In Search of the Elusive Ground Truth: The Internet's AS-level Connectivity Structure

Ricardo Oliveira
UCLA
rveloso@cs.ucla.edu

Dan Pei
AT&T Labs Research
peidan@research.att.com

Walter Willinger
AT&T Labs Research
walter@research.att.com

Beichuan Zhang
University of Arizona
bzhang@arizona.edu

Lixia Zhang
UCLA
lixia@cs.ucla.edu

ABSTRACT

Despite significant efforts to obtain an accurate picture of the Internet's actual connectivity structure at the level of individual autonomous systems (ASes), much has remained unknown in terms of the quality of the inferred AS maps that have been widely used by the research community. In this paper we assess the quality of the inferred Internet maps through case studies of a set of ASes. These case studies allow us to establish the ground truth of AS-level Internet connectivity between the set of ASes and their directly connected neighbors. They also enable a direct comparison between the ground truth and inferred topology maps and yield new insights into questions such as which parts of the actual topology are adequately captured by the inferred maps, and which parts are missing and why. This information is critical in assessing for what kinds of real-world networking problems the use of currently inferred AS maps or proposed AS topology models are, or are not, appropriate. More importantly, our newly gained insights also point to new directions towards building realistic and economically viable Internet topology maps.

Categories and Subject Descriptors

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

General Terms

Measurement, Verification

Keywords

Internet topology, BGP, inter-domain routing

1. INTRODUCTION

Many research projects have used a graphic representation of the Internet, where nodes represent entire autonomous systems (ASes) and two nodes are connected if and only if the two ASes are engaged in a business relationship to exchange data traffic. Due to the Internet's decentralized architecture, however, this AS-level construct is not readily available and obtaining accurate AS maps has remained an active area of research. A common feature of all the AS maps that have been used by the research community is that they have been inferred from either BGP-based or traceroute-based data. Unfortunately, both types of measurements are more a reflection of what we can measure than what we really would like to measure, resulting in fundamental limitations as far as their ability to reveal the Internet's true AS-level connectivity structure is concerned.

While these limitations inherent in the available data have long been recognized, there has been little effort in assessing the degree of completeness, accuracy, and ambiguity of the resulting AS maps. Although it is relatively easy to collect a more or less complete set of ASes, it has proven difficult, if not impossible, to collect the complete set of inter-AS links. The sheer scale of the AS-level Internet makes it infeasible to install monitors everywhere or crawl the topology exhaustively. At the same time, big stakeholders of the AS-level Internet, such as Internet service providers and large content providers, tend to view their AS connectivity as proprietary information and are in general unwilling to disclose it. As a result, the quality of the currently used AS maps has remained by and large unknown. Yet numerous projects have been conducted using these maps of unknown quality, causing serious scientific and practical concerns in terms of the validity of the claims made and accuracy of the results reported.

In this paper we take a first step towards a rigorous assessment of the quality of the Internet's AS-level connectivity maps inferred from public BGP data. Realizing the futility of attempting to obtain the complete global AS-level topology, we take an indirect approach to address the problem. Using a small number of different types of ASes whose complete AS connectivity information can be obtained, we conduct case studies to compare their actual connectivity with that of what we call the "public view" – the connectivity structure inferred from all the publicly available and commonly-used BGP data source (i.e., routing tables, updates, looking glasses, and routing registry). The case studies enable us to understand and verify what kinds of AS links are adequately captured by the public view and what kinds of (and how many) AS links are missing from the public view. They also provide valuable new insights into where the missing links are located within the overall AS topology.

More specifically, this paper makes the following original contributions. First, in Section 2 we define what we mean by "ground

SIGMETRICS'08, June 2–6, 2008, Annapolis, Maryland, USA.

truth" of AS-level Internet connectivity as far as a single AS and its neighbors are concerned. Second, we report in Section 3 on a series of case studies which highlight the difficulties in establishing the desired ground truth. What makes the desired ground truth so elusive is that for most ASes, the data sources necessary to unambiguously establish their AS-level connectivity are not publicly available. Nevertheless, by roughly classifying ASes into the a few major types, we can explore what fractions of what types of connections are missing from currently used AS maps, and we can typically even identify the reasons why they are missing. Lastly, motivated by this detailed understanding of the inherent shortcomings of inferred AS maps, we argue in Section 4 for a new approach to generating realistic AS-level topologies. To this end, we sketch a construction that, while being informed by the available measurements, results in AS maps annotated with AS relationships that are instantiations of fully functional and economically viable AS-level connectivity structures and could plausibly represent the Internet's actual AS-level ecosystem or a close approximation thereof.

The main findings of our search for the elusive ground truth of AS-level Internet connectivity can be summarized as follows. First, inferred AS maps based on single-period snapshots of publicly available BGP-based data are typically of low quality. The percentage of missing links ranges from 10-20% for Tier-1 and Tier-2 ASes to 85% and more for large content networks. Second, the quality of the inferred AS maps can be significantly improved by including historic data of BGP updates from all existing sources. For example, links on backup paths can be revealed by routing dynamics over time, but the time period required to collect the necessary information can be several years. Third, through the use of data collected over long enough time periods, the public view captures all the links of Tier-1 ASes and most customer-provider links at all tiers in the Internet. Fourth, due to the *no-valley* routing policy and the lack of monitors in stub networks, the public view misses a great number of peer links at all tiers except tier-1. It can miss as much as 90% of peer links in the case of large content provider networks, which have aggressively added peer links in recent years.

The paper concludes with a discussion in Section 4 of some of the main lessons learned from our case studies, a brief review of related work in Section 5, and a summary detailing our future research plans in Section 6.

2. SEARCHING FOR THE GROUND TRUTH

This section gives a brief background on inter-domain network connectivity, defines its *ground truth*, and describes the various data sets that we used to infer the inter-domain connectivity.

2.1 Inter-domain Connectivity and Peering

The Internet consists of more than 26,000 networks called "Autonomous Systems" (AS). Each AS is represented by a unique numeric ID known as AS number and may advertise one or more IP address prefixes. ASes run the Border Gateway Protocol (BGP) [36] to propagate prefix reachability information among themselves. As a path-vector protocol, BGP includes in its routing updates the entire AS-level path, which is used as basic ingredient for inferring the AS-level topology. Projects such as RouteViews [12] and RIPE-RIS [11] host multiple data collectors that establish BGP sessions with hundreds of operational routers, which we term *monitors*, to obtain their BGP forwarding tables and routing updates over time. In the rest of the paper, we call the connection between two ASes an *AS link* or simply a *link*.

BGP routing decisions are largely based on routing polices, in which the most important factor is the business relationship between neighboring ASes. Though the relationship can be fine-grained, in general there are three major types: *customer-provider*, *peer-peer* and *sibling-sibling*. In a customer-provider relationship, the customer pays the provider for transiting traffic from and to the rest of the Internet, thus the provider usually announces all the routes to the customer. In a peer-peer relationship, which is commonly described as "settlement-free," the two ASes exchange traffic without paying each other. The catch is that only the traffic originated from and destined to the two peering ASes or their downstream customers is allowed on the peer-peer link; traffic from their providers or other peers are not allowed. Therefore an AS does not announce routes containing peer-peer links to its providers or other peers. When an AS has multiple neighbors which all announce a path to reach the same destination, the AS usually prefers the path announced by a customer over a peer and over a provider. This is referred to as the **no-valley-and-prefer-customer** policy [22] and is believed to be a common practice in today's Internet. The sibling-sibling relationship usually happens between two ASes that belong to the same organization. Since sibling-sibling relationship are relatively rare in today's Internet, we do not consider them in this paper.

Among all the ASes, about 20-30% are transit networks, and the rest are stub networks. A transit network is an Internet Service Provider (ISP) whose business is to provide packet forwarding service between other networks. Stub networks, on the other hand, do not forward packets for other networks. In the global routing hierarchy, stub networks are at the bottom or at the edge, and they need transit networks as their providers to gain access to the rest of the Internet. The transit networks may have their own providers and peers, and are usually described by being at different tiers, *e.g.*, regional ISPs, national ISPs, and global ISPs. At the top of this hierarchy are about a dozen tier-1 ISPs, which form the core of the global routing infrastructure and connect to each other to produce a fully meshed core graph. The majority of stub networks multi-home with more than one provider, and some stub networks also peer with each other. In particular, *content networks*, e.g. networks supporting search engines, e-commerce, and social network sites tend to peer with a large number of other networks.

