On the Spectrum and Structure of Internet Topology Graphs *

Danica Vukadinović, Polly Huang, and Thomas Erlebach

Computer Engineering and Networks Laboratory (TIK),
Swiss Federal Institute of Technology (ETH), Zurich, Switzerland
{vukadin, huang, erlebach}@tik.ee.ethz.ch.

Abstract. In this paper we study properties of the Internet topology on the autonomous system (AS) level. We find that the normalized Laplacian spectrum (*nls*) of a graph provides a concise fingerprint of the corresponding network topology. The *nls* of AS graphs remains stable over time in spite of the explosive growth of the Internet, but the *nls* of synthetic graphs obtained using the state-of-the-art topology generator Inet-2.1 is significantly different, in particular concerning the multiplicity of eigenvalue 1. We relate this multiplicity to the sizes of certain subgraphs and thus obtain a new structural classification of the nodes in the AS graphs, which is also plausible in networking terms. These findings as well as new power-law relationships discovered in the interconnection structure of the subgraphs may lead to a new generator that creates more realistic topologies by combining structural and power-law properties.

1 Introduction

Significant research efforts have recently been invested in the analysis of the Internet topology. The current Internet is the result of rapid, distributed growth without controlled planning by a central authority. Therefore, its topology reflects in great parts the choices and decisions made by individual organizations whose subnetworks form the Internet. As a consequence, the characteristics of the Internet topology can only be investigated by analyzing the available data about the current connectivity of routers or autonomous systems or snapshots of that connectivity taken at an earlier time.

Gaining additional knowledge about the properties of the Internet topology is important for several reasons. In particular, optimization problems related to resource allocation, call admission control, routing, and Distributed Denial of Service (DDoS) attack prevention (see [13]) that are provably difficult to solve for general topologies might allow efficient solutions for a class of networks containing the real Internet. Furthermore, a good understanding of the Internet topology can lead to improvements in network topology generators in order to generate "Internet-like" networks of various sizes for simulations. Network

* Partially supported by European Commission - Fet Open project COSIN IST-2001-33555, with funding provided by BBW Switzerland.

simulations with realistic topologies can again help to design, tune and evaluate new protocols, applications, and algorithms.

1.1 Related Work: Topology Models and Generators

Until 1999, one of the most popular network generators was GT-ITM [3], a generator that combines the hierarchical models called Transit-stub and Tiers with popular random graph models such as Waxman's model [18]. However, a major new insight into properties of the real Internet topology was gained by Faloutsos et al. [7]. They found four power-laws[1] that appear to hold for various relations between popular graph metrics in the Internet (both on the router level and on the AS level): node degree vs. node rank, degree frequency vs. degree, number of nodes within a certain number of hops vs. number of hops, and 20 largest eigenvalues of the the adjacency matrix vs. their ranks. Thus it became clear that realistic topology generators must produce graphs satisfying these power-laws.

Exploring the power-law degree distribution in WWW and Internet graphs, Barabási and Albert [1] proposed *incremental growth* – the fact that the nodes are added incrementally – and *preferential connectivity* – which means that the probability of connecting a new node to node i is proportional to the degree of i – as two main reasons for the appearance of power-laws. Based on this model, the BRITE topology generator was created [12].

Jin et al. [11] proposed a model called Inet. For a given number of nodes and percentage of nodes with degree 1, the power-law exponents from the real AS Internet graphs are used to determine the degree distribution of the resulting graph. A spanning tree using only nodes with degree at least two is created. The degree 1 nodes are then attached to the tree with proportional probability.

A generalization of the linear preference in the Barabási–Albert model and a comparison of different power-law topology generators can be found in [2].

Tangmunarunkit et al. in [15] compared structural generators based on hierarchical models (such as GT-ITM) to "purely" power-law degree-based generators (such as BRITE or Inet). Using different graph-theoretic metrics, they argued that the degree-based generator models are more realistic than structural ones and that, surprisingly, a certain hierarchical structure is present even in the degree-based generator models.

A simple model for the Internet topology consisting of five layers determined by node degrees is given in [16]. They noticed a power-law in the connection of degree-one nodes to their neighbors, which is related to our observation of other power-laws in the structure of the Internet topology (see Section 4).

A specific behavior of the spectral density of different "real-world" graphs has been noticed in [8]. They also propose spectral analysis as a promising tool for network topology classification. Correlations among nodes in the real AS Internet graphs have been studied in [14].

