A Survey of Techniques for Internet Topology Discovery

Reza Motamedi, Reza Rejaie
University of Oregon
{motamedi,reza}@cs.uoregon.edu

Walter Willinger
Niksun Inc.
wwillinger@niksun.com

ABSTRACT

Capturing an accurate view of the Internet topology is of great interest to the networking research community as it has many uses ranging from the design and evaluation of new protocols and services to the vulnerability analysis of the network's infrastructure. However, the scale of today's Internet coupled with its distributed and heterogeneous nature makes it very challenging to acquire a complete and accurate snapshot of the topology.

The purpose of this survey is to examine the main research studies that have been conducted on topics related to Internet topology discovery in the last 15-20 years and present some of the main lessons learned from these past efforts. To this end, we classify these prior studies according to the "resolution" or "level" of the topology; that is, interface-level, router- level, PoP-level and AS-level. For each resolution, we describe the main techniques and tools used for data collection, identify their major limitations and issues, and discuss the key implications that these limitations have on the quality of the collected data. In the process, we present the latest efforts in modeling the Internet's topology at the different levels and report on the role that geographic characteristics play in this context. We present the lessons learned as a checklist that every researcher working on Internet topology discovery-related problems should consult to minimize the risk of repeating some of the same or similar mistakes that have been made in the past and as a result have hampered progress in this important area of Internet research.

1. INTRODUCTION

Composed of approximately 50,000 networks or *Autonomous Systems (AS)*, the Internet reigns as the ultimate network of networks. Most of these networks are separately owned and managed (however companies that own and manage multiple ASes do exist [1]), cover different geographic areas, build their own physical infrastructures, and serve different purposes. For example, an AS can be a Network Service Provider (NSP), an Internet Service Provider (ISP), an education network, a content provider (CP), a Content Distribution Network (CDN) or can provide any combination of these or other services. The diversity in network type and business along with their autonomous management makes it clear why individual ASes use network equipment from different vendors to build and operate different infrastructures, possibly with greatly varying physical topologies and why, in turn, not all ASes deploy the same intra-domain routing protocol but instead use the one(s) that best support(s) their operational needs. It also explains why one of the critical features of the Border Gateway Protocol (`BGP`) is its expressiveness – the ability to let ASes with potentially competing business interests express different policies for interconnecting with one another, presumably for the purpose of enabling the smooth and economically viable exchange of traffic.

As a result of this diversity, its scale, and its distributed and heterogeneous nature, mapping the Internet's global topology is inherently difficult and enormously challenging. For one, since the decommissioning of the NSFNET [2], there exists no entity or organization that has a complete picture (i.e., "ground truth") of the entire Internet or its individual constituents or ASes. Moreover, there exists no protocol or service whose sole purpose is the discovery of the network topology[3, 4]. In fact, the measurement tools that are most often used for topology discovery are merely "engineering hacks" that researchers have proposed to collect information about the Internet topology. In particular, the two most commonly used techniques for topology discovery, namely `traceroute` and BGP, have originally been designed for entirely different purposes – `traceroute` as originally introduced by V. Jacobson is a network debugging tool [5, 6, 7] and `BGP` is the de-facto standard inter-domain routing protocols in today's global Internet that indicates reachability rather than connectivity of individual ASes [8].

These difficulties and challenges notwithstanding, the study of the Internet's topology has fascinated both networking and non-networking researchers for the last 15-20 years. While non-networking researchers view the Internet or its topology as a prime example of a complex and large-scale technological network and are mainly interested in studying its structural properties and predicting its behavior, the network research community's interest is in general motivated by more practical concerns. For example, various topological properties of the Internet affect the performance of network protocols, network applications and services. Thus, a better understanding of the Internet topology and its main char-

acteristics would enable network researchers to design better network protocols or services and evaluate them under more realistic conditions. Moreover, an accurate map of the Internet would be very helpful for network engineers and operators who are constantly trying to improve or optimize the allocation of network resources such as proxies, replica servers, and data centers. Similarly, having a detailed and complete map of the Internet's topology, preferably annotated with attributes such as the exact geographic location of certain network equipment, could inform the study of a wide range of security-related problems and protocols such as backtracking malicious traffic or assessing the vulnerability of the Internet to blackouts or attacks on parts of its physical infrastructure.

The purpose of this survey is twofold. First, by viewing the Internet's topology at different well-defined resolutions or levels of detail (i.e., interface-level, router-level, Point-of-Presence or PoP-level, and AS-level), we present a systematic assessment of the main studies that have dealt with measuring and/or modeling the Internet's topology and have been published in the last 15-20 years. Second, by describing in detail the data collection techniques and tools, types of collected datasets, and inference methods used in these different studies, we provide a checklist that every researcher interested in working on Internet topology-related problems should consult before using these tools, datasets, or methods in their own work. Importantly, this checklist collects in one place our current understanding of the main limitations that these tools, datasets or methods have when used in the context of Internet topology research. In a nutshell, by being aware of these limitations and understanding their root causes, researchers will be able to answer for themselves whether or not the used tools, datasets, or methods are of sufficient quality to successfully tackle the particular research problem they are interested in. In this sense, this survey reports on lessons learned from 15-20 years of Internet topology research that will hopefully prevent researchers from repeating some of the same or similar mistakes that have been made in the past and that have negatively impacted progress in this important area of networking research.

The rest of this survey is organized as follows: Section 2 presents the notion of different Internet topology resolutions which defines our taxonomy. Sections 3, 4, 5, and 6 cover Internet topology at interface-level, router-level, PoP-level, and AS-level, respectively. In each section, we discuss the main data collection techniques and tools, types of collected datasets, and inference methods that have been used to study the topology at the corresponding resolution. We conclude our survey in Section 8 with a discussion of the main lessons learned and mention some of the exciting open research problems that will require new and creative solution methods.

2. TAXONOMY

The Internet's topology is often presented as a graph. However, different communities use the term "Internet graph" to refer to different structures. The latter range from the graph structure of the World Wide Web (WWW) and other overlay networks such as P2P systems or Online Social networks to the Internet's physical infrastructure and the more logical or virtual constructs that are enabled by the layered architecture of the network. The focus of this survey is the Internet's physical topology, where nodes represent meaningful network entities and links represent relations between those entities. However, even with this definition in place, a physical topology of the Internet can still mean different things to different interested parties.

Internet Graphs at Different Resolutions: To further disambiguate the meaning behind the notion of a physical Internet topology, we rely on the following taxonomy that considers the *resolutions* of the Internet topologies that have been studied in the past [3, 9, 10, 11]. In particular, we view Internet topology at four different granularity or resolution levels, organized from finest to coarsest as follows:

I) Interface level: At this level, a node represents a network interface with a designated IP address. An interface belongs to a host or a router and there is a one-to-one mapping between nodes and IPs [12, 13]. At the same time, a link between two nodes shows a direct network layer connectivity between them. This implies that the topology at this level ignores devices functioning at OSI layers lower than the network or IP layer (*e.g.*, hubs and switches).

II) Router level: The topology at this level is often the result of grouping interfaces that belong to the same router [14]. At this level, a node represents an IP-compliant network device (*e.g.*, a host or a router with multiple interfaces). Two nodes are connected by an edge if the corresponding devices have interfaces that are on the same IP broadcast domain.

III) PoP level: A PoP (Point of Presence) is a concentration of routers that belong to the same AS [15, 16]. ASes commonly build their physical networks in a more or less pronounced hierarchical manner; that is, an AS's PoPs are interconnected to form the AS's "backbone" and are also the locations where the AS connects to the PoPs of other ASes [17] and where it provides access to its customers or end users. In this sense, a node in the topology at this level represents a PoP that belongs to an AS. A link between two PoPs indicates that there is physical connectivity between the routers of the two PoPs.

IV) AS level: As opposed to the previous constructs, the AS-level topology represents a more logical view of the Internet [18, 19]. A node at this level represents an AS identified by a 16-bit (recently also a 32-bit) AS number. A link in the AS-level topology represents a business relationship between two ASes. Theses relationships reflect who pays whom when traffic is exchanged between the ASes in question and are key to a properly functioning and financially viable Internet ecosystem [3, 20]. Traditionally, these relationships are categorized into a) customer-provider (C2P), b)

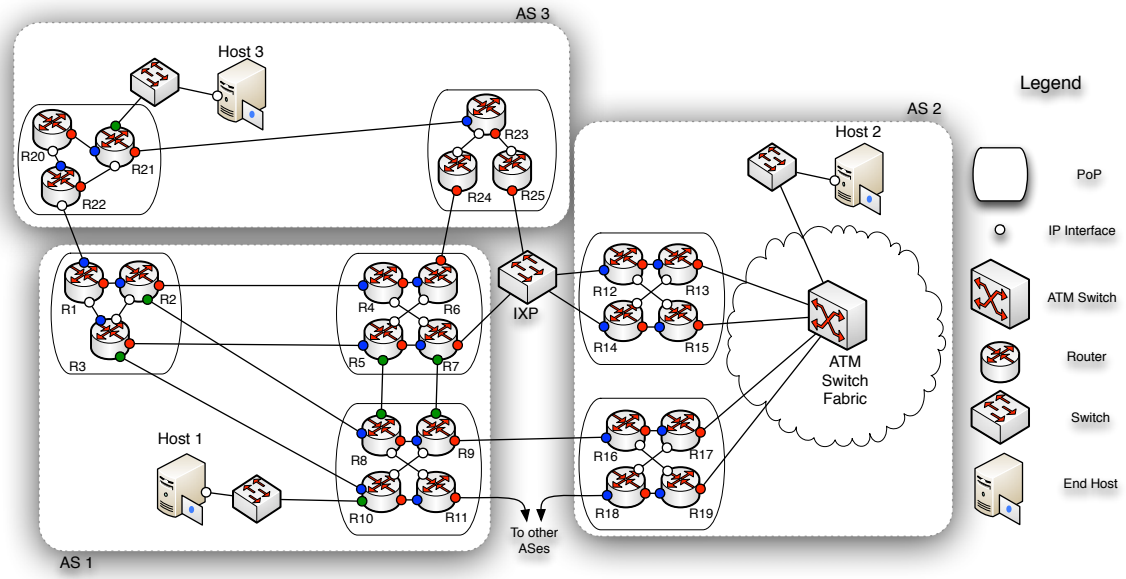


Figure 1: A detailed toy topology representing the Internet topology at different granularities

peer-peer (P2P), and c) sibling relationships, but other forms of relationships are known to exist as well. Since ASes typically cover entire geographic regions, with different PoPs in those regions' major cities, the physical connectivity between two ASes that have an established business relationship often occurs at multiple locations. Thus, an AS link is virtual rather than physical in nature in the sense that it is an abstraction and typically represents multiple physical connections between the two ASes [21].

In this survey, we make use of this taxonomy to categorize prior Internet topology studies. Moreover, for the different studies concerned with one and the same resolution, we (i) provide a detailed assessment of the limitations of the techniques employed to collect data, as well as an assessment of the quality of the collected data that is used to study the topology at the given resolution, and (ii) carefully examine the geographic characteristics of the inferred topology and the extent to which the topology at the given resolution is annotated with geographic attributes.