Peering is a delicate issue in managing inter-domain connectivity. Networks have incentives to peer with other networks to reduce the traffic that has to be sent to providers, hence saving operational costs. But peering also comes with its own issues. For ISPs, besides additional equipment and management cost, they also do not want to peer with potential customers. Therefore ISPs in general are very selective in choosing their peers. Common criteria include number of common locations, ratio of inbound and outbound traffic, and certain requirements on prefix announcements [2, 1]. In recent years, with the fast growth of content that is available in the Internet, content networks have been keen on peering with other networks directly to bypass their providers. Content networks do not have the concern with transit traffic or potential customers, thus they usually maintain an open peering policy and are willing to peer with a large number of other networks.

Peering can be implemented in two ways: *private peering* and *public peering*. A private peering is a dedicated router-to-router layer-2 connection between two networks. Private peering provides dedicated bandwidth, is easier to troubleshoot problems, but has higher cost. Recently there is a trend to migrate private peerings to public peerings since the latter costs less and its bandwidth capacity is increasing. Public peering usually happens at the Internet Exchange Points (IXPs), which are third-party maintained physical infrastructures that enable physical connectivity between their

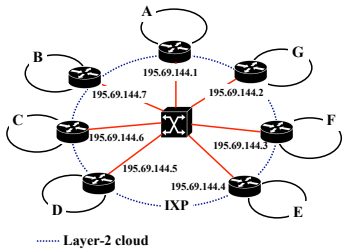

Figure 1: A sample IXP. ASes A through G connect to each other through a layer-2 switch in subnet 195.69.144/24.

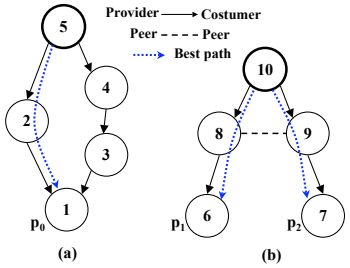

Figure 2: (a) Hidden Links, (b) Invisible Links

member networks[1]. Currently most IXPs connect their members through a common layer-2 switching fabric (or layer-2 cloud). Figure 1 shows an IXP that interconnects ASes A through G in the subnet 195.69.144/24. Though IXPs enable physical connectivity between all participants, whether to establish BGP peering sessions on top of the physical connectivity is up to individual networks. It is possible that one network may only peer with some of the other participants in the same IXP.

2.2 Ground Truth vs. Observed Map

Before starting the study of AS-level connectivity, we should define clearly what constitutes an inter-AS link. *A link between two ASes exists if the two ASes have a contractual agreement to exchange traffic over one or multiple BGP sessions.* The **ground truth** of AS-level connectivity is the *complete* set of AS links. As the Internet evolves, the AS-level connectivity also changes over time. We use $G_{real}(t)$ to denote the ground truth of the entire Internet AS-level connectivity at time t, and $A_{real}(t)$ to denote the ground truth of an individual AS A's connectivity at time t.

Ideally if ISPs maintain an up-to-date list of their AS links and make the list accessible, obtaining the ground truth would be trivial. However, such a list is proprietary and often not available, especially for large ISPs, who have a large and changing set of links. In this paper, we derive the ground truth of several individual networks from their router configurations, syslogs, BGP command outputs, and personal communications with operators.

From router configurations, syslogs and BGP command outputs, we can infer whether there is a working BGP session, *i.e.*, a BGP session that is in the *established* state as specified in RFC 4271 [36]. We assume there is a link between two ASes if there is at least one working BGP session between them, and there is no link if no working BGP session exists between them. If all the BGP sessions between two ASes are down at the moment of data collection, even though they have a valid agreement to exchange traffic, the link may not appear in the ground truth of that particular day, but it will show up in the ground truth of a later day when at least one session is established. Since we have continuous daily data going back for years, the problem of missing links in our inferred ground truth should be negligible.

Since the ground truth is usually not available, the most commonly used AS-level maps are built from observation data collected by remote monitors. We denote an observed global AS topology at time t by $G_{obsv}(t)$. Though BGP and traceroute use different mechanisms to collect data, neither was designed to distribute or measure AS-level connectivity. Inevitably, $G_{obsv}(t)$ typically provides only a partial view of the ground truth.

[1]Note that private and public peering can happen in the same physical facility.

There are two types of missing links when we compare G_{obsv} and G_{real}: **hidden links** and **invisible links**. Given a set of monitors, a hidden link is one that has not yet been observed but could possibly be revealed at a later time. An invisible link is one that is impossible to be observed by the set of monitors. For example, in Figure 2(a), AS5 is the monitor[2], and between the two customer paths to reach prefix p_0, it picks the best one, [5-2-1]. Given this selection, we would only be able to observe the existence of AS links 2-1 and 5-2. However, the three missing links, 5-4, 4-3, and 3-1, are hidden links because they will be revealed whenever AS5 switches to path [5-4-3-1] due to a failure in the primary path [5-2-1]. In Figure 2(b), the monitor AS10 uses paths [10-8-6] and [10-9-7] to reach prefixes p_1 and p_2, respectively. The missing link 8-9 is invisible, because this peer link will not be announced to AS10 under any circumstances due to the no-valley policy. It simply cannot be observed by the current monitor AS10.

Hidden links are typically revealed if we build AS maps using continuous data (*e.g.*, BGP updates) collected over an extended period. However, a problem of this approach is the introduction of potentially stale links; that is, links that existed some time ago but are no longer present. Therefore we need to set a timeout to remove possible stale links as suggested in [34]. To discover invisible links, we would need additional monitors at the place where the links are allowed to be announced by routing policy. These intrinsic limitations are shared by both BGP and traceroute measurements.

2.3 Data Sets

We use several different types of data to infer the AS-level connectivity and the ground truth of individual ASes.

BGP data: The **public view (PV)** of the AS-level connectivity is derived from all public BGP data at our disposal. These data include BGP forwarding tables and updates from ∼700 routers in ∼400 ASes provided by Routeviews, RIPE-RIS, Abilene [14], and the China Education and Research Network [3], BGP routing tables extracted from ∼80 route servers, and "show ip bgp sum" outputs from ∼150 looking glasses located worldwide. In addition, we use "show ip bgp" outputs from Abilene and Geant [5] to infer their ground truth. Note that we currently do not use AS topological data derived from traceroute measurements due to issues in converting router paths to AS paths, as extensively reported in previous work [18, 30, 23, 34].

IXP data: There are a number of websites, such as Packet Clearing House (PCH) [8], Peeringdb [9], and Euro-IX [4] that maintain a list of IXPs worldwide and also provide a list of ISP participants in some IXPs. Though the list of IXP facilities is close to be complete [10], the list of ISP participants at the different IXPs may be incomplete or outdated since it is inputted by the ISPs on a vol-

[2]Either a BGP monitoring router or a traceroute probing host

Presences (AS-IXP pairs)	Peeringdb	Euro-IX	PCH
Listed on source website	2,203	2,478	575
Inferred from reverse DNS	2,878		3,613
Unique within the source	4,092	2,478	3,870
Total unique across all sources	6,084		

Table 1: IXP membership data, July 2007.

untary basis. However, most IXPs publish the subnets they use in their layer-2 clouds, and best current practice [6] recommends that each IXP participant keeps reverse DNS entries for their assigned IP addresses inside the IXP subnet and no entries for unassigned addresses. Based on this, we adopted the method used in [43] to infer IXP participants. The basic idea is to do reverse DNS lookups on the IXP subnet IPs, and then infer the participant ISPs from the returned domain names. From the aforementioned three data sources, we were able to derive a total of 6,084 unique presences corresponding to 2,786 ASes in 204 IXPs worldwide. A *presence* means that there exists an AS-IXP pair. For example, if two ASes peer at two IXPs, it will be counted as two presences. Table 1 shows the breakdown of the observed presences per data source. Even though we do not expect that our list is complete, we noticed that the total number of presences we obtained is very close to the sum of the number of participants in each IXP disclosed on the PCH website.

IRR data: The Internet Routing Registry (IRR) [7] is a database to register inter-AS connectivity and routing polices. Since registration with IRR is done by ISP operators on a voluntary basis, it is well known that its information is incomplete and many records are outdated. We carefully filtered IRR records by ignoring all entries that had a "Last Modified" date that was more than one year old.

Proprietary Router Configurations and Syslogs: This is a major source for deriving the ground truths of a Tier-1 and a Tier-2 ISP[3]. The data include historical configuration files of more than one thousand routers in these two networks, historical syslog files from all routers in the Tier-1 network, and "show ip bgp sum" outputs from all routers in the Tier-2 network. We also have access to iBGP feeds of several routers in these two networks.

Other Proprietary Data: To obtain the ground truths for other types of networks, we had conversations with the operators of a small number of content providers. Since large content providers are unwilling to disclose their connectivity information in general, in this paper we present a factious content provider whose numbers of AS neighbors, peer links, and IXP presences are consistent with the data we collected privately. We also obtained the ground truths of AS-level connectivity for four stub networks directly from their operators.

