

# Where the Sidewalk Ends

## Extending the Internet AS Graph Using Traceroutes From P2P Users

Kai Chen David R. Choffnes Rahul Potharaju Yan Chen

Fabian E. Bustamante Dan Pei† Yao Zhao

Northwestern University

†AT&T Labs – Research

{kchen,drchoffnes,rpo219,jingo}@u.northwestern.edu, {ychen,fabianb}@northwestern.edu

†peidan@research.att.com

### ABSTRACT

An accurate Internet topology graph is important in many areas of networking, from deciding ISP business relationships to diagnosing network anomalies. Most Internet mapping efforts have derived the network structure, at the level of interconnected autonomous systems (ASes), from a limited number of either BGP- or traceroute-based data sources. While techniques for charting the topology continue to improve, the growth of the number of vantage points is significantly outpaced by the rapid growth of the Internet.

In this paper, we argue that a promising approach to revealing the hidden areas of the Internet topology is through active measurement from an observation platform that scales with the growing Internet. By leveraging measurements performed by an extension to a popular P2P system, we show that this approach indeed exposes significant new topological information. Based on traceroute measurements from more than 992,000 IPs in over 3,700 ASes distributed across the Internet hierarchy, our proposed heuristics identify 23,914 new AS links not visible in the publicly-available BGP data – 12.86% more *customer-provider* links and 40.99% more *peering links*, than previously reported. We validate our heuristics using data from a tier-1 ISP and show that they correctly filter out all false links introduced by public IP-to-AS mapping. We have made the identified set of links and their inferred relationships publicly available.

### Categories and Subject Descriptors

C.2.3 [Computer-Communication Networks]: Network Operations - Network monitoring

### General Terms

Measurement, Management

### Keywords

Internet measurement, AS-level topology, Traceroute, BGP

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

CoNEXT'09, December 1–4, 2009, Rome, Italy.

Copyright 2009 ACM 978-1-60558-636-6/09/12 ...\$10.00.

### 1. INTRODUCTION

An accurate Internet topology graph is important in many areas of networking, from deciding ISP business relationships to diagnosing network anomalies. Appropriately, several research efforts have investigated techniques for measuring and generating such graphs [1–6].

Most Internet mapping efforts have derived the network structure, at the AS level, from a limited number of data sources for either BGP paths or traceroute traces. The advantage of using BGP paths is that they can be gathered passively from BGP route collectors and thus require minimal measurement effort for obtaining a large number of Internet paths. Unfortunately, the publicly available BGP paths do not cover the entire Internet due to issues such as visibility constraints, route aggregation, hidden sub-optimal paths and policy filtering. In contrast, traceroute measurements provide the ability to infer the data paths that packets take when traversing the Internet. Because they are active measurements, traceroutes can be designed to potentially cover every corner of the Internet given sufficient numbers of vantage points (VPs).<sup>1</sup> However, all existing traceroute-based projects are restricted by their limited number of VPs. Furthermore, the traceroute measurements provide an IP-level map while our interest is the AS-level map. Converting an IP-level topology to an accurate AS-level one remains an open area of research [7].

In this paper, we argue that a promising approach to revealing the hidden areas of the Internet topology is through active measurement from an observation platform that scales with the growing Internet. Our work makes the following key contributions. First, we collect and analyze the diversity of paths covered by traceroutes gathered from hundreds of thousands of peer-to-peer (P2P) users worldwide (Section 2). Specifically, the probes are issued from over 992,000 P2P user IPs in 3,700 ASes, making our measurement study the largest-ever in terms of the number of VPs and network coverage.

Second, we provide a thorough set of heuristics for inferring AS-level paths from traceroute data (Section 3). To this end, we present a detailed analysis of issues that affect the accuracy of traceroute measurements and how our heuristics address these problems. Our proposed techniques for correcting IP-to-AS mappings are generic and work for the scenarios where traceroute VPs are poorly correlated with public BGP VPs. Furthermore, we validate our heuristics using data from a tier-1 ISP as ground-truth and show that they filter out all of the false links introduced by public IP-to-AS mapping for this ISP.

Third, we characterize the new links discovered by our P2P mea-

<sup>1</sup>Vantage points can be defined as locations with distinct network views. Because this paper focuses on AS topologies, we use *vantage point* to refer to a unique AS.