Figure 1 shows three resolutions of the topology. At the finest level, the router-topology is presented. The PoP-level topology is generated when PoPs and the connections between them are considered. Finally, the AS-level topology is obtained when we look only at the ASes and the links between them.

Data Types and Data Collection: The nature of the data and the type of data collection techniques are two other elements that we use to classify prior Internet topology studies. Regarding the nature of the collected data, measurements can be performed in the *control plane* or the *data plane*. In terms of measurements performed in the control plane, the collected data reveals information about routing in the Internet. For instance, `BGP` tables store the AS paths to reach different prefixes and they are classic examples of control plane data. In contrast, data plane measurements aim to discover the actual paths that packets travel along. The simplest measurement of this form is `Ping`. It measures the reachability of a target IP address and also reports the Round Trip Time (RTT) between the target IP and a source, based on the route that the probe packets take in the Internet. Regarding the collection technique, a measurement can be either *active* or *passive*. In active measurements, actual packets ($i.e.$, probe messages) are sent into the network and the replies are collected. On the other hand, passive measurements only tap into a wire and collect the information that is already flowing over that wire. `traceroute` and `BGP` monitors are examples of active and passive measurement techniques, respectively. A list of commonly-employed data sources and measurement techniques used for studying the Internet topology at each resolution is provided in Table 1.

Geographic Attributes of The Topology: Although a main element of a topology is connectivity, *geography* is another element that, when appropriate, can be added to the topology to increase its usability. However, the definition of a geographically annotated topology depends on the different resolutions of the Internet topology. Interfaces, routers and PoPs are entities that can in theory be geographically mapped to an exact location on a map. A geographical Internet map at these three levels of resolution involves assigning a pair of longitude and latitude coordinates to each entity. Therefore, the topology graph consists of points on the map and the links that connect those points together.

Table 1: Different resolutions of Internet topology and the commonly used data sources to capture the topology in addition to the corresponding limitations and challenges

Resolution	Tools & techniques	Limitations & challenges
Interface-level	`traceroute`	Router response inconsistency
		Opaque Layer 2 clouds
		Load balance routers
		Probe message filtering
Router-level	Subnet discovery	Router response inconsistency
		Probe message filtering
	Alias Resolution	Scalability
		Inaccurate (false positive and false negative)
	`SNMP`	Only applicable to one AS
		Requires administrative authorities over the AS
	`MRINFO`	Only applicable to ASes with DVMRP multicast-ready routers
	Aggregation techniques	Mapping IP to Geo is inaccurate
		DNS name to Geo is not always applicable
		DNS misnaming can add more error
	Delay based techniques	Sensitive to knowledge of geography and placement of candidate PoPs
	Online data sources	Public online data is not always up-to-date
AS level	`BGP`	Reachability announcement protocol with built in information hiding
	`traceroute`	Mapping IP to AS number is not trivial
		Using private IPs and other interface level inconsistencies add more complexities
	Internet Routing Registries	Obsolete data

However, in the case of the AS-level topology, geography is a more subtle notion and typically refers to the geographic region covered by an AS. In such a view, an AS is shown as a colored area on a map that represents its coverage, and different ASes covering parts of that same region are stacked vertically and are shown in different colors. In such a representation, AS relationships can be represented by connections between the differently-colored regions and can be further refined by incorporating the ASes' PoPs and showing the inter-AS connections at the PoP-level. This picture is further complicated by the existence of Internet eXchange Points (IXP), where multiple networks connect at one (or a few close-by) physical locations through a multipoint connection. As a result, the complete geo-annotated AS-level topology should be viewed as a hyper-graph [3] where nodes cover areas and links are annotated by the locations of the their corresponding PoPs.

Figure 2 depicts a sample AS-level topology graph, where each AS covers a certain area. Inter AS connections through vertical lines suggest AS level connectivity. As the figure shows, two ASes might be connected at different locations.

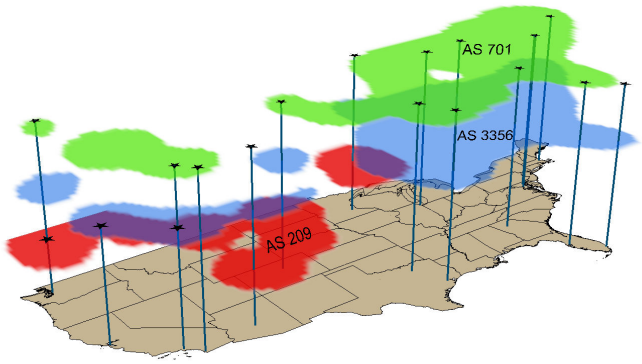


Figure 2: Sample view of geo-footprint for multiple ASes. The vertical lines indicate the city where PoP for and is located.

3. INTERFACE-LEVEL

The interface-level abstraction of the Internet topology portrays the network layer connectivity of its IP interfaces. IP interfaces of routers and end-hosts are represented as nodes. Having in general multiple interfaces, each router appears as multiple nodes, while normal end-hosts with one interface are depicted as a single node. The topology is typically simplified by ignoring end-hosts, therefore nodes only represent router interfaces. Links represent direct network layer connectivity between nodes. However, not all these links are point-to-point. For instance, layer 1 and layer 2 clouds can be traversed, although the connectivity is represented as a single link.

`traceroute` is the most widely used tool to map the topology of the Internet at this resolution. Based on the nature of the technique and the type of data it produces, it is an active measurement method performed in the data plane [8, 22]. It uses limited Time-To-Live (TTL) probes. The `traceroute` probes launched from a source to a target successively discover the IP addresses of IP-compliant router interfaces along the forward path, and at each hop measured RRT values are also reported. Multiple probe messages with the same TTL

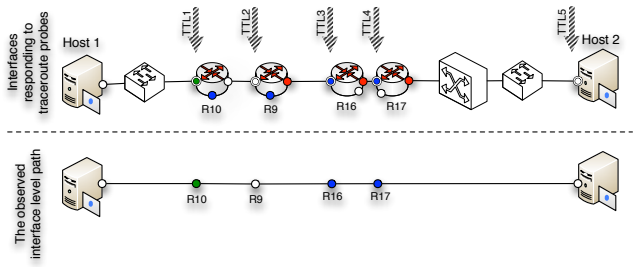


Figure 3: `Traceroute` **from** $Host1$ **towards** $Host2$ **and the corresponding interface-level path.**

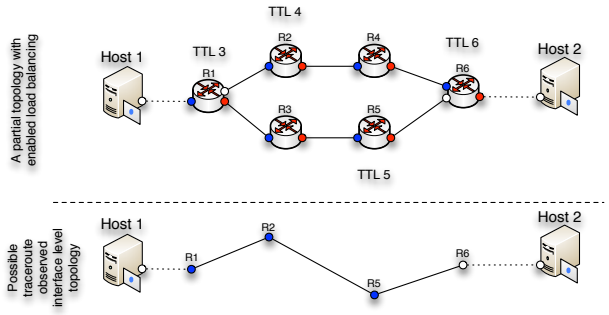


Figure 4: **False links inferred by** `traceroute` **in the presence of load balanced routers**

can be used to discover the IP at the same hop. In the perfect scenario, probes for the same hop would initiate a response from the same IP, but each would produce a slightly different RTT due to the inherently dynamic nature of network traffic. In the rest of this survey we assume that a single probe message is used for each hop discovery. Figure 3 shows the conducted `traceroute` from *Host1* to *Host2* and the observed interface-level path. Only one IP address per hop is identified, and the result does not indicate any layer 2 infrastructures.

Each individual `traceroute` measurement reveals one IP path composed of multiple IP segments. To discover the topology at the interface-level, the outcome of many `traceroute` measurements should be merged. `traceroute`-based techniques require a number of `traceroute` capable hosts (vantage points), and a list of target IPs. During a measurement campaign, a set of vantage points launch `traceroute` probes towards a given set of targets. The overall observed interface-level topology is generated from the union of all the IP paths, each measured by a `traceroute`.

In the following, we first describe `traceroute` in more detail, mainly because it is the most commonly-used active measurement tool, and then discuss its limitations. We also provide an overview of some of the main measurement-based studies that use active measurements to infer the interface-level for Internet topology and discuss a number of more recent proposals for collecting interface-level data.

3.1 `traceroute`

3.1.1 Basic Technique & Variants

`traceroute` involves actively sending probes into the network, rather than merely monitoring it. It is the most widely used active measurement tool to obtain a map of the physical Internet. V. Jacobson's `traceroute` – the first implementation of this tool – uses ICMP packets as probes [23]. However, other versions of `traceroute` exist that use other types of probe messages, for instance UDP and TCP packets [5].

UDP `traceroute` reveals the IP hops from a source to a destination by sending packets with limited TTLs and large (destination) port numbers. When an intermediate router receives such a probe with TTL equal to zero, it responds back to the source with an "ICMP time exceeded" message. The source progressively increases the TTL until the probe packet reaches the target. In this fashion, this technique identifies with each TTL one segment of the IP route in addition to its corresponding RTT. An "ICMP port unreachable message" indicates that the message was successfully received by the target, and using large destination port numbers minimizes the chance of randomly probing an open port on the target. In addition, the port number is used to match the probes and responses. Unix-like operating systems use this UDP `traceroute` by default, with port numbers between 33435 and 33534. The port number is incremented after each probe, thus enabling the source to identify the hop distance of the received responses.

ICMP `traceroute` also uses limited TTL but sends "ICMP echo requests". Since ICMP messages do not have port numbers, the matching of the probes and responses is done using an ICMP id/sequence that is part of the ICMP packet header. ICMP `traceroute` is the default setting for Microsoft Windows.

The main limitation of UDP and ICMP `traceroute` is that both UDP messages to high ports and ICMP messages are prone to be filtered by firewalls [24]. To bypass firewalls, TCP `traceroute` has been designed and uses TCP-SYN probes' well-known ports (*e.g.*, port 80). However, some firewalls are configured to filter TCP packets when no host behind the firewall accepts the TCP connection at the well-known port. This is especially the case at the edge of the network.

A comparison of using the UDP, ICMP and TCP `traceroute` techniques for topology discovery shows that ICMP `traceroute` reaches targets more successfully. Moreover, while UDP `traceroute` identifies more IP links, it is the least successful technique in terms of reaching the targeted IP [25].

The Internet is designed to route packets based on the destination IP. However, network administrators often employ load balancing techniques at certain routers to increase the utilization of their resources. They achieve this goal using "equal cost path" in their implementation of the inter-domain

routing protocols OSPF [26] or IS-IS [27]. Per packet and per flow load balancing are the two types of load balancing techniques that network administrators typically use. In per packet load balancing, each packet is individually load balanced, while in the per flow case, packets from the same flow are routed along the same path. Routers use IP headers to identify flows, and these headers include fields such as Source IP Address, Destination IP Address, Protocol, Source Port Number, Destination Port Number, the IP Type of Service (TOS), the ICMP Code and the Checksum. Note that in the case of the traditional `traceroute` technique, the values in some of these fields vary for different probes so as to be able to match the probes and the responses.

