Characterizing the Global Impact of P2P Overlays on the AS-Level Underlay

Amir Hassan Rasti[1], Reza Rejaie[1], and Walter Willinger[2]

[1] University of Oregon {amir,reza}@cs.uoregon.edu
[2] AT&T Labs Research walter@research.att.com

Abstract. This paper examines the problem of characterizing and assessing the global impact of the load imposed by a Peer-to-Peer (P2P) overlay on the AS-level underlay. In particular, we capture Gnutella snapshots for four consecutive years, obtain the corresponding AS-level topology snapshots of the Internet and infer the AS-paths associated with each overlay connection. Assuming a simple model of overlay traffic, we analyze the observed load imposed by these Gnutella snapshots on the AS-level underlay using metrics that characterize the load seen on individual AS-paths and by the transit ASes, illustrate the churn among the top transit ASes during this 4-year period, and describe the propagation of traffic within the AS-level hierarchy.

Key words: Overlay networks, AS-level topology, BGP simulation

1 Introduction

The large volume of traffic associated with Peer-to-Peer (P2P) applications has led to a growing concern among ISPs which need to carry the P2P traffic relayed by their costumers. This concern has led researchers and practitioners to focus on the idea of reducing the volume of external P2P traffic for edge ISPs by localizing the connectivity of the P2P overlay (for recent work, see for example [1,2]). However, such an approach only deals with the local effect of an overlay on individual edge ASes. Even though the volume of P2P traffic on the Internet is large and growing, assessing the *global* impact of a P2P overlay on the individual ASes in the network, which we call the *AS-level underlay*, remains a challenging problem and is not well understood. This is in part due to the fact that investigating this problem requires a solid understanding of an array of issues in two different domains: *(i)* design and characterization of overlay-based applications, and *(ii)* characterization of AS-level underlay topology and BGP routing in this underlay. Another significant challenge is dealing with inaccurate, missing, or ambiguous information about the AS-level underlay topology, AS relationships and tier properties, and BGP routing policies.

This paper investigates the problem of assessing the load imposed by a given overlay on the AS-level underlay. We show that assessing this impact requires tackling a number of challenging problems, including *(i)* capturing accurate snapshots of the desired overlay, *(ii)* estimating the load associated with individual overlay connections, and *(iii)* determining the AS-path in the underlay that corresponds to individual overlay connections. Toward this end, this paper makes two main contributions. First, we present a methodology for assessing the impact of an overlay on the AS-level underlay. Our methodology incorporates a collection of the best known practices for capturing accurate snapshots of a P2P overlay and, more importantly, for determining the AS-path corresponding to each overlay connection. We rely on snapshots of the AS-level Internet topology provided by CAIDA where each link between two ASes is annotated with the relationship between them. Using a BGP simulator called *C-BGP* [4], we perform a detailed simulation of BGP routing over these annotated snapshots of the AS-level underlay to infer the corresponding AS-path for each overlay connection and determine the aggregate load crossing individual ASes. To assess the propagation of overlay traffic through the AS-level hierarchy, we also infer the tier information for individual ASes using the *TierClassify* tool [3].

Second, we illustrate our methodology by characterizing the impact of four snapshots of the *Gnutella* overlay that were captured over four successive years on the AS-level underlay snapshots of the Internet taken on the same dates the Gnutella overlay snapshots were obtained. We characterize the load imposed by these overlays on the corresponding underlay in a number of different ways: *(i)* observed load on individual AS-paths and its diversity, *(ii)* observed load on individual transit ASes, *(iii)* AS-path length, and *(iv)* the propagation of overlay traffic through the AS-level hierarchy. Our analysis provides valuable insight into how changes in overlay connectivity and underlay topology affect the mapping of load on the AS-level underlay.

The rest of this paper is organized as follows. In Section 2, we further elaborate on the problem of mapping an overlay on the AS-level underlay, describe the challenges involved, and present our methodology. Section 3 describes our datasets and presents our characterization of the load imposed by the Gnutella overlay on the corresponding AS-level underlay, spanning a 4-year period. We conclude the paper and sketch our future plans in Section 4.