Project	# unique machines	# unique ASes
Routeviews/RIPE	790	438
Skitter	24	$\leq 24$
iPlane	192	$\leq 192$
DIMES	8,059	200
<b>Ono</b>	<b>600,000</b>	<b>6,000</b>

**Table 1: Approximate numbers of VPs for topology-gathering projects at the time of publication.**

surements (Section 4). We find that some common assumptions about the visibility of paths according to AS relationships are routinely violated. For example, while we have found 40.99% more *peering* links, we further observe that a VP can even miss some of its upstream *peering* links. More importantly, we reveal 12.86% more *customer-provider* links than what can be found in the publicly-available BGP data.

Fourth, we derive a number of root causes behind the identified missing links, presenting a detailed analysis of their occurrences, and quantify the number of missing links due to each of those reasons (Section 5). Interestingly, many of the missing links (75.02% in our dataset) are missing due to multiple, concurrent reasons.

We discuss limitations of this work in Section 6, review closely related research in Section 7 and conclude in Section 8.

## 2. P2P FOR TOPOLOGY MONITORING

Understanding and characterizing the salient features of the ever-changing Internet topology requires a system of observation points that grows organically with the network. Because ISP interconnectivity is driven by business arrangements often protected by nondisclosure agreements, one must infer AS links from publicly available information such as BGP and traceroute measurements. The success of either approach ultimately depends on the number of measurement VPs.

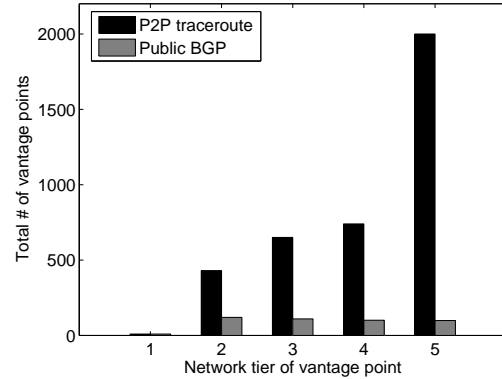
To achieve broad coverage, it is essential to use a platform built upon large-scale emergent systems, such as P2P, that grow with the Internet itself. By piggybacking on an existing P2P system, one can eliminate the need to place BGP monitors in each ISP; rather, each participating host in the system can contribute to the AS topology measurement study simply by performing traceroute measurements.

As a first step toward this goal, we use data gathered from Ono [8], an extension to the Vuze BitTorrent client. The software has been installed more than 600,000 times by hosts located in over 40,000 routable prefixes, spanning more than 6,000 ASes and 192 countries. Ono collects traceroute measurements between connected hosts to ensure that the software meets its goal of improving download performance while reducing cross-ISP traffic. Volunteers report this data to our central servers for offline analysis.<sup>2</sup> This platform constitutes the most diverse set of measurement VPs and is the largest set of traceroute measurements collected from end hosts to date. Table 1 compares the number of unique machines and VPs in our study and in a set of related efforts including Routeviews [9], RIPE/RIS [10], iPlane [11], DIMES [12] and Skitter [13]. For Ono it is difficult to determine the number of unique machines, so we use the number of times the software was installed.

As we show in Section 4, about 23,914 new links are discovered through these traceroute measurements. These new links include

<sup>2</sup>Users are informed of the diagnostic information gathered by the plugin and are given the chance to opt out. In any case, no personally identifiable information is ever published.

26 ASNs (AS numbers) that do not appear in the publicly-available BGP data and thus are truly “dark networks” when viewed through the lens of the public BGP servers. Thus the view of the network from P2P users contributes a vast amount of information about network topology unobtainable through other approaches such as BGP table dumps and strategic active probing from dedicated infrastructure.



**Figure 1: Distribution of VPs with respect to their network tiers.**

Figure 1 shows the distribution of VPs across hierarchical tiers for the publicly-available BGP data and the actually used P2P traceroutes. Note that P2P traceroutes have significantly more VPs compared to the publicly-available BGP data, especially in lower-tier networks. This unique perspective allows us to view previously hidden regions of the network and determine their impact on properties of the Internet topology. The following sections present our methodology for AS-level topology inference and report on our study of missing links.

## 3. METHODOLOGY

In this section, we describe the datasets that we use in this study, present a systematic approach to addressing the challenges associated with accurately inferring AS-level paths from traceroute data and discuss how we validate our resulting topologies. Finally, we explain the algorithms used for inferring properties of the AS topology.

