

Power Laws and the AS-Level Internet Topology

Georgos Siganos, *Student Member, IEEE*, Michalis Faloutsos, *Member, IEEE*, Petros Faloutsos, and Christos Faloutsos, *Member, IEEE*

Abstract—In this paper, we study and characterize the topology of the Internet at the autonomous system (AS) level. First, we show that the topology can be described efficiently with power laws. The elegance and simplicity of the power laws provide a novel perspective into the seemingly uncontrolled Internet structure. Second, we show that power laws have appeared consistently over the last five years. We also observe that the power laws hold even in the most recent and more complete topology with correlation coefficient above 99% for the degree-based power law. In addition, we study the evolution of the power-law exponents over the five-year interval and observe a variation for the degree-based power law of less than 10%. Third, we provide relationships between the exponents and other topological metrics.

Index Terms—Network modeling, power laws.

I. INTRODUCTION

IN THIS paper, we study the topology of the Internet and we identify several power laws. Furthermore, we discuss multiple benefits from understanding the topology of the Internet. Our work is motivated by questions such as the following: “What does the Internet look like?,” “Are there any topological properties that do not change in time?,” “How will it look a year from now?,” and “How can I generate Internet-like graphs for my simulations?”

Modeling the Internet topology is an important open problem despite the attention it has attracted recently. Paxson and Floyd consider this problem as a major reason why we do not know how to simulate the Internet [21]. An accurate topological model can have significant impact on network research. First, we can design more efficient protocols that take advantage of its topological properties. Second, we can create more accurate artificial models for simulation purposes. Third, we can derive estimates for topological parameters (e.g., the average number of neighbors within h hops) that are useful for the analysis of protocols and for speculations of the Internet topology in the future.

In this paper, we propose the use of power laws to describe the topology of the Internet at the autonomous system (AS) or interdomain level. Power laws are expressions of the form $y \propto x^a$, where a is a constant, x and y are the measures of interest,

Manuscript received March 14, 2001; revised February 21, 2002; approved by IEEE/ACM TRANSACTIONS ON NETWORKING Editor E. Zegura. This research was supported by the Defense Advanced Research Projects Agency (DARPA) under Grant N660001-00-1-8936 and by the National Science Foundation under Career Grant ANIR 9985195.

G. Siganos and M. Faloutsos are with the Department of Computer Science, University of California, Riverside, CA 92521 USA (e-mail: siganos@cs.ucr.edu; michalis@cs.ucr.edu).

P. Faloutsos is with the Department of Computer Science, University of California, Los Angeles, CA 90095 USA (e-mail: pfal@cs.ucla.edu).

C. Faloutsos is with the Department of Computer Science, Carnegie Mellon University, Pittsburgh, PA 15213 USA (e-mail: christos@cs.cmu.edu).

Digital Object Identifier 10.1109/TNET.2003.815300

and \propto stands for “proportional to.” Conceptually, our work has three main thrusts: 1) defining and identifying the power laws; 2) studying their evolution; and 3) relating power-law exponents and other graph metrics. Our work can be summarized in the following points.

First, we identify several power laws that describe the distribution of topological metrics such as node degree. We also show that the three power laws are tightly related theoretically. In addition, we introduce a graph metric to quantify the density of a graph and propose a power-law approximation of that metric.

Second, we study the evolution of the power laws between November 1997 and February 2002. The power laws hold for 1253 instances, with good linear fits in log–log plots; the correlation coefficient of the fit is at least 96% and typically above 98%. We note that their existence is persistent, and they hold even in the most recent and more complete topology [10].

Third, we present new and known relationships between power-law exponents and other graph metrics. We list mechanisms that create power laws and discuss their plausibility and their efficiency in creating graphs for practical purposes.

Our Work in Perspective: Power laws are a first step in understanding the Internet topology. The evidence of their existence is too strong to be dismissed as coincidence. We monitor and analyze the Internet over a period of five years, during which the size of the network quadrupled. The contributing sources for the data collection changed significantly in number and location [41]. Additionally, we analyzed the more recent and complete topology [10]. These observations exclude by and large the possibility of the power laws being the result of coincidence. Therefore, the power laws appear as a necessary though not sufficient condition for a topology to be realistic. There may be more topological properties of the Internet topology that are not captured by our power laws [45], [54].

The rest of this paper is structured as follows. In Section II, we present some definitions and previous work on measurements and models for the Internet. In Section III, we present our Internet instances and provide useful measurements. In Section IV, we present our three observed power laws and our power-law approximation. In Section V, we present the time evolution of the exponent of the power laws that we presented in the previous section. In Section VI, we present the models used to generate power-law graphs. In Section VII, we conclude this paper and discuss future directions.

II. BACKGROUND AND PREVIOUS WORK

The Internet can be decomposed into subnetworks that are under separate administrative authorities. These subnetworks are called *domains* or *autonomous systems*. This way, the topology of the Internet can be studied at two different

TABLE I
DEFINITIONS AND SYMBOLS

| Symbol | Definition |
|-----------|--|
| G | An undirected graph. |
| N | Number of nodes in a graph. |
| E | Number of edges in a graph. |
| δ | The diameter of the graph. |
| d_v | Degree of node v . |
| \bar{d} | Average degree of the nodes of a graph: $\bar{d} = 2 E/N$ |

granularities. At the **router level**, we represent each router by a node [43]. At the **inter-domain level**, each domain is represented by a single node [22] and each edge is an inter-domain interconnection. The study of the topology at both levels is equally important. The Internet community develops and employs different protocols inside a domain and between domains. An intra-domain protocol is limited within a domain, while an inter-domain protocol runs between domains treating each domain as one entity. Here, we focus on the autonomous system level and represent the topology of the Internet by an undirected graph.

Network Analysis Before Power Laws: Before 1999, the metrics that were used to evaluate network models were mainly the node degree and the distances between nodes. Given a graph, the *degree* of a node is defined as the number of edges incident to the node (see Table I). The distance between two nodes is the number of edges along the shortest path between the two nodes. Most studies report minimum, maximum, and average values and plot the degree and distance distribution. We denote the number of nodes of a graph by N , the number of edges by E , and the diameter of the graph by δ . Using these metrics, Govindan and Reddy [22] study the growth of the inter-domain topology of the Internet between 1994 and 1995. The graph is sparse, with 75% of the nodes having degrees less than or equal to two. Pansiot and Grad [43] study the topology of the Internet in 1995 at the router level. The distances they report are approximately two times larger compared with those of Govindan and Reddy.

For graph generation purposes, Waxman introduced a popular network model [55]. The link creation probabilities depend upon the Euclidean distance between the nodes. This model was successful in representing small early networks such as the ARPANET. As the size and the complexity of the network increased more detailed models were needed [8], [16]. Zegura *et al.* [60] reviewed these generation methods using a more expansive set of metrics, including some that are driven by uses, like multicast routing. Based on the limitations they found, they introduced a comprehensive model that includes several previous models.

Power Laws—A Ubiquitous Presence: Pareto was among the first to introduce power laws in 1896 [44]. He used power laws to describe the distribution of income where there are few very rich people, but most of the people have a low income. Another classical law, the Zipf law [61], was introduced in 1949, for the frequencies of the English words and the population of cities. Power laws have been found in numerous diverse fields span-

ning geological, natural, sociological, and biological systems. Some interesting examples of power-law distributions are the movie actor collaboration network [7], the human respiratory system [34], automobile networks [19], the size and location of earthquakes, stock-price fluctuations [6], the web of human sexual contacts [17], biological cellular networks [25], and the scientific citation network [50]. More details about the historical aspects of power laws can be found in Mitzenmacher [38] and an extensive presentation of power laws in many diverse fields in [3].

Network Analysis Using Power Laws: More recently, power laws have been observed in communication networks. First, power laws have been observed in network traffic [13], [30], [46], [56]. In addition, the topology of the World Wide Web [4], [28] can be described by power laws. Furthermore, power laws describe the topology of peer-to-peer networks [39] and properties of multicast trees [12], [47], [37], [57]. Among these properties, the Chuang–Sirbu law states that the size of the multicast tree follows a power law with respect to the number of group members with exponent 0.8.