2 The Problem and Our Methodology

Our goal is to map the traffic associated with a P2P overlay to the AS-level underlay. The input to this process is a representation of a P2P overlay structure consisting of the IP addresses (and port numbers) of the participating peers together with their neighbor lists. The output is the aggregate load on all affected ASes and between each pair of affected ASes that have a peering link with one another (in each direction). Our methodology to tackle this problem consists of the following intuitive steps:

1. Capturing the topology of a P2P overlay,
2. Estimating the load on individual connections in the overlay,
3. Inferring the AS-paths associated with individual overlay connections,
4. Determining the aggregate load on each AS and between connected ASes (in each direction separately).

In this section, we discuss the challenges posed by each step, clarify our assumptions, and describe our approach for each step.

2.1 Capturing the Overlay Topology

Capturing a snapshot of the overlay topology for a P2P application is feasible if the list of neighbors for individual peers can be obtained. For example, in Gnutella it is possible to query individual peers and retrieve their neighbor lists. Therefore, a Gnutella-specific crawler can be developed to progressively collect this information until a complete snapshot of the overlay is captured.

In our earlier work, we have developed a fast P2P crawler that can capture accurate snapshots of the Gnutella network in a few minutes [5]. Using this crawler, we have captured tens of thousands of snapshots of the Gnutella overlay topology over the past several years. In this study, we use a few of these snapshots for the top-level overlay of Gnutella (an overlay consisting of Gnutella *Ultrapeers*). While other P2P applications such as BitTorrent are responsible for a significantly larger volume of traffic over the Internet than Gnutella and would therefore provide a more relevant P2P system for this study, we are not aware of any reliable technique to capture accurate snapshots of the corresponding overlays. Since accuracy of the overlay topology is important in this study, we focus on Gnutella. However, our methodology is not restricted to this application and can be used with other P2P systems.

2.2 Estimating the Load of Individual Overlay Connections

The load of individual overlay connections depends on the subtle interactions between several factors including: *(i)* the number of peers that generate traffic (*i.e.*, sources), the rate and pattern of traffic generation by these peers, and their relative location in the overlay, *(ii)* the topology of the overlay, and *(iii)* the relaying (*i.e.*, routing) strategy at individual peers. Capturing these factors in a single model is a non-trivial task and could be application-specific. For example, the load of individual connections for live P2P video streaming is more or less constant, whereas the load of individual BitTorrent connections may vary significantly over time.

In the absence of any reliable model for per-connection traffic, without loss of generality, we assume in our analysis that all connections of the overlay experience the same average load in both directions. This simplifying assumption allows us to focus on the mapping of the overlay topology on the underlying AS-level topology. If a more reliable model for the load of individual connections is available, it can be easily plugged into our methodology by assigning proper weights (one in each direction) to each connection of the overlay. In this paper, we simply assume that the weight for all connections in both directions is one.

2.3 Inferring AS-Paths for Individual Overlay Connections

For each connection in the overlay, determining the corresponding AS-path in the underlay is clearly the most important and most challenging part of our methodology. We use a popular BGP simulator to determine the AS-path between any given pair of ASes, but note that carefully-designed measurement-based approaches may provide viable alternatives. Our simulation-based method consists of the following steps:

Mapping Peers to ASes: We use archived BGP snapshots from RouteViews [6] to map the IP addresses of individual peers to their corresponding ASes that we call edge ASes. Therefore, determining the AS-path for the overlay connection between two peers translates into determining the path between their corresponding edge ASes.

