Weighted Spectral Distribution for Internet Topology Analysis: Theory and Applications

Damien Fay, Hamed Haddadi, Andrew Thomason, Andrew W. Moore,
Richard Mortier, Almerima Jamakovic, Steve Uhlig, Miguel Rio

Abstract—Comparing graphs to determine the level of underlying structural similarity between them is a widely encountered problem in computer science. It is particularly relevant to the study of Internet topologies, such as the generation of synthetic topologies to represent the Internet's AS topology. We derive a new metric that enables exactly such a structural comparison, the *weighted spectral distribution*. We then apply this metric to three aspects of the study of the Internet's AS topology. (*i*) we use it to quantify the effect of changing the mixing properties of a simple synthetic network generator. (*ii*) we use this quantitative understanding to examine the evolution of the Internet's AS topology over approximately 7 years, finding that the distinction between the Internet core and periphery has blurred over time. (*iii*) we use the metric to derive optimal parameterizations of several widely used AS topology generators with respect to a large-scale measurement of the real AS topology.

Index Terms—Internet topology, Topology generation, Spectral graph theory, Graph metrics

I. INTRODUCTION

Graph comparison is a problem that occurs in many branches of computing, from vision to speech processing to systems. Many techniques exist for graph comparison, e.g., the edit distance [1] (the number of link and node additions required to turn one graph into another), or counting the number of common substructures in two graphs [2]. However, for large graphs such as the AS topologies examined here, these methods are computationally too expensive. In addition, they are inappropriate for dynamic graphs, resulting in varying edit distances or substructure counts. Instead, we require a metric which reflects the *structure* of large graphs in some meaningful sense. Typical currently used "metrics" include the clustering coefficient, the assortativity coefficient, the node degree distribution and the k-core decomposition. However, these are not metrics in the mathematical sense, but rather are measures, e.g., two graphs may have the same clustering coefficient but hugely different structures. This distinction is important as *a measure cannot be used to determine unique differences between graphs* : two graphs with the same measures may not in fact be the same.

In this paper we present a new metric, the *weighted spectral distribution* (WSD), which compares graphs based on the distribution of a decomposition of their structure. Specifically, the WSD is based on the spectrum of the normalised Laplacian matrix and is thus strongly associated with the distribution of *random walk cycles* in a network (as will be shown in Section III). The probability of randomly walking N steps from a node such that we return to that node, indicates the connectivity of that node. Hence, a low probability indicates high connectivity (there are many routes, few of which return) while a high probability indicates high clustering (many of the routes lead back to the original node).

The WSD is computationally inexpensive and so can be applied to very large graphs (more than 30,000 nodes and 200,000 edges). Also, it expresses the graph structure as a simply plotted curve which can be related to specific properties of AS graphs: hierarchy and local connectivity. Given that the WSD is a metric in the mathematical sense, as we show in Section III, several applications become possible: parameter estimation for topology generators with respect to a target dataset (Section V-C), direct comparison among topology generators using these optimal parameters (Table II), and quantification of change in the underlying structure of the Internet as it evolves over time (Section V-B). This metric is also a useful tool to evaluate the graphs that describe synthetic workloads generated from trace data. In such cases the generated graphs should *not* exactly match the original trace data, but should share some common structure with them. Such situations are encountered in workload generation and Internet topology generation.

In this paper we focus on applications of the WSD to the study of the Internet's AS topology. An AS represents a single network which can apply its own operational and peering policy. An Internet Service Provider (ISP) may use 1 or more ASes. The Internet contains over 28,000 ASes, each in a set of relationships with its neighbors, who are either its customers, providers or peers. In the Internet's core there is a full mesh formed between the ASes of the various tier-1 ISPs. However, at the edges there are a huge number of smaller ISPs and customer networks which connect through upstream providers and local public exchange points. These smaller ISPs and customer networks may have only one upstream provider, or may have many for resilience and performance reasons. In addition, the Internet constantly evolves; new networks are added, old ones disappear and existing ones grow and merge.

Links between ASes depend on business relationships which can and do change, sometimes rapidly, making any interpretation of the Internet as a *static* structure inaccurate. This rich and dynamic structure makes it difficult for researchers to provide either a single, representative topological model, or a single graph metric that captures all characteristics of any topology. However, such a metric would make it possible to generate realistic synthetic topologies improving the accuracy of Internet-wide protocol simulations, and perhaps enabling the prediction of the future evolution of the Internet's topology.

Many attempts to capture one or more characteristics have been made, resulting in several topology generators which each synthesize Internet-like topologies using different models and parameters. Unfortunately, validating these models is an *ad hoc* affair that typically means matching several of the measures given above and hoping that this will also ensure a matching structure. Users often select default parameters for these models based on specific datasets measured at particular times, and which no longer represent today's Internet. However, as noted previously, these measures cannot be used to estimate the optimum parameters for a model given a target topology.

This paper makes the following contributions:

(*i*) a spectral metric and a strawman model for comparing the structure of large graphs;

(*ii*) the analysis of more than 7 years of the evolution of the Internet AS topology seen from two different measurement techniques;

(*iii*) a comparison among the outputs of five major Internet topology generators and a measured dataset; and

(*iv*) optimal parameter estimation of said topology generators with respect to the measured dataset using our metric.

We proceed in Section II by reviewing related work on graph matching, spectral analysis of networks and Internet analysis. In Section III we present the necessary theoretical background, introducing the concepts on which we base our metric, before deriving the metric itself. In Section V we demonstrate use of the weighted spectral distribution in three distinct applications: comparing the structure of large graphs, quantifying the evolution of the AS topology over 7 years, and comparing and estimating optimal parameters for 5 widely used topology generators. We conclude in Section VI.

II. RELATED WORK

In this section we outline related work, classifying it into three groups: spectral graph analysis and the closely related WSD, evolution of the AS topology, and analysis of the clustering features of the AS topology.

The graph spectrum has been used for a variety of purposes in addition to characterization of Internet topologies, including space comparison [3], graph matching [4], cluster identification [5] and topology generator tuning [6]. Gkantsidis *et al.* [7] perform a comparison of clustering coefficients using the eigenvectors of the k largest eigenvalues of the adjacency matrices of AS topologies. k is chosen to retain the strongest eigenvectors discarding most of the others. Those retained are then shown to represent finer elements of the Internet structure.

The rest of the spectrum is considered unimportant, even though other works have shown that the eigenvalues of the adjacency matrix or the normalized Laplacian matrix can be used to accurately represent a topology [8], and some specific eigenvalues provide a measure of properties such as robustness of a network to failures [9].

Vukadinovic *et al.* [10] were the first to investigate the properties of the AS topology based on the normalized Laplacian spectrum. They observe that the normalized Laplacian spectrum can be used to distinguish between synthetic topologies generated by Inet [11] and AS topologies extracted from BGP data. These results indicates that the normalized Laplacian spectrum reveals important structural properties of the AS topology. However, as noted by Haddadi *et al.* [6], the spectrum *alone* cannot be used directly to compare graphs as it contains too detailed information about the network structure. We expand on this work by demonstrating how appropriate weighting of the eigenvalues can reveal the structural differences between two topologies. Perhaps the closest metric to the WSD is the fast graph kernel method proposed by Vishwanathan*et al.* [12]. Graph kernels are similar to the WSD in that they count random walks in networks but differ in the means by which they do so.

Shyu *et al.* [13] study the evolution of a set of topological metrics computed on a set of observed AS topologies. The authors rely on monthly snapshots extracted from BGP RouteViews from 1999 to 2006. The topological metrics they study are the average degree, average path length, node degree, expansion, resilience, distortion, link value, and the Normalized Laplacian Spectrum. They find that the metrics are not stable over time, except for the Normalized Laplacian Spectrum.

Latapy and Magnien [14] address the question of studying the relation between the size of a measurement sample and the corresponding topological properties. Based on AS topologies built from IP-level measurements from Skitter for a period from January 2005 to May 2006, they observe an increase in the average degree and the clustering coefficient when a larger dataset is used.

Wang and Loguinov [15] propose the Wealth-Based Internet Topology (WIT) model. Interestingly, central to their model is the notion that each AS picks its connections to maximise local random walks. This characteristic of the structure of the AS topology is particularly targeted by the WSD. However, as this model is not publicly available it is not included in our comparisons (Section V).

Wool and Sagie [16] propose several clustering algorithms to explore the AS topology using just a snapshot of the Skitter data. They focus on identification of the dominant clusters, although their result is sensitive to parameters such as the minimum cluster size. Our technique, the WSD, differs in that it focuses on random cycles instead of clusters and does not require any parameter estimation. In addition, we use the k-core decomposition to analyse the core of the Internet AS topology.

Li *et al.* [17] perform a similar study to the one presented here. In their work they use several different clustering methods to identify the distribution of clustering features

throughout a network. Interestingly, their clustering metric gives similar results for the Skitter and UCLA datasets, while WSD shows differing results reflecting directly the differing sampling characteristics of these two measurement techniques. The WSD also allows us to obtain a "best fit" prior to comparing topology generators with the observed datasets.

III. THEORETICAL BACKGROUND

We now derive our metric, the *weighted spectral distribution*, relating it to another common structural metric, the clustering coefficient, before showing how it characterises networks with different mixing properties.

