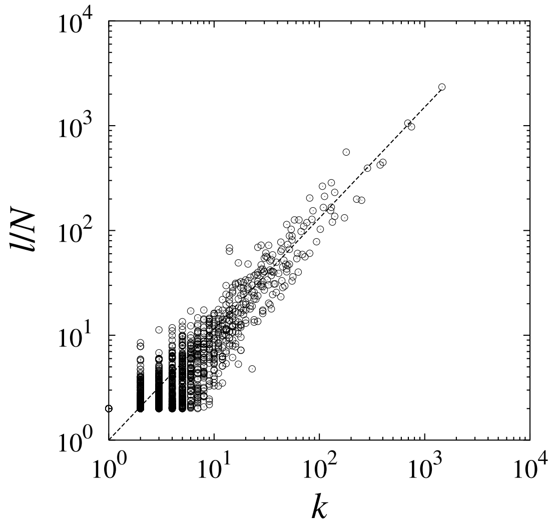

Fig. 9. The scatter plot of the load versus the degree of each node for the AS map as of January 2000. The slope of the dashed line is 1.06, drawn for the eye.

The load of a node is highly correlated with its degree. The Pearson correlation coefficient between the two quantities is as high as 0.98. This suggests a scaling relation between the load and the degree of a node as

$$\ell \sim k^{\eta} \tag{17}$$

and the scaling exponent η is estimated as $\eta = 1.06 \pm 0.03$ for January 2000 AS map (Fig. 9). In fact, if one assumes that the ranks of each node for the degree and the load are the same, then one can show that the exponent η depends on γ and δ as $\eta = (\gamma - 1)/(\delta - 1)$ with $\gamma \approx 2.1$ and $\delta \approx 2.0$, and we have $\eta \approx 1,1$, which is consistent with the direct measurement.

Since data packets are normally transferred along the shortest pathway between two nodes in the Internet, it would be interesting to investigate the generic topological feature of the shortest pathways between them. Along the shortest pathways, we count the total number of nodes $\mathcal{M}(d)$ lying on the shortest pathways separated by the distance d, averaged over all pairs of nodes separated by the same distance d. Adopting from the fractal theory, $\mathcal{M}(d)$ is called the "mass-distance" relation. We find that the mass, the total number of nodes lying on the pathways, increases linearly with increasing d, but the increasing rate is roughly 4 larger than 1. If the shortest pathway is singly connected, then the proportionality coefficient would be 1. The fact that it increases linearly in d means that the shortest pathway topology is effectively tree, but that its coefficient is significantly larger than 1. That means that there are many alternative pathways as shown in Fig. 10(b).

The time evolution of the load at each AS is also of interest. Practically, how the load scales with the total number of AS (the size of the AS map) is an important information

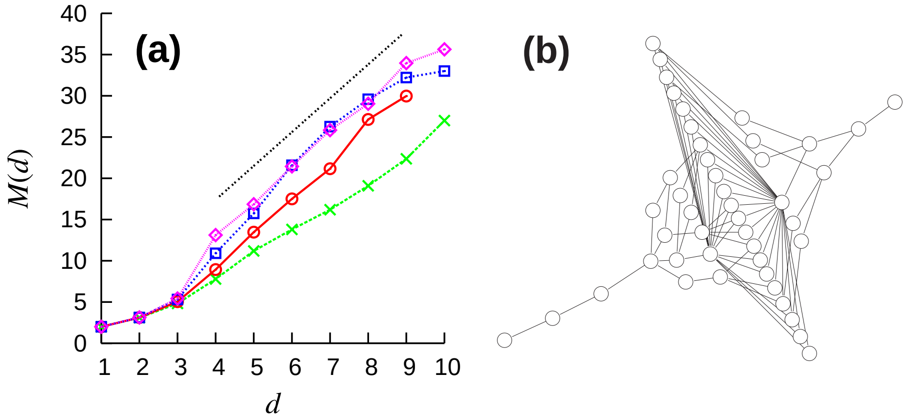

Fig. 10. The mass-distance relation $\mathcal{M}(d)$ for the Internet in the years 1998 ($\times$), 2000 ($\circ$), 2002 ($\square$), and 2004 ($\diamond$) (a) and the snapshot of the shortest pathways of length 10 as of January 2000 (b)