Capturing AS-level Topology and Inter-AS Relationships: In this study, we rely on the AS-level topologies provided by CAIDA [7]. These topologies have been widely used in the past, even though more recent work has shown that the provided topologies are missing a significant portion of peering links between lower-tiered ASes [8,9]. Note that our approach is not tied to using the CAIDA-provided AS-level topologies, and any more complete AS-level topology can be incorporated once it becomes available. To properly simulate BGP routing, we need to determine the AS relationship between connected ASes in the AS-level topology. Toward this end, we use the fact that CAIDA's snapshots of the AS-level topology [7] are annotated with the inferred relationships between each pair of connected ASes. In these snapshots, AS relationships are inferred using the algorithm initially proposed by Gao [10] and extended by Dimitropoulos *et al.* [11]. This algorithm, mainly based on the concept of "valley-free routing" in BGP (along with some other intuitive assumptions), categorizes the AS relationships into three categories: *(i)* Customer-Provider, *(ii)* Peer-Peer, or *(iii)* Sibling-Sibling.

Simulating BGP: We determine the AS-path between any pair of edge ASes that host connected peers in the overlay (*i.e.*, infer the corresponding AS-path) by simulating BGP over the annotated AS-level topology using the *C-BGP* simulator [4]. C-BGP abstracts the AS-level topology as a collection of interconnected routers, where each router represents an AS. It simulates the desired BGP routing policies for each relation between connected ASes. We use a set of intuitive BGP policies for each type of AS relationships that are specified by C-BGP. In particular, these policies *(i)* ensure that the routes through one's customers have the highest preference and those passing through its providers have the lowest preference, and *(ii)* prevent ASes with multiple providers from acting as transit node among their providers. We noticed that some characteristics of CAIDA's annotated AS-level topology, in particular the presence of circular provider-costumer relationships among a group of ASes, prevent our C-BGP simulations to converge with the above policies. To resolve these problems, we systematically change a small number of relationships (*e.g.*, to break a cycle in customer-provider relationships). Further details of this process are described in our related technical report [12]. We select snapshots of both the AS-level topology and the overlay topology of the same dates so as to minimize any potential error due to asynchrony in the snapshots.

Clearly, representing each AS by a single router results in inferring only one AS-path between each pair of ASes. This implies that multiple AS-paths that may exist in practice between two ASes [13] are not accounted for in our simulations. While this assumption simplifies the problem in a way that is not easily quantifiable, we are not aware of any existing technique that can reliably capture and account for this subtle behavior of BGP routing.

Assessing AS Tiers: To characterize the propagation of P2P traffic through the AS-level hierarchy, we first need to assess the location of each AS in this hierarchy. We use the "TierClassify" tool [3] to identify the *tier* of each individual AS. The algorithm used in this tool relies mainly on the assumption that all tier-1 ASes should be interconnected with one another. Therefore it tries to find a clique among the ASes with highest degrees. Once the tier-1 clique is identified, the algorithm simply follows provider-customer relationships and classifies other ASes such that each tier n AS can reach the tier-1 clique in $n-1$ hops.

2.4 Determining Aggregate Load on and between Individual ASes

Given the corresponding AS-path for each overlay connection, we can easily determine the aggregate load (in terms of the number of connections) that passes through each AS, as well as the transit load (in each direction) between each pair of connected ASes in the topology.

3 Effect of Overlays on the Underlay

In this section, we characterize the effect of a P2P overlay on the AS-level underlay using four snapshots of the Gnutella top-level overlay. We broadly divide ASes into two groups: *Edge ASes* that host peers in an overlay, and *Transit* (or *Core*) *ASes* that provide connectivity between edge ASes. We first describe our datasets (*i.e.*, the snapshots of overlay and the corresponding AS-level underlay topologies), and then we characterize the imposed load on the underlay using the following measures: *(i)* diversity and load on individual AS-paths, *(ii)* load on individual transit ASes, *(iii)* identity and evolution of the top transit ASes, *(iv)* AS-path length, and *(v)* propagation of traffic through the AS-level hierarchy.

Datasets: We use four snapshots of the top-level Gnutella overlay that were collected in four consecutive years starting in 2004. Examining overlay snapshots over time enables us to assess some trends that are associated with the evolution of the AS-level topology.