3. CASE STUDIES

In this section we infer the local ground truth of networks from which we have operational data, and compare it to the connectivity derived from BGP data to find out what links are missing from public view and why they are missing.

3.1 Tier-1 Network

In order to apply our definition of AS-level connectivity ground truth we need to know at each instant what are the BGP sessions that are in the established state for all routers of the network. The straightforward way to do this is to launch the command "show ip bgp summary" in all the routers simultaneously. An example output

[3]Here by Tier-2 we refer to a transit provider that is a direct customer of a Tier-1 AS.

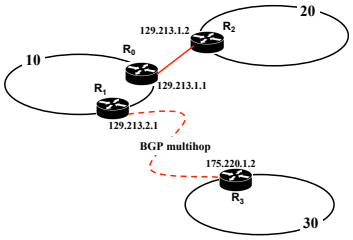

Figure 4: Configuring remote BGP peerings. R_0 and R_2 are physically directly connected, while R_1 and R_3 are not.

of this command is shown in Figure 3. The state of each BGP session can be inferred just by looking at the column "Sate/PfxRcd" (last column). In this case, all connections are in the *established* state except for the session with neighbor 64.125.0.137, which is in the *idle* state[4].

Due to large scale of the studied Tier-1 network, it is infeasible to repeatedly run the "show ip bgp sum" command in all the routers of the network to obtain data for a long study period, and it is also impossible to obtain any historic "show ip bgp sum" data for a past period during which this command was not run. Therefore, instead we resort to an alternative way to infer the connectivity ground truth - analyze routers' configuration files. Routers' configuration files are a valuable source of information about AS level connectivity. Before being able to set up a BGP session to a remote AS, each router needs to have a minimum configuration state. As an example, in Figure 4, in order for router R_0 in AS10 to open a BGP session with R_2 in AS20, it needs to have a "neighbor 129.213.1.2 remote-as 20" entry in its configuration. But even before that, in order to have IP connectivity between R_0 and R_2, R_0 needs to have configured a route to reach R_2, and R_2 needs to have configured a route to reach R_0.

The IP connectivity between the two routers of a BGP sesion can be accomplished in two different ways:

- **Single-hop**: two routers are physically directly connected, as R_0 and R_2 are in Figure 4. More specifically R_0 can (1) define a *subnet* for the local interface at R_0 that includes the remote address 129.213.1.2 of R_2, *e.g.* " ip address 129.213.1.1 255.255.255.252" (where 255.255.255.252 refers to the net mask) or (2) set a *static route* in R_0 to the remote address 129.213.1.2 of R_2, *e.g.* "ip route 129.213.1.0 255.255.255.252 Serial4/1/1/24:0" (in this case Serial4/1/1/24:0 refers to the name of the local interface at R_0).

- **Multi-hop**: two routers (such as R_1 and R_3 in Figure 4) are not physically directly connected. Instead, they are connected via other routers. To configure such a multi-hop BGP session, R_1 configures *e.g.* "neighbor 175.220.1.2 ebgp-multihop 3" (in this case 3 refers to the number of IP hops between R_1 and R_3); R_1 reaches R_3 by doing longest prefix matching of 175.220.1.2 in its routing table.

Ideally, we would like to check whether the IP connectivity is configured correctly on both sides of a session. However, it is usually impossible to get the router configs of the neighbor ASes. We thus limit ourselves to check only the IP connectivity of routers

[4]Whenever this column shows a numeric value, it refers to the number of prefixes received from the neighbor for the session, and it is implied that the BGP state is *established*.

```
Neighbor        V    AS MsgRcvd MsgSent   TblVer  InQ OutQ Up/Down  State/PfxRcd
4.68.1.166      4  3356  387968     6706  1652742    0    0 4d15h     231606
64.71.255.61    4   812  600036     6706  1652742    0    0 4d15h     230964
64.125.0.137    4  6461       0        0        0    0    0 never    Idle
65.106.7.139    4  2828  466128     6706  1652742    0    0 4d15h     232036
```

Figure 3: Output of "show ip bgp summary" command.

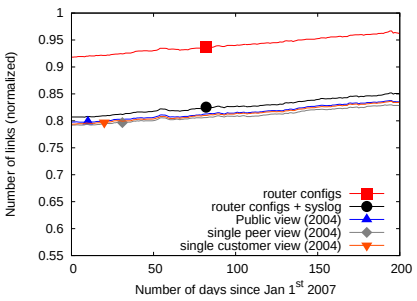

Figure 5: Connectivity of the Tier-1 network (since 2004).

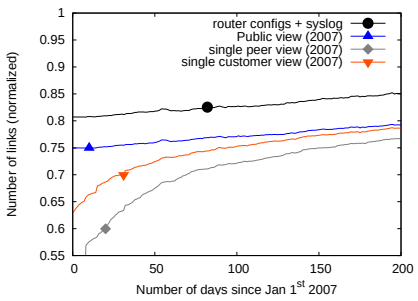

Figure 6: Connectivity of the Tier-1 network (since 2007).

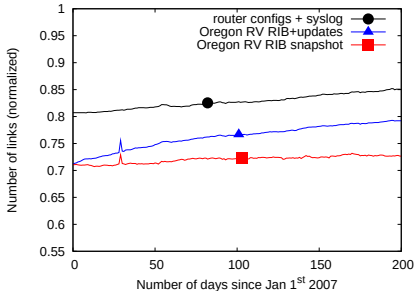

Figure 7: Capturing the connectivity of the Tier-1 network through table snapshots and updates.

belonging to the Tier-1 network. We noticed there were several entries in the router config files that did not satisfy the minimal BGP configuration described above, probably because the session was already inactive, and these sessions should be discarded. After searching systematically through the historic archive of router config files, we ended up with a list of neighbor ASes that have at least one session with a valid BGP configuration. The "router configs" curve in Figure 5 shows the number of neighbor ASes in this list over time[5].

However, even after this filtering, we still noticed a considerable number of neighbor ASes that appeared to be "correctly configured", but did not have any established BGP session. This is probably because routers on the other side of the sessions do not correctly configure the session or the connectivity underlying the session. Given that we do not have the configurations for those neighbor routers, we utilize router syslog data to further filter out the possible stale entries in the Tier-1's router configs. Syslog records include information about BGP session failures and recoveries, indicating at which time each session comes up and down. We use the following two simple rules to further filter the previous list of neighbors:

1. If the last message of a session occurs at day t and the content was "session down"', and if there is no other message from the session in the period [t, t + 1 month], then assume the session was removed at day t (*i.e.* we need to wait at least one month before discarding the session).

2. If a session is seen in a router config file at day t, but does not appear in syslog for the period [t, t + 1 year], then assume session was removed at day t (*i.e.* we need to wait at least 1 year before discarding the session).

Note that the above thresholds were empirically selected to minimize the number of false positives and false negatives in the inferred ground truth. A smaller value for the above thresholds would increase the number of false negatives (*i.e.* sessions that are prematurely removed by our scheme while still in the ground truth), whereas a higher value would increase the false positives (*i.e.* sessions that are no longer in the ground truth, but have not been removed yet by our scheme). Even though these threshold values worked well in this case, depending on the stability of links and routers' configuration state, other networks may require different values. Note also that these two rules are for individual BGP sessions only. An AS-level link between the Tier-1 ISP and a neighbor AS will be removed only when *all* of the sessions between them are removed by the above two rules. The sessions between the Tier-1 ISP and its peers tend to be stable (probably due to better equipments) with infrequent session failures [41], thus it is possible that a session never fails within a year. But our second rule above is not likely to remove the AS-level link between the Tier-1 ISP and its peer because there are usually multiple BGP sessions between them and the probability that none of the sessions have any failures for an entire year is very small. Similarly, this argument is true for large customers who have multiple sessions with the Tier-1 ISP. On the other hand, small customers tend to have small number of sessions with the Tier-1 ISP (maybe 1 or two), and the sessions tend to be less stable thus have more failures and recoveries. Thus if such a session is still valid, the above two rules will not filter them out since some syslog session up or down messages will be seen. For similar reasons, the results are not significantly affected by the fact that syslog messages might be lost during transmission due to unreliable transport protocol (UDP).

Using the two simple rules above, we removed a considerable number of entries from the config files, and obtained the curve "router configs+syslog" in Figure 5[6]. Once we achieved a good approximation of the ground truth, we compared it to the BGP-derived connectivity. For each day t, we compared the list of ASes in the inferred ground truth $T_{tier1}(t)$ obtained from router configs+syslog, with the list of ASes seen in public view (defined in section 2.3) as connected to the Tier-1 network up to day t. The "Public view (2004)" curve is obtained by accumulating public view BGP-derived connectivity since 2004. Comparing this curve with the "router configs+syslog" curve we note that there is an almost constant small gap, which is in the order of some tens of links (3% of the total links in "router configs+syslog"). We manually investigated these links, and found that there are three main causes for why they do not show up in the public view: (1) ASes that only ad-

[5]Note that the number is normalized for non-disclosure reasons.