[1] A power-law holds between two properties y and x if y is roughly proportional to x^c for some constant exponent c. If (x, y) data pairs are plotted with both axes in logarithmic scale, the resulting points lie close to a straight line with slope c.

1.2 Topology Data: Real and Synthetic

The Internet topology is usually represented as a graph. On the router level, individual hosts and routers are the nodes and physical connections between them are the edges, but it is difficult to obtain accurate snapshots of the Internet on this level.

On a more abstract level, the AS level, each node of the graph represents an autonomous system (AS, see [10]), i.e., a subnetwork under separate administrative control. An edge between two nodes means that there is at least one direct connection between the two AS domains.

AS-level topology data of the Internet can be inferred from BGP routing tables and is available on the NLANR website (http://moat.nlanr.net/AS/). In this work, we used snapshots of the AS topology from November 8, 1997 (a graph with 3015 nodes and 5156 edges) to March 16, 2001 (10515 nodes and 21455 edges) taken roughly every 3 months. We downloaded the corresponding files (ASconnlist.*) from the NLANR website and treat the graphs as simple, undirected graphs (i.e., we remove parallel links).

We are aware that the data from the NLANR website is potentially incomplete and inaccurate. However, we believe that the results of our analysis would not change drastically for more complete AS graphs. An alternative approach that determines the AS-level topology using router-level path traces was recently proposed in [4], but the coverage of their graphs is only around 60%. Since we are interested in using real AS graphs with large coverage, we found it more appropriate to use the AS-level topologies obtained from BGP routing data as explained above.

In order to compare properties of the AS graphs with graphs produced by a state-of-the-art network topology generator, we selected Inet-2.1. For each of the AS graphs we generate an Inet graph with the same number of nodes. Inet-2.1 allows to specify the fraction of vertices with degree 1. We specified this fraction identical to the one measured for the corresponding AS graph. Nevertheless, the Inet-2.1 generator produces graphs with a small amount of parallel edges. We removed those parallel edges since we deal specifically with simple, undirected graphs. As an effect, the fraction of nodes with degree 1 in these normalized Inet graphs was slightly higher than specified.

1.3 Outline

The remainder of the paper is structured as follows: In Section 2 we give the definitions and basic properties of the normalized Laplacian spectrum of a graph. Then we derive a lower bound on the multiplicity of eigenvalue 1 that turns out to be close to the real value on the AS graphs and Inet graphs. In Section 3, the quantities used in the computation of this lower bound lead us to a new structural classification of AS graphs that can be explained also in networking terms. In Section 4, statistics and comparisons based on the structural model are presented, and first steps towards a hybrid graph generation model are proposed. Finally, in Section 5, we summarize our results and discuss future work.

2 The Normalized Laplacian Spectrum

Previous studies in the context of network models mostly have considered the largest eigenvalues of the adjacency matrix of a graph, but it was noted in [12] that these eigenvalues seem to satisfy a power-law relationship for many different topologies. We propose to look not only at the largest eigenvalues, but at the (multi-)set of all eigenvalues, called the *spectrum*. In addition, we do not use the standard adjacency matrix, but the normalized Laplacian of the graph [5]. Among other reasons, this has the advantage that all eigenvalues are contained in the interval $[0, 2]$ (see [5]) so that it becomes easy to compare the spectra of different graphs even if the graphs have very different sizes.

Let $G = (V, E)$ be an undirected, simple graph, where V is the set of vertices and E is the set of edges. Let $\|V\| = n$, $\|E\| = m$, and d_v be the degree of node v.

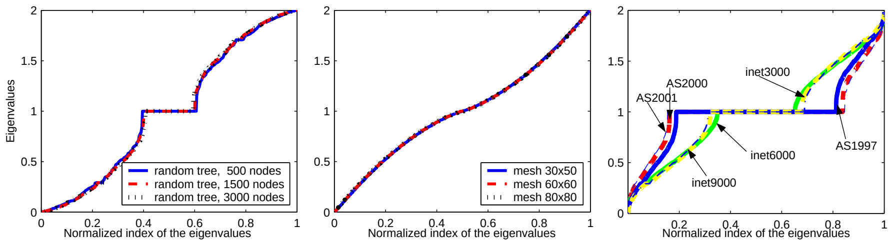
Definition 1. *The normalized Laplacian of the graph G is the matrix $\mathcal{L}(G)$ defined as follows:*

$$\mathcal{L}(G)(u,v) = \begin{cases} 1 & \textit{if } u = v \textit{ and } d_v \neq 0, \\ -\frac{1}{\sqrt{d_u d_v}} & \textit{if } u \textit{ and } v \textit{ are adjacent}, \\ 0 & \textit{otherwise}. \end{cases}$$