Denote an undirected graph as $G = (V, E)$ where V is the set of vertices (nodes) and E is the set of edges (links). The adjacency matrix of G, $A(G)$, has an entry of one if two nodes, u and v, are connected and zero otherwise

$$A(G)(u,v) = \begin{cases} 1, & \text{if } u, v \text{ are connected} \\ 0, & \text{if } u, v \text{ are not connected} \end{cases} \quad (1)$$

Let d_v be the degree of node v and $D = diag(sum(A))$ be the diagonal matrix having the degrees along its diagonal. Denoting by I the identity matrix $(I)_{i,j} = 1$ if $i = j, 0$ otherwise, the Normalised Laplacian L associated with graph G is constructed from A by normalising the entries of A by the node degrees of A as

$$L(G) = I - D^{-1/2}AD^{-1/2} \quad (2)$$

or equivalently

$$L(G)(u,v) = \begin{cases} 1, & \text{if } u = v \text{ and } d_v \neq 0 \\ -\dfrac{1}{\sqrt{d_u d_v}}, & \text{if } u \text{ and } v \text{ are adjacent} \\ 0, & \text{otherwise} \end{cases} \quad (3)$$

As L is a real symmetric matrix there is an orthonormal basis of real eigenvectors $e_0, \ldots, e_{n-1}$ (i.e., $e_i e_j^T = 0, i \neq j$ and $e_i e_i^T = 1$) with associated eigenvalues $\lambda_0, \ldots, \lambda_{n-1}$. It is convenient to label these so that $\lambda_0 \leq \ldots \leq \lambda_{n-1}$. The set of pairs (eigenvectors and eigenvalues of L) is called the spectrum of the graph. It can be seen that

$$L(G) = \sum_i \lambda_i e_i e_i^T \quad (4)$$

The eigenvalues $\lambda_0, \ldots, \lambda_{n-1}$ represent the strength of projection of the matrix onto the basis elements. This may be viewed from a statistical point of view [18] where each $\lambda_i e_i e_i^T$ may be used to approximate $A(G)$ with approximation error inversely proportional to $1 - \lambda_i$. However, for a graph, those nodes which are best approximated by $\lambda_i e_i e_i^T$ in fact form a cluster of nodes. This is the basis for spectral clustering, a technique which uses the eigenvectors of L to perform clustering of a dataset or graph [19]. The first (smallest) non-zero eigenvalue and associated eigenvector are associated with the main clusters of data. Subsequent eigenvalues and eigenvectors can be associated with cluster splitting and also identification of smaller clusters [5]. Typically, there exists what is called a *spectral gap* in which for some k and j, $\lambda_k \ll \lambda_{k+1} \approx 1 \approx \lambda_{j-1} \ll \lambda_j$. That is, eigenvalues

$\lambda_{k+1}, \ldots, \lambda_{j-1}$[1] are approximately equal to one and are likely to represent links in a graph which do not belong to any particular cluster. It is then usual to reduce the dimensionality of the data using an approximation based on the spectral decomposition. However, in this paper we are interested in representing the global structure of a graph (e.g., we are interested in the presence or absence of many small clusters), which is essentially the spread of clustering across the graph. This information is contained in all the eigenvalues of the spectral decomposition.

Let $x = (x_0, \ldots, x_{n-1})$ be a vector. From (3) we see that

$$xLx^T = \sum_{uv \in E} (x_u/\sqrt{d_u} - x_v/\sqrt{d_v})^2 \quad (5)$$

Now, the eigenvalues cannot be large because from (5) we obtain

$$\begin{aligned} xLx^T &\leq \sum_{uv \in E} (x_u/\sqrt{d_u} - x_v/\sqrt{d_v})^2 \\ &\qquad + (x_u/\sqrt{d_u} + x_v/\sqrt{d_v})^2 \\ &= 2\sum_u x_u^2 = 2xx^T \end{aligned} \quad (6)$$

and so $\lambda_i = e_i L e_i^T \leq 2$. What is more, the mean of the eigenvalues is 1 because

$$\sum_i \lambda_i = \text{tr}(L) = n \quad (7)$$

by (3), where $\text{tr}(L)$ is the *trace* of L.

To summarize: the eigenvalues of L lie in the range 0 to 2 (the smallest being 0), i.e., $0 = \lambda_0 \leq \ldots \leq \lambda_{n-1} \leq 2$, and their mean is 1.

The distribution of the n numbers $\lambda_0, \ldots, \lambda_{n-1}$ contains useful information about the network, as will be seen. In turn, information about this distribution is given by its moments in the statistical sense, where the Nth moment is $1/n \sum_i (1 - \lambda_i)^N$. These moments have a direct physical interpretation in terms of the network, as follows. Writing B for the matrix $D^{-1/2}AD^{-1/2}$, so that $L = I - B$, then by (3) the entries of B are given by

$$(D^{-1/2}AD^{-1/2})_{i,j} = \frac{A_{i,j}}{\sqrt{d_i}\sqrt{d_j}} \quad (8)$$

Now the numbers $1 - \lambda_i$ are the eigenvalues of $B = I - L$, and so $\sum_i (1-\lambda_i)^N$ is just $\text{tr}(B^N)$. Writing $b_{i,j}$ for the (i,j)-th entry of B, the (i,j)-th entry of B^N is the sum of all products $b_{i_0,i_1}b_{i_1,i_2} \ldots b_{i_{N-1}i_N}$ where $i_0 = i$ and $i_N = j$. But $b_{i,j}$, as given by (8), is zero unless nodes i and j are adjacent. So we define an N-cycle in G to be a sequence of vertices $u_1 u_2 \ldots u_N$ with u_i adjacent to u_{i+1} for $i = 1, \ldots, N-1$ and with u_N adjacent to u_1. (Thus, for example, a triangle in G with vertices set $\{a, b, c\}$ gives rise to six 3-cycles abc, acb, bca, bac, cab and cba. Note that, in general, an N-cycle might have repeated vertices.) We now have

$$\sum_i (1 - \lambda_i)^N = \text{tr}(B^N) = \sum_C \frac{1}{d_{u_1} d_{u_2} \ldots d_{u_N}} \quad (9)$$

---

[1] i.e., the eigenvalues at the centre of the spectrum.

the sum being over all N-cycles $C = u_1 u_2 \ldots u_N$ in G. Therefore, $\sum_i (1 - \lambda_i)^N$ counts the number of N-cycles, normalised by the degree of each node in the cycle.

The number of N-cycles is related to various graph properties. The number of 2-cycles is just (twice) the number of edges and the number of 3-cycles is (six times) the number of triangles. Hence $\sum_i (1 - \lambda)^3$ is related to the clustering coefficient, as discussed below. An important graph property is the number of 4-cycles. A graph which has the minimum number of 4-cycles, for a graph of its density, is quasi-random, i.e., it shares many of the properties of random graphs, including, typically, high connectivity, low diameter, having edges distributed uniformly through the graph, and so on. This statement is made precise in [20] and [21]. For regular graphs, (9) shows that the sum $\sum_i (1 - \lambda)^4$ is directly to the number of 4-cycles. In general, the sum counts the 4-cycles with weights: for the relationship between the sum and the quasi-randomness of the graph in the non-regular case, see the more detailed discussion in [22, Chapter 5]. The right hand side of (9) can also be seen in terms of random walks. A random walk starting at a vertex with degree d_u will choose an edge with probability $1/d_u$ and at the next vertex, say v, choose an edge with probability $1/d_v$ and so on. Thus the probability of starting and ending randomly at a vertex after N steps is the sum of the probabilities of all N-cycles that start and end at that vertex. In other words exactly the right hand side of (9). As pointed out in [15], random walks are an integral part of the Internet AS structure.

The left hand side of Equation (9) provides an interesting insight into graph structure. The right hand side is the sum of normalised N-cycles whereas the left hand side involves the spectral decomposition. We note in particular that the spectral gap is diminished because eigenvalues close to one are given a very low weighting compared to eigenvalues far from one. This is important as the eigenvalues in the spectral gap typically represent links in the network that do not belong to any specific cluster and are not therefore important parts of the larger structure of the network.

Next, we consider the well-known clustering coefficient. It should be noted that there is little connection between the clustering coefficient, and cluster identification, referred to above. The clustering coefficient, $\gamma(G)$, is defined as the average number of triangles divided by the total number of possible triangles

$$\gamma(G) = 1/n \sum_i \frac{T_i}{d_i(d_i - 1)/2}, d_i \geq 2 \qquad (10)$$

where T_i is the number of triangles for node i and d_i is the degree of node i. Now consider a specific triangle between nodes a, b and c. For the clustering coefficient, noting that the triangle will be considered three times, once from each node, the contribution to the average is

$$\frac{1}{d_a(d_a - 1)/2} + \frac{1}{d_b(d_b - 1)/2} + \frac{1}{d_c(d_c - 1)/2} \qquad (11)$$

However, for the weighted spectrum (with $N = 3$), this particular triangle gives rise to six 3-cycles and contributes

$$\frac{6}{d_a d_b d_c} \qquad (12)$$

So, it can be seen that the clustering coefficient normalises each triangle according to the total number of possible triangles while the weighted spectrum (with $N = 3$) instead normalises using a product of the degrees. Thus, the two metrics can be considered to be similar but not equal. Indeed, it should be noted that the clustering coefficient is in fact not a metric in the strict sense. While two networks can have the same clustering coefficient they may differ significantly in structure. In contrast, the elements of $\sum_i (1 - \lambda)^3$ will only agree if two networks are isomorphic.