[6]Note that our measurement actually started in 2006-01-01, but we used an initial 1-year window to apply the second syslog rule.

vertise prefixes longer than /24 which are then aggregated, thus the Tier-1 AS never sends any routes with such a neighbor's AS number in the path. This contributes to about half of the missing links; (2) there is one special purpose AS number (owned by the Tier-1 ISP) which is only used by the Tier-1 ISP; (3) false positives, *i.e.* ASes that were wrongly inferred as belonging to $T_{tier1}(t)$, including stale entries, as well as newly allocated ASes whose sessions were not up yet (thus not removed by the second syslog rule). The false positive contributes to about half of the "missing links" (which should be not really called "missing"). One additional note is that all the Tier-1 ISP's links to its peers and sibling ASes are captured by the public view. The complete coverage of peer-peer links is because such a link is not invisible as long as there is a monitor in either AS's customer or customer's customers, and so on, which is apparently true for tier1's peer-peer links given the small number of tier-1 networks and the fairly large set of monitors in public view.

Figure 6 shows similar curves using the same vertical scale as in Figure 5, but this time the public view BGP data collection is started in the beginning of 2007. When comparing "Public view (2007)" and "router configs+syslog" we note the gap is higher, which indicates that some of the entries in "router configs+syslog" did not show up in public view after 2007, but they did show up before, which likely means they are stale entries (false positives).

The "Single customer view" and "Single peer view" curves in both Figures 5 and 6 represent the Tier-1 connectivity as seen from a single router in a customer of the Tier-1 ISP and a single router in a peer of the ISP, both publicly available. In this case the single peer view captures slightly less links than the single customer view, corresponding to about $\sim$1.5% of the total number of links of the Tier-1 network. Further analysis revealed that this small delta corresponds to the peer links of the Tier-1, which are included in routes advertised to the customer but not advertised to the peers, which is expected and consistent with the no-valley policy. We also note that the "Single peer view" and "Single customer view" curves in Figure 6 show an exponential increase in the first few days of the x-axis, which is caused by the revelation of hidden links, as explained in Section 2.2. However, the nine months of the measurement should be enough to reveal the majority of the hidden links [34]. In addition, note that in both figures, the "Single customer view" curve is very close to the public view curve, which means that the connectivity seen by the customer is representative of what is visible from the public view.

Figure 7 shows the difference between using single routing table snapshots (RIB) versus initial RIB+updates from all the routers at Oregon RouteViews (a subset of 46 routers of the entire public BGP view). Note that in each day, the number of links in the curves "Oregon RV (RouteViews) RIB snapshot" and "Oregon RV RIB+updates" represent the overlap with the set of links in the inferred ground truth represented by the curve "router configs+syslog", *i.e.* , those links not in "router configs + syslog" are removed from the two "Oregon RV" curves. Even though both curves start in the same point, after more than nine months of measurement, "Oregon RV RIB+upates" reveals about 10% more links that those revealed by "Oregon RV RIB snapshot", which were likely revealed by alternative routes encountered during path exploration as described in [33]. Note that the difference between the two curves are all customer-provider links, and all the Tier-1 ISP's links to the peers are captured by the "Oregon RV RIB snapshot" given the large number of routes that go through these peer-peer links.

Summary:

- A single snapshot of the Oregon RV RIB can miss nontrivial percentage (e.g., 10%) of the Tier-1's AS-level links, all of them customer-provider links, when compared to using RIBs+updates accumulated in several months.
- The Tier-1 AS's links are covered fairly completely by the public view over time. All the peer-peer and sibling links are covered; the small percentage (e.g., 1.5%) of links missing from public view are those invisible ones with customers who only announce prefixes longer than /24.
- The Tier-1 AS's links are covered fairly completely by a single customer (as long as the historic BGP tables and updates are used), which can be considered representative of the public view.
- The Tier-1 AS's links are covered fairly completely by a single peer (as long as the historic BGP table and updates are used), while there are about 1.5% missing links, all of which are peer-peer links.

3.2 Tier-2 Network

The case of the Tier-2 network is different from the previous Tier-1 case. First of all, not being a Tier-1 network, the Tier-2 has providers. Second, it has considerably more peers than the Tier-1 network, and it is considerably smaller in size. In fact, even though the studied Tier-1 network peers exclusively through private peering, the Tier-2 network had close to $\frac{2}{3}$ of its peers in IXPs. We do an analysis similar to the Tier-1 case, except that now we do not have access to syslog data.

The "router configs" curve in Figure 8 shows the number of neighbors obtained from router configurations over time. Let us assume for now this is a good approximation of the ground truth of the Tier-2 network connectivity. We include in the figure single router views from a single router in a customer of the Tier-2 network, and single router in a provider of the Tier-2 network, both publicly available. Note that this time we started the measurement in March 2007, when the BGP data for the customer router first became available in the public view. Also note that the same customer router became unavailable on August 13, 2007, which is the reason the single customer view curve is chopped off at the end in the figure. Figure 8 shows that the provider view misses a significant number of links that are captured by the customer. In fact, the provider is missing more than 12% of the connectivity captured by the customer, which corresponds to the peer links of the Tier-2 network, and is consistent with the no-valley policy. For comparison, we also included the public view curve, starting at March 10^{th} 2007. Note that there is a very small number of neighbors present in public view, but not present in the customer view. We discovered that most of the links in this gap were revealed by routes with several levels of AS prepending originated by customers of the Tier-2. Due to the path inflation caused by the AS prepending, the Tier-2's customer we used was not picking these routes, but due to the prefer-customer policy, routers inside the Tier-2 network were picking them, including a router that was also in our public view set.

From Figure 8 we also note that the connectivity captured by the public view is $\sim$85% of that inferred from router configs, which hints that there might be a high number of false positives in the ground truth inferred solely from router configs. To eliminate these false positives, we launched a "show ip bgp summary" command on all the routers of the network in 2007-09-03, based on which we keep only those BGP sessions that were in the established state. The number of neighbors with at least one such session is shown in Figure 8 by the "show ip bgp sum" point, which has only 80% of the connectivity inferred from router configs. This means that about 20% of the connectivity extracted from router configs were

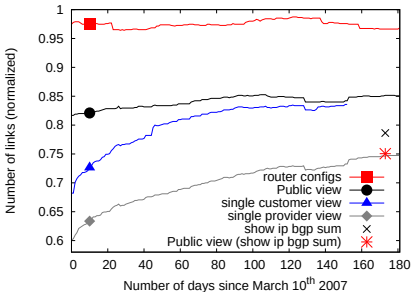

Figure 8: Tier-2 network connectivity.

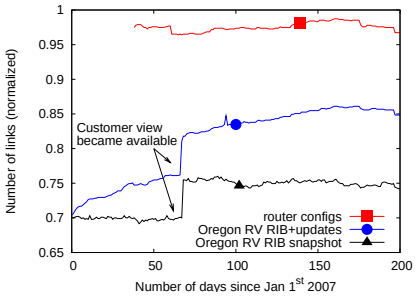

Figure 9: Capturing Tier-2 network connectivity through table snapshots and updates.

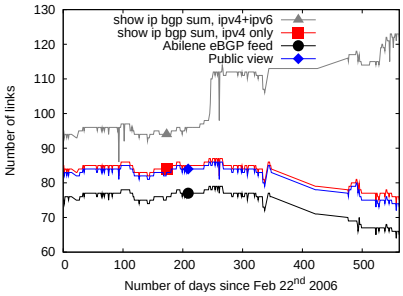

Figure 10: Abilene connectivity.

false positives. The point "Public view (show ip bgp sum)" in the figure represents the intersection between the set of neighbors extracted from "show ip bgp sum" and the set of neighbors seen so far in the public view. Note that public view is missing ∼7% of the links given by "show ip bgp sum", which in absolute numbers amounts to a few tens of links. One of these links was the RouteViews passive monitoring feed, other were internal AS numbers, and the remaining were cases of longer than /24 routes, which were being aggregated. Note also that the fairly complete coverage of the Tier-2 network's connectivity is due to the fact we have a public view monitor in a customer of the Tier-2, and as we explained in the Tier-1's study, a peer-peer link is not invisible if there is a monitor in either AS's customer or customer's customers, and so on.