Note that if A is the adjacency matrix of the graph G (where $a_{ij} = 1$ if there is an edge between v_i and v_j, and $a_{ij} = 0$ otherwise) and D is a diagonal matrix having $d_{ii} = d_{v_i}$, then $\mathcal{L}(G) = D^{-\frac{1}{2}}(D - A)D^{-\frac{1}{2}}$. The normalized Laplacian spectrum (nls) is the set of eigenvalues of $\mathcal{L}(G)$, i.e., all values λ such that $\mathcal{L}(G)u = \lambda u$ for some $u \in \mathbb{R}^n, u \neq 0$. More about its characteristics can be found in [5].

A first natural question is: What is the normalized Laplacian spectrum for simple topologies such as stars, chains, grids and random trees on n nodes? The nls of a star S_n is 0, 1 (with multiplicity $n - 2$), 2, and the nls of a chain P_n is $1 - \cos(\frac{\pi k}{n-1})$, $k = 0, \ldots, n - 1$. For grids and trees, plots of the numerically computed spectrum are shown in Fig. 1 (a) and (b). To generate our nls plots, we compute all n eigenvalues with MATLAB, sort them in non-decreasing order, and plot them so that the i-th smallest eigenvalue λ_i, $1 \leq i \leq n$, is drawn at (x, y) with $x = (i - 1)/(n - 1)$ and $y = \lambda_i$. In this way, the plot is always within $[0, 1] \times [0, 2]$ and it becomes convenient to compare the nls of graphs with different numbers of nodes.

We found remarkably similar plots of the nls for all real Internet AS-level snapshots from November 1997 to March 2001 (see Fig. 1(c)). The same consistency was detected for Inet graphs with different numbers of nodes, but the nls of Inet graphs was clearly different from the nls of real AS graphs (also shown in Fig. 1(c)), in particular with respect to the multiplicity of eigenvalue 1. (There are other differences as well, but the investigation of these differences is work in progress.) Together with the known spectra of chains and stars and the plots of the nls of grids and random trees (Fig. 1 (a) and (b)), this indicates that the nls can be used as a kind of "fingerprint" for network topologies (or other arbitrary large graphs that are difficult to compare directly).



(a) (b) (c)

Fig. 1. (a) The *nls* of random trees that were generated by starting with one node and then repeatedly adding a new node and making it adjacent to an existing node chosen uniformly at random. (b) The *nls* of mesh graphs (grids). (c) The *nls* of AS and Inet-2.1 graphs.

2.1 A lower bound on the multiplicity of eigenvalue 1

AS graphs and Inet graphs both have a relatively large multiplicity of eigenvalue 1, but the multiplicity is considerably higher for AS graphs. This has motivated our interest in bounding the multiplicity of eigenvalue 1 in terms of structural properties of graphs.

We will use a technique of [9] (where the standard adjacency matrix is considered instead of $\mathcal{L}(G)$) to find a lower bound on the multiplicity of eigenvalue 1. Denote by $P(G) = \{v \in V \mid d_v = 1\}$ the set of leaves in G, called *pendants*, and by $Q(G) = \{v \in V \mid \exists w, (v,w) \in E, w \in P(G)\}$ the set of the neighbors of the leaves, called *quasi-pendants*. Let $R(G) = V \setminus (Q(G) \cup P(G))$ be the set of nodes that are not leaves and that are not neighbors of leaves, called *inners*. Let p, q, r respectively be the cardinalities of the sets $P(G)$, $Q(G)$, and $R(G)$. We call the subgraph of G induced by $R(G)$ $Inner(G)$. By *inn* we denote the number of isolated vertices in $Inner(G)$. Let $m_G(1)$ denote the multiplicity of the eigenvalue 1. Then we obtain the following lower bound.