We now formally define the *weighted spectrum* as the normalised sum of N-cycles as

$$W(G, N) = \sum_i (1 - \lambda_i)^N \qquad (13)$$

However, calculating the eigenvalues of a large (even sparse) matrix is computationally expensive. In addition, the aim here is to represent the *global* structure of a graph and so precise estimates of *all* the eigenvalue values are not required. Thus, the distribution[2] of eigenvalues is sufficient. In this paper the distribution of eigenvalues $f(\lambda = k)$ is estimated using pivoting and Sylvester's Law of Inertia to compute the number of eigenvalues that fall in a given interval. To estimate the distribution we use K equally spaced bins [3]. A measure of the graph can then be constructed by considering the distribution of the eigenvalues as

$$\omega(G, N) = \sum_{k \in K} (1 - k)^N f(\lambda = k) \qquad (14)$$

where the elements of $\omega(G, N)$ form the *weighted spectral distribution*:

$$WSD : G \rightarrow \Re^{|K|} \{k \in K : ((1 - k)^N f(\lambda = k))\} \qquad (15)$$

In addition, a metric can then be constructed from $\omega(G)$ for comparing two graphs, G_1 and G_2, as

$$\Im(G_1, G_2, N) = \sum_{k \in K} (1 - k)^N (f_1(\lambda = k) - f_2(\lambda = k))^2 \qquad (16)$$

where f_1 and f_2 are the eigenvalue distributions of G_1 and G_2 and the distribution of eigenvalues is estimated in the set K of bins $\in [0, 2]$. Equation (16) satisfies all the properties of a metric (see Appendix A).

We next wish to test if the WSD for graphs generated by the same underlying process vary significantly (to show that the WSD is stable). To do this, we generate a set of graphs that have very similar structure and test to see if their WSDs are also similar. The results of an empirical test are shown in Figure 1. This plot was created by generating 50 topologies using the AB [23] generator with the (fixed) optimum parameters determined in Section V-C, but with

---

[2] The eigenvalues of a given graph are deterministic and so *distribution* here is not meant in a statistical sense.

[3] K can be increased depending on the granularity required.
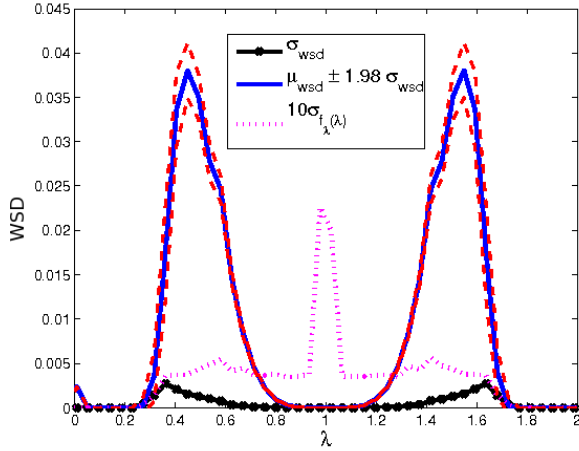
Fig. 1. Mean and standard deviations for WSD and (unweighted) spectrum for the AB model over 50 simulations.
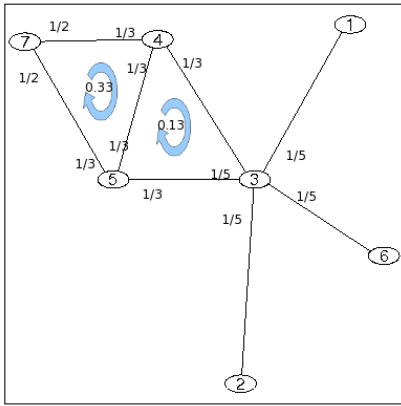

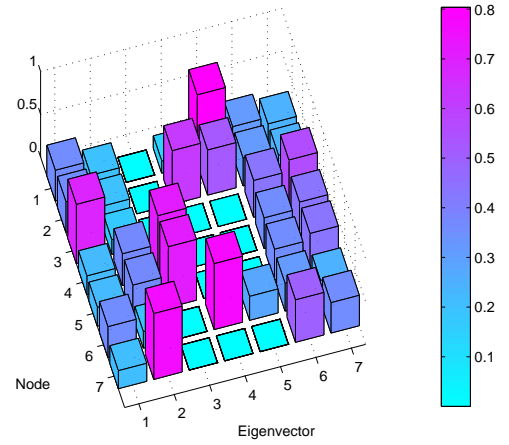
Fig. 2. A simple example network G_1.

Fig. 3. Eigenvectors of the simple example network.

TABLE I
EIGENVALUES, WSD AND DOMINANT NODES OF EXAMPLE NETWORK.

e_7	Eigenvector	λ	$1 - \lambda$	$(1 - \lambda)^3$	Dominant nodes
0.2500	1	1.8615	-0.8615	-0.6394	3,1,2,6
0.2500	2	1.3942	-0.3942	-0.0612	7,4,5
0.5590	3	1.3333	-0.3333	-0.0370	4,5
0.4330	4	1.0000	0.0000	0.0000	6,2
0.4330	5	1.0000	0.0000	0.0000	1,2,6
0.2500	6	0.4110	0.5890	0.2043	7,3
0.3536	7	0.0000	1.0000	1.0000	3,4,5,7
	$\sum_{i=1}^{7}(1-\lambda_i)^3$			0.4667	

different initial conditions [4]. For each run the spectral and weighted spectral distributions are recorded yielding 50×50 bin values which are then used to estimate standard deviations. As the underlying model (i.e. the AB generator) is the same for each run, the *structure* might be expected to remain the same and so a "structural metric" should be insensitive to random initial conditions. As can be seen the standard deviation[5] of the (unweighted) spectrum, $\sigma_{f_\lambda}(\lambda)$, is significantly higher at the centre of the spectrum. However, for the WSD, the standard deviation, σ_{wsd}, peaks at the same point as the WSD; the noise in the spectral gap has been suppressed. The evidence seems to show that the WSD is therefore stable.

IV. EXAMPLES

After the fairly theoretical previous section, we aim at giving the reader a better intuition behind the WSD with a simple example. Figure 2 shows a small network, called G_1, with 7

---

[4]We found similar results for other parameters and topology generators.

[5]Multiplied by a factor of ten for clarity.

nodes and 8 links. As can be seen there are 2 cycles of length 3 in this network and one of length 4. We will take $N = 3$ in this example for convenience and without loss of generality. The random walk probabilities are labeled in Figure 2. For example, node 3 has a degree of 5 resulting in a probability of $1/5^{th}$ for each edge. The total probability of taking a random walk around each 3-cycle is: $6 \times 1/2 \times 1/3 \times 1/3 = 0.33$, also shown.[6]

Figure 3 shows a 3-D plot of the absolute value (for clarity) of the eigenvectors of the normalized Laplacian. The corresponding eigenvalues are shown in Table I.

As is well known, the eigenvectors of the normalised Laplacian perform a partitioning of the nodes in a graph. In this example nodes 4 and 5 are grouped into eigenvector 3, nodes 1,2 and 6 into eigenvectors 4 and 5, node 7 into eigenvector 2 and node 3 into eigenvector 1 (Figure 3). Note that for each partition the nodes in the partition are the same; i.e. we could swap the labels between nodes 4 and 5 and the network would not change (i.e. an isomorphism). Eigenvector and eigenvalue 7, e_7 and $\lambda_7 = 0$, are special and partitions all the nodes in the network with the most central nodes having the highest coefficients (see Table I, column 1). In general the number of eigenvalues that are zero is equal to the number of components, arguably the most important structural property

---

[6]The six comes from the fact that the random walk can start in one of three nodes and go in one of two directions. It can be viewed in our case as really just a nuisance scaling factor.
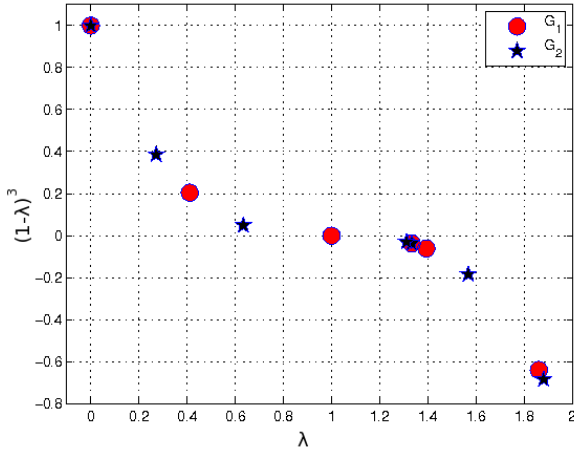
Fig. 4. WSD of the example network.

Fig. 5. The second example network, G_2.

in a graph. This graph contains 1 connected component and so has a single zero eigenvalue (λ_7). Note that the highest possible weighting in the WSD is given at zero (i.e. 1 = 1-0); the number of components in the graph.
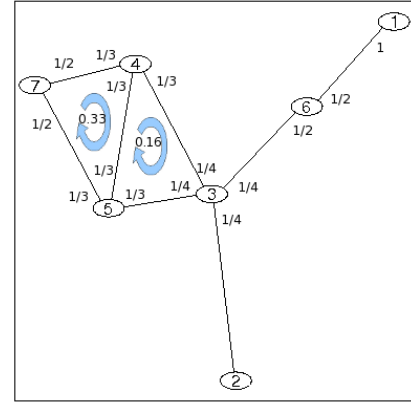
Note that the sum of the eigenvalues taken to the power of N is indeed the same as the sum of the probabilities of taking N random walk cycles in the graph. This is shown in Table I, last row, $\sum_{i=1}^{7}(1-\lambda_i)^3 = 0.4667$ which can be easily verified by adding the cycle probabilities from Figure 3 $(0.3333 + 0.1333 = 0.467)$. What is interesting is how this sum is constructed. In Table I the main contributions to the sum are from eigenvalues 1,2,3 and 6 (we ignore eigenvalue 7 as it merely reflects that the graph is connected) which are dominated by the nodes which form the cycles; 3, 4, 5 and 7.