Figure 9 shows the difference between using single RIB snapshot versus initial RIB+updates from Oregon RouteViews, using the same vertical scale as in Figure 8. In this case, using updates reveals ∼12% more links than those revealed by single router snapshots in the long run. Note that there is a lack of configuration files for the first days of 2007, hence the missing initial part on the curve "router configs". The jump in the figure is caused by the appearance of the customer AS in the Oregon RV set, which revealed the peer links of the Tier-2 network.

Summary:

- A single snapshot of the Oregon RV RIB can miss nontrivial percentage (e.g., 12%) of the Tier-2's AS-level links, all of them customer-provider links, when compared to using RIBs+updates accumulated in several months.
- The Tier-2 AS's links are covered fairly completely by a single customer over time (RIBs +updates), which can be considered representative of the entire public view. The very small percentage of hidden links (<1%) corresponding to long AS prepending cases.
- A single provider view can miss nontrivial percentage (e.g., 12%) of the Tier-2's links, and all the missing links are peer-peer links.
- A Tier-2 AS's links are covered fairly completely by the public view over time if there is a monitor in its customer or its customer's customers, in which case all the peer-peer links are covered. The small percentage (e.g., 7%) of links missing from the public view are those invisible ones with customers who only announce prefixes longer than /24 or those ASes dedicated for internal use.

3.3 Abilene and Geant

Abilene: Abilene (AS11537) is the network interconnecting universities and research institutions in the US. The Abilene Observatory [14] keeps archives of the output of "show ip bgp summary" for all the routers in the network. Using this data set, we built a list of Abilene AS neighbors over time, which is shown in the "show ip bgp sum, ipv4+ipv6" curve in Figure 10. Even though Abilene does not provide commercial-to-commercial transit, it enables special arrangements where its customers might inject their prefixes to commercial providers through Abilene, and receive routes from commercial providers also through Abilene. The academic-to-commercial service is called Commercial Peering Service (or CPS) versus the default academic-to-academic Research & Education (R&E) service. These two services are implemented by two different VPNs that are both layered on top of the Abilene backbone. BGP sessions for both services are included in the output of "show ip bgp summary". We compare Abilene connectivity ground truth with that derived from a single router eBGP feed (residing in Abilene) containing only the R&E sessions. In addition, we do a similar comparison with our public view, which should contain both CPS and R&E sessions (since public view contains eBGP+iBGP Abilene feeds, as well as BGPs from commercial providers of Abilene). However, since there are a considerable number of neighbors in Abilene that are using only ipv6, and since the BGP feeds in our data set are mostly ipv4-only, we decided to place the ipv4-only neighbors in a separate set. The curve "show ip bgp sum, ipv4 only" in Figure 10 shows only the AS neighbors that have at least one ipv4 session connected to Abilene[7]. Contrary to the "show ip bgp sum, ipv4+ipv6" curve which includes all sessions, the ipv4-only curve shows a decreasing trend. We believe this is because some of the ipv4 neighbors have been migrating to ipv6 over time. When comparing the "show ip bgp sum, ipv4 only" curve with the one derived from the eBGP feed, we find there's a constant gap of about 10 neighbors. A close look into these cases revealed that these AS numbers belonged to commercial ASes with sessions associated with the CPS service. The small gap between the public view and the ipv4-only curve corresponds to the passive monitoring session with RouteViews (AS6447).

Geant: Geant (AS20965) is the European research network connecting 26 R&E networks representing 30 countries across Europe. In contrast to Abilene where the philosophy is to focus on establishing academic-to-academic connectivity, Geant enables its members to connect to the commercial Internet using its backbone. We inferred Geant connectivity ground truth by running the

[7]Note that there was a period of time between days 350 and 475 for which there was no "show ip bgp sum" data from Abilene.

command "show ip bgp sum" in all its routers through its looking glass site [5]. We were able to find a total of 50 AS neighbors with at least one session in the established state. When comparing Geant ground truth with the connectivity revealed in public view, we found a match on all neighbor ASes except two cases. One of the exceptions was a neighbor which was running only ipv6 multicast sessions, and therefore hidden from public view which consists mostly of ipv4-only feeds. The second exception was what seems to be a passive monitoring session to a remote site, which explains why its AS number was missing from BGP feeds.

Summary: In Abilene and Geant, the public view matches the connectivity ground truth (no invisible or hidden links), capturing all the customer-provider and peer links. Abilene represents a special case, where depending on the viewpoint there can be invisible links. For instance, some Abilene connectivity may be invisible to its customers due to the academic-to-commercial special arrangements.

3.4 Content provider

Content networks are fundamentally different from transit providers such as the Tier-1 and Tier-2 cases we studied above. Content networks do not have to transit traffic between networks, thus they only have peers and providers, but no customers. They usually try to maximize the number of peers, and by doing so, they reduce the amount of (more expensive) traffic that is forwarded to providers. Therefore, content networks usually have a heavy presence in IXPs, where they can peer simultaneously with multiple different networks. Another difference is that while two transit providers usually peer at every location where they have a common presence in order to disperse traffic to closer exit-points, peering in content networks is more "data-driven" (versus "route-driven"), and may happen in only a fraction of the IXPs where two networks have common locations. Based on this last observation, we estimate the connectivity of a representative content provider C, and compare it to the connectivity observed from our BGP-derived public view. We assume that in each IXP where C has presence, it connects to a fixed fraction q of the networks also colocated at that IXP, *i.e.* if C has n common locations with another network X, then the chances that C and X are connected in at least one IXP are given by $1-(1-q)^n$. More generally, the expected number of peer ASes of C, P_C is given by $P_C = \sum_i (1-(1-q)^{n_i})$, where i represents all networks that have at least one common presence with C, and n_i is the number of IXPs where both C and i have presence. In our data set, C has presence in 30 IXPs worldwide, which is very close to the number that was disclosed to us by the operators of C. Furthermore, we know that the number of providers of C is negligible compared to the number of peers, and that more than 95% of its peerings are done at IXPs. Therefore it is reasonable to represent the AS-level connectivity of C by its peerings at IXPs.

Figure 11 shows the projection of the number of neighbor ASes of C as a function of the connection probability q at each IXP. For comparison purposes, we also include the number of neighbor ASes of C as inferred from the public view over a window of 6 months. From discussions with the network operators of C, we know that at each IXP, C peers with about 80-95% of the participants (parameter q), and that the total number of BGP sessions of C is close to 3,000, even though we do not know the total number of unique peer ASes[8]. In view of these numbers, the projection in Figure 11 seems reasonable, even taken in account that our IXP membership data is incomplete. The most striking observation is

[8]Note that the number of unique neighbor ASes is less than the total number of BGP sessions, as there exist multiple BGP sessions with the same neighbor AS.

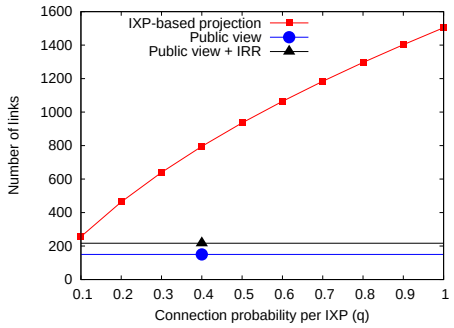

Figure 11: Projection of the number of peer ASes of a representative content provider.

the vast amount of connectivity missed from BGP-derived public view, on the order of thousands of links representing about 90% of C's connectivity. This is not entirely surprising, however, given that the content pprovider C will not announce its peer-peer links to anyone due to no-valley policy, and a peer-peer link is visible only when the public view has a monitor in the peer or a customer of the peer, and the number of such monitors is much smaller than the projected total number of peer-peer links of C. We believe the same is true for other large content providers, search engines, and content distribution networks.

Trying to close this gap, we looked for additional connectivity in the IRR, as described in Section 2.3. We discovered 62 additional neighbor ASes for C that were not present in the initial set of 155 ASes seen in public view. Even though we increased the number of covered neighbor ASes of C to 217, it still represents only about 15% of the AS-level connectivity of C.

Summary: Even accumulating public view over 6 months, we are still missing about 90% of C's connectivity, most of which are invisible peer-peer links at IXPs. Using IRR information can slightly reduce the missing connectivity to 85%. The public BGP view's inability to catch these peer-peer links is due to the no-valley policy and the absence of monitors in the peers or the customers of peers of the content network.

3.5 Stub networks

Stub networks correspond to stub ASes that are only connected to providers (with no peers or customers), typically representing small companies/institutions. Even though their degree is usually small (<4), they represent most of the networks in the Internet. We obtained the AS-level connectivity ground truth of 4 stub networks by directly contacting the operators. Table 2 shows for each network the number of neighbor ASes in the ground truth as reported by the operators, as well as the number of neighbor ASes captured by the BGP-derived public view. Note that for public view we use 6 month worth of BGP RIB and updates to accumulate the topology to account for hidden links that take time to be revealed [34]. Network D is the only case where there is a perfect match between ground truth and public view. For network A, there are two neighbors included in public view that were disconnected during the 6-month window (false positives). For network B, the public view was missing a neighbor due to a special peer-peer like agreement in which the routes learned from the neighbor are not announced to B's provider. Finally, for network C there was an extra neighbor in public view that was never connected to C, but appeared in routes during one day in the 6-month window. We believe this case was originated either by a misconfiguration or a malicious false link attack.