Our initial work [20] on power laws has generated significant follow-up work. Various researchers verified our observations with different datasets [23], [24], [33]. In addition, significant work has been devoted in understanding the origin [36], and generating power-law topologies [26], [35], [36], [42], [54], [58]. We discuss these approaches for generating power laws in Section VI. More recently, several works have focused on describing the topology in a qualitative way [31], [32], [52], [53].

III. OUR INTERNET INSTANCES

In this section, we present the Internet instances we study in this work. We use topologies from two sources. First, we use the Oregon route views project [41]. The information is collected by a route server from BGP¹ [49] routing tables of multiple geographically distributed BGP routers. This is the *only* archival repository we could find in order to study the evolution of the topology. However, the Oregon data does not identify all possible links between ASs [10]. For this reason, we use a second data set from 2001 [10], which is the superset of Oregon and several other routing repositories. This data is currently considered as the most comprehensive AS topology, although it is almost certainly not complete. Unfortunately, there is a limited number of these instances, which span only nine weeks, starting from March 2001, and thus, it does not lend itself to an evolution study.

The Oregon dataset contains 1253 daily instances. These instances span an interval of 1600 days, more than five years, from November 8, 1997, to February 28, 2002. Note that we filter the data to remove incomplete data files because they do not correctly represent the topology. We identify and remove the instances that have less than 50% of the nodes found in the previous instance. For example, we removed the reported topology on August 29, 1999, which has 103 nodes, since the files on the previous and next days have more than 5600 nodes. Among the 1253 instances, we selected the instance of May 26, 2001,

¹BGP stands for the Border Gateway Protocol, and is the inter-domain routing protocol.

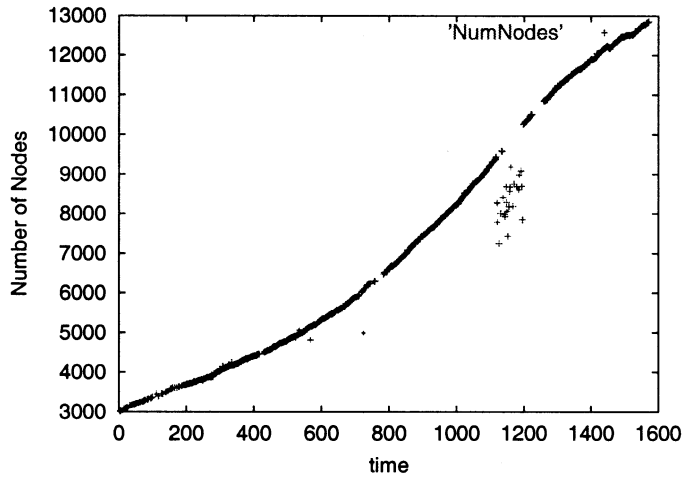


Fig. 1. Growth of the Internet: the number of domains versus time between the end of 1997 and the start of 2002.

to demonstrate the power laws, so that we can compare the results we have with the more complete topology. For the rest of this paper, we will refer to the instance from Oregon as **Oregon**, and the instance which represents the more complete topology as **Multi**.

Note that the remaining 1252 instances also follow the power laws. Furthermore, the size of the topology in the five-year period quadrupled (see Fig. 1). The change is significant, and it ensures that our instances reflect different snapshots of an evolving network.

IV. POWER LAWS OF THE INTERNET

In this section, we observe three power laws of the Internet topology. We propose and measure graph properties, which demonstrate a regularity that is unlikely to be a coincidence. The exponents of the power laws can be used to characterize graphs. In addition, we introduce a graph metric that is tailored to the needs of the complexity analysis of protocols. The metric reflects the density or the connectivity of nodes, and we offer a rough approximation of its value through a power law. Finally, using our observations and metrics, we identify a number of interesting relationships between important graph parameters.

The goal of our work is to find metrics that quantify topological properties and describe concisely skewed data distributions. Previous metrics, such as the average degree, fail to do so. First, metrics that are based on minimum, maximum, and average values are not good descriptors of skewed distributions; they miss a lot of information, probably including the “interesting” part that we would want to capture. Second, the plots of the previous metrics are difficult to quantify, and this makes the comparison of graphs difficult.

To express our power laws, we introduce several graph metrics, given in Table II. We define D_d to be the complementary cumulative distribution function (CCDF) of a degree d , which is the percentage of nodes that have degree greater than the degree d . If we sort the nodes in decreasing degree sequence, we define *rank* r_v to be the index of the node in the sequence, while ties in sorting are broken arbitrarily. We define the number of pairs of nodes $P(h)$ to be the total number of pairs of nodes within less

TABLE II
NOVEL DEFINITIONS AND THEIR SYMBOLS

| Symbol | Definition |
|-----------|---|
| D_d | The Complementary Cumulative Distribution Function or CCDF, of a degree, is the percentage of nodes that have degree greater than the degree d . |
| r_v | The rank of a node, v , is its index in the order of decreasing degree. |
| $P(h)$ | The <i>number of pairs</i> of nodes is the total number of pairs of nodes within less or equal to h hops, including self-pairs, and counting all other pairs twice. |
| $NN(h)$ | The average number of nodes in a neighborhood of h hops. |
| λ | The eigenvalue of an $N \times N$ matrix A : $X : X \in \mathcal{R}^N \setminus \{0\}$ and $AX = \lambda X$. |
| i | The <i>order</i> of λ_i in the sorted sequence $\lambda_1 \geq \lambda_2 \dots \geq \lambda_N$ of the eigenvalues of a matrix. |

than or equal to h hops, including self-pairs, and counting all other pairs twice. The use of this metric will become apparent later. We also define $NN(h)$ to be the average number of nodes in a neighborhood of h hops. Finally, we recall the definition of the eigenvalues of a graph, which are the eigenvalues of its adjacency matrix.

We use linear regression to fit a line in a set of two-dimensional points [48]. The technique is based on the least-square errors method. The validity of the approximation is indicated by the correlation coefficient which is a number between -1.0 and 1.0 . For the rest of this paper, we use the absolute value of the correlation coefficient (ACC). An ACC value of 1.0 indicates perfect linear correlation, i.e., the data points are exactly on a line.

A. Rank Exponent \mathcal{R}

In this section, we study the degrees of the nodes. We sort the nodes in decreasing order of degree d_v and plot the (r_v, d_v) pairs in log–log scale. The plots are shown in Fig. 2. The measured data is represented by points, while the solid line represents the least-squares approximation.

A striking observation is that the plots are approximated well by linear regression. The correlation coefficient is 0.97 for the Oregon and 0.978 for the Multi topology. This leads us to the following power law and definition.

Power Law 1 (Rank Exponent): Given a graph, the degree d_v of a node v is proportional to the rank of the node r_v to the power of a constant \mathcal{R} :

$$d_v \propto r_v^{\mathcal{R}}.$$

Definition 1: Let us sort the nodes of a graph in decreasing order of degree. We define the rank exponent \mathcal{R} to be the slope of the plot of the degrees of the nodes versus the rank of the nodes in log–log scale.

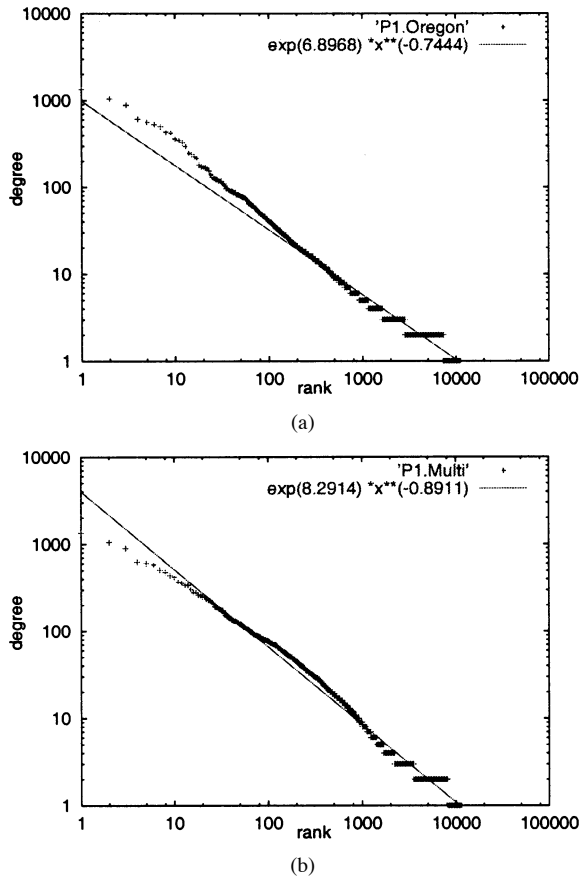


Fig. 2. Rank plot. Log–log plot of the degree d_v versus the rank r_v in the sequence of decreasing degree. (a) Oregon. (b) Multi.