As a result, per flow load balancing may result in the routing of probes of the same `traceroute` measurement through different paths. Put differently, when measuring a load balanced route, some `traceroute` techniques will infer the existence of an IP segment that does not exist in the actual topology. Figure 4 shows a possible `traceroute` when it travels through a load balanced path. *R1* is a load balanced router. Probe messages can either visit *R2* or *R3* based on the load balancing decision made at *R1*. In our example, for TTL 3, 4, and 5, the visited routers are *R1*, *R2*, and *R5*, respectively. As a result, a false link between *R2* and *R5* is inferred. Paris `traceroute` [7] has been designed to address this issue by using probes that are routed similarly when per flow load balancing is in use. By manipulating the ICMP headers in the probes, Paris `traceroute` ensures that all the packets of a `traceroute` measurement take the same path. Paris `traceroute` resolves the flow based load balancing anomalies in the observed route, but anomalies due to per packet load balancing are not resolved.

3.1.2 Limitations & Issues

`traceroute` is used predominantly in network troubleshooting. In fact, it has been designed as a generic reachability diagnosis tool, and using it for discovering the interface-level topology of the Internet is both an after-thought and a less-than-perfect heuristic [28]. In general, the limitations and issues concerning the use of `traceroute` for interface measurements have to do with the nature of the measurement method itself and with the inherent difficulties of using large-scale distributed measurement platforms for performing Internet-wide `traceroute` campaigns. In the following, we summarize the most important limitations and issues when using `traceroute` for the purpose of discovering the interface-level topology of the Internet.

Measurement Limitations: First and foremost, there exists no unique setting for a router's response to a TTL zero probe. The router configuration determines the response, and network operators are in charge of, among other tasks, configuring routers. With respect to responding to TTL zero probes, network operators typically choose one of the following five policies: (i) *Null interface routers* means to remain reticent to the probes. For these routers, `traceroute` detects their existence, but not their interface address ("anonymous routers") [29]. In this case, the RTT is also not reported. (ii) *Probed interface routers* means to respond with the IP address of the probed interface. This configuration is most common when the router is directly probed. (iii) *Incoming interface routers* says to respond with the IP address of the interface from which the probe message was received. This configuration is reported to be the most common setting when the router is probed with indirect TTL-limited messages [30]. (iv) *Shortest-path interface routers* says to respond with the IP address of the interface that is closest to the source. Note that because of asymmetric routing in today's Internet, the incoming interface and shortest-path interface are not necessarily the same. (v) *Default interface routers* says to respond always with a designated fixed interface IP address ($i.e.$, irrespective of the probed interface). (vi) *Default IP routers* says to respond with a randomly-selected IP address. This IP can be configured indifferent to the IPs assigned to any of the router's interfaces. In addition to these router configuration settings, firewalls can also be configured to prevent probed routers from responding. In short, a `traceroute` probe packet suggests the existence of one interface per router in the foreword path, at best.

A second limitation of `traceroute` is that the IP address it records at each hop is not necessarily a valid IP address. This can occur due to (mal)practices in assigning IP addresses to router interfaces. (Mis)configured IP addresses sometimes suggest the appearance of private non-routable addresses and carrier-grade NAT (large scale NAT) addresses. Such addresses can lead to routing loops or other anomalies because they can be used by multiple ASes. In addition, these IPs cannot be mapped to a single router or an AS and cannot be used to map the location of the interface because of the one-to-many relationship between the IP and the assigned interfaces.

Third, the RTT value reported at each hop cannot be used to accurately measure the delay to and from the target. `traceroute` is a foreword route diagnostic tool, and a rule of thumb in Internet routing is that routes between two IPs are not always symmetric. Hence, the path taken by a `traceroute` probe may differ from the path taken by its response. In fact, variations in the delay between two consecutive hops could be due to congestion on the link, variable delays in the router's queues, or asymmetric routing.

A fourth limitation that has become increasingly more relevant in today's Internet is that layer 2 clouds are generally opaque to a `traceroute`. These clouds have the explicit purpose of hiding the network infrastructure from the IP layer. For example, ATM (Asynchronous Transfer Mode) clouds are completely hidden from `traceroute`. From the perspective of `traceroute`, an AS using ATM switches provides direct connectivity between its IP routers, although in reality the IP interfaces are interconnected via a collection of ATM switches. For instance, in the observed topol-

ogy of AS2 in Figure 1, routers directly connected to the ATM cloud have a mesh-like interconnectivity. A more popular layer 2 technology is MultiProtocol Label Switching (MPLS), and it is commonly used to configure tunnels passing through multiple routers. It has been reported that at least 30% of the paths tested in a recent study traverse an MPLS tunnel [31, 32]. Routers using MPLS can be configured to either decrement the TTL (MPLS opaque option), as `traceroute` requires, or ignore the TTL field completely. If MPLS routers are configured to respond back to ICMP `traceroute` messages, extra tags (*e.g.*, `MPLS Label=1048 Exp=7 TTL=1 S=0`) appear in the resulting `traceroute` measurement and reveal the existence of an MPLS tunnel. Although it is be possible to detect MPLS tunnels from `traceroute` measurements [31, 32], the inference methods are known to be imperfect and are very specific to MPLS tunnels.

Large-Scale Measurements Issues: Clearly, the choice of vantage points and targets impacts the observable interface-level topology. For example, the probability of sampling an IP segment is directly related to the placement of the vantage points and the type of IP segment. In particular, back-up inter-AS routes are hard to discover, and IP segments representing inter-AS peer-to-peer relationships are among the least discoverable ones [8]. To deal with these and similar issues, two approaches have been proposed. First, Eriksson *et al.* [33, 34] suggested a statistical approach to infer the unseen components of the Internet. By proposing to map the problem to a statistical "unseen species problem", they first estimate the number of unseen components using incomplete observations. Next, matrix completion techniques are used to infer the components and the connectivity between the inferred components and the rest of the topology. The inferred topology is then validated by adaptive targeted probing. The second approach relies on targeted probing to discover less visible IP segments. In this case, domain experts use their knowledge of the topology and routing policies to devise targeted mapping experiments. The rationale behind this approach is that doing more measurements does not compensate for the measurement bias [8]. Instead, this bias can be addressed by making informed decisions about the locations of the vantage points and targets in relation to the IP segments in question. For instance, Augustin *et al.* [35] use targeted probing with `traceroute` to discover peering links at Internet Exchange Points (IXPs) that are otherwise hard to detect.

Given a platform with a set of vantage points and targets, orchestrating a large measurement campaign often imposes a high load on the network as a whole and the measurement infrastructure in particular. The measurement load is higher closer to the vantage points and the set of targets as these segments are redundantly sampled. The high probe traffic may be even be viewed and identified as a Denial of Service (DoS) attack by Intrusion Detection Systems (IDS) [36]. The redundant measurements are classified in [37] into two different types. "Intra-monitor redundancy" occurs close to a vantage point. An individual vantage point redundantly measures the IP segments in its vicinity due to the tree-like structure of routers rooted at the vantage point. "Inter-monitor redundancy" occurs close to targets. Similar to the former type of redundancy, the tree-like structure of routers close to a target causes these routers to be redundantly probed by multiple vantage points.

Different methods to reduce the overhead resulting from such redundancies have been proposed in the literature. On the one hand,"far probes" [37] are proposed to address the intra-monitor redundancy. In this case, when the topology close to the vantage point is fully discovered, instead of using `traceroute` with probes starting with TTL 1, a higher TTL value is chosen. On the other hand, "top set" (collaborative probing) [37, 38] aims to address the inter-monitor redundancy. Consider two vantage points running `traceroute` to the target t. The idea is that if the corresponding routes merge at an intermediate router, they will follow the same path toward t due to destination based routing. Therefore, a per target stop list is required to halt the measurement from one vantage point when the rest of the route is already discovered from former measurements conducted by the other vantage point. Beverly *et al.* [22] used high frequency measurement with adaptive probing techniques to limit the imposed measurement load, while keeping the discovery rate high. In each cycle, their "interface set cover" algorithm minimizes the `traceroute` load while maintaining a high discovery rate. To maximize the gain from each `traceroute`, "subnet centric probing" selects targets to reveal the maximum information from the inside of a network.

3.1.3 Large-Scale `traceroute` *Campaigns*

Obtaining the Internet-wide interface-level topology hinges on the idea of performing `traceroute` measurements between many different vantage points and targets, *i.e.*, collecting data from a large-scale `traceroute` campaign. In the following, we discuss in more detail the pre-requisites for performing such campaigns and using the resulting data for inferring the interface-level Internet topology; that is, the availability of appropriate measurement platforms and a solid understanding of the coverage and completeness of the obtained data.

Measurements Platforms: Starting with the original paper published in 1998 by Pansiot and Grad [39], there have been many `traceroute`-based Internet topology studies that have gradually improved our understanding of the Internet's topology. These studies have either used a single vantage point (*e.g.*, Pansiot and Grad [39]), a moderate number of dedicated instrumentation boxes located across the network (*e.g.*, Skitter [40] or its successor Archipelago [41]), or relied on a publicly available general-purpose experimentation platform like PlanetLab (*e.g.*, iPlane [13], RocketFuel [14] and [38]).

The use of public `traceroute` servers, also known as *looking glasses*, to conduct active measurements has also

gained much attention, mainly due to the large coverage in term of the placement of vantage points. However, as publicly available resources that have been deployed with the network operator community in mind, these `traceroute` servers impose limits on the rate at which active measurements can be performed. As a result, they are mainly used for small-scale measurement experiments and validation (*e.g.*, RETRO [42] and [35]). Note that in addition to `traceroute`, many looking glasses also have the capability to issue other network-related debugging commands, especially in support of `BGP`, and their use for collecting `BGP` data will be described in Section 6 below.

Although using dedicated boxes or relying on PlanetLab are still very common approaches to conducting Internet-wide active measurement campaigns, more recent studies have started to deploy platforms that support "crowd-sourcing measurement campaigns"; that is, use of software agents to collect measurements from a large number of vantage points (*e.g.*, Scriptroute [43], Dimes [12], Bitprobe [44]). By asking end users to download a simple measurement plug-in, the idea is to turn massive numbers of unpredictable end-users (in terms of their availability and capabilities) at the edge of the Internet into vantage points and not rely on a small number of dedicated machines in well provisioned networks (*e.g.*, PlanetLab).

These newer platforms use either an altruistic model (*e.g.*, Dimes [12]) whereby individual users are encouraged to participate in the platform and serve as a measurement node for the good of science, or deploy incentive-based models (*e.g.*, Ono [45] and Dasu [46]). Based on recent experience with such incentive-based platforms that aim to ensure that the measurements conducted by the software agents are beneficial for both the users and the experimenter, they are able to attract and retain end users in larger numbers than their altruistic counterparts. As such, they have the potential of growing into Internet measurement platforms that will consist of an unprecedented number of powerful vantage points. However, as already alluded to in Section 3.1.2, performing, for example, crowd-sourcing `traceroute` campaigns on such platforms requires extra care in their design and instrumentation due to concerns over excessive network loads and security issues (for more details, see for example [46]).