Network	# of neighbor ASes in ground truth	#of neighbor ASes in public view
A	8	10
B	7	6
C	3	4
D	2	2

Table 2: Connectivity of stub networks.

Summary: The 6-month accumulated public view captured all the customer-provider links of the stub networks studied. In total, the public view has 1 false negative (invisible link) and 3 false positives, the later being possible to eliminate by reducing the interval of the observation window of public view.

4. LESSONS LEARNED

The premise of this paper is that the efforts intended to infer the complete maps of the Internet's AS-level connectivity structure from observation data are doomed. In this section, we back that premise by summarizing the classes of topological information that are captured and necessarily missed in the public view, and illustrate with a number of concrete examples some of the consequences of relying on incomplete AS topologies. We also explain how our newly gained insights can help identify the types of studies whose results are insensitive to the current limitations of the inferred topology maps. Lastly, we sketch a concrete alternative for dealing with the completeness problem by outlining a practical construction for generating fully functional and economically viable AS topologies.

4.1 "Public view" vs. ground truth

We use Figure 12 as an illustration to summarize the public view's quality in terms of completeness of the observed topology. Note that our observations here are the natural results of the *no-valley-and-prefer-customer* policy, and some of them have been speculated briefly in previous work. However, they are quantified and verified for the first time in this paper by means of comparing the ground truth with the observed topology. Though the few classes of networks we have examined are not necessarily exhaustive, we believe the observations drawn from these case studies are representative in the Internet in general.

First, if an AS has a monitor inside its network, the public view should be able to capture all of its direct links, including customer-provider and peer links. However, not all the links of the AS may show up in a snapshot of the observed topology. It takes some time, which can be as long as a few years, to have all hidden customer-provider links exposed by routing dynamics. Second, a monitor at a provider network should be able to capture all the provider-customer links between its downstream customers, and a monitor in a customer network should be able to capture all the customer-provider links between its upstream providers. For example, in Figure 12, a monitor at AS2 is able to capture not only its direct provider-customer links (2-6 and 2-7), but also the provider-customer links between its downstream customers (6-8, 6-9, 7-9, and 7-10). AS5, as a peer of AS2, is also able to capture all the provider-customer links downstream of AS2 since AS2 will announce its customer routes to its peers. Again, it can take quite a long time to actually discover all the hidden links. Third, a monitor cannot observe a peer link at a lower tier or non-direct peer links at the same tier [9]. For example, a monitor at AS5 will not be able to capture the peer link 6-7 or 1-2, because a peer route is not announced to providers or other peers according to the *no-valley* policy. Fourth, to capture a remote peer link, we need a monitor at a downstream customer of one of the peer ASes incident to the link. For example, a monitor at AS9 would be able to capture peer links 6-7 and 5-2, but not the peer link 1-3 since AS9 is not a downstream customer of either AS1 or AS3.

The current public view has monitors in all Tier-1 ASes but one. Even for the only Tier-1 AS that does not have a monitor, there is a monitor in one of its direct customers. Together with the above observations, we can summarize and generalize the quality of the public view coverage as follows.

- **Coverage of Tier-1 links**: The public view contains all the links of all the Tier-1 ASes.
- **Coverage of customer-provider links**: There are no invisible customer-provider links. Over time the public view can reveal all the customer-provider links in the Internet topology, *i.e.* , the number of hidden customer-provider links will gradually approach zero after an observation period long enough. This is supported by our empirical findings that in all our cases studies we were able to discover all customer-provider links using BGP data collected over several years.
- **Coverage of peer links**: The public view potentially misses a large number of peer links, especially in the lower tier of the Internet routing hierarchy. The public view will not capture a peer link A–B unless the public view has a monitor installed in either A or B, or in a downstream customer under A or B. However, the public monitors are in about 400+ ASes out of a total of 26,000+ existing ASes, which gives a rough perspective on the percentage of peer links missing from the public view. Peer links between stub networks (*i.e.* , links 8-9 and 9-10 in Figure 12) are among the most difficult ones to capture. Unfortunately, with the recent growth of content networks, it is precisely these links that are rapidly increasing in numbers.

4.2 Using "public view" in practice

Different research projects and studies that involve Internet-wide simulations and rely on inferred AS topologies are likely to be impacted differently by the deficiencies of the used AS maps as revealed by our case studies. In the following, we illustrate with some concrete examples some of the problems that can arise.

Stub AS growth rates and network diameter: Given that the public view captures almost all the AS nodes and customer-provider links, it is an adequate data source for studies of AS-topology metrics such as network diameter; growth rates and trends for the number of AS stub networks; and quantifying customer multihoming.

Other graph-theoretic metrics: Given that the public view is by and large inadequate as far as the coverage of peer links is concerned, and given that these peer links typically allow for shortcuts in the data plane, relying on the public view can clearly cause major distortions when studying generic graph properties such as node degrees, path lengths, clustering, etc.

Impact of prefix hijacking: Prefix hijacking is currently a serious security threat in the Internet and happens when an AS announces prefixes that it does not own. Recent work on this topic [24, 46, 15, 42] relies on evaluations based on inferred AS topologies that have one or more of the very limitations elaborated on in the previous sections. Depending on the exact hijack scenario, the impact can be either underestimated or overestimated. Figure 13 shows an example of a hijack simulation scenario, where AS2 announces prefix p belonging to AS1. Because of the invisible peer

[9]We are assuming the provider-customer links do not form a circle, which, if exists, should be very rare.

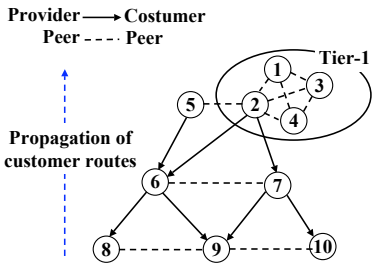

Figure 12: Customer-provider links can be revealed over time, but downstream peer links are invisible to upstream monitors.

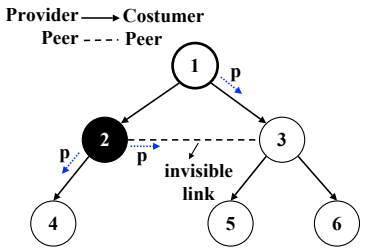

Figure 13: Example of a prefix hijack scenario where AS2 announces prefix p belonging to AS1. Because of the invisible peer link AS2–AS3, the number of ASes affected by the attack is underestimated.

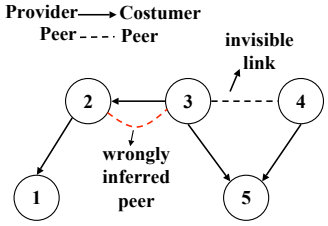

Figure 14: Example of AS relationship inference, where the invisible peer link 3–4 observed in a path 1–2–3–4 invalidates the inferred peer link 2–3 since it creates a no-valley violation.

link 1–2, the number of impacted ASes is underestimated, *i.e.* ASes 3,5 and 6 are believed to pick the route originated by AS1, whereas in reality they would pick the more preferred peer route coming from the hijacker AS2. On the other hand, with an incomplete topology, a simulation could also overestimate the impact of a hijack scenario. For example, since the peers of content network C considered in Section 3 have direct routes to C, they are not likely to be impacted by a hijack by a remote AS, so missing 90% of the C's peer links would significantly overestimate the impact of such a hijack. Note that a different scenario where C is the hijacker would result in an underestimation of the impact.

Relationship inference/path inference: Several studies [22, 39, 29] have addressed the problem of inferring the relationship between ASes based on observed routing paths. Figure 14 shows an example where the customer-provider link 2–3 is wrongly inferred as a peer link based on the observed set of paths, creating a no-valley violation. Knowledge of the invisible peer link 3–4 present in path 1–2–3–4 would have avoided the previous error. The path inference heuristics [29, 31, 32] (relying on relationship or not) are also impacted by the incompleteness problem, mainly because they a priori exclude all paths that traverse invisible peer links.

Routing resiliency to failures: Studies that address robustness properties of the Internet under different failure scenarios (e.g., see [21, 42]) also depend strongly on the assumption of a complete and accurate AS-level topology on top of which failures are simulated. We can easily envision scenarios where two parts of the network are thought to become disconnected after a failure, even though there are invisible peer links connecting them. Knowing that currently used inferred AS maps tend to miss a substantial number of peer links, robustness-related claims that are based on taking these maps at face value need to be viewed with a grain of salt.