Intuitively, Power Law 1 reflects a principle of the way domains connect; the linearity observed in 1253 graph instances is unlikely to be a coincidence.

Extended Discussion—Applications: We can estimate the proportionality constant for Power Law 1 if we require that the minimum degree of the graph is m ($d_N = m$). This way, we can refine the power law as follows.

Lemma 1: In a graph where Power Law 1 holds, the degree d_v of a node v is a function of the rank of the node r_v and the rank exponent \mathcal{R} as follows:

$$d_v = \frac{m}{N^{\mathcal{R}}} r_v^{\mathcal{R}}.$$

Proof: We can estimate the proportionality constant C for Power Law 1 if we require that the degree of the N th node is m , $d_N = m$

$$\begin{aligned} d_N = CN^{\mathcal{R}} &\Rightarrow \\ C = m/N^{\mathcal{R}}. &\quad (1) \end{aligned}$$

We combine Power Law 2 with (1) and conclude the proof. ■

Finally, using lemma 1, we relate the number of edges with the number of nodes and the rank exponent.

Lemma 2: In a graph where Power Law 1 holds, the number of edges E of a graph can be estimated as a function of the number of nodes N and the rank exponent \mathcal{R} as follows:

$$E \approx \frac{1}{2(\mathcal{R} + 1)} \left(1 - \frac{1}{N^{\mathcal{R}+1}}\right) N.$$

Proof: The sum of all the degrees for all the ranks is equal to two times the number of edges, since we count each edge twice

$$\begin{aligned} 2E &= \sum_{r_v=1}^N d_v \Rightarrow \\ 2E &= \sum_{r_v=1}^N (r_v/N)^{\mathcal{R}} = (1/N)^{\mathcal{R}} \sum_{r_v=1}^N r_v^{\mathcal{R}} \Rightarrow \\ E &\approx \frac{1}{2N^{\mathcal{R}}} \int_1^N r_v^{\mathcal{R}} dr_v. \quad (2) \end{aligned}$$

In the last step above, we approximate the summation with an integral. Calculating the integral concludes the proof. ■

Note that Lemma 2 can give us the number of edges as a function of the number of nodes for a given rank exponent. For an additional discussion on estimates using this formula, see [20].

B. Degree Exponent \mathcal{D}

In this section, we study the distribution of the degree of the nodes. We plot the D_d versus the degree d in log–log scale in Fig. 3. The major observation is that the plots are approximately linear. The correlation coefficient is 0.996 for the Oregon and 0.991 for the Multi topology. As in the previous power law, the slope of the exponent is different, something which is expected since the Multi topology has many more links. Note that in [10] it is argued that this power law does not hold for the Multi topology, without trying to approximate it using linear regression. Their conclusion is arguable, since we have a correlation coefficient of 0.991. Again, as in the last power law, we checked to see if the power law holds for all the instances we had. We found that the power law holds for all the instances, and the correlation coefficient was always higher than 0.99. This leads us to the following power law and definition.

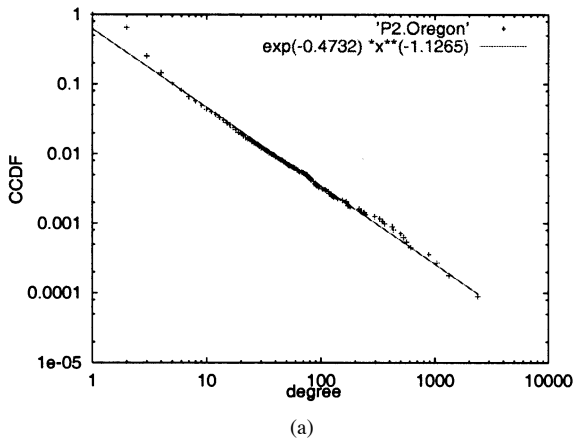
Power Law 2 (Degree Exponent): Given a graph, the CCDF D_d of a degree d is proportional to the degree to the power of a constant \mathcal{D} :

$$D_d \propto d^{\mathcal{D}}.$$

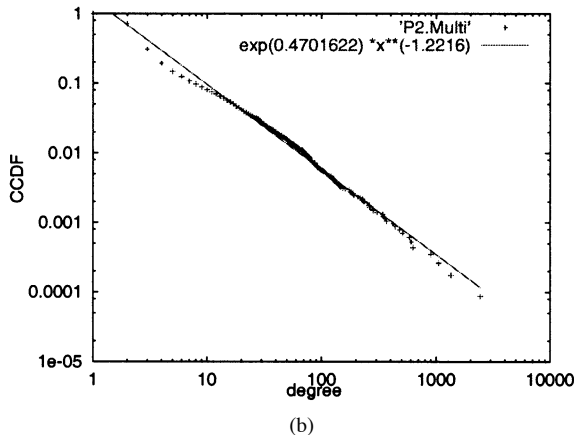
Definition 2: We define the degree exponent \mathcal{D} to be the slope of the plot of the cumulative degree of the degrees versus the degrees in log–log scale.

The intuition behind this power law is that the distribution of the degree of Internet nodes is not arbitrary. The qualitative observation is that degrees range over several orders of magnitude in a scale-invariant way. As a result, there is a nontrivial probability of finding nodes with very high degree. Our power law manages to quantify this observation with the degree exponent. This way, we can test the realism of a graph with a simple numerical comparison. If a graph does not follow Power Law 2, or if its degree exponent is considerably different from the real exponents, it probably does not represent a realistic topology.

Comment: Note the degree power law that we present here is different than the one presented in our earlier work [20]. They both refer to the same distribution; their difference is that the previous power law uses the probability distribution func-



(a)



(b)

Fig. 3. Log-log plot of D_d versus the degree for the Oregon and Multi topologies. (a) Oregon. (b) Multi.

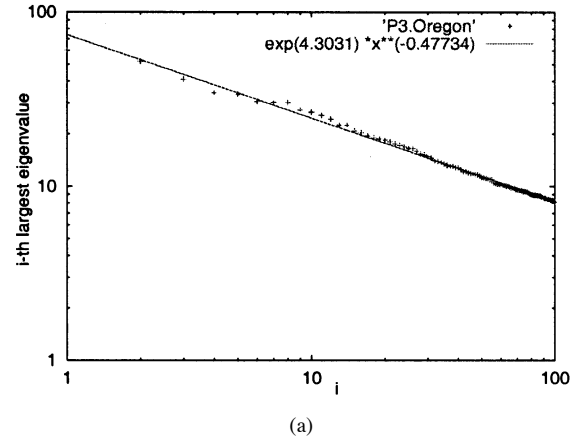
tion,² while the power law here uses the cumulative distribution function. As a result, the exponents of the different power laws differ approximately by one. Theoretically, the difference should be exactly one, since the cumulative distribution can be obtained by integrating the probability distribution. In practice, we see that the difference is not equal to one, due to approximations like the use of curve-fitting to find the slope. The cumulative distribution is preferable since it can be estimated in a statistically robust way.

Relationship of the Rank and Degree Power Laws: Both the rank and the degree power laws characterize the degree distribution from different angles. It can be shown that the exponents of the two power laws are related [11], [29], [61]. More specifically, in a perfect power-law distribution, the slope of the rank power law is equal to $\mathcal{R} = 1/D$. For the Multi topology, the rank slope is 0.89. Using the above formula and the degree slope, we find the rank slope to be equal to 0.81. This discrepancy can be attributed to measurement imperfections and inaccuracies. In this regard, we think that it is useful to report both exponents when characterizing a topology.