### 3.1 Data Collected

#### 3.1.1 P2P traceroutes

The traceroutes in our dataset are collected by P2P users recording the result of the `traceroute` command provided by their operating system. Because the software performing the measurements is cross-platform, there are multiple traceroute implementations that generate data for our study. The vast majority of the data that we gather comes from the Windows traceroute implementation.

The measurement is performed using default settings except that the timeout for router responses is 3 seconds and no reverse DNS lookups are performed. Each peer running our software performs at most one measurement at a time; after each traceroute completes, the peer issues another to a randomly selected destination from the set of connections it has established through BitTorrent.<sup>3</sup>

<sup>3</sup>Note that Ono biases BitTorrent connections towards nearby

There are three measurements for each router hop; the ordered set of hops is sent to our central data-collection servers along with the time at which the measurement was performed. We use the data collected between Dec 1, 2007 and Sep 30, 2008, which consists of 541,023,742 measurements containing over 6.2 billion hops. The data was collected from 992,197 distinct peer IPs<sup>4</sup> in 3,723 unique ASes. Together, these peers probe more than 84 million distinct destination IPs.

### 3.1.2 BGP feeds

The BGP data used in this study includes a collection of BGP routing tables from 790 BGP speaking routers in 438 unique ASes. Specifically, we combine several BGP feeds: Routeviews [9] collected at route-views.oregon-ix.net, which is the most widely used BGP archive so far, six other Oregon route servers and 16 route collectors of RIPE/RIS [10]. We use 10 months of data gathered between Dec 1, 2007 and Sep 30, 2008, the same time period for our P2P traceroute data. Furthermore, we download AS links from the UCLA IRL lab [14] which also contain those links collected from route servers, looking glasses, and IRR [15]. Because the UCLA data does not include BGP AS paths, nor information from new VPs added near the time of publication, we combine all of these sources of AS links to obtain the most complete set of AS links. For the rest of this paper, we will refer to this dataset as the “public view” [2, 3]. According to Oliveira et al. [2–4], 10 months of the public view data should be enough to cover “all” the hidden links<sup>5</sup>

### 3.1.3 Ground-truth data

To validate our inferred AS links, we use router configurations and syslogs from a tier-1 ISP as ground-truth connectivity information. The data includes historical configuration and syslog files for more than 800 routers in this network. We simply leverage the heuristics in [2] to process these files and extract the ground-truth AS links that can be used as baseline for our validation.

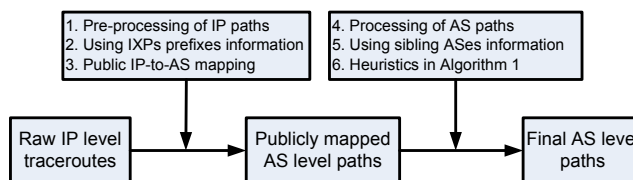
## 3.2 Using Traceroutes

While traceroute probes can provide detailed network topology information, there are a number of issues that prevent their widespread use in AS topology generation. For one, the number of probe sources and targets required to reveal new topological information grows with the size of the Internet. As we discussed in Section 2, we address this issue through measurements from P2P users. Another limitation is that traceroutes provide IP-level views of the topology and the IP-to-AS mappings gathered from publicly available information are incomplete and potentially incorrect. Finally, traceroute measurements are subject to the constraints of the routers they visit, which can drop probes, silently forward them without altering the TTL or even erroneously modify the TTL in ways that affect the inferred path. When using traceroutes as a telescope for viewing the AS topology, one must expect a blurry lens with many artifacts. In this section, we discuss a systematic approach for sharpening and clarifying this view by addressing these limitations.

peers, so there is a slightly higher probability that traceroutes will be issued to them. We posit that this assists the discovery of new AS links because these nearby peers are often located in access networks at lower tiers of the AS topology.

<sup>4</sup>The number of unique installs and the number of distinct IPs are not equal because each user is often assigned dynamic IP addresses and some users disable traceroute probes.

<sup>5</sup>Hidden links are those policy-allowed links which do not always show up in public view. For example, links only on sub-optimal paths do not show up in public view unless the primary paths fail.



**Figure 2: High-level architecture for converting IP paths to AS paths.**

Figure 2 illustrates the steps we take to convert traceroute IP paths into their corresponding AS-level paths. In the next subsection, we discuss the processing we perform on IP-level paths (Steps 1–3). Then, after obtaining traceroute AS-level paths based on public IP-to-AS mappings, we adjust the paths to correct for inconsistencies with the corresponding BGP AS paths (Steps 4–6).