Evaluation of new inter-domain protocols: The evaluation of new inter-domain routing protocols relies strongly on the accuracy of the AS-level underlay on top of which the protocol is supposed to run. For instance, [40] proposes a new protocol where Tier-1 ASes run a path-vector protocol between them, while all ASes under each Tier-1 run link-state routing. The assumption is that customer trees of Tier-1 ASes are disjoint, therefore they can be considered to be contained in link-state domains, and violations of this assumption are treated by the protocol as exceptions. However, in view of our findings, there are a substantial number of invisible peer links interconnecting ASes at the edge of the network, and therefore connectivity between different customer trees becomes the rule rather than the exception. We would imagine the performance of the proposed protocol under complete and incomplete topologies to be different, possibly quite significantly.

4.3 Inherent limitations of the "public view"

In the absence of a central authority or agency to provide the connectivity structure of a fully decentralized and distributed large-scale system such as the Internet's AS-level ecosystem, engineers typically rely on creative alternatives or hacks. In the quest for obtaining the Internet's actual AS-level connectivity structure, the hack consists of using BGP-based measurements. By its very nature, BGP – while serving as the de facto standard inter-domain routing protocol deployed in today's Internet – is **not** a mechanism by which ASes distribute their connectivity. Instead, it is a protocol used by ASes to distribute the reachability of their networks via a set of routing paths that have been chosen by other ASes in accordance with their policies. While this BGP-derived reachability information in the form of routing tables and routing updates and as collected by projects such as RouteViews and RIPE-RIS is undoubtedly useful for inferring AS-level connectivity, it also has some inherent limitations.

As illustrated by our case studies and summarized in Section 4.1, a main limitation of inferring topology from BGP data is due to the location of the monitors. On the one hand, we have verified that when aggregated over time (i.e., using routing updates), the inferred connectivity structure at the level of Tier-1 ASes is rather complete, mainly because of the existence of a monitor in almost all the Tier-1 ASes. The percentage of missing links is typically in the low single digits and is largely due to invisible connections (e.g., using internal ASes or long prefixes). On the other hand, our case studies also provide concrete evidence that even when aggregated over time, the inferred AS-level connectivity at lower tiers tends to be very incomplete when there is no monitor in the lower tier's domain or its customer base. This shortcoming is best highlighted by our case study of a major content provider network where we showed that the measurements miss the bulk (i.e., about 90%) of the links, the majority of them being invisible peer links at IXPs.

Since the placement of monitors is largely an administrative rather than a technical issue, the problems caused by the inherently limited coverage of the existing monitors is unlikely to go away. Even if future developments would favor a more optimal (in the sense of minimizing the number of missing links) placement of monitors, the dynamic nature of the Internet's actual AS-level ecosystem may render such an effort useless. Launching traceroute probes from

multiple vantage points [13, 37, 26] does not change the fundamental limitation in gathering actual AS topology data, that is, the inability to uncover all peer links short of installing a probe in all the stub ASes. Assuming the existence of a measurement box M, the AS-level links revealed by traceroute probes launched from M over time is at best a subset of those revealed by using M as a BGP monitor (assuming the discrepancy between routing and data planes is negligible), because traceroute is likely to miss those links that can be revealed during BGP path exploration [33].

4.4 Moving beyond the "public view"

Given the technical and practical difficulties that stand in the way of collecting the data necessary for obtaining a complete and accurate map of the AS-level Internet, what alternative other than relying on AS maps of questionable quality is available? Clearly, generative methods that rely on some underlying mathematical model of the Internet's AS-level topology are by and large not helpful as the assumed models are derived from the very same measurements whose quality is being questioned in the first place [28, 27]. Instead, we argue in this paper for a more pragmatic approach that avoids the model-fitting part all together and results in AS maps annotated with AS relationships that, while being informed by the available data, are not obviously inconsistent when viewed from either an engineering, administrative, or economic perspective. In this sense, we advocate here an approach that results in instantiations of AS topologies that could plausibly represent the Internet's actual AS-level ecosystem.

While the details of this new approach to constructing realistic and viable AS maps will appear in a forthcoming paper, we illustrate in the following the type of heuristics we envision for "filling in" BGP-derived AS maps; that is, deliberately augmenting them with a target percentage of new links of a certain type and/or modifying existing links (i.e., changing AS relationships) so that the resulting connectivity structures represent fully functional and economically viable AS-level ecosystems, consistent with real-world peering practices. Suppose we want to augment a given inferred AS map that is missing a significant number of peer links with, say, 50% more links of the peer type. From our case studies, we know that public peering happens predominately at IXPs and that BGP-derived AS maps miss the majority of IXP-related peer links. In fact, a necessary condition for two ASes to publicly peer in an IXP is that both of them have a Point-of-Presence (PoP) in that IXP, i.e., a router of each ASes is co-located there. Moreover, it is reasonable to assume that the more PoPs two ASes have in common, the more likely it is for them to peer. Using IXP co-location information, these observations can be incorporated into a simple biased urn model for determining whether two ASes peer with one another, based solely on the number of PoPs they have in common. Initial results of using this heuristic to boost the number of peer links are encouraging and show good agreement with the ground truth available for the Tier-2 AS in Section 3.2. However, to be more flexible and apply across a wider spectrum of ASes, a slightly more complicated model may be needed that can account for AS-specific metrics other than the number of PoPs (e.g., see [19]).

5. RELATED WORK

Studies focusing on the Internet AS-level topology have become an important component in Internet research. Three main types of data sets have been available for AS-level topology inference: (1) BGP tables and updates, (2) traceroute measurements, and (3) Internet Routing Registry (IRR) information. BGP tables and updates have been collected by the University of Oregon RouteViews project [12] as well as by RIPE-RIS in Europe [11]. Traceroute-based datasets have been gathered by CAIDA as part of the Skitter project [13], by researchers in the EU-project called Dimes [37], and more recently by the iPlane project [26]. Other efforts have extended the above measurements by including data from the Internet Routing Registry [17, 38, 43]. However the studies that rely critically on the topology measurement data rarely examined the data quality in detail, thus the results' (in)sensitivity to the known or suspected deficiencies in the measurements goes largely unnoticed.

Chang *et al.* [17, 20, 16] were among the first to study the completeness of commonly used BGP-derived topology maps, and later studies [44, 35, 43], using different data sources, yielded similar results confirming that at least 40% or more AS links may exist in the actual Internet but are missed by the commonly-used BGP-derived AS maps. He *et al.* [43] report an additional 300% of peer links in IRR compared to those extracted from widely used BGP views, however this percentage might be very inflated since they only considered RIB snapshots from 35 of the $\sim$700 routers providing BGP feeds to RouteViews and RIPE-RIS. The problems associated with inferring AS-level connectivity from traceroute measurements have been detailed in [18, 30, 23], and the inaccuracy has been recently quantified to some extent in [34]. All these efforts have in common that they try to *incrementally* close the completeness gap, without first quantifying the degree of (in)completeness of currently inferred AS maps. Our paper relies on the ground truth of AS-level connectivity of different types of actual ASes to shed light on *how much* and *what* is missing from the commonly-used AS maps and *why*.

Within the context of the existing vast body of literature on AS topology modeling, our proposed approach is a departure from more traditional efforts that attempt to describe the Internet's AS-level topology with the help of generic graph-theoretic constructions (see for example [28, 27]). Instead, similar to [25], we argue in this paper that the Internet AS-level ecosystem is a large-scale decentralized virtual infrastructure, consisting of a diverse set of networks or businesses that interact with one another via well-defined business relationships or contracts. As a result, we forego in this paper the notion of AS topology modeling in the traditional sense and instead propose a first-principles type approach that is informed by the available measurements but relies on a set of economically-motivated, policy-driven, or technology-based heuristics for turning an incomplete and inaccurate inferred AS map into a realistic and viable AS-level topology. The development of these heuristics and their validation is part of our future work and will be described in detail in a forthcoming paper.

6. CONCLUSION

Assessing the quality of inferred AS-level Internet topology maps is an important and difficult problem. There have been generally accepted notions that the public view is good at capturing customer-provider links but may miss peering links. However, there has been no systematic effort to provide hard evidence to either confirm or dismiss these notions. This paper represents a first step towards addressing this challenging problem. Recognizing that it is impractical to obtain a complete AS topology through currently pursued data collection efforts, we approach the problem from a new and different angle: obtaining the ground truth of sample ASes' connectivity structures and comparing them with the AS connectivity inferred from publicly available data sets. A key benefit we derive from this new way of tackling the problem is that we gain a basic understanding of not only what parts of the actual topology may be missing from the inferred ones, but also how severe the incompleteness problem may be.