Coverage & Completeness: Early studies such as [40] or [41] have suggested the utilization of a few vantage points and a large set of targets that are well distributed across the network. The claim was that the gain from adding vantage points increases only marginally by adding more vantage points [47]. However, later studies reported that despite the diminishing return of extra vantage points, the observed topology is more complete [48]. This discussion of the quantity (*i.e.*, number) vs. the quality (*i.e.*, location) of the vantage points of a measurement platform is in need of yet another revision with the recent discovery of massive amounts of peering links at IXPs, the vast majority of which being completely invisible to past `traceroute` campaigns [49]. This finding confirms earlier observations that only purposefully-placed new vantage points have a chance to detect certain types of IP segments when relying on data plane measurements only [8] and serves as an important reminder that networking researchers have a long way to go before being able to claim to have a complete map of the interface-level Internet topology.

In their quest to produce a more complete picture of the interface-level topology, researchers have not only increased the number of vantage points and targets [45, 46] but also the duration of the period over which the measurements are performed [22]. While the former can increase the scope of the captured topology but depends critically on the placement of the vantage points, the latter can also reveal a more complete view by exploiting the dynamic nature of the topology (*e.g.*, measurement probes launched at certain times may take rarely used back-up routes due to, for example, router failure-induced route changes). However, the drawback of this solution is that it cannot easily distinguish between routes that have been seen in the past but no longer exist. In general, it is not easy to account for such an inherent churn when allowing longer measurement periods, and there are currently only error-prone heuristics in place to deal with the problems caused by this "solution."

3.2 Other Approaches

Although `traceroute` is the most commonly used method for obtaining the interface-level topology, its limitations have expedited the proposal of other approaches to collect additional connectivity information. While `traceroute` with different types of probe messages mainly attempts to penetrate through firewall filters, other active measurement techniques are used to address its other limitations.

3.2.1 IP Options

IP options are fields in the IP packet header that provide additional information for the packet's routing. Packets with enabled IP options are processed according to the type of enabled IP option by intermediate routers. As a result, these packets may be routed differently than other packets, or additional information can be registered in the packets. To obtain a more accurate and complete topology, IP options have been widely employed to enrich the collected data with more information when possible.

The completeness of a captured topology is correlated with the number of vantage points performing the `traceroute` measurements. The cost and the complexity of the deployment of these vantage points may limit the observed view of the interface-level topology. "Source Routing" (SR) offers more flexibility to discover network topology because it allows the sender to specify at most 9 routers that a given packet should go through before reaching the destination. The intermediate routers should also have this option enabled. When used in conjunction with `traceroute`, source routing increases the scope of the discovered topology. This

can be used to direct the probes to a route that is not usually taken by packets. In essence, source routed probes allow the vantage point to observe an additional view of the network. Although the number of SR-capable routers is a small fraction of all routers in the Internet (around 8%), Govindan *et al.* [30] show that this number is enough to capture 90% of the topology in a sparse random graph using simulation. However, this number seems very optimistic for `traceroute` measurements, due to the sensitivity of the observation to the placement of source route enabled routers and the fact that the Internet topology is not random. Augustin *et al.* [35] have also exploited the IPv4 "Loose Source Record Route" (LSRR) option to increase the coverage of their vantage points without increasing their number. Although they found LSRR-capable routers in many different ASes, they report that routers ignore `traceroute` probe messages with LSSR option much more frequently than regular `traceroute` probes. In effect, source routing is used only very infrequently in the context of topology measurement.

The asymmetric nature of Internet routing implies that the discovered routes are only forward routes from the vantage points to the targets. Reverse `traceroute` [50] uses the "Record Route" (RR) option and "IP Timestamp" to detect the interfaces on the reverse routes as well. An RR enabled probe stores the router interfaces it encounters. The IP standard limits the number of stored interfaces to 9. If the distance from the vantage point to the target is shorter than 9 hops, then the probe will return interfaces observed on the reverse path. A probe with IP timestamp option stores up to four ordered IP addresses. The probe queries the router by specifying its IP to record the timestamp if the previously specified IP addresses on the list are already stamped. This method can be used to validate the existence of a sequence of routers with specified IPs on the same route.

While using IP options can provide information that is not available using simple `traceroute`, it increases the chances for processing delay, being discarded, or triggering an alarm at IDSs [51].

3.2.2 *Subnet Discovery*

In the subnet discovery, the idea is to map the subnet view of Internet topology. A subnet is a link layer (layer 2) concept. It is a logical grouping of connected network interfaces that are all in the same broadcast domain. All IPs in a subnet are addressed with a common most-significant bit-group (IP prefix). Studying this topological structure of the Internet map has two advantages. First, it improves our understanding of the interface-level topology. Second, applications that require disjoint route segments can benefit from this view of the Internet. In the subnet graph, each subnet is a node and subnets adjacent to one router are connected via an edge. Figure 5 shows the topological structure of a sample network. Corresponding subnets are depicted as clouds.

Subnet level discovery tools such as `XNET` [52] aim to reveal all ping-able IP addresses on a subnet. `XNET` identifies boundaries associated with the IP prefix of a subnet with a series of tests on IPs that can potentially be in one subnet. The methodology is developed based on the fact that all IP addresses in one subnet share a prefix and have at most one-hop distance difference from a vantage point. The problem is that the size of the subnet is in general unknown. Given IP address t that is n hops away from a vantage point, `XNET` probes IPs in the prefix that includes t starting from the smallest /31 prefix (mate-31). If the probes to all IPs in this prefix travel through the same route and their hop distances to the vantage point are within the boundaries that support their existence in the same subnet as t, then the target prefix is expanded and IPs in this expanded prefix are subjected to the same tests. `XNET` incrementally expands the prefix until at least one IP fails the tests. At this point the last successfully tested prefix identifies the subnet that includes t.

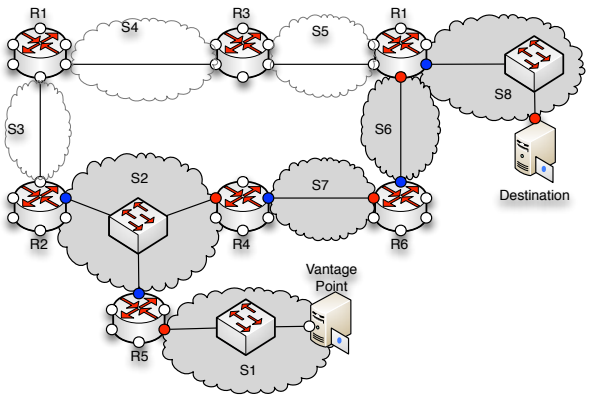


Figure 5: An example topology and corresponding subnets represented by clouds. Subnets identified by `tracenet` are marked grey.

`tracenet` [53] uses the same principles as `XNET` to find subnets along a path. It runs `XNET` on IP addresses discovered by `traceroute` from a vantage point to a destination. Figure 5 shows the application of `tracenet` on a sample topology and identified subnets are depicted in grey. In this figure, Interfaces discovered by `traceroute` are marked as red circles, and blue circles represent interfaces discovered by the `XNET` component of the tool. If `traceroute` returns the incoming interface of each visited router, `tracenet` is able to identify the corresponding subnets along the route from the vantage point to the destination. The principal assumption in `tracenet` is that routers are configured with an incoming interface response setting. However, if a router is configured with another setting, `XNET` discovers an invalid subnet on the path. For instance, in Figure 5, if *R1* responds with its green interface, *S5* is discovered instead of *S6* as the fourth subnet on the route.

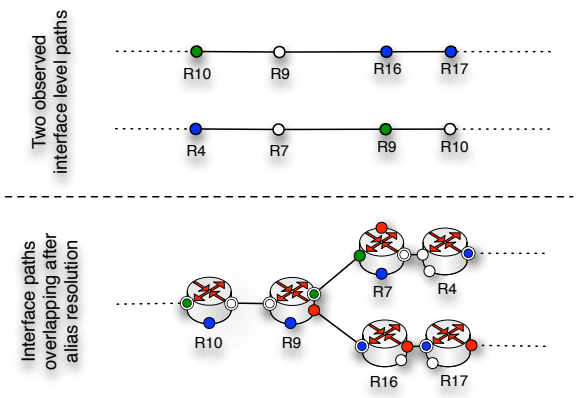


Figure 6: Two partial `traceroutes` with no common hops. Resolving IP aliases shows that the paths overlap.

4. ROUTER-LEVEL

The router-level topology shows the routers and the interconnectivity among their interfaces in the Internet. At this resolution of the Internet topology, nodes represent end-hosts (with one interface) or routers (typically with multiple interfaces) and links show layer 3 connectivity between these devices. The topology at this level can be viewed as the outcome of the aggregation of IP interfaces that belong to a single router. The following two main techniques are considered for collecting the router-level topology:

- **Alias Resolution:** Alias resolution [30, 54] or router disambiguation [47] is a set of techniques used to identify the IP interfaces that belong to the same router. Such disambiguation is necessitated by the aggregation of `traceroute` data that underlies the inference of the router-level topology from the interface-level topology. The main challenge consists of relating different interfaces of a router that were discovered in different `traceroute` measurements.

- **Recursive Router Discovery:** Another class of techniques employed for obtaining the router-level topology relies on a router's capability to be queried for its neighbor on each interface. The Simple Network Measurement Protocol (`SNMP`) and the Internet Group Management Protocol (`IGMP`) are two methods that can be used to discover the neighboring routers of a queried router in an intranet and the Internet, respectively.

4.1 Alias Resolution

Typically, routers have multiple interfaces, each with a different IP address. Two IPs are referred to as aliases if they are assigned to the interfaces of a single router. Alias resolution is the process of grouping IP addresses that belong to the same router. In theory, the true router-level topology can be obtained or derived from the interface level topology, as the result of this process. Figure 6 shows two partial interface paths observed from `traceroute` measurements in the topology of Figure 1, the first from *Host1* to *Host2* and another from *Host3* to *Host1*. The measurements do not have any IP hop in common. However, resolving alias IPs shows that the two measurements visit two different interfaces of *R9* and *R10*. In the context of alias resolution, a false positive detects interfaces belonging to multiple routers as aliases. On the other hand, in the case of a false negative, alias resolution falls short in relating two alias interfaces. We next list and discuss the most widely-used alias resolution methods.

Common Source Address: This technique was proposed and used by Pansiot *et al.* [39] in their original `traceroute`-based topology study and was also implemented in Meractor [55]. When resolving the alias of the IP address A, Meractor sends a TCP or a UDP alias probe towards an unused port number of A that replies with an ICMP "port unreachable" message. This message typically has the IP address of the router's shortest-path interface as its source address. If the source IP address of the reply message is different from A, these two IPs are aliases of the same router. This method is prone to the router response configuration problems discussed in Section 3.1.2.

Common IP-identification Counter: The packet ID in the IP header is used for packet reassembly after fragmentation. This technique assumes that a router has a single IP ID counter. For such a router, consecutive packets generated from the router have consecutive IP IDs, regardless of the interface from which the packet left the router. The Ally tool described in [14] and used in the Rocketfuel project implements this mechanism to detect aliases. It sends a UDP probe packet with a high port number to two potential alias IPs. The ICMP "Port Unreachable" responses are encapsulated within separate IP packets and each includes an ID (x and y) in the IP header. Then, it sends the third packet to the address that responded first. Assuming that z is the ID of the third response, if $x < y < z$ and $z - x$ is small (*e.g.*, smaller than 200 in case of Ally [14, 56]), the addresses are assumed to be aliases [14].