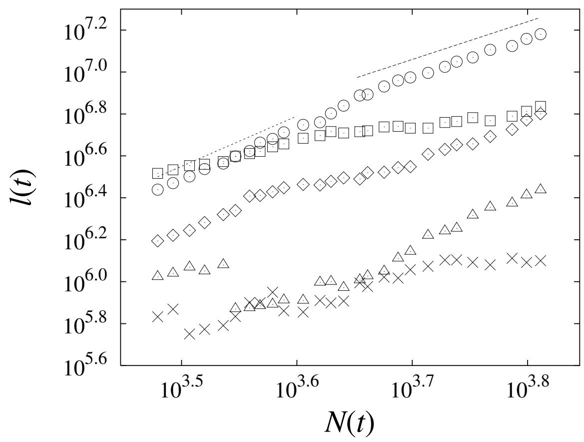

Fig. 11. Time evolution of the load versus $N(t)$ at the ASes of degree-rank 1($\bigcirc$), 2 ($\square$), 3 ($\Diamond$), 4 ($\triangle$), and 5 ($\times$). The dashed line for larger N has slope 1.8, drawn for the eye.

for the network management. In Fig. 11, we show $\ell_i(t)$ versus $N(t)$ for 5 ASes with the highest rank in degree, i.e., 5 ASes that have largest degrees at $t = 0$. The data of $\{\ell_i(t)\}$ shows large fluctuations in time. Interestingly, the fluctuation is moderate for the hub, implying that the connections of the hub is rather stable. The load at the hub is found to scale with $N(t)$ as $\ell_h(t) \sim N(t)^{\mu}$, but the scaling shows a crossover from $\mu \approx 2.4$ to $\mu \approx 1.8$ around $t \approx 14$, as it did for the degree.

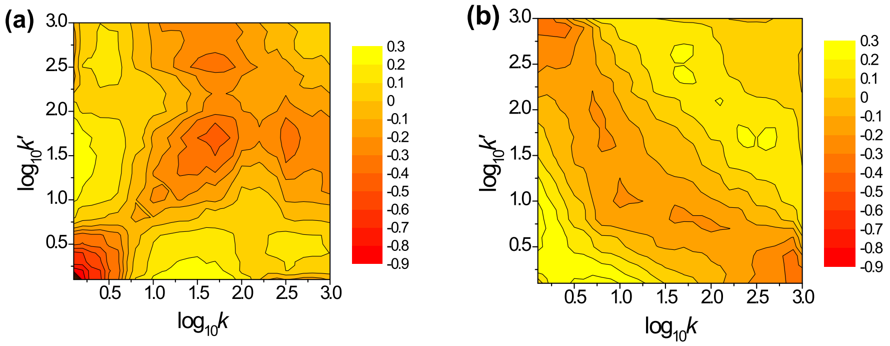

Fig. 12. The plot of the joint probabilities $P(k', k)$ for the Internet as of January, 2000 (a) and the adaptation model (b). The concentration denotes the value of $\log_{10}[P(k,k')/P_{\rm random}(k,k')]$, where $P_{\rm random}(k,k')$ is the joint probability of the randomized networks generated for the purpose of reference. They are generated by the switching method [24] conserving the degree sequence.

7 Summary

We have studied the temporal evolution of the Internet AS map and showed that it can be described in the framework of multiplicative stochastic process for the degree growth dynamics. We measured the values of relevant parameters from the history of the AS map monitored by the Oregon Route Views project. With those values, the AS number growth rate $\alpha = 0.029$, the effective degree growth rate $g_{\rm eff} = 0.019$, and its effective standard deviation $\sigma_{\rm eff} = 0.14$ are obtained. Using these values, we were able to predict the degree exponent γ as $\gamma \approx 2.1$, which is in excellent agreement with the previously reported empirical values $\gamma_{\rm measured} = 2.1 \sim 2.2$. We also studied the degree-degree correlation of the Internet in terms of the quantity $\langle k_{\rm nn}\rangle(k)$, the mean degree of the neighbors of nodes with degree k and the assortativity coefficient. The modularity of the Internet was also examined by measuring the clustering function $C(k)$. We found that the degree-degree correlation and the modularity cannot be explained by the HA model only with the multiplicative process. We modified the HA model by introducing the adaptive rewiring of links in the process of attaching links. While the adaptation model seems to be successful for reproducing $C(k)$ and $\langle k_{\rm nn}\rangle(k)$ of the Internet AS roughly, it is not satisfactory yet to describe the real world Internet. For example, the joint probability $P(k,k')$ that two nodes with degrees k and k' are connected via an edge cannot be reproduced by the adaptation model as shown in Fig. 12. The concentration profile of the joint probability for the adaptation model shows different pattern compared to that of the real Internet. Thus, further refinement of the adaptation model reflecting the evolution of the Internet map is needed.