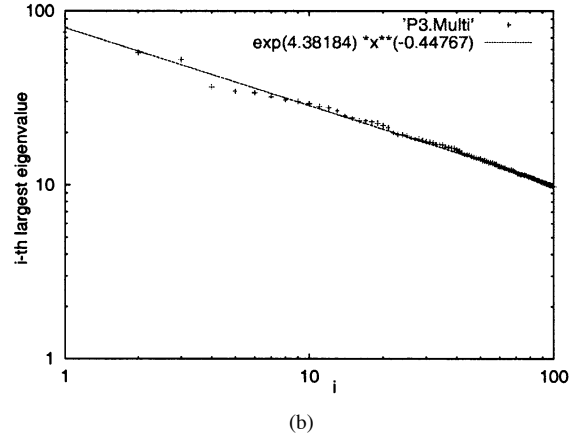
C. Eigen Exponent \mathcal{E}

In this section, we identify properties of the eigenvalues of our Internet graphs. Recall that the eigenvalues of a graph are

²The actual law stated that the frequency f_d of a degree d is proportional to the degree to the power of a constant \mathcal{D} where the frequency f_d of a degree is the number of nodes with degree d .



(a)



(b)

Fig. 4. Plot of the 100 largest eigenvalues for the Oregon and Multi topologies. (a) Oregon. (b) Multi.

the eigenvalues of its adjacency matrix. We plot the eigenvalue λ_i versus i in log-log scale for the first 100 eigenvalues. Recall that i is the order of λ_i in the decreasing sequence of eigenvalues. The result is shown in Fig. 4. The eigenvalues are shown as points in the figures, and the solid lines are approximations using a least-squares fit. Similar observations with equally high correlation coefficients were observed for all the other instances. We observe that the plots are practically linear with a correlation coefficient of 0.996 for both plots. The eigen exponent is -0.477 for the Oregon and -0.447 for the Multi topology.

It is rather unlikely that such a canonical form of the eigenvalues is purely coincidental, and we therefore conjecture that it constitutes the following empirical power law of the Internet topology.

Power Law 3 (Eigen Exponent): Given a graph, the eigenvalues λ_i are proportional to the order i to the power of a constant \mathcal{E} :

$$\lambda_i \propto i^{\mathcal{E}}.$$

Definition 3: We define the eigen exponent \mathcal{E} to be the slope of the plot of the sorted eigenvalues versus their order in log-log scale.

Eigenvalues are fundamental graph metrics. There is a rich literature that proves that the eigenvalues of a graph are closely related to many basic topological properties such as the diameter, the number of edges, the number of spanning trees, the number of connected components, and the number of walks of a

certain length between vertices, as we can see in [14]. All of the above suggests that the eigenvalues intimately relate to topological properties of graphs. However, it is not trivial to explore the nature and the implications of this power law.

Relationship of the Degree and the Eigenvalue Power Laws: A surprising relationship exists between the two exponents: the eigenvalue exponent is approximately half of the degree exponent. In more detail, Mihail *et al.* [40] show that if the degrees d_1, \dots, d_n of a graph follow a power law, then the nonincreasing sequence of the largest eigenvalues λ_i has the following one-to-one correspondence: $\lambda_i = \sqrt{d_i}$. It is worth noting that this is an asymptotic limit of the eigenvalues. If we take the logarithm of the previous equation, it follows that the two exponents differ by a factor of two. In practice, the exponents obey adequately the mathematical relationship, although the match is naturally not perfect. For example, the degree exponent for the Oregon topology is 1.12 and the eigenvalue exponent 0.47, which yields a ratio of 0.52 instead of 0.5.

D. Hop-Plot Exponent \mathcal{H}

In this section, we quantify the connectivity and distances between the Internet nodes in a novel way. We choose to study the size of the neighborhood within some distance, instead of the distance itself. Namely, we use the total number of pairs of nodes $P(h)$ within h hops, which we define as the total number of pairs of nodes within less than or equal to h hops, including self-pairs, and counting all other pairs twice.

In Fig. 5, we plot the number of pairs $P(h)$ as a function of the number of hops h in log-log scale. The data is represented by points. We want to describe the plot by a line in least-squares fit, for $h \ll \delta$, shown as a solid line in the plots. We approximate the first four hops in the inter-domain graphs. The correlation coefficient is 0.9765 and 0.9784 for the Oregon and Multi topology, respectively.

Unfortunately, four points is a rather small number to verify or disprove a linearity hypothesis experimentally. However, even this rough approximation has several useful applications, as we show later in this section. It is worth mentioning that Philips *et al.* [47] state that the neighborhood growth is exponential and not a power law. In Fig. 6, we plot again the number of pairs in log-lin for the Multi topology. We approximate the first four hops and found a correlation coefficient of 0.918, which is much lower than the previous correlation. From this, it seems that we can approximate the hop-plot better with a power law than with an exponential function.

Approximation 1 (Hop-Plot Exponent): The total number of pairs of nodes $P(h)$ within h hops is proportional to the number of hops to the power of a constant \mathcal{H} :

$$P(h) \propto h^{\mathcal{H}}, \quad h \ll \delta.$$

Definition 4: Let us plot the number of pairs of nodes $P(h)$ within h hops versus the number of hops in log-log scale. For $h \ll \delta$, we define the slope of this plot to be the *hop-plot exponent* \mathcal{H} .

Extended Discussion—Applications: We can refine Approximation 1 by calculating its proportionality constant. Let us recall the definition of the number of pairs $P(h)$. For $h = 1$, we

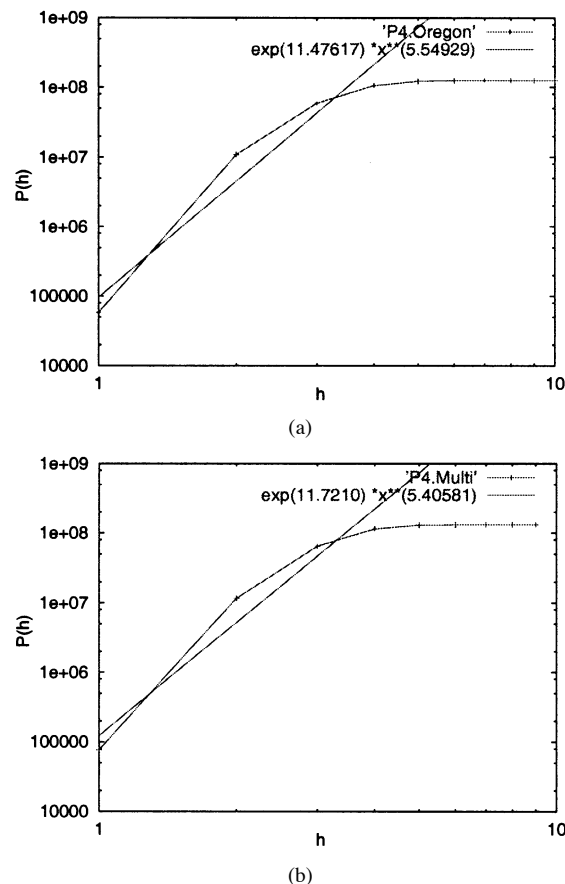


Fig. 5. Hop-plot: Log-log plots of the number of pairs of nodes $P(h)$ within h hops versus the number of hops h . (a) Oregon. (b) Multi.

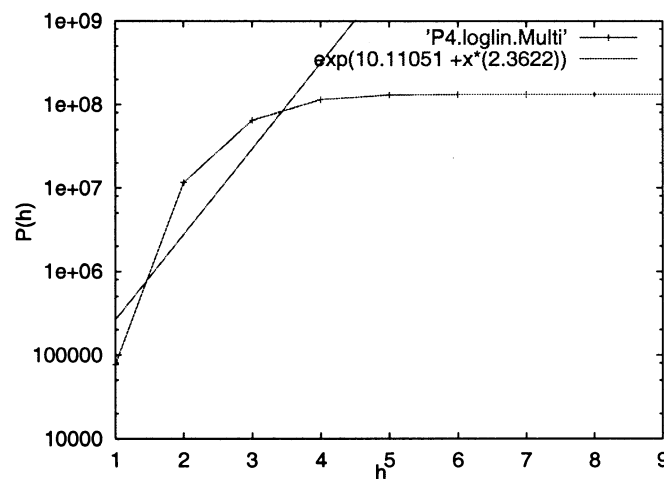


Fig. 6. Approximating the hop-plot with an exponential function. This is a linear-logarithmic plot of the number of pairs of nodes $P(h)$ within h hops versus the number of hops h .

consider each edge twice and we have the self-pairs, therefore, $P(1) = N + 2E$. We demand that Approximation 1 satisfies the previous equation as an initial condition.