Alias resolution based on the ID fingerprint is prone to false negatives due to ID increment settings on routers that are larger than one [57], the absence of a global IPID counter for some router [57], or unexpected jitter in the delivery of probe messages. False positives can also occur as a result of randomly synchronized ID counters of two routers, but this problem can be mitigated by running more tests after a wait period. The other major drawback of this ID-based technique is the overhead of running it on a large set of discovered interfaces; its complexity is $O(n^2)$ for a set of n interfaces. In the case of Ally, some heuristics have been proposed to improve the efficiency of the tool by restricting the possible alias candidates using delays and TTLs [14]. The idea is that alias candidates should have similar TTLs from

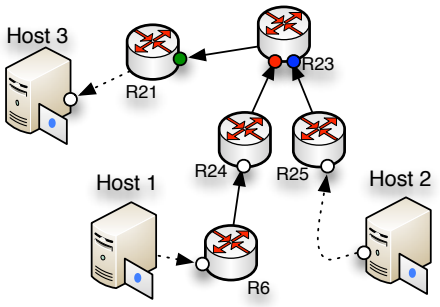


Figure 7: Graph based alias resolution; The green interface succeeds the blue and the red interface in two `traceroutes` so red & blue would be aliases.

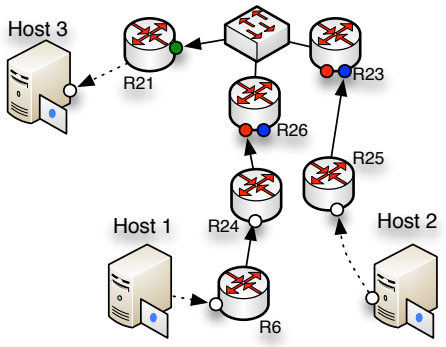


Figure 8: False positive in graph based alias resolution due to the presence of a layer 2 switch; The green interface succeeds the blue and the red interface in two `traceroutes` so red & blue are inferred to be aliases.

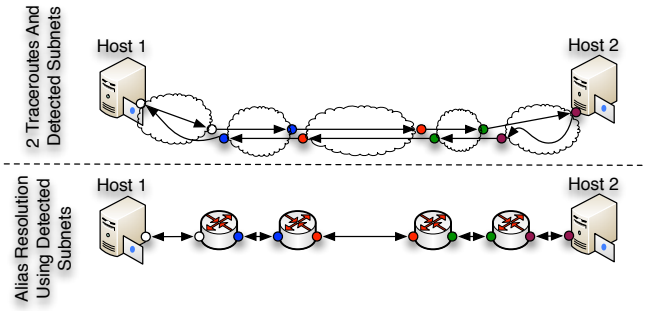


Figure 9: Analytical Alias Resolution for detecting IP aliases on a symmetric path segment.

different vantage points. Thus, the list of candidate aliases can be pruned based on the difference in the hop count distance from common vantage points.

RadarGun [58] mitigates the limitations of Ally by modeling the changes in the packet ID counter. Instead of directly testing each pair of IP addresses separately, it iteratively probes the list of IP addresses at least 30 times. Two IPs are inferred to be aliases if the "velocity" of their corresponding ID counters are consistent in all their responses. The probe complexity of RadarGun is $O(n)$. The main drawback of this technique is the potential of error when used on a large list of IPs. Since routers use a 16-bit counter for the packet ID, counter wrap-arounds can occur during the measurement period. If the probes to the same IP are separated by a period of 40 seconds or longer due the large number of IPs on the list, multiple wrap-arounds are likely to occur. Although the designers of RadarGun have accounted for the possibility of a single wrap-around, the accuracy of the technique diminishes in the presence of multiple wrap-arounds.

DNS-Name: The similarities in DNS names associated with router interfaces can also be used to infer aliases [14, 54], but this approach also has a number of limitations. For one, this technique only works when an AS uses a clear naming convention for assigning DNS names to router interfaces. Second, the complexity of the naming conventions may require human intervention to resolve aliases which limits the scalability of this method. Lastly, the technique is known to be highly inaccurate at the AS borders. The interfaces of border routers usually belong to different ASes which are likely to use different naming conventions. This observation complicates the use of this technique for performing alias resolution at the AS borders [30].

Graph-Based Resolution Heuristics: `traceroute` measurement can offer heuristics on alias inference [54]. Graph-based alias resolution constructs a directed graph by overlaying an individual `traceroute` measurement as illustrated in Figure 1. The "common successor" heuristic suggests which two IP addresses may be aliases. This heuristic relies on the prevalence of routers that respond to `traceroute` probes with the incoming interface. When two `traceroute` paths merge, the common IP belongs to the second router on the shared path. IP addresses prior to the common IP should belong to different interfaces of a single router and hence would be aliases. Figure 7 shows a partial view of a `traceroute` measurement from $Host1$ and $Host2$ toward $Host3$ in our toy example. In this example, the black interface succeeds the red interface in one `traceroute` and succeeds the blue interface in another `traceroute`. The heuristic suggests that the blue and the red interfaces are aliases.

This heuristic falsely infers aliases in the presence of layer 2 switches or multiple-access clouds. Figure 8 depicts an alternate topology to Figure 7. The `traceroute` view in both figures are similar, hence the heuristic infers $R26$'s red interface and $R23$'s blue interface are aliases.

The "same `traceroute`" heuristic identifies IP addresses that can not be aliases. Since each packet visits a router only once, this heuristic states that two IPs occurring on the same `traceroute` can not be aliases.

Analytical Alias Resolution: Given a set of `traceroute`-derived paths, Analytical Alias Resolver (AAR) [59] uti-

lizes the common IP address assignment scheme to infer IP aliases within two opposite paths, one from A to B and the other from B to A. It first identifies the subnets that are linking the routers (as discussed in 3.2.2). Then it aligns the two traceroute paths using the discovered subnets. Alias IPs are easily resolved when point-to-point links are used and the route is symmetric. To illustrate this technique, consider the `traceroute` measurements between *Host1* and *Host2* shown in Figure 9. The top view shows the two `traceroute` paths and the identified subnets. The bottom view depicts how the detected subnets can be used to align the two `traceroute` paths and resolve aliases.

The Analytic and Probe-based Alias Resolver (APAR) [60] consists of an analytical and probe-based component. The analytical component uses the same scheme as ARR, while the probe-based component increases the accuracy of mapping with limited probing overhead. The probe-based component uses `ping`-like probes to determine the distance to each observed IP and mitigates false positives. Any two interfaces can be aliases only if their hop distance differs by at most one hop from a single vantage point. This `ping`-like probe also helps to identify aliases when the source address of the reply is different from the probed IP (*i.e.*, the Common Source Address approach).

Record Route Option: The DisCarte tool [61] uses the standard `traceroute` with enabled Record Route (RR) IP option to detect IP aliases. For the first nine hops, two interfaces are captured, one in the forward path and one in the reverse path. Although the technique sounds intuitive, it is difficult to use effectively in practice because of inconsistent RR implementations by routers and the complexity of aligning RR data with `traceroute` data. DisCarte uses Disjunctive Logic Programming (DLP) to intelligently merge RR and `traceroute` data. However, its implementation does not scale to large datasets. For instance, the application of DisCarte on traces between 379 sources and 376,408 destinations is reported to be so complex that is in fact intractable.

4.2 Progressive Router Discovery

In some networks, routers store information about their neighboring routers. Using this information, the topology can be discovered progressively. In a local area network with `SNMP`-enabled routers, a list of neighboring interfaces can be identified from the "`ipRoute Table MIB`" entry of the router [62]. This technique can be used recursively to discover new routers and the connectivity between them. Although accurate, the use of this technique is limited to the interior of an AS and can only be used by the network administrators with adequate privileges.

More recently, `MRINFO` has been used to discover the topology at the router-level using `IGMP` messages with a similar incremental method [63, 64]. Upon receipt of an `IGMP` "`ASK NEIGHBORS`" message, an IPv4 multicast-capable router replies with an `IGMP` "`NEIGHBORS REPLY`" message that lists all its interfaces and the directly connected interface of the neighboring router. The applicability of this technique is however limited to DVMRP multicast-enabled routers, and their number in today's Internet is small.

4.3 Modeling

The most-cited work on Internet topology modeling is by Faloutsos *et al.* [65]. In their paper, they relied on the `traceroute` data collected by Pansiot *et al.* [39] in mid-1995 which consisted of the inferred router-level paths taken by packets in the Internet and produced an observed router topology. One of their main observations was the scale-free structure of the inferred router topology; that is, the power-law degree distribution of routers. Intuitively, this finding implies the existence of a small number of high-degree core routers and a large number of lower degree edge routers. This paper fueled many of the subsequent studies on modeling the Internet's router-level topology (*e.g.*, [66]) that aimed at reproducing the observed scale-free structure of the inferred topology.

Although the observations reported in [65] seem plausible, many domain experts argued that they are indeed erroneous [67]. For one, no publicly available router topologies exhibit the claimed scale-free structure. For example, in the publicly available maps of Internet2, there is no evidence of a few highly-connected core routers. Second, technology constraints and engineering intuition rule out the existence of high-degree core routers in real-world networks. When configuring a router, network operators are limited by the tradeoff between traffic volume vs. degree. In particular, a core router that processes a large volume of traffic on each interface cannot have a large number of interfaces. On the other hand, routers at the edge of the network carry less traffic per interface and are capable of having more interfaces. These constraints suggest that while router topologies can in theory exhibit degree distributions that are consistent with the reported power-law behavior, the high-degree routers must necessarily be at the edge of the network and not in its core. Ironically, because of the measurement platforms used, none of the `traceroute` campaigns performed in the past would be able to detect those high-degree nodes at the network edge. Third, there is a clear mismatch between the observed scale-free topology and the design philosophy of the Internet. An important requirement of the original DARPA network design was that "Internet communication must continue despite loss of networks or gateways" [17]. However, in a scale-free topology, a failed high-degree central router can lead to a partitioning of the network as shown by Albert *et al.* [66], an alleged property that became well-known as the Internet's "Achilles' heel". Lastly, it has been shown that the errors in the router-level topology considered in [65] are a result of the afore-mentioned limitations of the alias resolution methods and the fact that the inferred high-degree nodes are an artifact of `traceroute`'s inability to penetrate opaque layer-2 clouds—the observed topology of a group of routers at the edge of a layer-2 cloud appears as a

mesh-like ($e.g.$, complete graph) interconnection among all routers and automatically results in the appearance of high-degree nodes.