In the second part of the paper, we introduced a quantity called the load as the accumulated sum of the fraction of packets that passes through a given node when a unit packet is transmitted between each pair of nodes along the shortest paths between the two. It can be thought of as the amount of traffic that a node should handle to keep the whole network away from the unwelcome congestion and maintain free-flow state, giving a measure of desired capacity of the nodes. The load distribution also follows a power law. The load and the degree of an AS are highly correlated with each other. The analysis of the temporal change of the load reveals that the load at the hub scales with the system size as $N^{1.8}$. Finally, we note that the contents of this article partly overlaps with our previous studies published in [19, 22, 25].

This work is supported by the KOSEF grants No. R14-2002-059-01000-0 in the ABRL program.

References

1. R. Pastor-Satorras and A. Vespignani, *Evolution and Structure of the Internet* (Cambridge University Press, Cambridge, 2004).
2. R. Albert and A.-L. Barabási, Rev. Mod. Phys. **74**, 47 (2002).
3. S. N. Dorogovtsev and J. F. F. Mendes, *Evolution of Networks* (Oxford University Press, Oxford, 2003).
4. M. E. J. Newman, SIAM Rev. **45**, 167 (2003).
5. M. Faloutsos, P. Faloutsos, and C. Faloutsos, Comput. Commun. Rev. **29**, 251 (1999).
6. A.-L. Barabási, R. Albert, and H. Jeong, Physica A **272**, 173 (1999).
7. A.-L. Barabási and R. Albert, Science **286**, 509 (1999).
8. R. Pastor-Satorras, A. Vázquez, and A. Vespignani, Phys. Rev. Lett. **87**, 258701 (2001).
9. H. Jeong, Z. Néda, and A.-L. Barabási, Europhys. Lett. **61**, 567 (2003).
10. B. A. Huberman and L. A. Adamic, e-print (http://arxiv.org/abs/cond-mat/9901071) (1999).
11. C. W. Gardiner, *Handbook of Stochastic Methods* (Springer-Verlag, Berlin, 1983).
12. D. Meyer, *University of Oregon Route Views Archive Project* (http://archive.routeviews.org).
13. The NLANR project sponsored by the National Science Foundation (http://moat.nlanr.net).
14. P. Krapivsky and S. Redner, Phys. Rev. E **63**, 066123 (2001).
15. M. E. J. Newman, Phys. Rev. Lett. **89**, 208701 (2002).
16. K. A. Eriksen, I. Simonsen, S. Maslov, and K. Sneppen, Phys. Rev. Lett. **90**, 148701 (2003).
17. E. Ravasz, A. L. Somera, D. A. Mongru, Z. N. Oltvai, and A.-L. Barabási, Science **297**, 1551 (2002).
18. E. Ravasz and A.-L. Barabási, Phys. Rev. E **67**, 026112 (2003).
19. K.-I. Goh, B. Kahng, and D. Kim, Phys. Rev. Lett. **87**, 278701 (2001).
20. M. E. J. Newman, Phys. Rev. E **64**, 016132 (2001).
21. U. Brandes, J. Math. Sociol. **25**, 163 (2001).

22. K.-I. Goh, E. Oh, H. Jeong, B. Kahng, and D. Kim, Proc. Natl. Acad. Sci. USA **99**, 12583 (2002).
23. A. Vázquez, R. Pastor-Satorras, and A. Vespignani, Phys. Rev. E **65**, 066130 (2002).
24. S. Maslov, K. Sneppen, and A. Zaliznyak, e-print (http://arxiv.org/abs/cond-mat/0205379).
25. K.-I. Goh, B. Kahng, and D. Kim, Phys. Rev. Lett. **88**, 108701 (2002).