We use the labels G-xx to refer to the snapshot taken in year 20xx. The left columns of Table 1 (labeled "Gnutella snapshots") summarize the capture date, number of peers and edges for these overlay snapshots. The table shows that the population of Gnutella peers in the top-level overlay and their pairwise connections have both increased by $\approx 600\%$ during this four-year period.

We also use daily snapshots of the BGP routing table retrieved from the RouteViews archive collected at the same dates as our overlay snapshots. The middle columns in Table 1 (labeled "BGP snapshots") give the number of IP

Table 1. Data profile: Gnutella snapshots, BGP snapshots and mapping overlay connections to the underlay. Imp. AS-paths are those with +100 overlay connections.

		Gnutella Snapshots		BGP Snapshots		AS-Paths	
Snapshot	Date	#Peers	#Conn.	#Prefixes	#ASes	#Unique	%Important
G-04	04-11-20	177k	1.46M	165k	18.7k	192k	2.0
G-05	05-08-30	681k	5.83M	185k	20.6k	384k	2.9
G-06	06-08-25	1.0M	8.64M	210k	23.2k	605k	2.8
G-07	07-03-15	1.2M	9.80M	229k	24.9k	684k	2.7

prefixes and the total number of ASes in each BGP snapshot. These numbers show that the AS-level topology has also grown significantly during this four-year period.

Diversity and Load on Individual AS-Paths: One way to characterize the impact of an overlay on the underlay is to determine the number of unique AS-paths that all overlay connections are mapped on as well as distribution of load among those AS-paths. The right columns of Table 1 (labeled "AS-paths") show the number of unique AS-paths for all connections of each overlay along with the percentage of those paths that carry more than 100 overlay connections. The number of unique AS-paths is growing over time but at a lower pace compared to the number of overlay connections. This suggests that there is more similarity in AS-paths among overlay connections as the overlay grows in size over time.

To examine the mapping of overlay connections to AS-paths more closely, Figure 1(a) depicts the CCDF of the number of overlay connections that map to individual AS-paths in log-log scale for all four overlay snapshots. The skewed shape of these distributions indicates that a small number of AS-paths carry a large fraction of load. For example, whereas around 10% of paths carry more than 10 connections, only 1% of the paths carry more than 200 connections. Interestingly, the distributions of overlay connections that map to AS-paths are very similar across different snapshots despite significant changes in the identity of peers and in the topologies of overlay and underlay.

Observed Load on Individual Transit ASes: Since we assumed that all overlay connections have the same load, we simply quantify the load on each transit AS by the number of overlay connections crossing that AS. Figure 1(b) depicts the number of overlay connections that cross each transit AS in log-log scale, where ASes are ranked (from high to low) based on their overall observed load. The figure shows that the load on transit ASes is very skewed. A small number of them carry a large volume of traffic while the load on most transit ASes is rather small. Again, we observe that the overall shape of the resulting curves is very similar for all four snapshots, except for the outward shift in the more recent snapshots caused by the increasing size of the overlay over time. This similarity in the skewness of the observed load on transit ASes despite significant changes in the overlay and underlay topologies over time could be due to the dominance of one the following factors: *(i)* the stability over time of the top-10