Alternatively, Heuristically Optimal Topology (HOT) models have been proposed to model the Internet topology. These models are based on the method of reverse-engineering and rely on domain knowledge as an alternative resource as compared to using data in the form of `traceroute` measurements to drive the modeling effort. HOT models are comprised of the following three main components: (i) An objective function that captures the ISP's business goals, (ii) technology constraints that dictate that the basic design of the ISP's router topology reflect a tradeoff that has to be made between cost and efficiency, and (iii) the uncertainty in the environment in the form of the traffic demands imposed on the network. When combining all these ingredients, *constraint optimization* can be used to construct an optimal router topology for an ISP with the stated objective and demands. The construction of such an optimal solution may be NP-hard, but HOT models are not concerned with optimality. Instead, they are concerned with the construction of heuristically optimal solutions that result in "good" performance [17]. A hallmark of the resulting ISP router topologies is the presence of a pronounced backbone consisting of low-degree but high-capacity routers. Moreover, such a backbone is fed by tree-like access networks that are built from high-degree but low-capacity routers, with additional links added for redundancy and resilience.

5. POP LEVEL

The PoP-level is the ideal resolution to study the connectivity of an AS when the objective is identifying all the locations where, at least in theory ($i.e.$, ignoring routing policies), the AS can exchange traffic with its neighbors. As a result, the topology at this level is also very useful for potential customers of an AS who may be interested in the geographic coverage of the AS or in knowing the locations where they can connect.

5.1 Terminology & Approaches

The term PoP (Point of Presence) is a loosely defined term within the Internet community. Internet service providers use PoP to refer to either a physical building with a specific address where they keep their routers, or a metropolitan area where customers can reach their services. In the research community, however, a PoP usually means a collection of tightly connected routers owned by an AS that by design work as a group to provide connectivity to users or to other PoPs of that same AS or other ASes. Therefore, PoPs are the reflection of a hierarchical design principle that many ASes apply when designing their physical infrastructure. Adhering to such a design achieves scalability and facilitates maintainability of a network.

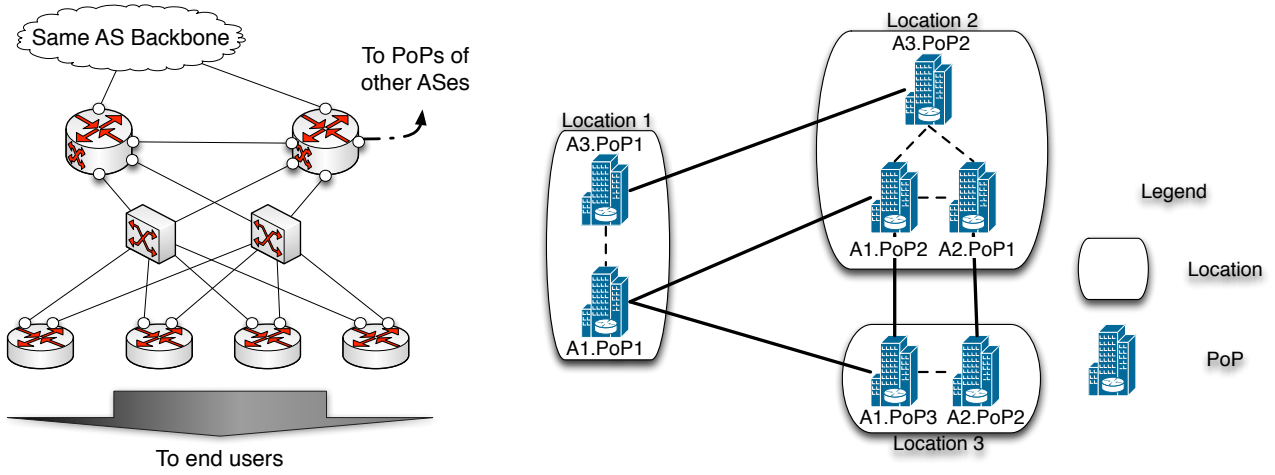
Network operators often apply "cookie cutter" patterns when designing PoPs [68, 69]. This modular design strategy simplifies network debugging and management. Figure 10(a) depicts an example of such a cookie cutter pattern applied to the design of a PoP. Certain patterns are explicitly recommended by some network equipment vendors and show how their products are best used for the buildup of PoPs of certain sizes and with desirable properties ($e.g.$, redundancy, scalability). The design typically ensures redundant within-PoP connectivity, access for customers of the AS that owns the PoP, and connectivity to the rest of the Internet. As a result of this practice, the PoPs of many different ASes have similar internal ($e.g.$, router-level) structure and are found as repeated patterns across the global Internet.

A node in the PoP-level topology of the Internet is the PoP of a given AS and is ideally tagged with the PoP's owner ($i.e.$, AS) and geographical information ($i.e.$, location). Inter-PoP links can be categorized into two types. While *core* or *backbone links* connect two PoPs of the same AS, *peering* links connect PoPs of different ASes. Figure 10(b) shows the PoP-level topology corresponding to the network in Figure 1. Each PoP is identified by its AS and its location. Although $AS1.PoP1$ and $AS3.PoP1$ are in the same location (building), each one is represented by a PoP or node. Backbone links are represented by lines, and dotted lines show peering links.

Prior studies in this area have considered three different basic approaches for obtaining the PoP-level topology of the Internet. The first and the most common approach has been to identify PoPs by aggregating data collected from `traceroute` measurements. This method receives either an interface-level or a router- level topology as input and groups nodes that belong to one PoP. Relevant studies are discussed in Section 5.2.

The second approach is delay-based, but instead of using the per-hop RTT information from `traceroute`, it relies on delay estimates obtained from `ping` measurements. Yoshida *et al.* [16] used this technique to detect the PoPs of four major ISPs in Japan. They argue that the information about an ISP's core network ($e.g.$, routers, DNS names) that is obtained through `traceroute` is unreliable. Instead, they used their own Japan-wide measurement platform to perform large-scale `ping` campaigns. Based on a model that relates the measured end-to-end delays to the sum of the delays between consecutively traversed PoPs, they inferred the presence of PoPs.

The third approach relies on information that is published by different ISPs on their websites. Figure 11 shows one example of an AS's PoP-level topology ($i.e.$, Cogent Communications) that is available online. The map depicts cities that have a PoP of this AS and also shows the interconnection among PoPs of the AS. Topology Zoo [70] is a collection of about 200 topology maps taken from online pages published by a range of different ASes. Since this data is published by the provider itself, it should be more accurate than maps generated by measurement-based techniques. However, obtained maps from online resources are prone to errors due



(a) Cookie cutter design used in the PoP of an AS [68, 69]

(b) The PoP-level topology our example

Figure 10: PoP level topology.

Figure 11: The PoP-level topology of Cogent available online at `http://www.cogentco.com/en/network/network-map`.

to the out-dated data. Moreover, these maps typically only show the connectivity within an AS and do not reveal AS peerings. The Internet Atlas is another research project that aims at providing a map of physical connectivity of the Internet [71]. Nodes in this Atlas (map) represent buildings (*e.g.*, hosting facilities, data centers or colocation buildings), and links show interconnectivity between them. Atlas is built using resources such as online maps and other publicly available information from different repositories or databases.

5.2 Aggregation Methods

In the following, we discuss prior studies that focused on interface and router aggregation to unravel PoP-level topology. Due to the importance of geography at this resolution, we also discuss the studies that examined geographical characterization of PoPs.

The first study that focused on the discovery of PoPs was Rocketfuel [14]. It tried to infer the structure of an AS using `traceroute` measurement and used the PoP-level topology to visualize an AS infrastructure. Rocketfuel first identified alias IPs using Ally's packet ID counter method. It then leveraged the inferred DNS naming conventions used by an AS to geolocate the discovered IPs using a tool called `UNDNS`. `UNDNS` uses a large set of regular expressions to extract city and airport codes embedded in DNS names and infer the geographical location of an interface. In the end, Rocketfuel groups interfaces that are mapped to one and the same geographical location into a PoP.

iPlane [13] extends the approach advanced by Rocketfuel. First, a Meractor-like [55] alias resolution is used to identify routers. Additionally, iPlane uses a mate-30 heuristic similar to AAR [59] and identifies subnets to find candidate alias pairs. A Packet ID fingerprinting technique is used on the candidate alias pairs to infer aliases [14]. Next, DNS names are used to geo-locate routers and group them into PoPs. However, this step is riddled with issues. For one, for some routers, there is no DNS name assigned to any of their interfaces. Also, there is no guarantee that assigned DNS names contain any relevant geography-related information. Furthermore, DNS misnaming can introduce error to this mapping process. DNS names are voluntarily assigned by network administrators and interface misnaming

is fairly common especially due to relocating routers and using old assigned DNS names [72]. In a final step, iPlane considers all routers that it has not been able to map to a location and assigns them a location using a clustering approach that is based on a notion of similarity between interfaces with respect to routing and performance. To this end, iPlane probes all interfaces with `ICMP echo` probes from different Planet Lab nodes. Each interface is assigned a vector in which the i^{th} element is the length of the path from the i^{th} vantage point. Hence the PoP detection problem is translated into a clustering problem involving these measurements, and interfaces in one cluster are assumed to belong to the same PoP.

Note that both projects rely heavily on the capability to extract information about the location of a router from the DNS name assigned to it. The structure of DNS names was recently revisited by Chabarek *et al.* [73]. Their study shows that aside from geographical information, DNS names may include information about interface types, bandwidth, and router manufacturers. However, meaningful encodings are more common in the core of the Internet [73], and the naming structure tends to be strongly tied to the AS that owns the router [74].

Another popular approach is to use geo-IP databases to assign a location to an IP address. Tian *et al.* [38] use these databases in conjunction with a heuristic approach to locate router interfaces. They initially rely on existing geo-IP databases to annotate the given interface level topology graph with geographic information. The resulting annotated graph contains some clusters corresponding to each city. Their heuristic technique re-annotates an interface to a new location if the new annotation results in more coherent groupings, where more links are inside a group. Each group is detected as a PoP. One basic problem with this approach is the well-known inaccuracy of the freely or commercially available geo-IP databases [75, 76, 77].

A PoP consists of a set of routers with high interconnectivity among them. Links inside a PoP are usually very short, implying small delays in general. These properties were used by Feldman *et al.* [15, 78] to propose a more automatic approach for detecting PoPs. In their graph-based approach, network "motifs" are used to detect repeated patterns in `traceroute`-derived interface-level topologies collected by DIMES [12]. These repeated patterns are used to identify tightly connected interfaces. To this end, they ignore all links with a delay above a certain threshold (5 ms); these links are likely to be long-haul connections between distant PoPs. This step generates a graph with disconnected components, each of which is a candidate to represent either a single PoP or multiple PoPs. Different refinement techniques are applied to either split one component or merge different components to detect the PoPs based on graph motifs. To geolocate the inferred PoPs, they use several geolocation services including the MaxMind GeoIP [79]. Finally, they validate their PoP-level topology using a DNS name-based geo-localization data base and two geo-IP data bases. Their claim is that by not using the DNS names as part of their methodology, this information can be used as "ground truth" to validate the accuracy of their technique. Unfortunately, the accuracy of using DNS names to infer geographical location of an interface is questionable [72].