Theorem 1. *The multiplicity of eigenvalue* 1 *of the normalized Laplacian is bounded from below by the sum of the number of pendants, the number of isolated inner nodes, and the negative of the number of quasi-pendants:*

$$m_G(1) \geq p - q + inn \tag{1}$$

Proof. We can assume the following labeling of the nodes, because the eigenvalues are independent of the labeling: $v_1, ..., v_n$ where $v_1, \cdots, v_r \in R(G)$, $v_{r+1}, \cdots, v_{r+q} \in Q(G)$, and $v_{r+q+1}, \cdots, v_n \in P(G)$. Also, we can assume that $(v_{r+i}, v_{r+q+i}) \in E, i = 1, \ldots, q$. Then, the structure of the normalized Laplacian is

$$\mathcal{L}(G) = \begin{pmatrix} R & rQ & 0 \\ rQ^T & Q & qP \\ 0 & qP^T & I_p \end{pmatrix}$$

Here R is an r-by-r matrix, rQ is r-by-q, Q is q-by-q, qP is q-by-p and I_p is the p-by-p identity matrix. From the basic equations $\lambda u = \mathcal{L}(G)u$, we obtain that $m_G(1) = \text{nullity}(\mathcal{L}(G) - I_n)$, where I_n is the n-by-n identity matrix.

Using the labeling assumptions, we observe that qP contains a principal submatrix D_q which is diagonal, having $-\frac{1}{\sqrt{d_{v_{r+i}}}}, i = 1, \ldots, q$ on the diagonal. Now let $LI(G) = \mathcal{L}(G) - I_n$. Using D_q and elementary transformations that do not change the rank (adding a multiple of one row to another row, or the same for columns), we obtain a new matrix $LI'(G)$ from $LI(G)$:

$$LI'(G) = \begin{pmatrix} R - I_r & 0 & 0 & 0 \\ 0 & 0 & D_q & 0 \\ 0 & D_q^T & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Now it is enough to prove that $\text{nullity}(LI'(G)) = n - \text{rank}(LI'(G)) \geq p - q + inn$. We have that $\text{rank}(D_q) = q$, thus $n - \text{rank}(LI'(G)) \geq n - 2q - \text{rank}(R - I_r) = p - q + r - \text{rank}(R - I_r)$. Now, if inn is the number of isolated vertices in $Inner(G)$, each row that contains an isolated vertex will have 0 at the first r columns of $LI'(G)$, thus $\text{rank}(R - I_r) \leq r - inn$ and the statement follows. □

3 A new structural classification of AS nodes

The lower bound of Theorem 1 is given in terms of pendants, quasi-pendants, and isolated inner nodes. We found that this lower bound is close to the real multiplicity observed in the AS graphs and Inet graphs. Therefore, we classify the nodes of the graphs into sets P, Q, R and I as follows, and investigate their cardinalities. A node is in P if its degree is 1 (i.e., if it is a leaf) and in Q if it has at least one neighbor in P. Let $Inner(G)$ be the subgraph of G induced by nodes not being in P or Q. A node in $Inner(G)$ is in I if it is an isolated node in $Inner(G)$ and in R otherwise (i.e., if it is contained in a connected component of $Inner(G)$ with at least 2 nodes).

3.1 Physical interpretation

The classes P, Q, R and I are defined in graph-theoretic terms motivated by Theorem 1. To relate these notions to ASs in the real Internet, we now propose plausible interpretations of the four node sets in networking terms. Further evidence supporting our interpretations can be derived from AS name tables and is given in the technical report [17].

Q nodes, best-connected nodes of the Internet. The class Q contains only a small number of nodes compared to the size of the whole graph and to the sizes of the other classes, but the best-connected nodes (largest degree) belong to Q. The subgraph induced by Q nodes has a similar structure for all observed graphs: it contains a big connected component with a characteristic *nls* (see [17]), and about 5% of isolated nodes.

We interpret the nodes in the big connected component of Q as core nodes. Note that Q nodes have leaf neighbors by definition. The isolated nodes, which have no Q neighbors, can be explained as exchange points serving to connect P, R, and I nodes.

***R* nodes, core and alliances.** The subgraph induced by R consists of a larger number of connected components. Their size and frequency exhibit power-law relationships (see Section 4). The biggest connected component dominates by its size and node degrees. We interpret the connected components of R nodes, with the exception of the biggest one, as *AS alliances*. Alliances can be built on national, regional, commercial, or other grounds.

In order to gain further insights, we investigated the k-cores of the AS graphs. The k-core of a graph is defined to be the subgraph obtained by recursively deleting all nodes with degree less than k. Intuitively, the deeper cores of an AS graph (i.e., the k-cores for larger values of k, say $k \geq 5$) should correspond roughly to more well-connected "backbones" of the Internet at that time. We found that most of the nodes in the deeper cores are in Q and some are members of the biggest R component. This fact motivates an interpretation of the biggest component of R as being made up partly of AS domains belonging to the core and partly of multi-homed stubs or alliances. Note that core nodes in R do not have any leaf neighbors.