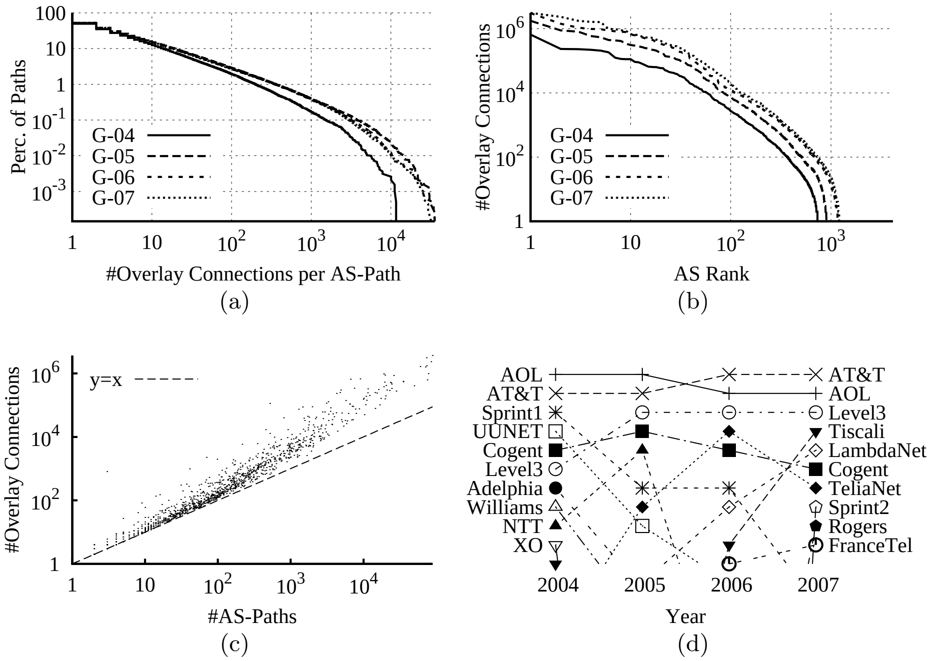

Fig. 1. (a) Distribution of load across AS-paths, (b) Overlay connections passing through transit ASes, (c) Scatterplot of number of relevant AS-paths vs. load, (d) Identity and evolution of top-10 transit ASes carrying the largest number of overlay connections.

ASes that host most peers, and *(ii)* the constraint imposed by valley-free routing over the hierarchical structure of the AS-level underlay.

To further investigate the underlying causes for the observed skewed nature of observed load on transit ASes, we examine the distribution of the number of unique AS-paths (associated with overlay connections) that pass through each transit AS. The shape of this distribution is very similar to Figure 1(b) (not shown), suggesting that the number of crossing connections for individual ASes is primarily determined by the underlay shape and routing rather than connectivity and footprint of the overlay. Figure 1(c) validates this observation by showing a scatterplot of the number of crossing AS-paths (x-axis) and number of overlay connections (y-axis) through each transit AS. This figure essentially relates the previous two distributions and confirms that the observed load on individual transit ASes depends primarily on the number of unique AS-paths crossing those ASes. Note that once the number of cross AS-paths exceeds a certain threshold (a few hundreds), the observed load increases at a much faster pace.

Identity and Evolution of Transit ASes: To investigate the observed load by transit ASes from a different angle, we examine and present the identity of the top-10 transit ASes that carry the highest number of crossing overlay

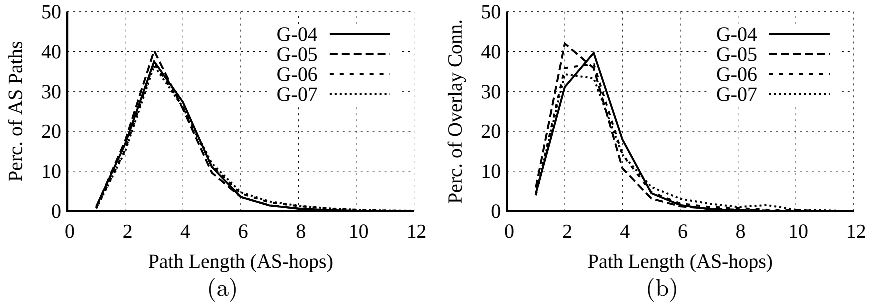

Fig. 2. (a) Distribution of AS-path length between connected edge ASes, (b) Distribution of AS-path length for all overlay connections.

connections (and their evolution over time) in Figure 1(d). For each of the four overlay snapshots, the transit ASes are rank-ordered (highest load first), and the figure depicts their standings in these rank-ordered lists over time. We observe that only four transit ASes (*i.e.*, AT&T, AOL, Level3, and Cogent) remain in the top-10 list across all four snapshots and that the changes in the other transit ASes is more chaotic. This is due to the fact that ranking of transit ASes is affected by a combination of factors including changes in the topology of AS-level underlay, in routing policies, and in the location of peers. Disentangling these different factors and trying to identify the root causes for the observed churn among the top-10 transit ASes over time remains a challenging open problem.