Lemma 3: The number of pairs within h hops is

$$P(h) = \begin{cases} ch^{\mathcal{H}}, & h \ll \delta \\ N^2, & h \geq \delta \end{cases}$$

where $c = N + 2E$ to satisfy initial conditions.

In networks, we often need to reach a target without knowing its exact position [9], [51]. In these cases, selecting the extent of our broadcast or search is an issue.³ On the one hand, a small broadcast may not reach our target. On the other hand, an extended broadcast creates too many messages and takes a long time to complete. Ideally, we want to know how many hops are required to reach a “sufficiently large” part of the network. In our hop-plots, a promising solution is the intersection of the two asymptote lines: the horizontal one at level N^2 and the asymptote with slope \mathcal{H} . We calculate the intersection point using Lemma 3, and we define the following.

Definition 5 (Effective Diameter): Given a graph with N nodes, E edges, and \mathcal{H} hop-plot exponent, we define the effective diameter δ_{ef} as

$$\delta_{ef} = \left(\frac{N^2}{N + 2E} \right)^{1/\mathcal{H}}.$$

Intuitively, the effective diameter can be understood as follows. Any two nodes are within δ_{ef} hops from each other with high probability. We verified the above statement experimentally. The effective diameters of our inter-domain graphs was slightly over four. Rounding the effective diameter to four, approximately 80% of the pairs of nodes are within this distance. The ceiling of the effective diameter is five, which covers more than 95% of the pairs of nodes. The above confirms that the effective diameter manages to capture the majority of the distances. Furthermore, it argues indirectly that the hop-plot exponent as a metric seems useful.

An advantage of the effective diameter is that it can be calculated easily, when we know N and \mathcal{H} . Recall that we can calculate the number of edges from Lemma 2. Therefore, given estimates the hop-plot and rank-plot exponents, we can calculate the effective diameter of future Internet instances of a given size [20].

Furthermore, we can estimate the average size of the neighborhood $NN(h)$ within h hops using the number of pairs $P(h)$. Recall that $P(h) - N$ is the number of pairs without the self-pairs

$$NN(h) = \frac{P(h)}{N} - 1. \quad (3)$$

Using (3) and Lemma 3, we can estimate the average neighborhood size.

Lemma 4: The average size of the neighborhood $NN(h)$ within h hops as a function of the hop-plot exponent \mathcal{H} for $h \ll \delta$ is

$$NN(h) = \frac{c}{N} h^{\mathcal{H}} - 1, \quad h > 0$$

where $c = N + 2E$ to satisfy initial conditions.

The average neighborhood is a commonly used parameter in the performance of network protocols. Our estimate is an im-

³This problem has direct practical importance. The Internet has a built-in mechanism for limiting the number of hops a packets makes. The time-to-live field of a packet is a counter that is decreased at each hop until it reaches zero, at which point the packet is not forwarded further.

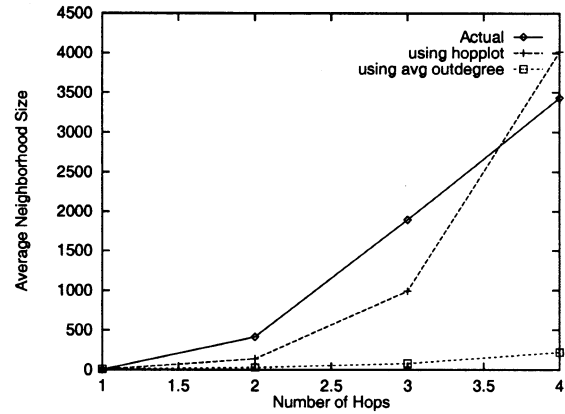


Fig. 7. Average neighborhood size versus number of hops. Actual and estimated size using hop-plot exponent and average degree.

provement over the commonly used estimate that uses the average degree [51], [59], which we call **average-degree estimate**

$$NN'(h) = \bar{d}(\bar{d} - 1)^{h-1}.$$

In Fig. 7, we plot the actual and the two estimates of the average neighborhood size versus the number of hops using an instance from 1998. The superiority of the hop-plot exponent estimate is apparent compared to the average-degree estimate. The discrepancy of the average-degree estimate can be explained if we consider that the estimate does not comply with the real data; it implicitly assumes that the degree distribution is uniform. In more detail, it assumes that each node in the periphery of the neighborhood adds $\bar{d} - 1$ new nodes at the next hop. Our data shows that the degree distribution is highly skewed, which explains why the use of the hop-plot estimate gives a better approximation.

The most interesting difference between the two estimates is qualitative. The average degree based estimate considers the neighborhood size exponential in the number of hops. Our estimate considers the neighborhood as an \mathcal{H} -dimensional sphere with radius equal to the number of hops, which is a novel way to look at the topology of a network. Our data suggests that the hop-plot exponent-based estimate gives a closer approximation compared to the average-degree-based metric.

V. PERSISTENCE OF POWER-LAW EXPONENTS

We examine the evolution of power-law exponents in the five-year span from November 1997 to February 2002. We want to stress that the main observation is that the power laws hold for every instance. The evolution of the slope is a secondary issue.

Evolution of Rank Exponent \mathcal{R} : In Fig. 8, we examine the time evolution of the slope of the rank exponent. We plot the rank exponent versus the day that the instance of the graph was collected. The rank exponent \mathcal{R} power law holds for all the instances over a period of five years. The correlation coefficient of the law seems to decrease over time, and for the last instance it is close to 0.9654. This indicates that the rank exponent for the topologies from Oregon should be treated with care in the

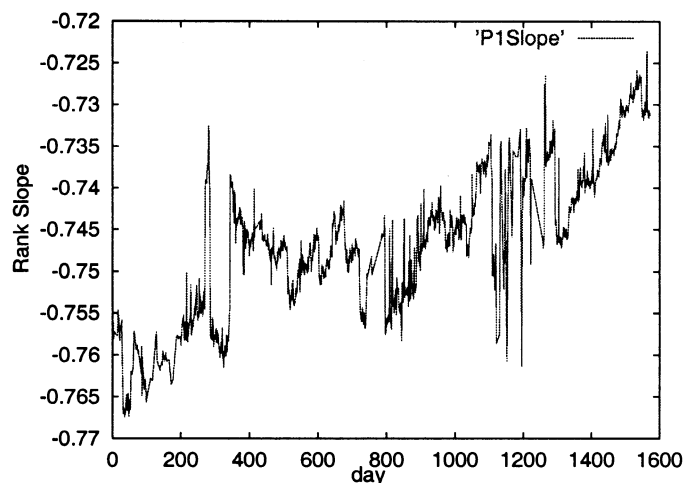


Fig. 8. Evolution of the slope of the rank exponent.

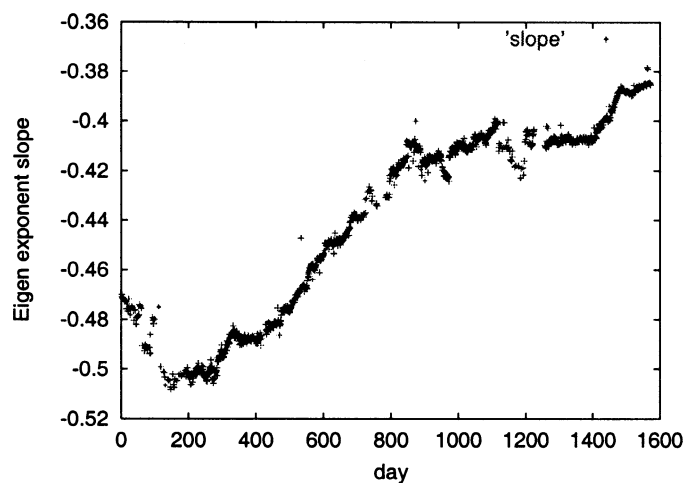


Fig. 10. Evolution of the slope of the eigen exponent.