### 3.2.1 IP-level Adjustments

**Step 1.** Before performing IP-to-AS mappings, we inspect each IP-level path. First, we search for those measurements that contain repeated, consecutive IP addresses in the path. When this occurs, the repeated IP is likely to be upstream from a router that is not decrementing the traceroute probe’s TTL. Such routers are effectively hidden from our measurement and could lead to falsely inferred AS links. There are other known problems such as load balancing, zero-TTL forwarding and address rewriting of gateway routers that would cause routing loops [16]. To avoid the potential problems of these issues, we conservatively remove the entire path from our analysis.

**Step 2.** As explained by Mao et al. [7], paths that traverse Internet eXchange Points (IXPs) can lead to falsely inferred AS links. Using a list of known IXP prefixes, such as PCH [17], PeeringDB [18] and Euro-IX [19], we remove from each path any hop that belongs to an IXP. This allows us to correctly infer direct links between the ASes that connect to each other at an IXP. However, we cannot rely on the publicly available information to completely eliminate such kind of false links because they are known to be incomplete. Our heuristics in the next subsection will address the remaining problems of IXPs at the AS level.

**Step 3.** After the first two steps, we simply convert IP-level paths to AS-level ones by directly using the AS mappings provided by Team Cymru [20], which incorporates both publicly available and private BGP information.

### 3.2.2 IP-to-AS Mapping

The next phase in our analysis is to address the issues contained in the conversion from IP-level paths into AS-level ones. While previous work has investigated the problem of accurate IP-to-AS mappings in networks where BGP data is available [7], our study is the first to address the problem for an arbitrary (and large) set of networks. The proposed techniques will consider the scenario where the traceroute VPs are not the same as the BGP VPs, which makes them generically applicable. Furthermore, unlike previous work using traceroutes, we expect to see a significant number of new links compared to the public view because our software monitors a larger portion of the Internet. The key challenge that we address in this section is how to distinguish the real new links from those that are falsely inferred due to incorrect IP-to-AS mappings. To evaluate the quality of our heuristics, we compare our results with ground-truth from a tier-1 ISP.

In the first phase of our analysis, we simply convert IP-level paths to AS-level ones (*i.e.*, Step 3). The authors in [7] identify

	Problem	Symptom				Filtering heuristic(s)
		Loop	Missing hop	Substitute hop	Extra hop	
Incomplete paths	Unresolved hops within an AS					Steps 1, 4
	Unmapped hops between ASes MOAS hops at the end					Step 4 Step 4
False AS links	Internet exchange points (IXPs)				✓	Steps 2, 4, 6
	Sibling ASes	✓	✓	✓	✓	Steps 5, 6
	Unannounced IP addresses	✓	✓	✓	✓	Step 6
	Using outgoing interface IPs		✓	✓	✓	Step 6
	Private peering interface IPs		✓			Step 6

**Table 2: Problems within traceroute-inferred AS-level paths, symptoms for these problems, and the step(s) we take to solve them. Note that we do not consider the symptoms for “incomplete paths” because they are addressed in [7]. Reading the last row of the table, “private peering interface IPs” will cause missing hop problem, and we address this problem in Step 6 of our techniques.**

several patterns of discrepancies between traceroute and BGP paths (as shown in Table 2), each of which entails a difference of at most one AS hop (e.g., an AS is missing from the path, an extra AS appears in the hop, or a substitute AS appears in the path). To account for these discrepancies while still preserving true new AS links discovered by traceroute measurements, we mark a new link to be *pending* if it could be corrected by techniques used by Mao et al. [7]; otherwise, we assume that the new link is real. In our implementation, we conservatively modify all the *pending* links such that they are consistent with the corresponding BGP paths. We emphasize that this approach prevents false positives, but may filter out real links not present in BGP. Note that unlike the work in [7], we only correct the AS-level paths generated by traceroutes so that we can confidently infer new links. Correcting the IP-to-AS mappings is beyond the scope of this paper.

We show in Table 2 that our implementation for converting IP paths to AS paths can address most of the well-known problems identified by Mao et al. [7]. Since their work addressed the problem of incomplete paths, we directly apply their techniques to our dataset (Steps 1 & 4). However, identifying and modifying falsely mapped AS links is a significant challenge that we address in this work.