AS-Path Length: One way to quantify the impact of an overlay on the AS-level underlay is to characterize the length of AS-paths for individual overlay connections. Figure 2(a) shows the empirical density of the length of all AS-paths between edge ASes for each of the four snapshots. We observe that around 40% of the paths are three AS-hops long, while 80% of the paths in each overlay are at most 4 AS-hops long.

Figure 2(b) depicts the empirical density of AS-path length across all *overlay connections* for each of the four snapshots. In essence, this plot can be viewed as a *weighted* version of Figure 2(a) described above where the length of each path is weighted by the number of overlay connections crossing it. The figure shows a very similar pattern across all overlay snapshots despite the changes in the number of peers and their connections. The two figures are very similar, however the average path length across the overlay connections is slightly shorter indicating that a slightly higher fraction of connections are associated with shorter paths. (*e.g.*, for G-07, the average length of all AS-paths is 3.2 hops while the average path length across overlay connections is 3.7 hops.)

Propagation of Traffic through the AS-Level Hierarchy: An interesting way to quantify the load that an overlay imposes on the AS-level underlay is to determine the fraction of load that is propagated upward in the AS-level hierarchy towards the top-tiered ASes. Table 2 gives the percentage of paths and percentage of overlay connections whose top AS is a tier-1, tier-2, and tier-3

Table 2. Percentage of paths/connections reaching each tier of AS hierarchy.

	Tier-1		Tier-2		Tier-3	
Snapshot	Path	Conn	Path	Conn	Path	Conn
G-04	51	84	46	16	2.4	0.0
G-05	59	73	38	27	3.0	0.0
G-06	52	64	38	36	10	0.0
G-07	55	63	41	37	3.6	0.1

AS, respectively, in each overlay snapshot. The columns marked "Path" give the percentage of the relevant AS-paths reaching each tier while the columns marked "Conn" represent the percentage of the overlay connections (*i.e.*, aggregate load) reaching each tier. We note that more than half of the paths reach a tier-1 AS, and roughly 40% of the paths peak at a tier-2 AS across all four snapshots.

The percentage of connections that reach a tier-1 AS is even higher than that for paths, indicating that a larger fraction of connections are mapped to these paths. At the same time, a lower percentage of connections reach a tier-2 AS (16% to 37%) compared to paths that peak in tier-2 ASes. Interestingly, the percentage of connections that reach a tier-1 AS decreases over time while the percentage of connections that peak in a tier-2 AS is increasing. A plausible explanation of this trend is the increasing connectivity over time between ASes in the lower tiers which reduces the fraction of connections that have to climb the hierarchy up to tier-1 ASes. A closer examination (not shown here) confirmed that this shift in traffic towards lower tiers is indeed primarily due to the presence of shortcuts between lower-tier ASes in the AS topology (*e.g.*, more aggressive peering at Internet exchange points over time). In particular, the observed shift has little to do with changes in the overlay topology, mainly because the connectivity of the Gnutella overlay has not become significantly more localized over time.

4 Conclusion and Future Work

In this paper, we studied the problem of quantifying the load that a particular overlay imposes on the AS-level underlay. We identified the challenging aspects of this problem and described existing techniques to address each of these aspects. Relying on an existing set of best practices, we presented a methodology for mapping the load of an application-level overlay onto the AS-level underlay. We illustrated our methodology with an example of a real-world P2P overlay (*i.e.*, Gnutella). While our study contributes to a deeper understanding of the interactions between application-level overlays and the AS-level underlay in today's Internet, a more detailed analysis of the sensitivity of our results to known overlay-specific issues, known underlay-related problems (*e.g.*, incomplete AS graph, ambiguous AS relationships), and known BGP-related difficulties (*e.g.*, , limitations of the C-BGP simulator) looms as important next step.