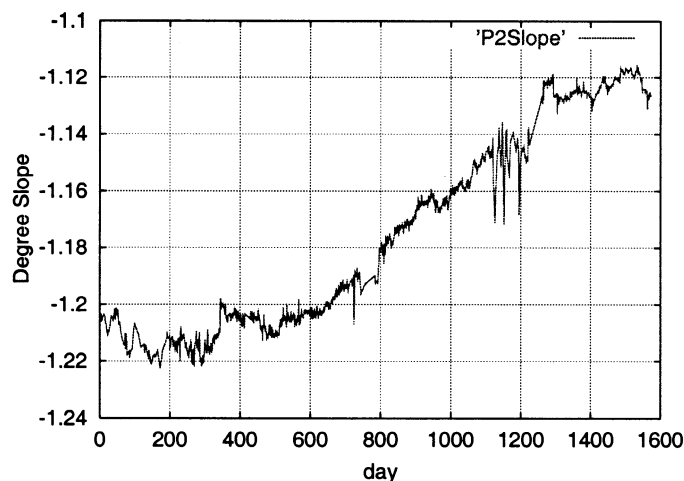


Fig. 9. Evolution of the slope of the degree exponent.

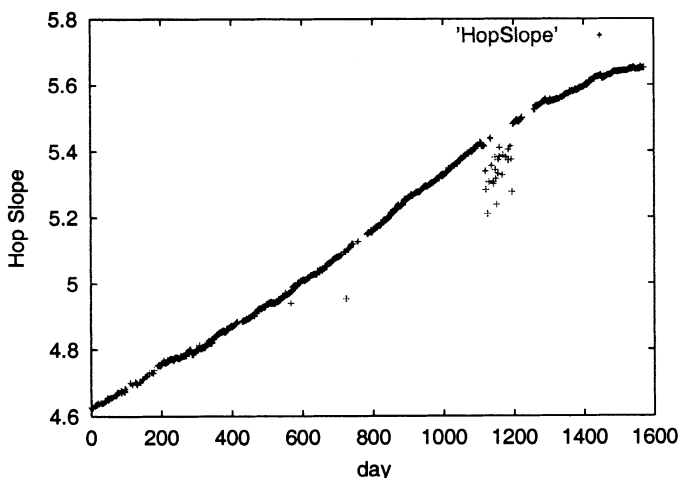


Fig. 11. Evolution of the slope of the hop-plot exponent.

future. However, we see in Fig. 2 that for the more complete topology we have a higher correlation coefficient.

Evolution of Degree Slope \mathcal{D} : We study the slope of the cumulative degree exponent and its evolution in time. In Fig. 9, we plot the degree exponent versus time. The degree exponent power law holds for all the instances with a correlation coefficient always higher than 0.99. We observe from the graph that the slope is between -1.12 and -1.22 , i.e., a variation of less than 9%.

Evolution of Eigen Exponent \mathcal{E} : In Fig. 10, we plot the time evolution of the eigen exponent. The power law holds for all the instances we have measured. As we can see from the graph, the value of the eigen exponent decreases for the first 150 instances and then starts to rise again for the rest of the instances. We do not have an intuitive explanation for this behavior. Note that the eigenvalues of a graph does not depend on the way the nodes are enumerated.

Evolution of Hop-Plot Exponent: In Fig. 11, we plot the time evolution of the hop-plot exponent's slope. The power law holds for all the instances with a correlation coefficient always higher than 0.97. We observe that the value of slope increases steadily. The initial value of the hop-plot exponent is 4.6 and for the latest instance is 5.7.

Understanding the Hop-Plot Increase: As we saw, the network size increases significantly, while the distances between nodes increase very little. In Table III, we list the percentage of nodes that we can reach as a function of the number of hops or the neighborhood of a node within h hops. We compare two graph instances, November 8, 1997, and our last instance, February 28, 2002. Although the size of the graph quadrupled, we reach approximately the same percentage of nodes with the same number of hops. In absolute numbers, the number of nodes we can reach in six hops increased from approximately 3000 to 13 000.

VI. GENERATION OF POWER LAWS

Why would such an uncontrolled⁴ entity like the Internet follow any statistical regularities? Note that the high correlation coefficients rule out the possibility of pure coincidence. Intrigued by the previous question, and by the appearance of power laws in many diverse fields, many scientists have tried to

⁴The term *uncontrolled* refers to the fact that the Internet is not governed by a central authority, and its growth and design is driven by many different optimization goals, such as financial, business, and performance related.

TABLE III
SIZE OF THE NEIGHBORHOOD OF A NODE (AS PERCENTAGE OF THE TOTAL)
AS A FUNCTION OF THE HOPS (RADIUS OF NEIGHBORHOOD)

| hops | 11-08-1997 | 02-28-2002 |
|------|------------|------------|
| 1 | 0.08% | 0.04% |
| 2 | 8.86% | 8.09% |
| 3 | 43.40% | 46.64% |
| 4 | 80.99% | 84.76% |
| 5 | 96.70% | 97.46% |
| 6 | 99.65% | 99.71% |

find the mechanism responsible for the creation of power-law graphs.

In this section, we will first give a small review of the most popular models that try to explain the appearance of power laws in networks. Later, we will briefly describe the status on the graph generation tools.

Scale-Free Networks: A very elegant growth model has been proposed by Barabasi and Reka. In their original work [7], their model states that the scale-free nature roots in two mechanisms, the addition of new nodes, and their preferential attachment. Their model grows a graph by adding nodes. The probability of a new node connecting with node i of degree d_i is proportional to its degree: $d_i / \sum d_j$, where $\sum d_j$ is the sum of the degrees of all current nodes. In a more recent work, the same authors propose a more general model that includes generation of edges between existing nodes and rewiring [2], that is, the removal of one edge and the creation of another between existing nodes. An extensive review on variations of the original idea that include more parameters can be found in [3].

There exists a number of real data studies based on the growth model proposed by Barabasi. In [10], they conclude that this theoretical model is not supported by the real data. Instead, they mention that rewiring occurs infrequently, and that new nodes express a greater preference for nodes with large degree than is represented by the simple linear preference model. On the other hand, in [3], [5], and [58], the authors show that the addition of nodes and edges follows the linear preferential model. This is controversial and more work is needed to compare the different approaches used.

Heuristically Optimized Tradeoffs: In [18], Fabrikant *et al.* propose “a simple and primitive model of Internet growth.” In their model, the power-law distributions root from the Internet growth in which two objectives are optimized simultaneously. The two objectives are: the connection costs (last mile), and the transmission delays measured in hops. Their model works as follows. They use a unit square plane, where nodes arrive and their place is chosen uniformly at random. Each node attaches itself on one of the previous nodes. They use two metrics in order to choose where the node should attach. The first metric is the Euclidean distance d_{ij} between the new node i and a node j . This metric captures the “last-mile” costs. The second metric is a measure of the centrality of a node j , h_j , if the new node attached to node j . This shows how close is the node to the center, and they mention that this captures the operation costs due to communication delays. Node i chooses to connect with node j

that minimizes the weighted sum $\min_{j < i} a d_{ij} + h_j$, where a is used to change the relative importance of the two objectives.

Highly Optimized Tolerance: In [27], Carlson and Doyle propose that power laws are the result of an optimization, either through natural selection or engineering design, to provide robust performance despite uncertain environments. Regarding the Internet, they mention that the survivability built in the Internet and its protocols could be the cause of the power laws.

Topology Generators: The introduction of power laws [20] brought a revision of the graph generation models in the networking community. The power laws can be used as one question in the “qualifying exam” for the realism of a graph. The early generators failed when tested against power laws, so after that a number of new generators was proposed [26], [35], [36], [42], [54], [58].

There are two kinds of graph generation tools. In the first one, we have the tools that take the power laws as given, and they do not attempt to emulate the process that leads to a power law [1], [26], [42]. In the second category [35], [36], [54], [58], we have generators which try to capture the actual process that governs the creation of power laws. All of them use variations of the preferential attachment model, described in [2] and [7].

In the most recent effort [54], Bu and Towsley proposed a new generator that generates more realistic Internet topologies. Furthermore, they used additional metrics found in small-world networks [15]. They showed that previous generators fail in some of these new criteria. They showed that by deviating from the linear preferential model by giving higher preferentiality to high degree nodes, more realistic topologies are generated.

VII. CONCLUSION

In this paper, we propose power laws as a tool to describe the Internet topology and examine their persistence in time. The power laws capture concisely the highly skewed distributions of the graph properties. Finally, we show how these exponents relate to each other and how they relate to other topological properties.