However, this does not mean that the information provided by the WSD is confined to N-cycles in the graph. For example in Figure 5 we take the edge linking nodes 1 and 3 and rewire it so that 1 and 6 are now connected. Note that while the right cycle is still in place its probabilities have now changed, as the degree of node 3 is now 4. The corresponding eigenvalues have also changed as seen in Figure 4. [7]

In conclusion, the WSD can roughly be seen as an amalgamation of *local* views (i.e. walks of length N) taken from all the nodes. As $(1 - \lambda_i) \leq 1 \ \forall i$, $(1 - \lambda_i)^N$ will suppress the smaller eigenvalues more and more as N increases [8]. We consider 3 and 4 to be suitable values of N for the current application: $N = 3$ is related to the well-known and understood clustering coefficient; and $N = 4$ as a 4-cycle represents two routes (i.e., minimal redundancy) between two nodes. For other applications, other values of N may be of interest. Also note that in section III we propose using the *distribution* of the eigenvalues for large networks; unfortunaty it is not instructive to talk about a distribution for a small number of eigenvalues (7 in this example).

V. APPLICATIONS

In this section, we use the WSD to better understand the mixing properties of networks, the Internet's AS topology evolution, and the behavior of topology generators with respect to observed Internet AS topologies.

A. Mixing properties of networks

The synthetic topology generator introduced in this section is a strawman (demonstration) tool that can be adjusted to show the effect of different parts of a topology on the resulting WSD. It generates a set of topologies using a simple model based on the existence of a network core and a periphery.

Figure 6 shows a small topology of 500 nodes. All M nodes within the graph are first assigned locations using a uniform distribution. Nodes within a circle of diameter D are then defined as the *core* and nodes outside a circle of diameter $D \times (1 - m)$ as the periphery, where $m \leq 1$ is a factor called the mixing factor. Connections are then assigned between the core nodes using a Waxman model:

$$P(u \rightarrow v) = \alpha_{core} \exp^{\frac{-d\beta_{core}}{D}} \qquad (17)$$

where P denotes the probability that node u is connected to node v, α_{core} and β_{core} are the Waxman coefficients for the core, and d is the distance between two nodes u and v. Subsequently, connections are also assigned in the periphery[9] using a Waxman model with different coefficients, α_{per} and β_{per}. After this process, isolated nodes are connected to their nearest neighbor.[10] Figure 7 shows the WSD (using $N = 4$) for a topology generated with $M = 2000$ nodes, $D = 0.25$, $\alpha_{core} = 0.08$, $\beta_{core} = 0.08$, $\alpha_{per} = 0.06$, $\beta_{per} = 0.7$, and $m = 0.95$ (i.e., 5% mixing), resulting in a small (relatively) meshed core with a less well connected periphery. There are several things to note from this figure.

---

[7]Note that if we had used the adjacency matrix instead of the normalised Laplacian the re-wiring would have no effect on the sum of the eigenvalues.

[8]This is closely related to the settling times in Markov chains which are often expressed in terms of the largest non-trivial eigenvalue. It differs in that the Walk Laplacian and not the normalised Laplacian is used.

[9]Note that nodes lying between D and $D \times (1 - m)$ are members of the core *and* the periphery and can be assigned connections by both processes.

[10]Note that there are likely to be some disconnected components in the resulting graphs giving asymmetrical spectra, but this does not affect the main results.
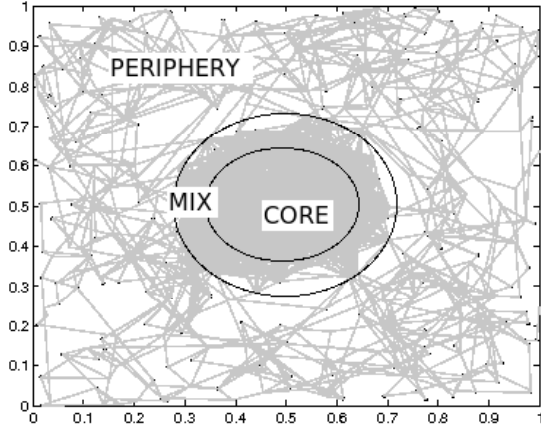
Fig. 6.   Synthetic topology example ($M = 2000$).

Fig. 7.   Synthetic topology spectra.

Ignoring the asymmetric part of the curve, which is due to a small number of disconnected components, the peak of the weighted spectrum of the periphery alone lies at $\lambda = 0.7$ while that for the core lies at 0.5. The spectrum for the overall network has *two peaks* at these points. This is a direct result of the fact that the spectrum of a graph is the union of the spectra of its disconnected subgraphs [22]. In terms of the WSD, the union of spectra is equivalent to a weighted average of the WSD. That is, for a graph $G + H$ composed of two disconnected subgraphs G and H:

$$\omega(G + H, N) = |G + H| \left(\frac{\omega(G, N)}{|G|} + \frac{\omega(H, N)}{|H|} \right) \quad (18)$$

where $|.|$ denotes graph volume (number of vertices). Although there is 5% mixing between the core and periphery $\omega(G + H, N)$ results in a close estimate of the network WSD (see Figure 7, denoted $\Sigma||E(1 - \lambda_i)^4||$). As $m \to 0$ (i.e., the core and periphery become less and less connected) this estimate becomes more accurate and is exact at $m = 0$.

Figure 8 shows the effect of increasing the mix between the periphery and the core.[11] As can be seen, the core becomes less distinct in the resulting spectrum, and has practically disappeared with 40% mixing. By increasing the mix we are effectively adding edges connecting the core and periphery, which results in a spreading of the eigenvalues and thus a spreading of the WSD, giving less distinct peaks. This result is a consequence of the following theorem from [24]:

Theorem 5.1: Let G be a weighted graph and H a subgraph on the vertices of G with t non-isolated vertices. If $\{\lambda_0 \leq \lambda_1 ... \leq \lambda_{m-1}\}$ and $\{\theta_0 \leq \theta_1 ... \leq \theta_{n-1}\}$ are the eigenvalues of $L(G)$ and $L(G + H)$ respectively, then for $k = 0, 1, \ldots, n - 1$ we have:

$$\lambda_{k+t-1} \leq \theta_k \leq \begin{cases} \lambda_{k-t+1}, & H \text{ is bipartite,} \\ \lambda_{k-t} & \text{otherwise} \end{cases} \quad (19)$$

In the current context, H is the subgraph containing the nodes connecting the core and periphery with t connecting edges

[11]Again the large peaks before 0.2 represent isolated subgraphs and are ignored.
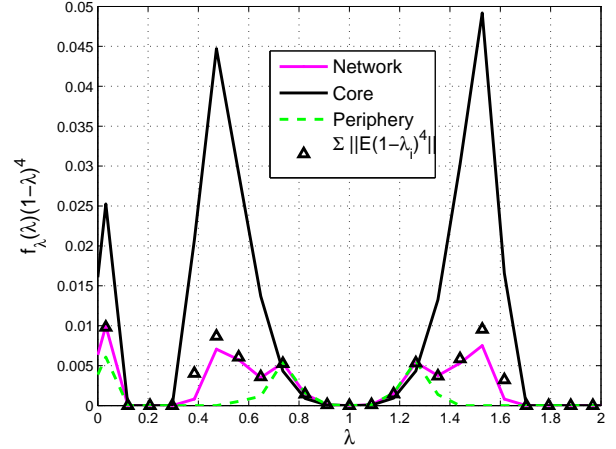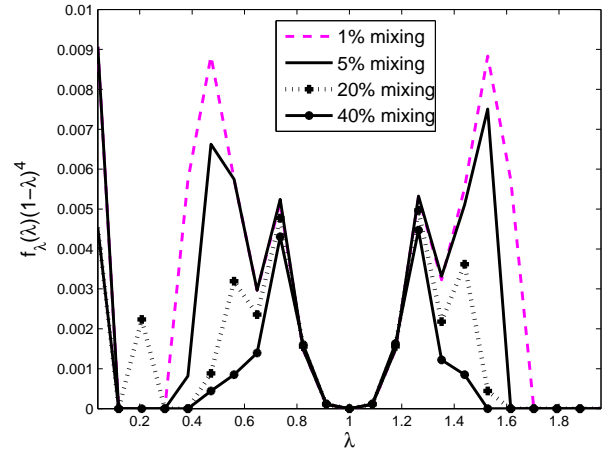

Fig. 8.   Effect of a change in m on the spectrum of the overall network.

(i.e. the mix). These edges cause the eigenvalues of $G - H$ to spread by by at most t places. Thus the core peak becomes less distinct. This reflects that the core is less distinct.

B. Evolution of the Internet

In this section two data sets measuring the Internet's AS topology over several years are compared using the WSD and standard measures. The first dataset we study consists of 7 years of traceroute measurements starting in January 2001 and ending in December 2007, collected by the CAIDA Skitter project [25]. Traceroutes are initiated from several locations in the world toward a large range of destination IP addresses. The IP addresses reported in the traceroutes are mapped to AS numbers using RouteViews BGP data. We use a monthly union of the set of all unambiguous links collected on a daily basis by the Skitter project[12].

The second dataset consists of 52 snapshots, one per month, from January 2004 to April 2008. This dataset, referred to as

[12]A link may be ambiguous for a variety of reasons, principally due to problems resolving an IP address to its AS number; we ignore such links.

the UCLA dataset, is available at the Internet topology collection[13] maintained by Oliveira *et al.* [26]. These topologies are built and updated daily using data sources such as BGP routing tables and updates from RouteViews, RIPE [14], Abilene[15] and LookingGlass servers.