***P* and *I* nodes, stub domains.** P nodes are leaves (nodes with degree 1) by definition. Therefore, they must be stub nodes (nodes that do not forward traffic that neither originates in that node nor is destined for that node). I nodes, whose degree is small in most cases (i.e., they have just two or three neighbors in Q), are mostly multi-homed stub domains. The percentage of I nodes is increasing in the AS graphs over time, and the percentage of P nodes is decreasing. Currently, the I class is the biggest part of the Internet. It became bigger than the P class at the beginning of 2001. This positive trend in the number of I nodes and negative trend in the number of P nodes agrees with the fact that more and more leaf domains want to become multi-homed for better fault-tolerance [6].

The number of I nodes with higher degree (degree about 10) is rather small. These nodes are mainly big companies or universities with multiple connections to the backbone, but not providing forwarding services.

In summary, the classes P and I represent the outskirts of the Internet, and in a sense they correspond to the stub domains in the Transit-Stub model. Although the smaller alliances in R are arguably also the outskirts of the Internet, for the sake of simplicity and since some nodes in the biggest R component are found in the deeper cores of the AS graphs, from now on we refer to $Q \cup R$ as the *core* and to $P \cup I$ as the *edge* of the Internet.

4 Statistics and Comparisons

Having proposed a structural classification of AS nodes in the previous section, now we investigate how the Internet evolves in this structural sense and how it compares to graphs generated by Inet. Through our analysis, we find that both structural and power-law properties are important. We also observe new power-laws in the internal structure of R and the interconnection between P and Q. These findings represent first steps towards a hybrid generator model that encompasses both power-laws and structural properties.

4.1 Methodology

The set of graphs undergone analysis has been described earlier in Section 1.2. To observe how AS graphs evolve structurally and how they compare to Inet graphs, we look at the following three sets of metrics.

1. Ratio of nodes in P, I, Q, R
2. Ratio of links connecting PQ, IQ, QQ, RQ, RR
3. Average node degree (number of neighbors) in the whole graph, average node degree of the edge nodes ($P \cup I$), and average node degree within the core (subgraph induced by $Q \cup R$).

It is not to our surprise that the numbers of nodes in each component and the numbers of links inter-connecting the components are increasing. Thus we focus our analysis on how the components expand or shrink relatively to the whole graph and omit discussions on absolute numbers.

Results are depicted in Figs. 4, 5, and 6, respectively, which can be found at the end of this paper. Each plot shows changes of the AS graphs in the corresponding metric (Y axis) in time (X axis). In the next two subsections, we highlight important trends in the evolution of AS graphs and significant differences to Inet graphs.

4.2 Evolution of AS graphs

Observation 1. We see from the top plots of Fig. 4 that the ratio of nodes in P is decreasing while that of I is increasing. That means, applying our interpretation of P and I, the area of single-homed leaf nodes is shrinking while that of multi-homed stub nodes is rapidly increasing.
Observation 2. The ratio of nodes in P and I combined has increased from approximately 67% to 74%. The ratio of nodes in R remains stable. The combined ratio of Q and R components, containing the core of the Internet, decreases from 33% to 26% (bottom plots of Fig. 4). In terms of ratio of nodes, the edge of the Internet is growing faster than the core.
Observation 3. Given that the ratio of nodes in P is decreasing, it is not surprising to see the ratio of links (edges) interconnecting QP decrease, and similarly to see the ratio of links QI increase as I increases (top plots of Fig. 5).

QP and QI combined increases from 54% to 60%. QQ, QR, and RR each remains relatively stable and the combined ratio decreases from 46% to 40%. This shows that in terms of the ratio of links, the edge of the Internet is again growing faster than the core.

Observation 4. More interestingly, the ratio of links in Q and R decreases by 6%, which is slightly lower than the 7% decrease in ratio of nodes. The ratio of links out-growing the ratio of nodes in the core indicates that the average node degree within the core (subgraph induced by $Q \cup R$) is increasing. This can be confirmed by the middle plot of Fig. 6.