A critical aspect of our search for the elusive ground truth of AS-

level Internet connectivity and of the proposed pragmatic approach to constructing realistic and viable AS maps is that they both treat ASes not as generic nodes but as objects with a rich, important, and diverse internal structure. Exploiting this structure is at the heart of our work. The nature of this AS-internal structure permeates our definition of "ground truth" of AS-level connectivity, our analysis of the available data sets in search of this ground truth, our detailed understanding of the reasons behind and importance of the deficiencies of commonly-used AS-level Internet topologies, and our proposed efforts to construct realistic and viable maps of the Internet's AS-level ecosystem. Faithfully accounting for this internal structure can also be expected to favor the constructions of AS maps that withstand scrutiny by domain experts. Such constructions also stand a better chance to represent fully functional and economically viable AS-level topologies than models where the interconnections between different ASes are solely determined by independent coin tosses. Validating the consistency of an approach to understanding the AS-level Internet that utilizes the network-intrinsic meaning of what a node and a link represents clearly requires extra efforts and creativity and will therefore feature prominently in our future research efforts in this area as discussed in Section 4.4.

Acknowledgements

We would like to thank Tom Scholl, Bill Woodcock, Ren Provo, Jay Borkenhagen, Jennifer Yates, Seungjoon Lee, Alex Gerber and Aman Shaikh for many helpful discussions. We would also like to thank a number of network operators who helped us gain insight into the Internet topological connectivity.

7. REFERENCES

[1] AOL peering requirements. http://www.atdn.net/settlement_free_int.shtml.
[2] AT&T peering requirements. http://www.corp.att.com/peering/.
[3] CERNET BGP feeds. http://bgpview.6test.edu.cn/bgp-view/.
[4] European Internet exchange association. http://www.euro-ix.net.
[5] Geant2 looking glass. http://stats.geant2.net/lg/.
[6] Good practices in Internet exchange points. http://www.pch.net/resources/papers/ix-documentation-bcp/ix-documentation-bcp-v14en.pdf.
[7] Internet Routing Registry. http://www.irr.net/.
[8] Packet clearing house IXP directory. http://www.pch.net/ixpdir/Main.pl.
[9] PeeringDB website. http://www.peeringdb.com/.
[10] Personal Communication with Bill Woodcock@PCH.
[11] RIPE routing information service project. http://www.ripe.net/.
[12] RouteViews routing table archive. http://www.routeviews.org/.
[13] Skitter AS adjacency list. http://www.caida.org/tools/measurement/skitter/as_adjacencies.xml.
[14] The Abilene Observatory Data Collections. http://abilene.internet2.edu/observatory/data-collections.html.
[15] H. Ballani, P. Francis, and X. Zhang. A study of prefix hijacking and interception in the Internet. In *Proc. of ACM SIGCOMM*, 2007.
[16] H. Chang. *An Economic-Based Empirical Approach to Modeling the Internet Inter-Domain Topology and Traffic Matrix*. PhD thesis, University of Michigan, 2006.
[17] H. Chang, R. Govindan, S. Jamin, S. J. Shenker, and W. Willinger. Towards capturing representative AS-level Internet topologies. *Elsevier Computer Networks Journal*, 44(6):737–755, 2004.
[18] H. Chang, S. Jamin, and W. Willinger. Inferring AS-level Internet topology from router-level path traces. In *SPIE ITCom*, 2001.
[19] H. Chang, S. Jamin, and W. Willinger. To peer or not to peer: modeling the evolution of the Internet's AS-level topology. In *Proc. of IEEE INFOCOM*, 2006.
[20] H. Chang and W. Willinger. Difficulties measuring the Internet's AS-level ecosystem. In *Annual Conference on Information Sciences and Systems (CISS'06)*, pages 1479–1483, 2006.
[21] D. Dolev, S. Jamin, O. Mokryn, and Y. Shavitt. Internet resiliency to attacks and failures under bgp policy routing. *Computer Networks*, 50(16):3183–3196, 2006.
[22] L. Gao. On inferring autonomous system relationships in the Internet. *IEEE/ACM Transactions on Networking*, 9(6):733–745, 2001.
[23] Y. Hyun, A. Broido, and kc claffy. On third-party addresses in traceroute paths. In *Proc. of Passive and Active Measurement Workshop (PAM)*, 2003.
[24] M. Lad, R. Oliveira, B. Zhang, and L. Zhang. Understanding the resiliency of Internet topology against false origin attacks. In *Proc. of IEEE DSN*, 2007.
[25] L. Li, D. Alderson, W. Willinger, and J. Doyle. A first-principles approach to understanding the Internet's router-level topology. In *Proc. of ACM SIGCOMM*, 2004.
[26] H. Madhyastha, T. Isdal, M. Piatek, C. Dixon, T. Anderson, A. Krishnamurthy, and A. Venkataramani. iPlane: an information plane for distributed services. In *Proc. of OSDI*, 2006.
[27] P. Mahadevan, C. Hubble, D. Krioukov, B. Huffaker, and A. Vahdat. Orbis: rescaling degree correlations to generate annotated internet topologies. *Proc. of ACM SIGCOMM*, 2007.
[28] P. Mahadevan, D. Krioukov, K. Fall, and A. Vahdat. Systematic topology analysis and generation using degree correlations. In *Proc. of ACM SIGCOMM*, 2006.
[29] Z. M. Mao, L. Qiu, J. Wang, and Y. Zhang. On AS-level path inference. In *Proc. SIGMETRICS*, 2005.
[30] Z. M. Mao, J. Rexford, J. Wang, and R. H. Katz. Towards an accurate AS-level traceroute tool. In *Proc. of ACM SIGCOMM*, 2003.
[31] W. Mühlbauer, A. Feldmann, O. Maennel, M. Roughan, and S. Uhlig. Building an AS-topology model that captures route diversity. In *Proc. of ACM SIGCOMM*, 2006.
[32] W. Mühlbauer, S. Uhlig, B. Fu, M. Meulle, and O. Maennel. In search for an appropriate granularity to model routing policies. In *Proc. of ACM SIGCOMM*, 2007.
[33] R. Oliveira, B. Zhang, D. Pei, R. Izhak-Ratzin, and L. Zhang. Quantifying Path Exploration in the Internet. In *ACM Internet Measurement Conference (IMC)*, October 2006.
[34] R. Oliveira, B. Zhang, and L. Zhang. Observing the evolution of Internet AS topology. In *ACM SIGCOMM*, 2007.
[35] D. Raz and R. Cohen. The Internet dark matter: on the missing links in the AS connectivity map. In *Proc. of IEEE INFOCOM*, 2006.
[36] Y. Rekhter, T. Li, and S. Hares. Border Gateway Protocol 4. RFC 4271, Internet Engineering Task Force, January 2006.
[37] Y. Shavitt and E. Shir. DIMES: Let the Internet measure itself. *ACM SIGCOMM Computer Comm. Review (CCR)*, 2005.
[38] G. Siganos and M. Faloutsos. Analyzing BGP policies: Methodology and tool. In *Proc. of IEEE INFOCOM*, 2004.
[39] L. Subramanian, S. Agarwal, J. Rexford, and R. Katz. Characterizing the internet hierarchy from multiple vantage points. In *INFOCOM*, 2002.
[40] L. Subramanian, M. Caesar, C. T. Ee, M. Handley, M. Mao, S. Shenker, and I. Stoica. HLP: a next generation inter-domain routing protocol. In *Proc. ACM SIGCOMM*, 2005.
[41] L. Wang, M. Saranu, J. M. Gottlieb, and D. Pei. Understanding BGP session failures in a large ISP. In *Proc. of IEEE INFOCOM*, 2007.
[42] J. Wu, Y. Zhang, Z. Mao, and K. Shin. Internet routing resilience to failures: analysis and implications. In *Proc. of ACM CoNext*, 2007.
[43] Y. He, G. Siganos, M. Faloutsos, S. V. Krishnamurthy. A systematic framework for unearthing the missing links: measurements and impact. In *Proc. of NSDI*, 2007.
[44] B. Zhang, R. Liu, D. Massey, and L. Zhang. Collecting the Internet AS-level topology. *ACM SIGCOMM Computer Comm. Review (CCR)*, 35(1):53–61, 2005.
[45] Y. Zhang, Z. Zhang, Z. M. Mao, C. Hu, and B. M. Maggs. On the impact of route monitor selection. In *ACM/USENIX IMC*, 2007.
[46] C. Zheng, L. Ji, D. Pei, J. Wang, and P. Francis. A light-weight distributed scheme for detecting IP prefix hijacks in real-time. In *Proc. of ACM SIGCOMM*, 2007.