Figure 9 shows the values of several graph measures for the Skitter and UCLA datasets. As can be seen there is an overlap between 2004 and 2007.

The number of ASes seen by Skitter exhibits abrupt changes during the first 40 months. At the end of those 40 months, changes were made in the way probing was performed.[16] The large increases in the number of ASes, observed during the first 40 months, are due to new monitors being added to the system. After each increase in the number of ASes a smooth decrease follows, corresponding to a subset of the IP addresses of the Skitter list that no longer respond to probes, e.g., because a firewall starts blocking the probes. The variations in the number of ASes seen by Skitter are probably not caused by changes in the AS topology itself, but rather by artifacts of the probing, e.g. firewalls that block probes or non-responding hosts.

The number of AS edges and the average node degree both follow the behavior of the number of ASes seen. We only observe a large increase in the number of links during the first few months, during which new monitors are added resulting in new regions of the Internet being covered by Skitter measurements. As the list of destinations used by Skitter does not cover the global set of ASes well [27], and the same list is shared by all monitors, a new monitor will typically discover new ASes close to its location. However, most of the AS edges close to the destination IP addresses have probably already been discovered by existing monitors [28].

The AS edges that Skitter no longer observes probably still exist but can no longer be seen by Skitter due to its shrinking probing scope. To be effective in observing topology dynamics, traceroute data collection must update destination lists constantly to give optimal AS coverage. This limitation of Skitter is visible in the decreasing average node degree. We would normally expect to see a net increase in the average node degree as ASes tend to add rather than remove peerings, and the results of the BGP data support this view. If the coverage of the Skitter measurements was not worsening, we should see an increasing node degree.

The lower three graphs of Figure 9 present the evolution of the clustering coefficient, the assortativity coefficient,[17] and the weighted spectrum with $N = 3$, $\omega(G, 3)$ (related to the topology's clustering)[18]. We observe that changes were made

to the way Skitter probes the Internet around month 40: the metrics take an unusual value, very small for the clustering and very high for assortativity. The values of the clustering and the assortativity coefficients fluctuate randomly over the 7 years, as if the sampling of the AS topology by Skitter is not stable. Neither the clustering nor the assortativity seem to decrease or increase over the 7 years. The value of $\omega(G, 3)$ shows a long-term increasing trend, similar to the decreasing trend in the average node degree. Although related to the clustering, $\omega(G, 3)$ gives different weights to different parts of the topology. The subset of the topology that corresponds to duplicated topological structures (e.g., the periphery) receives a smaller weight than the rest. The increasing $\omega(G, 3)$ is likely to be caused by the shrinking network sampled by Skitter, that contains more 3-cycles on average.

The UCLA AS topologies display a completely different evolution to the Skitter dataset, more consistent with expectation. As the three upper graphs of Figure 9 show, the number of ASes, AS edges, and the average node degree are all increasing, as expected in a growing Internet.

The increasing assortativity coefficient indicates that ASes increasingly peer with ASes of similar degree. The preferential attachment model seem to be less dominant over time. This trend towards a less disassortative network is consistent with more ASes bypassing the tier-1 providers through public IXPs [31], hence connecting with nodes of similar degree. Another possible explanation for the increasing assortativity is an improvement in the visibility of non-core edges in BGP data. Contrary to Skitter, $\omega(G, 3)$ for UCLA decreases over time. As a weighted clustering metric, $\omega(G, 3)$ indicates that the transit part of the AS topology is actually becoming sparser over time compared to the periphery. Increasing local peering with small ASes in order to reduce the traffic sent to providers decreases both the hierarchy induced by strict customer-provider relationships, and in turn decreases the number of 3-cycles on which $\omega(G, 3)$ is based.

To further investigate this result, we now introduce supporting evidence using k-cores. A k-core is defined as the maximum connected subgraph, H, of a graph, G, with the property that $d_v \geq k \ \forall v \in H$[19]. As pointed out by Alvarez-Hamelin *et al.* [32], the k-core exposes the structure of a graph by pruning nodes with successively higher degrees k and examining the maximum remaining subgraph. Note that this is not the same as simply pruning all nodes with degree k or less. Figure 10 shows the proportion of nodes in each k-core as a function of k for the Skitter topologies. There are 84 plots shown over the 7 years, but as can be seen there is little difference between each of them, demonstrating that the proportion of nodes in each core is not changing over time. This is not surprising due to the nature of the Skitter sampling process: the Skitter data set is composed of traceroutes rooted at a limited set of locations. The observed network is therefore tree-like and so the k-core typically just removes the leaves of the tree. [32] referred to this as being similar to *peeling the layers from an onion*. From an evolution point of view

---

[13]http://irl.cs.ucla.edu/topology/

[14]http://www.ripe.net/db/irr.html

[15]http://abilene.internet2.edu/

[16]These changes were subject to caveats and bugs affecting measurements, and, thus, the resulting metrics, at month 40. For more information we refer to http://www.caida.org/data/active/skitter_aslinks_dataset.xml/

[17]Assortativity is a measure of the likelihood of connection of nodes of similar degrees [29]. This is usually expressed by means of the *assortativity coefficient r*: assortative networks have $r > 0$ (disassortative have $r < 0$ resp.) and tend to have nodes that are connected to nodes with similar (dissimilar resp.) degree.

[18]See [30] for a detailed explanation on the mathematical measures and different datasets

[19]To take the k-core of a graph we first remove all nodes with degree 1. This might split the graph into two or more pieces, in which case we take the largest piece. We then remove all nodes of degree 2 and repeat k times.
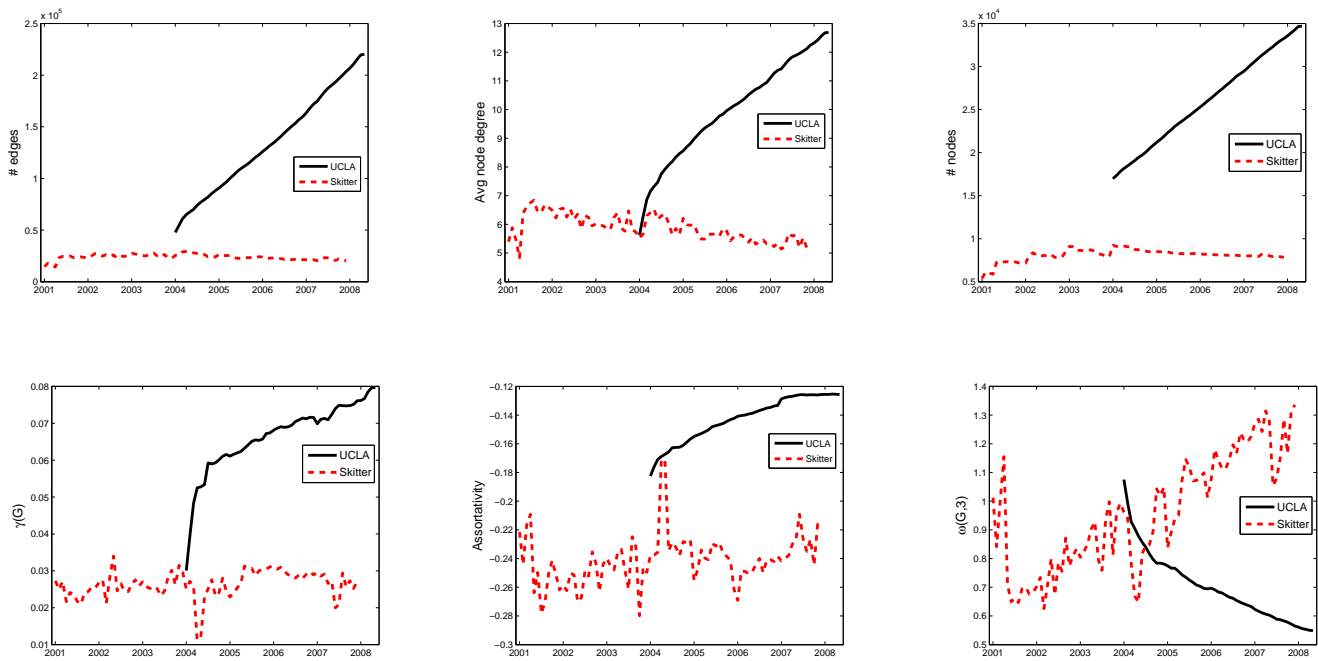
Fig. 9. Measures of the UCLA and Skitter AS topologies over several years.

this result shows that, although the number of nodes being sampled by Skitter is decreasing, the hierarchy of the Internet as observed by Skitter is not changing. This also implies that Skitter is not sampling the edge of the AS topology well and so cannot see evolutionary changes there.

Figure 11 shows the proportion of nodes in each k-core as a function of k for the UCLA data set. There are 52 plots (for more than 4 years) shown as a smooth transition between the first and last plots, emphasized. As can be seen, the distribution of k-cores moves to the right over time, indicating that the proportion of nodes with higher connectivity is increasing over time. This adds further weight to the conclusion that the UCLA dataset shows a weakening of hierarchy in the Internet, with more peering relationships between ASes. Note that the UCLA data set was not examined in [32].

Next we examine the WSDs for the two datasets; a subset showing their evolution is presented in figures 12 and 13. There are two peaks which evolve (Skitter) and disappear (UCLA). Comparing Figures 12 and 13 with Figures 7 and 8, we confirm the results given by the other graph measures. The WSDs show in the case of the UCLA dataset (Figure 12) a less dominant core (the core *peak* is disappearing) and less structured Internet (the WSD is getting lower and moving to the right), while Figure 13 shows a more dominant core for Skitter and a more structured Internet (the WSDs are increasing and moving to the left).