We note the persistence of power laws in time. They appear in more than 1200 daily instances over the span of more than five years from 1997 to 2002. In this interval, the network underwent significant changes in size (400%) and rate of growth. The monitoring infrastructure changed and evolved as well. This suggests that the appearance of power laws is unlikely to be a coincidence or an artifact. An orthogonal but also striking observation is that some of the exponents did not change more than 10%. Furthermore, the power laws seem to hold even in the most complete topology, which combines multiple sources. In fact, some of the power laws hold with higher correlation coefficients in this data set.

Some additional observations are summarized in the following points.

- Power-law exponents are a more efficient way to describe the highly skewed graph metrics compared with average values of real graphs.
- We propose the number of pairs $P(h)$ within h hops as a metric of the density of the graph and approximate it using the hop-plot exponent \mathcal{H} .

- We derive formulas that link the exponents of our power laws with graph metrics such as the number of nodes, the number of edges, and the average neighborhood size.
- Using power laws, we obtain better intuition, for example, we can see that the network becomes denser by observing the hop-plot exponent.

Apart from their theoretical interest, our power laws have practical applications. First, our power laws can assess the realism of synthetic graphs and enhance the validity of simulations. Second, they can help analyze the average-case behavior of network protocols. For example, we can estimate the message complexity of protocols using our estimate for the neighborhood size. Third, the power laws can help answer “what-if” scenarios such as “*What will be the diameter of the Internet, when the number of nodes doubles?*” and “*What will be the number of edges then?*”

VIII. FUTURE WORK

The topological power laws presented here form a critical step toward understanding and modeling the Internet. However, there are several open questions. First, we would like to explore further the meaning and the value of the exponents. We believe that such analysis could reveal interesting interplays and tradeoffs between the forces that govern the creation of the topology. Second, the power laws alone may not be sufficient in describing the topology in all its complexity. For example, we would like to develop more structural properties that will quantify the topology in a way that is easier to visualize. The goal of this direction is to develop a simple and intuitive model for the Internet topology.

ACKNOWLEDGMENT

The authors would like to thank M. Craven, D. Zappala, and A. Perrig for their help in earlier phases of this work, and V. Paxson and E. Zegura for their valuable feedback.

REFERENCES

- [1] W. Aiello, F. Chung, and L. Lu, “A random graph model for massive graphs,” in *Proc. STOC*, 2000, pp. 171–180.
- [2] R. Albert and A. Barabasi, “Topology of evolving networks: Local events and universality,” *Phys. Rev. Lett.*, vol. 85, pp. 5234–5237, 2000.
- [3] —, “Statistical mechanics of complex networks,” *Rev. Modern Phys.*, vol. 74, pp. 47–97, 2002.
- [4] R. Albert, H. Jeong, and A. L. Barabasi, “Diameter of the World Wide Web,” *Nature*, vol. 401, pp. 130–131, 1999.
- [5] A. Vázquez, R. Pastor-Satorras, and A. Vespignani, “Large-scale topological and dynamical properties of Internet,” *Phys. Rev. E*, vol. 65, June 2002.
- [6] P. Bak, *How Nature Works: The Science of Self Organized Criticality*. New York: Springer-Verlag, 1996.
- [7] A. Barabasi and R. Albert, “Emergence of scaling in random networks,” *Science*, vol. 8, pp. 509–512, Oct. 1999.
- [8] K. L. Calvert, M. B. Doar, and E. W. Zegura, “Modeling internet topology,” *IEEE Commun. Mag.*, vol. 35, pp. 160–163, June 1997.
- [9] K. Carlberg and J. Crowcroft, “Building shared trees using a one-to-many joining mechanism,” *ACM Comput. Commun. Rev.*, pp. 5–11, Jan. 1997.
- [10] Q. Chen, H. Chang, R. Govindan, S. Jamin, S. J. Shenker, and W. Willinger, “The origin of power laws in Internet topologies revisited,” in *Proc. IEEE INFOCOM*, New York, NY, June 2002, pp. 608–617.
- [11] H. Chou, “A note on power laws of Internet topology,” Harvard Univ., Cambridge, MA, arXiv Tech. Rep. 0012019, Dec. 2000.
- [12] J. Chuang and M. Sirbu, “Pricing multicast communications: A cost based approach,” *Telecommun. Syst.*, vol. 17.3, pp. 281–297, 2001.
- [13] M. Crovella and A. Bestavros, “Self-similarity in World Wide Web traffic, evidence and possible causes,” in *Proc. ACM SIGMETRICS*, 1996, pp. 160–169.
- [14] D. M. Cvetković, M. Boob, and H. Sachs, *Spectra of Graphs*. New York: Academic, 1979.
- [15] D. J. Watts and S. H. Strogatz, “Collective dynamics of small-world networks,” *Nature*, vol. 393, pp. 440–442, 1998.
- [16] M. Doar, “A better model for generating test networks,” in *Proc. IEEE Global Internet*, Nov. 1996.
- [17] F. Liljeros, C. Edling, L. A. N. Amaral, H. E. Stanley, and Y. Aberg, “The web of human sexual contacts,” *Nature*, vol. 411, pp. 907–908, 2001.
- [18] A. Fabrikant, E. Koutsoupias, and C. H. Papadimitriou, “Heuristically optimized trade-offs: A new paradigm for power laws in the Internet,” in *Proc. ICALP*, 2002, pp. 110–122. Springer-Verlag LNCS.
- [19] C. Faloutsos and I. Kamel, “Beyond uniformity and independence: Analysis of R-trees using the concept of fractal dimension,” in *Proc. ACM SIGACT-SIGMOD-SIGART PODS*, Minneapolis, MN, May 1994, pp. 4–13.
- [20] M. Faloutsos, P. Faloutsos, and C. Faloutsos, “On power-law relationships of the Internet topology,” in *Proc. ACM SIGCOMM*, Cambridge, MA, Sept. 1999, pp. 251–262.
- [21] S. Floyd and V. Paxson, “Difficulties in simulating the Internet,” *IEEE/ACM Trans. Networking*, vol. 9, pp. 392–403, Aug. 2001.
- [22] R. Govindan and A. Reddy, “An analysis of Internet Inter-domain topology and route stability,” in *Proc. IEEE INFOCOM*, Kobe, Japan, Apr. 1997, pp. 850–857.
- [23] R. Govindan and H. Tangmunarunkit, “Heuristics for Internet map discovery,” in *Proc. IEEE INFOCOM*, Tel Aviv, Israel, Mar. 2000, pp. 1371–1380.
- [24] S. Jamin, C. Jin, Y. Jin, D. Raz, Y. Shavitt, and L. Zhang, “On the placement of Internet instrumentation,” in *Proc. IEEE INFOCOM*, Tel Aviv, Israel, Mar. 2000, pp. 295–304.
- [25] H. Jeong, B. Tombler, R. Albert, Z. N. Oltvai, and A.-L. Barabasi, “The large-scale organization of metabolic networks,” *Nature*, vol. 407, pp. 651–654, 2000.
- [26] C. Jin, Q. Chen, and S. Jamin, “Inet: Internet topology generator,” Dept. EECS, Univ. Michigan, Ann Arbor, Tech. Rep. CSE-TR-433-00, 2000.
- [27] J. M. Carlson and J. Doyle, “Highly optimized tolerance: A mechanism for power laws in designed systems,” *Phys. Rev. E*, vol. 60, no. 2, pp. 1412–1427, 1999.
- [28] R. Kumar, P. Raghavan, S. Rajagopalan, D. Sivakumar, A. Tomkins, and E. Upfal, “The Web as a graph,” in *Proc. ACM Symp. Principles of Database Systems*, 2000, pp. 1–10.
- [29] L. A. Adamic. (2000) Zipf, power laws, and Pareto—A ranking tutorial. [Online]. Available: <http://www.parc.xerox.com/iea/>
- [30] W. E. Leland, M. S. Taqqu, W. Willinger, and D. V. Wilson, “On the self-similar nature of ethernet traffic,” *IEEE Trans. Networking*, vol. 2, pp. 1–15, Feb. 1994.
- [31] L. Gao, “On inferring autonomous system relationships in the internet,” *IEEE/ACM Trans. Networking*, vol. 9, pp. 733–745, Dec. 2001.
- [32] L. Subramanian, S. Agarwal, J. Rexford, and R. Katz, “Characterizing the Internet hierarchy from multiple vantage points,” in *Proc. IEEE INFOCOM*, June 2002, pp. 618–627.
- [33] D. Magoni and J. J. Pansiot, “Analysis of the autonomous system network topology,” *ACM Comput. Commun. Rev.*, pp. 26–37, July 2001.
- [34] B. Mandelbrot, *Fractal Geometry of Nature*. New York: Freeman, 1977.
- [35] A. Medina, A. Lakhina, I. Matta, and J. Byers, “Brite: An approach to universal topology generation,” in *Proc. MASCOIS*, 2001.
- [36] A. Medina, I. Matta, and J. Byers, “On the origin of power laws in Internet topologies,” *ACM Comput. Commun. Rev.*, vol. 30, no. 2, pp. 18–34, Apr. 2000.
- [37] P. Van Mieghem, G. Hooghiemstra, and R. van der Hofstad, “On the efficiency of multicast,” *IEEE/ACM Trans. Networking*, vol. 9, pp. 719–732, Dec. 2001.
- [38] M. Mitzenmacher, “A brief history of generative models for power law and lognormal distributions,” *Proc. Allerton Conf.*, 2001.
- [39] M. Jovanovic, “Modeling large-scale peer-to-peer networks and a case study of gnutella,” M.S. thesis, Univ. Cincinnati, Cincinnati, OH, 2001.
- [40] M. Mihail and C. H. Papadimitriou, “On the Eigenvalue Power Law,” presented at the *Random Workshop*, 2002.
- [41] Online data and reports. Univ. Oregon Route Views Project, Advanced Network Technology Center, Eugene, OR. [Online]. Available: <http://www.routeviews.org/>