6. AS-LEVEL

The Internet's topology at the AS-level is typically modeled using a simple graph where a node is an AS identified by an AS number. As previously described, an Autonomous System or AS is commonly defined as a collection of IP prefixes under the control of a single network operator that presents a common, clearly defined routing policy to the Internet [80]. In such an AS graph, links represent logical connectivity between two ASes and are labeled according to the type of connection; customer-provider, peer-peer, and sibling relationship. The logical connectivity between two ASes usually represents multiple physical connections that are established between PoPs of the two ASes, presumably to enable the efficient exchange of traffic between them.

This graph representation of the AS topology has a number of limitations. First, each AS has a geographical footprint that may overlap with the footprint of another AS. This feature cannot be illustrated using a simple node to represent an AS, unless the node is replaced by a region that covers the area in question. Second, ASes are widely considered to be coherent entities with a clearly defined routing policy. However, for historical reasons or due to their often global reach, some ASes use different policies in different parts of their network. In this context, Muhlbauer *et al.* [81] demonstrated that treating ASes as atomic structures is a severe over-simplification and negatively impacts our understanding of inter-domain routing. Third, the fact that many inter-AS links represent multiple geographically dispersed physical AS connections cannot be captured by a simple graph. Fourth, IXPs also complicate the AS-level topology by providing connectivity between many ASes, most commonly through layer 2 multiple access clouds. As a result, in a realistic AS topology graph, IXPs should be modeled as links that connect more than two ASes. Together, these issues suggest that a hyper-graph [3] provides a more detailed and informative structure of the Internet's AS topology. However, these numerous limitations notwithstanding, simple graph representations of the AS-level topology are considered to be useful and have been studied for the past two decades to a great extent.

6.1 AS Topology Data Sources

Techniques for discovering the AS-level topology rely mainly on the following three data sources: `BGP` information, `traceroute` measurements, and Internet Routing Registries (IRR) [82]. Below we discuss each type of data source and its limitations in more detail.

`BGP` **Information:** `BGP` is the de-facto standard inter-domain routing protocol of the Internet. `BGP` is a path vec-

tor protocol in which routing decisions are made based on reachability via the advertised AS paths and expressed network policies. The term "reachability protocol" has been used to emphasize this characteristic of BGP. BGP uses the AS number to specify the origin AS of a prefix and ASes along the path to reach the origin AS.

BGP data has been collected by various projects and has subsequently been used in different forms. BGP information can be obtained from various resources, including (i) BGP *archive*: Oregon RouteViews [83] and Reseaux IP Europeens (RIPE) Routing Information Service (RIS) [84] collect BGP route information through a set of route collectors also known as BGP monitors or vantage points. The original purpose of these projects was to help network operators with troubleshooting and debugging tasks, and for these purposes, the data has proved to be invaluable. Both services collect routing table dumps and route update traces on an ongoing basis. While BGP dumps show the best path to reach other ASes, the back-up links and the dynamic nature of BGP routes are more likely captured by "route updates". (ii) *Route Servers*: A route server is a BGP-speaking router that offers interactive login access via `telnet` or `ssh` and permits third parties to run many non-privileged router commands [85]. For example, BGP summary information can be obtained by executing the "`show BGP summary`" command. (iii) *Looking Glasses*: A looking glass is a web interface to a BGP router which often allows basic BGP data querying and supports limited use of debugging tools such as `ping` and `traceroute` [85].

BGP information was the first data source used to map the AS-level topology [86]. Two representative studies that use both BGP dumps and updates to capture the AS-level topology are [85, 87]. Although passive collections of BGP tables and updates have fueled many studies concerned with the AS-level topology, there also have been efforts that used active measurements of BGP. In this context, a BGP beacon [88, 8] is a router that actively advertises and withdraws prefixes. Observing the resulting announcements from the perspective of different route collectors within the larger Internet enables researchers to infer some of BGP's overall behavior (*e.g.*, protocol convergence time and the average AS distance an advertisement travels in the control plane). In a similar manner, BGP route poisoning prevents BGP announcements from reaching an AS. Bush *et al.* [8] used this technique to measure the prevalence of default routes in the Internet and explain the differences in the AS-level topologies obtained from control vs. data plane measurements.

Using BGP for inferring the AS-level topology has several advantages. First, compared to Internet registries, the data collected from BGP shows the actual reachability as seen from the perspective of the Internet control plane. Hence, the data is typically not prone to being stale, obsolete or incorrect. Second, BGP updates can be used to study the dynamic behavior of Internet routing which, in turn, can reveal otherwise hard-to-detect backup links. Third, engineering solutions such as the use of BGP beacons and route poisoning can be applied on top of BGP to improve our view of the topology.

Despite all its advantages, using BGP information to infer the AS-level topology is not without limitations. The main reason is that BGP is merely an information hiding protocol and only indicates reachability, not connectivity. More specifically, AS path announcements are primarily used for loop detection. For traffic engineering reasons, adding an AS in the announcements is not uncommon. Also, ASes may announce AS paths that do not correspond to real paths [81]. Moreover, as a path vector protocol, BGP does not announce information about every available path. As a result, back-up paths might never appear in the BGP dumps. In fact, since BGP only announces the best paths, many alternative AS paths remain hidden from any route collector. Since route collectors are normally deployed in larger ISPs and mostly in the US and Europe, their observed AS-level topology is biased to be more complete for these regions. Additionally, even if the route collectors were randomly placed in different ASes, the likelihood of discovery of an AS relationship is proportional to the number of ASes using that link [42, 8]. This finding proves a measurement bias in BGP-based AS topologies because P2P links are only used for traffic originating from the customers of any of the peering ASes. Hence P2P AS relations are in general much harder to discover than C2P AS relationships [42]. In fact, the majority of the missing AS links in AS-level topologies inferred from BGP data are known to be P2P links [4]. The severeness of this bias and the resulting degree of incompleteness of even the most-carefully inferred currently available AS-level topologies have recently come to light with the discovery of massive amounts of public peering links (*i.e.*, P2P AS connections) at a large European IXP [49], most of which have remained invisible in presently available BGP data.

`traceroute` Measurement: Another approach to discover the Internet's AS-level topology is to use the interface-level topology obtained from `traceroute` measurements. In this approach, each IP in a `traceroute` is mapped to its corresponding AS. BGP routing tables and IRR can be used to map an IP to an AS based on the IP prefixes that are announced by the AS [89]. Consecutive IPs that belong to two different ASes reveal the connectivity between ASes.

This technique has the advantage of revealing a potentially more detailed view of the AS-level topology. Recall that ASes can be connected at multiple locations. `traceroute`-based measurements allow us to distinguish between multiple inter-AS connections between two ASes. In addition, `traceroute` measurements often use more vantage points, mainly because deploying a `traceroute` vantage point is much easier than deploying a BGP route collector. As a result, the AS-level topologies inferred from data collected by large-scale `traceroute` measurement campaigns are generally considered to be more complete than those collected from BGP information [12, 45, 46].

Apart from the limitations of `traceroute` that we discussed in Section 3.1.2, active measurements in the data plane have other limitations when used for mapping the AS-level topology. First, IP-to-AS mapping is a non-trivial task. Prefix registries are often incomplete and using `BGP` for mapping IPs to AS numbers is not accurate due to BGP's information hiding characteristics. Second, discovering a false inter-AS connection is likely due to inconsistencies in router responses [45, 6]. Third, private IPs and IPs in the carrier-grade NAT (large-scale NAT) IP range may also appear in a `traceroute` which renders the IP-to-AS mapping impossible for these IPs[6].

Finally, it is worth mentioning that when measuring the AS-level topology using `BGP` and `traceroute` measurements respectively, what is really measured are the Internet control and data planes. While the control plane focuses on "reachability", the data plane is all about "connectivity". The inconsistencies in the data plane and the control plane measurement may result in different and inconsistent views of the Internet AS-level topology. In general, these issues stem from the limitations of the data that is used to infer the topology and the lack of knowledge about the effects of these limitations on the observed topology [8]. For instance, "default routing" limits the view of passive `BGP` measurements while it has the potential of enhancing the view of active measurements (*e.g.*, observe the route). The general consensus is that the AS-level topology inferred from measurements in the data plane results in a more accurate and complete view as compared to relying on measurements in the control plane [8, 17, 3]. However, [49] is a reminder of the caveat associated with this consensus.

Internet Routing Registries: The Routing Arbiter Database (RADb) provided by IRR is a group of look-up databases maintained by several organizations. These databases are designed to provide fundamental information about routing in the Internet, including documented routing policies, regulations, and peering information

The main advantage of using IRR is its simplicity. All the information is accessible via the `WHOIS` command and can be obtained through `FTP` servers. Being based on data provided by the different ASes themselves, this resource does not exhibit the sort of limitations that data obtained through measurements have. However, when using this resource, extra care is needed for different reasons. For one, since these registries are populated and maintained on a completely voluntary basis, the available data may be stale or incomplete due to confidentiality reasons, personnel changes in the different ASes, or because of the overhead of updating an external data store. For instance, reports that checked the accuracy of RIPE-provided data show inconsistencies in different IRR-provided databases [90].

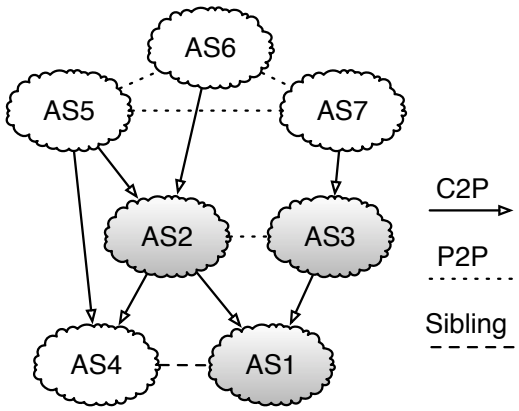


Figure 12: AS graph annotation with AS relations

6.2 AS Relationship & AS Tiers

Although the logical AS-level topology is interesting in itself, to be more useful in practice, the inter-AS routing policies should also be inferred. We recall that the business relations between connected ASes are broadly classified into [91] (1) Customer-Provider (C2P), (2) Peer-Peer (P2P), and (3) Sibling relations. From a financial perspective, in a C2P relation, the customer is billed for using the provider to reach the rest of the Internet. The other two types of relationship are in general settlement-free; that is, no money is exchanged between the two parties involved in a P2P relationship. A P2P relation helps, for example, two small ASes with high inter-AS traffic profiles to reduce their cost by directly exchanging traffic, hence reducing the traffic sent towards their providers. Sibling relations typically occur when business mergers happen or when multiple ASes are owned and operated by one and the same company or organization.

Early approaches to inferring AS relations used AS size and AS degree. Gao *et al.* proposed an algorithm based on the intuition that a provider typically has a larger size than its customers and that the size of an AS is typically proportional to its degree in the AS graph [92].

The commonly-used approach to infer inter-AS relationships is to use the observed routing paths and assume that the "valley-free property" holds without exceptions in the Internet [93, 92, 94]. For an AS path, if we number links as +1, 0, -1 for provider-to-customer, peer-to-peer and customer-to-provider, the valley-free property states that any valid path should only see a sequence of +1's, followed by at most one 0, followed by a sequence of -1's. The type of relationship assignment can be formulated as an optimization problem. Given an undirected graph representation of the AS topology and a set of AS-level paths, the aim is to assign policy labels to the links in such a way as to minimize the number of invalid routes. Although this problem is proven to be NP-hard, some approximation techniques have been presented in the literature [93].