4.3 Comparison to Inet graphs

Difference 1. While we see the ratios of I and QI expand in AS graphs, they remain stable in Inet graphs (top-right plot of Fig. 4 and 5). This is also reflected in the left plot of Fig. 6, where we observe that the average number of neighbors of a node in the edge components (P and I) is increasing in AS graphs whereas it remains stable in Inet graphs.

Difference 2. The ratio of R remains stable in AS graphs while it expands significantly in Inet graphs (bottom-right plot of Fig. 4). Similar contrasts can be found in link statistics in Fig. 5: QR (bottom-middle) and RR (bottom-right).

Difference 3. Inet graphs, similar to AS graphs, have rather stable Q and QQ ratios, but the ratio level is higher than that in AS graphs (bottom-left plot of Fig. 4 and 5).

Difference 2 and 3, contrary to the evolution of AS graphs, indicate that the core of Inet graphs is not only larger but also expanding while the edge is losing its ground.

Difference 4. It appears that in the core of Inet graphs the ratio of links is growing just as fast as the ratio of nodes, thus resulting in a rather constant average degree within the core (middle plot of Fig. 6). This is in contrast with the average degree within the core of AS graphs, which is increasing significantly over time.

These differences show that although successful in modeling some of the power-law properties, Inet fails to capture structural changes in AS graphs. In particular, the core of the Internet is becoming relatively smaller and the edge larger, but the evolution of Inet graphs shows the exact opposite. More interestingly, the Internet is structured so that the average node degree is increasing within the core as well as for edge nodes. On the contrary, these average degrees remain constant for Inet graphs.

4.4 New Power-Laws and First Steps Towards a Hybrid Model

These structural differences are potentially critical when studying properties of routing protocols, for example quality of path, route convergence, and fault tolerance. To be more concrete, one can expect global effectiveness of an alternative route computation algorithm to be different when evaluated using graphs with a higher ratio of multi-homed stub nodes (I) and better connected core (Q and R). Below we outline a structural model that will allow us to verify this conjecture in the future.

Our premise is to model AS graphs encompassing both statistical and structural properties. By that, we mean to 1) generate a degree sequence that obeys power-law properties observed earlier, 2) form P, I, Q, and R clouds based on the size dynamics, and finally 3) inter-/intra-connect these clouds with analysis of their connectivity.

First results show that a power-law exists not only in the overall degree sequence but also within the structure. For instance, while investigating ways to form the R cloud, we find that the sizes of the connected components in R have a power-law distribution. Fig. 2 shows the sizes of the connected components in R ranked from the largest to the smallest (rank-size plot), and occurrences of different sizes (size-occurrence plot).

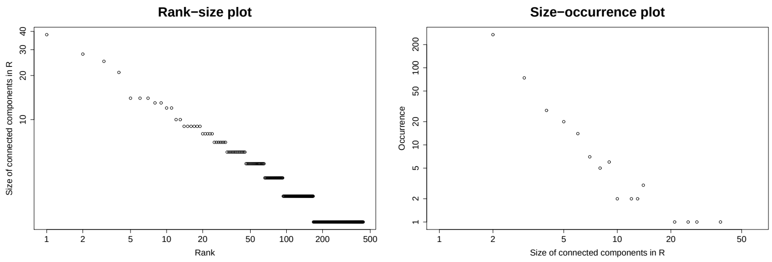


Fig. 2. Power-law in sizes of connected components in R

A power-law is also present in the way P and Q clouds are inter-connected. The left plot in Fig. 3 shows, for each node in Q, its rank on node degree (Y axis) and rank on number of degree-1 neighbors (X axis). Those Q nodes with one degree-one neighbor are ranked the 411th (lowest) in X axis; those with 2, 3, 4, and 5 degree 1 neighbors are ranked 275th, 191st, 147th, and 109th, gradually improving. Each column in the plot shows that a variety of Q nodes, with different node degrees, may have the same amount of P neighbors. The right plot of Fig. 3 illustrates the distribution of such Q nodes. The linear relationship in the log-log scale plot hints on a power-law distribution for Q nodes having one P neighbor.

This power-law property persists across different degree 1 neighbor ranks, until data points are too few to show a clear linear relationship in log-log plots. Further analysis verifies that the same power-law property between P and Q clouds exists in other snapshots of the Internet.