**Step 4.** Besides dealing with incomplete paths in this step, we further filter additional IXPs. While we have used the available IXP prefixes to delete the AS hops belonging to any IXP, our list of IXP prefixes is not complete. For those IXPs that do not make their prefixes publicly available, we still can identify them by using the IXP participating AS list we have. We pick out the AS hops in the middle of traceroute AS paths that are publicly mapped to multiple ASes and check if these multiple ASes are collocated in an IXP. This occurs when the shared infrastructure address is originated into BGP by multiple participating ASes. However, we cannot use this approach to identify IXP that use their own AS numbers – a limitation that we address in Step 6.

**Step 5.** A single organization may own and manage multiple sibling ASes. Among two sibling ASes, one AS may use some address blocks from another to number its equipment or during route propagation only one AS includes its AS number in the BGP AS path while the other does not. This would cause problems within the traceroute AS paths. To mitigate such problems, we download the known sibling ASes from CAIDA [21]. For a sibling AS pair ( $X, Y$ ), we may see the cases where traceroute AS path is [... $WXYZ$ ...] while a corresponding BGP AS path is [... $WXZ$ ...] or [... $WYZ$ ...]. For this case, we modify the traceroute AS path to be [... $W\{X, Y\}Z$ ...]; In our measurement, we also find instances where the traceroute AS path is [... $WYZ$ ...] while a corresponding BGP AS path is [... $WXZ$ ...]. In those cases we use the BGP AS

path to modify the traceroute AS paths. Again, publicly available sibling-AS information is limited. In the next step, we use heuristics that mitigate the remaining problems when sibling ASes cause discrepancies between traceroute AS paths and BGP AS paths.

<b>PROCEDURE</b> Address issues within traceroute AS paths
1 Initialization: set the DISTANCE of each AS link on the traceroute AS paths;
2 <b>foreach</b> AS link in the traceroute AS paths (e.g., use $B-C$ at the top of Figure 3 as illustration) <b>do</b>
3 <b>if</b> $DISTANCE(B, C) = 1$ <b>then</b>
4         AS link $B-C$ is considered <i>true</i> ;
5 <b>if</b> $DISTANCE(B, C) = 2$ <b>then</b>
6         Check the public view BGP AS paths;
7 <b>if</b> There exists an AS path ... $B X C$ ... <b>then</b>
8             Fix $B-C$ using $B-X-C$ and set DISTANCE of each of these two links as 1 (For multiple $X$ s, choose the longest matched one; For instance, both [ $A B X_1 C D$ ] and [ $A' B X_2 C D'$ ] exist, the first one matches the traceroute AS path [ $A B C D$ ] better, hence $X_1$ is preferred;
9 <b>if</b> There does not exist an AS path ... $B X C$ ... <b>then</b>
10 <b>if</b> ... $A X C$ ... (or ... $B X D$ ...) appears in BGP AS paths <b>then</b>
11                 Replace $B$ (or $C$ ) with $X$ and set the DISTANCE of each link as 1 (longest match for multiple $X$ s);
12 <b>else</b>
13 <b>if</b> $DISTANCE(A, C)=1$ (or $DISTANCE(B, D)=1$ ) <b>then</b>
14                     Delete $B$ (or $C$ ) and set the DISTANCE of link $A-C$ (or $B-D$ ) as 1 ;
15 <b>if</b> $DISTANCE(A, C) \neq 1$ and $DISTANCE(B, D) \neq 1$ <b>then</b>
16                         Mark $B-C$ as a real link and set DISTANCE( $B, C$ ) as 1 ;
17 <b>end</b>
18 <b>if</b> $DISTANCE(B, C) \geq 3$ <b>then</b>
19 <b>if</b> $DISTANCE(A, C)=1$ (or $DISTANCE(B, D)=1$ ) <b>then</b>
20             Delete $B$ (or $C$ ) and set the DISTANCE of link $A-C$ (or $B-D$ ) as 1 ;
21 <b>if</b> $DISTANCE(A, C) \neq 1$ and $DISTANCE(B, D) \neq 1$ <b>then</b>
22             Mark $B-C$ as a real link and set DISTANCE( $B, C$ ) as 1 ;
23 <b>end</b>
24 <b>end</b>
25 Return the traceroute AS path when DISTANCES for all links are 1 ;

**Algorithm 1: Heuristics in Step 6 of Figure 2.**

**Step 6.** Algorithm 1 addresses a variety of issues with traceroute AS paths that remain after the previous five steps. Below, we discuss how these heuristics apply each symptom, i.e., loops and missing/extra/substitute hops as shown in Table 2.

- **Loop.** Loops in traceroute AS paths can happen due to unan-