- [42] C. R. Palmer and J. G. Steffan, "Generating network topologies that obey power laws," in *Proc. IEEE Globecom*, Nov.–Dec. 2000, pp. 434–438.
- [43] J.-J. Pansiot and D. Grad, "On routes and multicast trees in the Internet," in *ACM Comput. Commun. Rev.*, vol. 28, Jan. 1998, pp. 41–50.
- [44] V. Pareto, "Cours d'économie politique," in *Dronz*, Geneva, Switzerland, 1896.
- [45] K. Park and H. Lee, "On the effectiveness of route-based packet filtering for distributed DoS attack prevention in power-law Internets," in *Proc. ACM SIGCOMM*, San Diego, CA, Aug. 2001, pp. 15–26.
- [46] V. Paxson and S. Floyd, "Wide-area traffic: The failure of Poisson modeling," *IEEE/ACM Trans. Networking*, vol. 3, pp. 226–244, June 1995.
- [47] G. Philips, S. Shenker, and H. Tangmunarunkit, "Scaling of multicast trees: Comments on the Chuang–Sirbu scaling law," in *Proc. ACM SIGCOMM*, Sept. 1999, pp. 41–51.
- [48] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C*, 2nd ed. Cambridge, U.K.: Cambridge Univ. Press, 1992.
- [49] Y. Rekhter and T. Li, Eds., "A Border Gateway Protocol 4 (BGP-4)," IETF, RFC 1771, 1995.
- [50] S. Redner, "How popular is your paper? An empirical study of the citation distribution," *Eur. Phys. J. B*, vol. 4, pp. 131–134, 1998.
- [51] S. Yan, M. Faloutsos, and A. Banerjee, "Qos-aware multicast routing for the Internet: The design and evaluation of Qosmic," *IEEE/ACM Trans. Networking*, vol. 10, pp. 54–66, Feb. 2002.
- [52] H. Tangmunarunkit, R. Govindan, S. Jamin, S. Shenker, and W. Willinger, "Network topology generators: Degree-based vs. structural," in *Proc. ACM SIGCOMM*, Aug. 2002, pp. 147–159.
- [53] L. Tauro, C. Palmer, G. Siganos, and M. Faloutsos, "A simple conceptual model for the Internet topology," in *Proc. Global Internet*, San Antonio, TX, 2001.
- [54] T. Bu and D. Towsley, "On distinguishing between Internet power law topology generators," in *Proc. IEEE INFOCOM*, New York, NY, June 2002, pp. 638–647.
- [55] B. M. Waxman, "Routing of multipoint connections," *IEEE J. Select. Areas Commun.*, vol. 6, pp. 1617–1622, Dec. 1988.
- [56] W. Willinger, M. Taqqu, R. Sherman, and D. V. Wilson, "Self-similarity through high variability: Statistical analysis of ethernet LAN traffic at the source level," *ACM Comput. Commun. Rev.*, vol. 25, pp. 100–113, 1995.
- [57] T. Wong and R. Katz, "An analysis of multicast forwarding state scalability," in *Proc. Int. Conf. Network Protocols*, 2000, pp. 105–115.
- [58] S. H. Yook, H. Jeong, and A. Barabasi, "Modeling the Internet's large-scale topology," *Proc. Nat. Acad. Sciences*, vol. 99, pp. 13382–13386, Oct. 2002.
- [59] D. Zappala, D. Estrin, and S. Shenker, "Alternate path routing and pinning for interdomain multicast routing," Univ. Southern California, Los Angeles, Tech. Rep. USC CS TR 97-655, 1997.
- [60] E. W. Zegura, K. L. Calvert, and M. J. Donahoo, "A quantitative comparison of graph-based models for internetworks," *IEEE/ACM Trans. Networking*, vol. 5, pp. 770–783, Dec. 1997.
- [61] G. K. Zipf, *Human Behavior and Principle of Least Effort: An Introduction to Human Ecology*. Cambridge, MA: Addison Wesley, 1949.



Georgos Siganos (S'00) received the B.Sc. degree in electrical and computer engineering from the Technical University of Crete, Crete, Greece, in 1999. He is currently working toward the Ph.D. degree at the University of California at Riverside.

His main research interests are in the areas of Internet topology and interdomain routing.

Mr. Siganos is a student member of the Association for Computing Machinery.



Michalis Faloutsos (M'99) received the B.S. degree from the National Technical University of Athens, Athens, Greece, in 1993 and the M.Sc. and Ph.D. degrees from the University of Toronto, Toronto, ON, Canada, in 1995 and 1999, respectively.

He is a Faculty Member with the Department of Computer Science, University of California at Riverside. He is actively involved in the community as a reviewer and a technical program committee member in many conferences and journals. His research interests include Internet protocols and measurements, multicasting, and mobile and ad-hoc networks.

Dr. Faloutsos has received a National Science Foundation CAREER Award, two Defense Advanced Research Projects Agency (DARPA) Awards, and grants from AT&T and Tata Consultancy Services.



Petros Faloutsos received the B.Eng. degree in electrical engineering from the National Technical University of Athens, Athens, Greece, in 1993 and the M.Sc. and Ph.D. degrees in computer science from the University of Toronto, Toronto, ON, Canada, in 1995 and 2002, respectively.

He is an Assistant Professor with the Department of Computer Science, University of California, Los Angeles (UCLA). His research interests are in the area of computer graphics and animation, focusing on physics-based animation techniques for anthropomorphic virtual characters. His work toward virtual characters resulted in a prototype virtual stuntman that received a Best Paper Award (Computer and Graphics) and international media attention. His work has been featured in international forums including CNN, Discovery Channel, and the New York Times. He is the founder and the director of the graphics laboratory at the Department of Computer Science at UCLA, called MAGIX (Modeling Animation and GraFX Laboratory). He is also interested in computer networks and the topology of the internet.

Dr. Faloutsos is a member of the Association for Computing Machinery and the Technical Chamber of Greece (Professional Engineers).



Christos Faloutsos (M'85) received the B.Sc. degree in electrical engineering from the National Technical University of Athens, Athens, Greece, in 1981 and the M.Sc. and Ph.D. degrees in computer science from the University of Toronto, Toronto, ON, Canada, in 1982 and 1985, respectively.

He is currently a Professor with the Department of Computer Science, Carnegie Mellon University, Pittsburgh, PA. He has published more than 90 refereed articles and one monograph and holds four patents. His research interests include data mining, indexing in relational and multimedia databases, and database performance.

Dr. Faloutsos has received the Presidential Young Investigator Award by the U.S. National Science Foundation in 1989, two Best Paper Awards (from SIGMOD'94 and VLDB'97), and four teaching awards. He is a member of the IEEE Computer Society.