Another power-law in the interconnection of P and Q (relating the number of degree-1 neighbors of a Q node to the rank of that node in the sorted list of all Q nodes) has been discovered independently in [16]. These power-law properties within or between sub-structures of the AS graphs could be the missing links

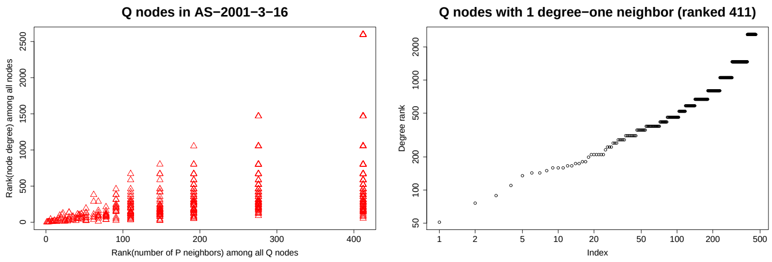


Fig. 3. Power-law in PQ inter-connection

between the state-of-the-art generators, i.e., Inet and GT-ITM. Each of them identifies one important aspect of the Internet topology — Inet for the statistical aspect and GT-ITM for the structural aspect — but unfortunately misses out on the other.

5 Conclusions and Future Work

We have investigated the normalized Laplacian spectrum of Internet topology graphs. It turned out that the nls can be used as a concise fingerprint of a graph. Real AS graphs from 1997 to 2001 produced nearly identical nls plots in spite of the significant difference in the number of nodes. Similarly, graphs generated with the Inet-2.1 generator had a characteristic nls, but different from the nls of real AS graphs, in particular with respect to the multiplicity of eigenvalue 1. We gave a lower bound on the multiplicity of eigenvalue 1 in terms of the cardinalities of different node sets P, Q, R, and I. For the real AS graphs, we found plausible interpretations for the nodes in P, Q, R, and I. In particular, the classification of nodes into these four types provides a structured view of the graph, featuring leaf domains, multi-homed stub domains, alliances, and core nodes.

Besides, while analyzing the subgraphs induced by P, Q, R and I and their interconnections we obtained new, previously unobserved power-law relationships with respect to the sizes of connected components in R and with respect to the degree rank of Q nodes with the same number of P neighbors.

Our results provide several immediate starting points for future work. First of all, we intend to use the insights we obtained from our analysis for improving the quality of the Inet-2.1 generator (with respect to the similarity to real AS graphs) and to explore the potential of a newly designed random topology generator based on our structural view in terms of P, Q, R and I nodes. We are aware that many aspects of the AS graphs are yet to be investigated thoroughly, for example the structure of the connections within Q and those between Q and R.

Furthermore, it will be interesting to identify additional characteristics of the nls (e.g. occurence of other multiplicities, convex and concave regions) that can be interpreted in terms of graph theory or networking concepts.

Finally, it is important to investigate the differences with respect to the *behavior* of a communication network (in terms of performance metrics such as throughput, delay, fault-tolerance) that arise from differences observed in the corresponding topology graphs. Thorough case studies could help to identify which of the graph properties have substantial effects on network performance and which are only of theoretical interest and do not affect performance.

References

1. A. Barabási and R. Albert. Emergence of scaling in random networks. *Science*, 286:509–512, 1999.
2. T. Bu and D. Towsley. On distinguishing between Internet power law topology generators. In *INFOCOM'02*, 2002. To appear.
3. K. Calvert, M. Doar, and E. W. Zegura. Modeling Internet topology. *IEEE Communications Magazine*, June 1997.
4. H. Chang, S. Jamin, and W. Willinger. Inferring AS-level Internet topology from router-level path traces. In *SPIE ITCom 2001*, Denver, CO, August 2001.
5. F. R. K. Chung. *Spectral Graph Theory*. Conference Board of the Mathematical Sciences, Providence, Rhode Island, 1997.
6. C. Diot. Sprint tier 1 IP backbone: Architecture, traffic characteristics, and some measurement results, 2001. http://talk.epfl.ch/talkdetail.php?talkId=42.
7. M. Faloutsos, P. Faloutsos, and C. Faloutsos. On power-law relationships of the Internet topology. In *SIGCOMM'99*, 1999.
8. I. Farkas, I. Dernyi, A. Barabási, and T. Vicsek. Spectra of "real-world" graphs: Beyond the semicircle law. *Physical Review E*, 64, August 2001.
9. R. Grone, R. Merris, and V. Sunder. The Laplacian spectrum of a graph. *SIAM Journal on Matrix Analysis and Applications*, 11(2):218–238, April 1990.
10. J. Hawkinson and T. Bates. RFC 1930: Guidelines for creation, selection, and registration of an autonomous system (AS). IETF, March 1996.
11. C. Jin, Q. Chen, and S. Jamin. Inet topology generator. Technical Report CSE-TR-433-00, EECS Department, University of Michigan, 2000.
12. A. Medina, I. Matta, and J. Byers. On the origin of power laws in Internet topologies. *ACM Computer Communication Review*, 30(2):18–28, 2000. April 2000.
13. K. Park and H. Lee. On the effectiveness of route-based packet filtering for distributed DoS attack prevention in power-law internets. In *SIGCOMM 2001*, August 2001.
14. R. Pastor-Satorras, A. Vázquez, and A. Vespignani. Dynamical and correlation properties of the Internet. *Physical Review Letters*, 87(25):258701–1, 2001.
15. H. Tangmunarunkit, R. Govindan, S. Jamin, S. Shenker, and W. Willinger. Network topologies, power laws, and hierarchy. Technical Report 01-746, Computer Science Department, University of Southern California, 2001.
16. L. Tauro, C. Palmer, G. Siganos, and M. Faloutsos. A simple conceptual model for the Internet topology. In *Global Internet*, San Antonio, Texas, November 2001.
17. D. Vukadinović, P. Huang, and T. Erlebach. Spectral analysis of Internet graphs. Technical Report 118, TIK Institute, ETH Zürich, 2001.
18. B. M. Waxman. Routing of multipoint connections. *IEEE Journal on Selected Areas in Communications*, 6(9):1617–1622, 1988.