The differences in AS topology evolution observed by Skitter and UCLA are likely to reflect different views of the Internet from the two datasets. The IP-level traceroute-based Skitter traverses paths and sees only a small portion of the AS-level topology, mainly at the core. In contrast, the BGP data of the UCLA dataset is provided by public feeds from monitors not located at the core. In conclusion, the WSD provides a
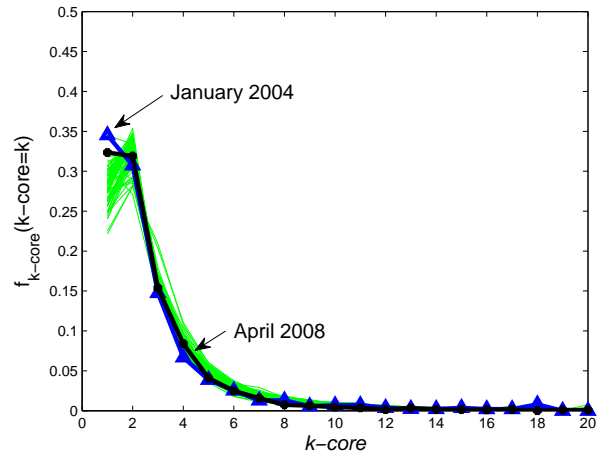

Fig. 10. k-core proportions, Skitter AS topology.

valuable tool together with more traditional graph measures for network characterisation.

C. Tuning Synthetic Topology Generators

This section examines the estimation of optimal parameters for several well known topology generators. Parameter estimation requires two elements: a metric and a target graph. The aim is then to estimate the parameters that give a best fit between the target graph and the topologies generated. Typically, topology generator parameters are tuned by matching measures such as the number of links, the node degree distribution or the clustering coefficient [33, 34]. However, as these are not metrics in the strict sense, obtaining a "good fit" does not mean that the topology generator is a good model of
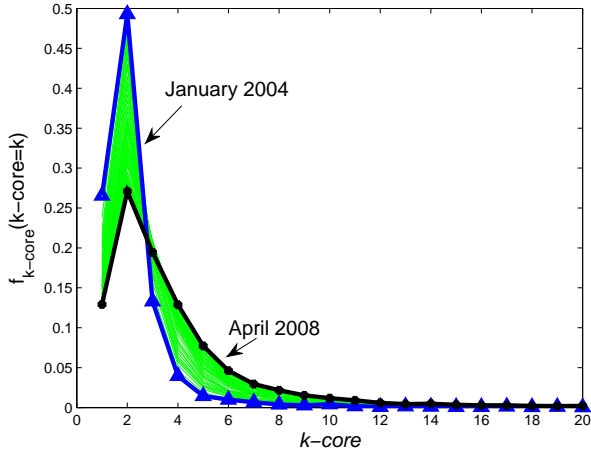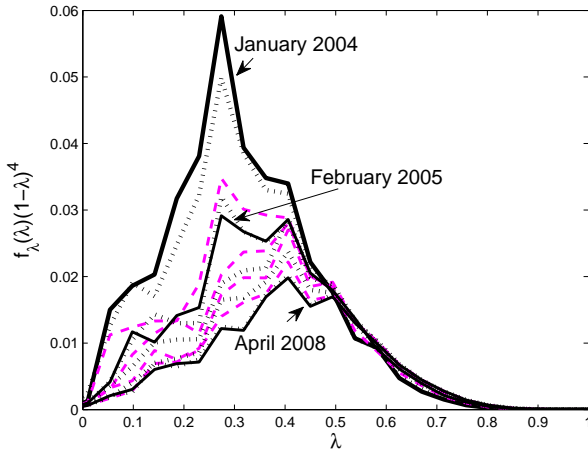
Fig. 11. k-core proportions, UCLA AS topology.

Fig. 12. Weighted Spectra Distribution, UCLA AS topology.
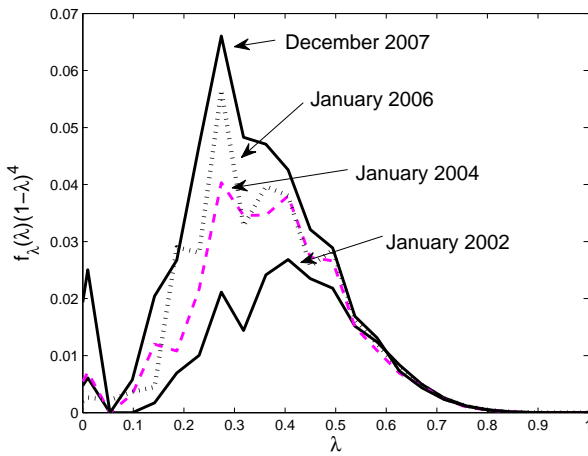

Fig. 13. Weighted Spectral Distribution, Skitter AS topology.

the target graph. For example, we can tune the AB model to match the clustering coefficient of our target graph exactly but then the degree distributions might differ. The core question

here is: is the target graph likely to have been generated by this topology generator? As the WSD is a metric, obtaining identical or close WSDs implies that the target is likely to have been generated by this topology generator. However, this can only happen if the topology generator is capable of mimicking the structure of the target graph. As will be seen, all the topology generators examined in this section fail to do so. For a more detailed examination of this application see Haddadi *et al.* [6].

The topology generators we consider are Waxman [35], AB [23], GLP [34], Inet [11] and PFP [36] (PFP does not have any parameters but is included for completeness of the comparison). For this study we use a particular sample AS topology, the Skitter graph from March 2004 as used by Mahadevan *et al.* [37]. The process we describe could easily be applied to other sampled AS topologies. The cost function used is the standard quadratic norm between the weighted spectral distributions using a value of $N = 4$ ($N = 3$ gives similar results)

$$J(G_s, G_i(\Theta)) = \sum_{k \in K} (1-k)^4 (f_s(\lambda = k) - f_i(\lambda = k))^2 \quad (20)$$

where the distribution of eigenvalues is estimated in the set K of bins $\in [0, 2]$, $J(G_s, G_i(\Theta))$ is the cost between the weighted spectra of the Skitter graph G_s and the graphs generated by the ith topology generator using a vector of parameters Θ.

To facilitate this comparison, grids are constructed over the possible values of the parameter space and $J(G_s, G_i(\Theta))$ is evaluated at each point. For example, Figure 14 shows the case for the AB model which has two parameters, $\Theta = [p\ q]^T$. The first thing to note is the dark region at the centre of the graph, which represents those parameter values that return the lowest values of $J(G_s, G_i(\Theta))$. Figure 14 shows that the WSD varies smoothly with changing structure. The AB model is known to exhibit scale free behavior below the line indicated in Figure 14 and exponential behavior above this line. It is satisfying to note that a clear minimum exists and also that the region with low cost (Region A) lies in the scale free region as expected for a model of the Internet. Similar results were also found for the other topology generators [6].

As an aside, a similar grid was also constructed using the spectrum, specifically the unweighted spectral distribution, as the basis for a cost function:

$$J'(G_s, G_i(\Theta)) = \sum_{k \in K} (f_s(\lambda = k) - f_i(\lambda = k))^2. \quad (21)$$

The resulting grid is shown for the AB model in Figure 15. As can be seen, $J'(G_s, G_i(\Theta))$ does not change smoothly with Θ, which implies that a small change in the structure can lead to a large change in the spectrum. In other words, the tuning of the topology generator parameters through a cost function based on the spectrum is too sensitive to changes in the graph structure. The main reason for this was shown in Figure 1: the variance in the spectral gap is too large.

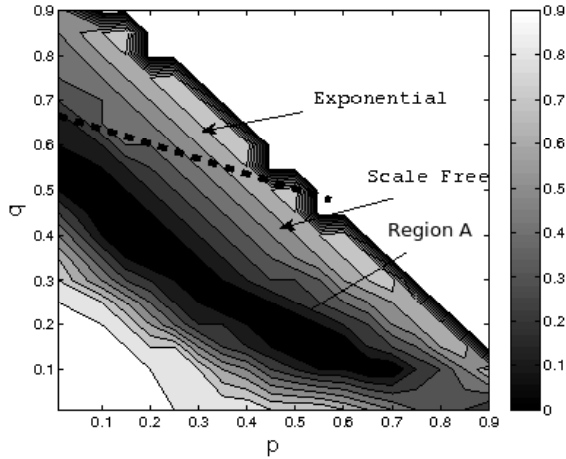In order to estimate the optimum parameters for the topol-

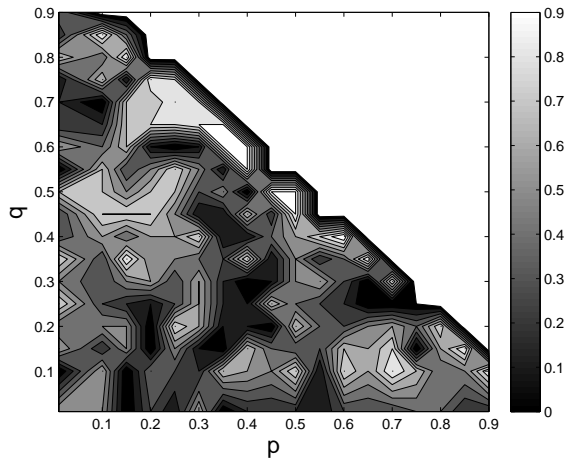Fig. 14. $J(G_s, G_{AB}(\Theta))$ for allowed values of $\Theta(= [p\ q]^T)$.