As part of our future work, we plan to investigate how changing the geographical location of peers and their connectivity affect the load imposed on the AS-level underlay. Furthermore, we plan to derive realistic traffic models for different P2P application and incorporate them into our methodology. Finally, we intend to examine pricing models that are used by ISPs to determine how structure and workload of an overlay affect the revenues of the various ISPs in the AS hierarchy of the underlay.

Acknowledgment

We would like to thank Matthew Roughan for helpful discussions, Daniel Stutzbach for sharing tools and data, and Bruno Quoitin and Wolfgang Muehlbauer for promptly answering all of our C-BGP-related inquiries. This work is supported in part by the National Science Foundation (NSF) under Grant No. Nets-NBD-0627202 and an unrestricted gift from Cisco Systems. Part of this work was performed while the first author was visiting IPAM (UCLA) in the Fall of 2008 as a participant of the "Internet Multi-Resolution Analysis" program.

References

1. V. Aggarwal, A. Feldmann, and C. Schneideler, "Can ISPs and P2P systems cooperate for improved performance?" *ACM SIGCOMM Computer Communication Review*, vol. 37, no. 3, pp. 29–40, 2007.
2. D. R. Choffnes and F. E. Bustamante, "Taming the torrent: A practical approach to reducing cross-ISP traffic in P2P systems," in *ACM SIGCOMM*, Aug. 2008.
3. Z. Ge, D. R. Figueiredo, S. Jaiswal, and L. Gao, "On the Hierarchical Structure of the Logical Internet Graph," in *SPIE ITCom*, Nov. 2001.
4. B. Quoitin and S. Uhlig, "Modeling the Routing of an Autonomous System with C-BGP," *IEEE Network*, vol. 19, no. 6, 2005.
5. D. Stutzbach, R. Rejaie, and S. Sen, "Characterizing Unstructured Overlay Topologies in Modern P2P File-Sharing Systems," in *ACM IMC*, Oct. 2005.
6. University of Oregon, "RouteViews Project," http://routeviews.org
7. CAIDA, "Cooperative Association for Internet Data Analysis," http://caida.org
8. R. Oliveira, D. Pei, W. Willinger, B. Zhang, and L. Zhang, "In search of the elusive ground truth: the internet's as-level connectivity structure," in *ACM SIGMETRICS*, Jun. 2008.
9. M. Roughan, S. J. Tuke, and O. Maennel, "Bigfoot, sasquatch, the yeti and other missing links: what we don't know about the as graph," in *ACM IMC*, Oct. 2008.
10. L. Gao, "On Inferring Autonomous System Relationships in the Internet," *IEEE/ACM Transactions on Networking*, vol. 9, pp. 733–745, 2000.
11. X. Dimitropoulos, D. Krioukov, M. Fomenkov, B. Huffaker, Y. Hyun, K. Claffy, and G. Riley, "AS Relationships: Inference and Validation," *ACM SIGCOMM Computer Communication Review*, vol. 37, no. 1, pp. 29–40, 2007.
12. A. H. Rasti, R. Rejaie, and W. Willinger, "Characterizing the Global Impact of the P2P Overlay on the AS-level Underlay," University of Oregon, Tech. Rep. CIS-TR-2010-01, Jan. 2010. http://mirage.cs.uoregon.edu/pub/tr10-01.pdf
13. W. Muhlbauer, A. Feldmann, O. Maennel, M. Roughan, and S. Uhlig, "Building an AS-topology model that captures route diversity," *ACM SIGCOMM Computer Communication Review*, vol. 36, no. 4, pp. 195–206, 2006.