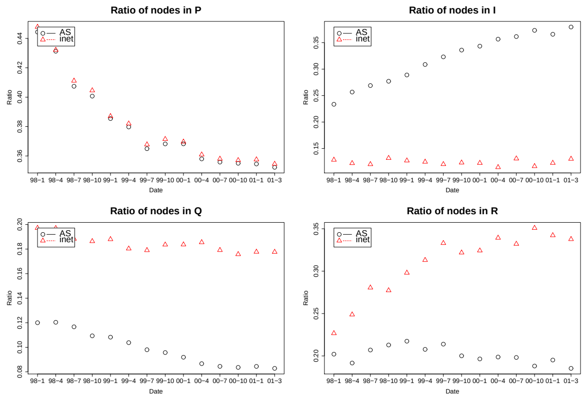


Fig. 4. Ratio of nodes in P, I (top), Q, R (bottom)

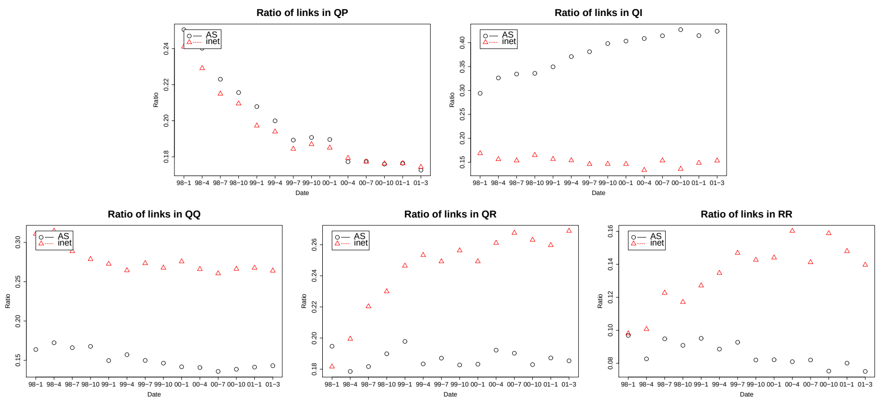


Fig. 5. Ratio of links in QP, QI (top), QQ, QR, RR (bottom)

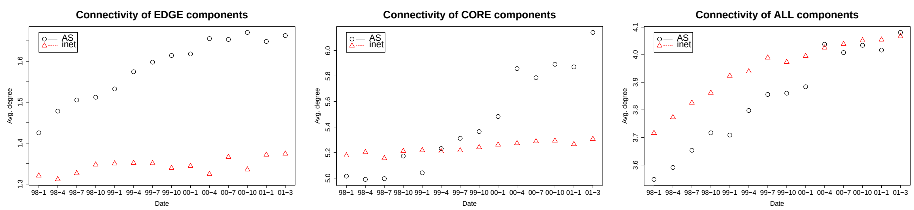


Fig. 6. Average node degree of edge nodes, within the core, and in the whole graph