Fig. 16. Best fit WSDs for topology generators relative to target Skitter data set.

Fig. 15. $J'(G_s, G_{AB}(\Theta))$ for allowed values of $\Theta(= [p\ q]^T)$.

ogy generators[20], we optimise Equation 20 using the Nelder Meade simplex search algorithm [38, 39].

Table II displays these optimum parameter values Θ of the topology generators when matching to the Skitter graph. In addition, Table II displays the values of $J(G_s, G_i(\Theta))$, which shows that PFP provides the closest fit followed by AB, GLP, Waxman and, finally, Inet. While these results are generally as expected, the ranking of Inet as the least desirable topology generator is not as it is one of the most recent models.

Figure 16 shows the WSD of the target graph and of the "best fit" WSDs for each of the topology generators. As can be seen, the main peak in the WSD for the Skitter data occurs at a value of $\lambda = 0.4$. The Waxman generator peak occurs at $\lambda = 0.6$, which is closer to 1 and demonstrates the greater amount of random structure in the Waxman topologies.

[20]For a given Θ, each run of a topology generator will lead to a different network because of random initial conditions. This interferes with the estimation of the gradients used in the search algorithm. Thus, to reduce the variance of gradient estimates, the averaged weighted spectra from ten runs are used, although empirical evidence suggests that the variance of these weighted spectra is very low (see Figure 1).
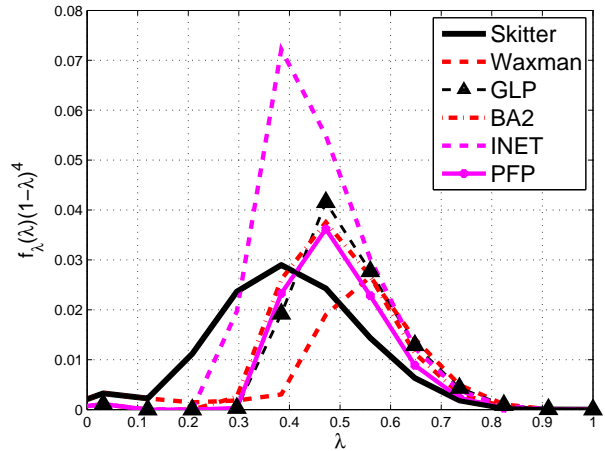
The INET generator is interesting in that it is the one generator that obtains a peak at the correct point. This is due to the nature of the generator: the parameter α to the INET generator specifies the percentage of nodes that are in a highly meshed core. While a similar structure exists in the Skitter dataset, the WSD shows that the INET core is too densely connected relative to the Skitter observations.

All the WSDs for the generated topologies lie to the right of the target WSD, showing that the Skitter data has far more structure than is mimicked by any of the topology generators. It is important to note that these are the closest approximations that the topology generators can make to the Skitter data. In other words, it is simply not possible to force the AB model for instance to have a peak at $\lambda = 0.4$ with amplitude of 0.03. It is our experience that each topology generator leads to topologies that cover only a small subset of the WSD space.

VI. CONCLUSIONS

In this paper we introduced a new metric, the *weighted spectral distribution* (WSD). The WSD differs from other graph measures such as the clustering and assortativity coefficients, the node degree distribution, *etc*, in that it is a metric in the mathematical sense, and so it can be used to measure the distance between two graphs.

The WSD has many applications, and in particular can be used for very large graphs because of its low computational requirements, making it a good choice for topology tuning and other applications that require multiple evaluations of a cost function. We presented three applications of the WSD, using it to understand: (*i*) the mixing properties of graphs, (*ii*) the evolution of the AS topology, and (*iii*) the tuning of Internet topology generators to match a target graph.

Observed evolution in the WSD of the Internet graph, supported by evidence using other common graph measures, suggests that the Internet is becoming more edge-centric. We created a similar effect using our strawman model supporting this hypothesis. In addition, we observed that the topology generators we examined fail to match the complex structure

TABLE II
OPTIMUM PARAMETER VALUES FOR MATCHING SKITTER TOPOLOGY SAMPLED IN MARCH 2004.

Waxman	$\alpha = 0.08$ (default $= 0.15$)	$\beta = 0.08$ (default $= -0.2$)	$J(G_{Waxman}, G_{Skitter}) = 0.0026$
AB	$p = 0.2865$ (default $= 0.6$)	$q = 0.3145$ (default $= 0.3$)	$J(G_{AB}, G_{Skitter}) = 0.0014$
GLP	$p = 0.5972$ (default $= 0.45$)	$\beta = 0.1004$ (default $= 0.64$)	$J(G_{GLP}, G_{Skitter}) = 0.0021$
Inet	$\alpha = 0.1013$ (default $= 0.3$)	$-$	$J(G_{INET}, G_{Skitter}) = 0.0064$
PFP	$-$	$-$	$J(G_{PFP}, G_{Skitter}) = 0.0014$

of the Internet. This result is important for future topology generators: the WSD could provide a valuable tool in their design and validation.

In conclusion, the WSD provides a new analytical tool augmenting other the ability of the graph analyst to obtain a more complete picture of a network's structure. Future avenues of research also include examining the WSD of large biological and social networks, among others.

APPENDIX A
METRIC DEFINITION

The metric we propose in this paper is

$$J(G_x, G_y) = \sum_{k \in K} (1 - k)^4 (f_x(\lambda = k) - f_y(\lambda = k))^2 \quad (22)$$

We now show that $\sqrt{J(G_x, G_y)}$ is a metric in the mathematical sense. The difference between $\sqrt{J(G_x, G_y)}$ and $J(G_x, G_y)$ is similar to the difference between the sum squared error and the root mean squared error. We prefer the sum squared error (i.e., $J(G_x, G_y)$) in this application as it provides the well known minimum variance-bias trade-off.

A metric satisfies the following four conditions:

(a) $J(G_x, G_y) \geq 0$ (non-negativity)
(b) $J(G_x, G_y) = 0 \Leftrightarrow x = y$ (identity of indiscernibles)
(c) $J(G_x, G_y) = J(G_y, G_x)$ (symmetry)
(d) $J(G_x, G_z) \leq J(G_x, G_y) + J(G_y, G_z)$ (triangle inequality)

(a) and (c) follow directly from (22). Noting that all the elements of the sum in $J(G_x, G_y)$ are positive $\implies$ $J(G_x, G_y) = 0$ if and only if $f_x(\lambda = k) = f_y(\lambda = k) \ \forall k$. Arranging (and increasing the number of bins if necessary) the k bins such that each bin contains at most 1 eigenvalue Requires G_x to be co-spectral and isomorphic to G_y. Two graphs may be *co-spectral*, i.e., they share the same spectrum but are not isomorphic. However, studies have shown [40] that the number of co-spectral graphs falls dramatically with the number of vertices in the graph. For example, only 0.05% of all graphs with 21 vertices are co-spectral and not isomorphic; this number is thought to decrease with increasing number of vertices [40]. Thus, condition (b) is true *almost certainly*, in the statistical sense.

$\sqrt{J(G_x, G_y)}$ defines the standard metric space R_2^K [41]. This can be seen by distributing the weights $(1 - k)^4$ as:

$$\sqrt{J(G_x, G_y)} = \left(\sum_{k \in K} (h_x(\lambda = k) - h_y(\lambda = k))^2 \right)^{1/2}$$

$$(23)$$

where

$$h_x(\lambda = k) = (1 - k)^2 f_x(\lambda = k) \quad (24)$$

and $h_y(\lambda = k)$ is similarly defined. The triangle inequality holds for (23). For a detailed proof see [41] Chapter 2, Section 5.

ACKNOWLEDGMENT

REFERENCES

[1] H. Bunke, "Graph matching: Theoretical foundations, algorithms, and applications," in *Proc. of the International Conference on Vision Interface*, May 2000, pp. 82–88.

[2] V. Kann, "On the approximability of the maximum common subgraph problem," in *Proc. 9th Annual Symposium on Theoretical Aspects of Computer Science*, 1992, pp. 377–388.

[3] S. Hanna, "Representation and generation of plans using graph spectra," in *6th International Space Syntax Symposium*, Istanbul, (2007).

[4] B. Luo and E. Hancock, "Structural graph matching using the em algorithm and singular value decomposition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 10, pp. 1120–1136, Oct. 2001.

[5] A. Ng, M. Jordan, and Y. Weiss, "On spectral clustering: analysis and an algorithm," in *Advances in Neural Information Processing Systems 14*, T. Dietterich, S. Becker, and Z. Ghahramani, Eds. MIT Press, (2002).

[6] H. Haddadi, D. Fay, S. Uhlig, A. Moore, R. Mortier, A. Jamakovic, and M. Rio, "Tuning topology generators using spectral distributions," in *Lecture Notes in Computer Science, Volume 5119, SPEC International Performance Evaluation Workshop*. Darmstadt, Germany: Springer, 2008.

[7] C. Gkantsidis, M. Mihail, and E. Zegura, "Spectral analysis of Internet topologies," in *Proceedings of IEEE Infocom 2003*, San Francisco, CA, Apr. (2003).

[8] S. Butler, "Lecture notes for spectral graph theory," *Lectures in Nankai University, Tianjin, China*, (2006).

[9] A. Jamakovic and S. Uhlig, "On the relationship between the algebraic connectivity and graph's robustness to node

and link failures," in *Next Generation Internet Networks, 3rd EuroNGI Conference on*, Trondheim, Norway, (2007).

[10] D. Vukadinovic, P. Huang, and T. Erlebach, "On the spectrum and structure of Internet topology graphs," in *IICS '02: Proceedings of the Second International Workshop on Innovative Internet Computing Systems*, (2002).

[11] J. Winick and S. Jamin, "Inet-3.0: Internet topology generator," University of Michigan Technical Report CSE-TR-456-02, Tech. Rep. CSE-TR-456-02, (2002).

[12] S. V. N. Vishwanathan, K. Borgwardt, and N. Schraudolph, "Fast computation of graph kernels," *In Advances in Neural Information Processing Systems 19 (NIPS 2006)*, 2006.

[13] L. Shyu, S.-Y. Lau, and P. Huang, "On the search of Internet AS-level topology invariants," in *Proceedings of IEEE Global Telecommunications Conference (GLOBECOM) 2006*, San Francisco, CA, USA, 2006, pp. 1–5.

[14] M. Latapy and C. Magnien, "Complex network measurements: Estimating the relevance of observed properties," in *Proceedings of IEEE Infocom 2008*, April 2008.

[15] X. Wang and D. Loguinov, "Wealth-based evolution model for the internet as-level topology," *INFOCOM 2006. Proceedings*, pp. 1–11, April 2006.

[16] A. Wool and G. Sagie, "A clustering approach for exploring the Internet structure," *Proceedings of 23rd IEEE Convention of Electrical and Electronics Engineers in Israel*, pp. 149–152, Sept. 2004.

[17] Y. Li, J.-H. Cui, D. Maggiorini, and M. Faloutsos, "Characterizing and modelling clustering features in as-level internet topology," *INFOCOM 2008. The 27th Conference on Computer Communications. IEEE*, pp. 271–275, April 2008.

[18] A. Seary and W. Richards, "Spectral methods for analyzing and visualizing networks: an introduction." in *Dynamic Social Network Modeling and Analysis*. National Academic Press, 2003, pp. 209–228.

[19] B. Nadler, S. Lafon, R. Coifman, and I. Kevrekidis, "Diffusion maps, spectral clustering and eigenfunctions of fokker-planck operators," in *Neural Information Processing Systems (NIPS)*, (2005).

[20] A. G. Thomason, "Pseudo-random graphs," *Random Graphs '85, North-Holland Mathematical Study*, vol. 144, pp. 307–331, 1987.

[21] F. R. K. Chung, R. L. Graham, and R. M. Wilson, "Pseudo-random graphs," *Combinatorica*, vol. 9, no. 4, pp. 345–362, 1989.

[22] F. R. K. Chung, *Spectral Graph Theory (CBMS Regional Conference Series in Mathematics)*. American Mathematical Society, (1997).

[23] R. Albert and A.-L. Barabasi, "Topology of evolving networks: local events and universality," *Physical Review Letters*, vol. 85, p. 5234, (2000). [Online]. Available: http://www.citebase.org/abstract?id=oai:arXiv.org:cond-mat/0005085

[24] S. Butler, "Interlacing for weighted graphs using the normalized laplacian," *Electronic Journal of Linear Algebra*, vol. 16, pp. 90–98, 2007.

[25] B. Huffaker, D. Andersen, E. Aben, M. Luckie, k Claffy, and C. Shannon, "The Skitter AS Links Dataset," 2001-2007. [Online]. Available: http://www.caida.org/data/active/skitter_aslinks_dataset.xml/

[26] R. Oliveira, B. Zhang, and L. Zhang, "Observing the Evolution of Internet AS Topology," in *Proceedings of ACM SIGCOMM 2007*, Kyoto, Japan, Aug. 2007.

[27] R. Bush, J. Hiebert, O. Maennel, M. Roughan, and S. Uhlig, "Testing the reachability of (new) address space," in *Proceedings of the 2007 SIGCOMM workshop on Internet network management (INM'07)*, 2007.

[28] A. Lakhina, J. Byers, M. Crovella, and P. Xie, "Sampling biases in IP topology measurements," in *Proceedings of IEEE Infocom 2003*, San Francisco, CA, Apr. 2003.

[29] M. Newman, "Assortative mixing in networks," *Physical Review Letters*, vol. 89, no. 20, pp. 871–898, 2002.

[30] H. Haddadi, D. Fay, A. Jamakovic, O. Maennel, A. W. Moore, R. Mortier, M. Rio, and S. Uhlig, "Beyond node degree: Evaluating AS topology models," University of Cambridge, Computer Laboratory, Tech. Rep. UCAM-CL-TR-725, July 2008.

[31] P. Gill, M. Arlitt, Z. Li, and A. Mahanti, "The flattening internet topology: Natural evolution, unsightly barnacles or contrived collapse?" in *Proceedings of Passive and Active Measurement Conference (PAM)*, April 2008.

[32] J. I. Alvarez-Hamelin, L. Dall'Asta, A. Barrat, and A. Vespignani, "k-core decomposition of Internet graphs: hierarchies, self-similarity and measurement biases," *Networks and Heterogeneous Media*, vol. 3, p. 371, 2008.

[33] E. W. Zegura, K. L. Calvert, and M. J. Donahoo, "A quantitative comparison of graph-based models for Internet topology," *IEEE/ACM Transactions on Networking (TON)*, vol. 5, no. 6, pp. 770–783, (1997).

[34] T. Bu and D. Towsley, "On distinguishing between Internet power law topology generators," in *Proceedings of IEEE Infocom 2002*, New York, NY, Jun. (2002).

[35] B. M. Waxman, "Routing of multipoint connections," *IEEE Journal on Selected Areas in Communications (JSAC)*, vol. 6, no. 9, pp. 1617–1622, Dec. (1988).

[36] S. Zhou, "Characterising and modelling the Internet topology, the rich-club phenomenon and the PFP model," *BT Technology Journal*, vol. 24, (2006).

[37] P. Mahadevan, D. Krioukov, M. Fomenkov, X. Dimitropoulos, k c claffy, and A. Vahdat, "The Internet AS-level topology: three data sources and one definitive metric," *SIGCOMM Computer Communication Review*, vol. 36, no. 1, pp. 17–26, (2006).

[38] J. Nelder and R. Mead, "A simplex method for function minimization," *Comput. J.*, vol. 7, pp. 308–313, (1965).

[39] J. Dennis and D. Woods, "Optimization in microcomputers: The nelder-meade simplex algorithm," in *New Computing Environments: Microcomputers in Large-Scale Computing*, A. Wouk, Ed. SIAM, (1987), pp. 116–122.

[40] P. Zhu and C. Wilson, "A study of graph spectra for comparing graphs," in *The 16th British Machine Vision Conference (BMVC)*, September 2005.

[41] A. N. Kolmogorov and S. V. Fomin, *Introductory Real*

Analysis. Dover Publication, 1975.

Damien Fay obtained a B.Eng from University College Dublin (1995), an MEng (1997) and PhD (2003) from Dublin City University and worked as a mathematics lecturer at the National University of Ireland (2003-2007) before joining the NetOS group, Computer Laboratory, Cambridge in 2007 as a research associate. He is currently a research associate at McGill University, Montreal. His research interests include applied graph theory and time series analysis.

Hamed Haddadi is a postdoctoral researcher at Max Planck Institute for Software Systems in Germany. He completed his Ph.D. research jointly at the Computer Laboratory, University of Cambridge and University College London. His research interests include network measurement and monitoring, online advertisement and online privacy.

Andrew Thomason is a professor in the department of Pure Mathematics at the University of Cambridge. His research interests lie in combinatorial mathematics and its applications in computer science.

Andrew W. Moore is a lecturer at the University of Cambridge, Computer Laboratory. His interests lie in addressing the Scalability, Usability, and Reliability of the Internet. He completed his Ph.D. with the Cambridge University Computer Laboratory in 2001 and prior to that took a Masters degree and an honours degree from Monash University in Melbourne. Australia. He is a chartered engineer with the IET and a member of the IEEE, ACM and USENIX.

Richard Mortier is co-founder at Vipadia Limited following six years as a Researcher in the Cambridge Systems and Networking group at Microsoft Research Cambridge where he investigated a variety of topics concerning networked system performance and management. Prior to joining Microsoft he spent six months working for Sprint Advanced Technology Labs in California. He completed his Ph.D. with the Systems Research Group at the University of Cambridge Computer Laboratory in 2001.

Almerima Jamakovic obtained the MSc (2004) and PhD (2008) degrees in Electrical Engineering from Delft University of Technology, the Netherlands. During her PhD studies she performed research in the field of complex networks, focusing on the robustness of real-world structures. After graduating she joined TNO Information and Communication Technology to work on the latest performance and robustness challenges of ICT systems and telecommunication networks. Her research interests include complex networks, sensor networks, network robustness and ICT performance.

Steve Uhlig is a researcher at Deutsche Telekom Laboratories (T-labs) and Technische Universitt Berlin, Germany. He obtained a PhD degree (2004) in Applied Sciences from the University of Louvain, Louvain-la-neuve, Belgium. After his PhD, he was a Postdoctoral Fellow of the Belgian National Fund for Scientific Research (F.N.R.S.), as well as a visiting scientist at Intel research Cambridge, UK, and at the University of Adelaide, Australia. Prior to joining T-labs he was an assistant professor at Delft University of Technology, the Netherlands. His main research interests are focused on the large-scale behavior of Internet routing and traffic, and their interactions through the network topology.

Miguel Rio is a lecturer at the Department of Electronic and Electrical Engineering of University College London, specialized in Internet Technology. He completed his PhD at the Computer Laboratory of the University of Kent at Canterbury. His research interests include Internet congestion controL, multimedia distribution using peer-to-peer networks, QoS routing and traffic measurement and analysis.