An alternative approach is to check the consistency of in-

ferred relations (using any of the above methods) with other measurements [18]. For instance, Muhlbauer *et al.* [81] used `traceroute` measurements to estimate the accuracy of the inference by comparing the inferred routes and the real routes. In their approach, they use multiple quasi-routers to capture route diversity within the ASes.

Traditionally, the AS-level topology is widely regarded to be hierarchical in nature, where ASes are categorized into different tiers [93, 94]. Tier-1 ASes are defined as those that don't buy transit from any other AS. These tier-1 ASes form a complete graph (*i.e.*, full mesh connectivity) at the highest tier. Tier-2 providers are customers of the tier-1 ASes using them for Internet transit. Additionally, tier-2 ASes use peer-peer relations with other tier-2 ASes to decrease the transit cost. This hierarchical structure can be extended to more levels. However, this perception is changing. First, many new ASes (*e.g.*, content providers and Content Distribution Networks (CDN)) are inherently different from the traditional ISPs and also tend to have many connections at various locations. These new types of ASes do not fit within the traditional tiered AS hierarchy. In addition, new studies explain this changing perception using the abundance of missing links and the limited observability of P2P connections in currently studied AS topologies [17]. Although the existence of large transit ASes at the highest tier remains valid, the tier-based hierarchical view is replaced by a flatter but more modular view. Figure 12 shows an example of an annotated AS graph. *AS1*, *AS2*, and *AS3* are the ASes in our previous examples. *AS5*, *AS6*, and *AS7* are tier-1 ASes forming a full mesh at the highest tier. However, there is no longer a pronounced hierarchical structure below these tier-1s.

6.3 Coverage & Completeness

As of 2011, diligently-inferred AS-level topologies consists of approximately 40,000 ASes and 115,000 to 135,000 edges, with 80,000-90,000 C2P links and the rest P2P links [17]. While such topologies seem to be complete with respect to their node sets (*i.e.*, ASes), their edge sets are typically inaccurate and miss a large number of AS-links, especially with respect to P2P links.

A great deal of research has been dedicated to studying the question of completeness of inferred AS-level topologies. The "Lord of the Links" study [42] compares `BGP` routing tables, Internet Routing Registries, and `traceroute` measurements, cross validates the topology captured from these various sources and captures a more complete view of the AS topology. The authors of this study also extract a significant amount of new information from the operational IXPs worldwide and use this information in their cross validation process.

The incompleteness of the Internet AS map has also been studied (*e.g.*, [4, 95]). Oliveira *et al.* [96] use ground truth data provided by a large tier-1 ISP to validate the accuracy of their derived AS maps for a few target ASes. The ground truth is built upon router configuration files, syslogs, `BGP` command outputs, and personal communications with the network operators. Oliveira *et al.* [4] categorize the missing links into *hidden* and *invisible* links. Invisible links are missing due to the limitations imposed by the placement of vantage points. Hidden links, on the other hand, can be found with further measurements. On the active measurement side, the importance of the distribution of `traceroute` vantage points is studied by Shavitt *et al.* [48]. Given a large set of vantage points, they use sensitivity analysis and measure the changes in the discovered topology using a different number of vantage points. They show that although increasing the number of vantage points can help reducing sampling bias, it can not overcome the bias due to their placement. They conclude that measuring from within a network is important for discovering more of its links, mainly for low-tier ASes.

More recently, the AS-level map underwent a major revamping due to the availability of ground truth data from one of the largest IXPs in Europe with some 400 AS members [49]. The main finding was that in this single IXP, there are more than 50,000 P2P links visible, which is more than the total number of P2P inferred Internet-wide. This finding suggests the total number of P2P links in the Internet is likely to be larger than 200,000. More importantly, this recent observation shows that the presently-used AS-level Internet topologies are far from complete, with much room for improvement.

6.4 Geolocation

Apart from prior studies on the geographic locations of PoPs of an AS, little has been done on mapping the geography of ASes (i.e. the geographical area that is served by an AS). The notion of the geography of an AS has become even more delicate with the emergence of newer types of ASes such as larger content providers, CDNs and and cloud providers. While the geography of a traditional AS like 7018 (AT&T North America) is well defined in the sense that it covers the US, defining the geography for ASes like 15169 (Google) or 20940 (Akamai) is more complicated as they cover roughly the whole globe.

Internet registries and directories such as PeeringBD [97] provide a plethora of information about the geography of ASes. PeeringDB for instance provides a list of public and private facilities where an AS has PoPs. Similar to other online resources, these directories are easy to use but can be out of date and incomplete.

The geographical footprint of eyeball ASes (ISP that serve residential costumers) was studied in [98]. Using large-scale measurements from Peer-to-Peer applications, the authors of this study identify a large set of end-host IPs. These IPs are then mapped to ASes. Next, the geographical coverage an of AS is estimated using the geo-density of a large number of its customers. Different IP-to-geolocation databases are used to find the location of an IP address, taking into account

the errors inherent in those databases. Since a large volume of customers are used to map the geo-footprint of an AS, the potential error in IP-to-geo mapping does not influence the final discovered coverage of the AS.

6.5 Modeling

Several studies have examined the presumed AS topology of the Internet from a graph-theoretic perspective and have proposed different graph-based network models. However, there is no consensus on which of the studied models is more relevant or realistic due to the incompleteness of the inferred topologies. Zhou *et al.* [99] propose a growth model with Positive-Feedback-Preference which reproduces many topological properties of inferred AS-level topologies. Their model, however, uses the Skitter [40] `traceroute` dataset to infer the target AS-level topology. As discussed earlier, this dataset suffers from well-known limitations of `traceroute`-based mapping efforts. For instance, the observed power-law degree distribution of this AS topology is known to be due to the bias in the measurement techniques [3, 17]. Mahadevan *et al.* [87] used the AS topologies inferred from multiple data sources that included BGP, `traceroute` and `WHOIS` measurements. They compared the resulting graphs from a graph-analysis perspective and reported that the "joint degree distribution" can be used to characterize the Internet AS graph. They also showed how the data collection peculiarities can explain differences in their graph comparison study.

The evolution of the Internet AS map has also been investigated. The main challenge with respect to the long-term evolution of the topology is to distinguish between topology changes and changes due to routing dynamics. Oliveira *et al.* [100] compose a model that distinguishes between the two different events. Their findings suggest that the impact of transient routing dynamics on topology decreases exponentially over time. Dhamdhere *et al.* [101, 102] take a different approach in characterizing the AS map evolution. They compare the AS maps collected during the past 12 years using `BGP` dumps. They report that the AS-level topology was growing exponentially until 2001, but this growth has settled into a slower exponential growth in terms of both ASes and inter-AS links. However, the average path length has remained the same. These measured graph properties can be used in topology generators to build AS-level models of the Internet.

In view of the latest understanding of the quality of the different inferred AS-level topologies that have been studied in the past, a recent common theme in AS topology modeling has been that a proposed model is only as good as the underlying data. Moreover, there has been increasing awareness that any strict graph-theoretic treatment of the AS-level Internet topology necessarily misses out on the key fact that this topology is a construct that is mainly driven by economic factors and decisions. [103, 104, 101] are early attempts at addressing these points. For example, Chang *et al.* [104] use a policy-based graph model, where policies are implemented in a simulated environment and effect how ASes decide to create new AS relations. Similar to the HOT modeling approach for the router-level topology, their model follows the reverse-engineering approach. As part of an AS's decision making process, they consider the gain from P2P links and C2P links, using simulated traffic demands. Using different profiles for ASes with different objectives, they model the behavior of these ASes and model the Internet using an evolutionary framework. To validate the model, they use publicly available measurements and perform their own measurement experiments to check for consistency of the model with the real-world Internet. Lodhi *et al.* [105] build on this initial attempt described in [103] and consider an agent-based network formation model for the AS-level Internet. The proposed model, called GENESIS, is based on realistic provider and peering strategies, with ASes acting in a myopic and decentralized manner to optimize a cost-related fitness function.

7. DISCUSSION

7.1 Examples of "Big (Internet) Data"

The Internet is arguably the largest man-made complex system. As such, it has attracted the attention of the larger scientific community, and the number of studies on topics related to measuring, analyzing, modeling, predicting and providing a basic understanding of the structure and behavior of this highly-engineered network has increased dramatically over the last two decades. Importantly, this increase in Internet-related publications started with the initial availability of large new datasets of Internet-specific measurements (*e.g.*, traffic traces [106, 107], routing data [83, 86]), and the subsequent explosive growth has been largely driven by "big Internet data"; that is, publicly available or proprietary datasets resulting from large-scale measurement experiments that tend to produce voluminous amounts of observations.[1] Typically, these observations have rich semantic content and often provide useful information about the Internet as a whole or about its individual components (*e.g.*, ASes, routers, protocols, services).

In general, the producers and owners of the various types of "big Internet data" are network researchers or operators, and while their reasoning for collecting data may vary, practical reasons (*e.g.*, for trouble shooting) almost always trump altruistic arguments (*e.g.*, for the good of science). For example, in the case of Internet topology-related big data which is the focus of this survey, the realization that the influence of the Internet's structure or topology on the network's functionality (and vice versa) is in general not well-understood but is the root cause of many encountered networking problems has been the main motivation for data collection projects

[1] As part of the Ark project [41] alone, some 10 billion `traceroute` measurements have been collected during September 2007 and January 2011.

such as Route Views [83] and RIPE RIS [84]. The measurements from these and similar efforts have proven to be invaluable for purposes such as network management, troubleshooting, and debugging. In addition, they have also informed the design of new protocols, applications, and services and have contributed to an increased awareness of the vulnerability of the Internet to a growing number of ever more potent cyber threats.

7.2 Lessons Learned & a Check List

A key lesson learned from surveying the existing literature on Internet topology discovery has been the realization that "more is not always better." That is, using more measurements from the same `traceroute` campaign or from the same set of BGP monitors is now viewed as a non-starter for solving the severe degree of incompleteness of all past and current inferred Internet topology maps at all four levels. In this sense, "big Internet data" is a reminder that the extraction of key information from big data cannot rely on big data analytics alone but is often intimately tied to applying detailed domain knowledge and hard-to-quantify engineering intuition.

A closely-related lesson is that in the context of data sources for Internet topology discovery, "less can actually be more" in the sense that a strategically-placed vantage point can have much better visibility in certain substrates of the Internet topology than a large number of vantage points that have been selected in an ad-hoc fashion or are tied to a fixed measurement platform. The recent IXP studies [49, 108, 109] are prime examples that highlight this point.

At a more technical level, the main lesson from getting to know the lay of the land with respect to Internet topology discovery is that "details matter." For example, using `traceroute` measurements "blindly" without knowing the technique's main idiosyncrasies detailed in Sections 3 and 4 is bound to lead to incorrect result, flawed claims or wrong findings about the Internet in general and its topologies at the different levels in particular. Similar comments apply to the "blind" use of BGP measurements or other data sources that have been tapped for Internet topology discovery.

Even though many of the datasets that have been used in the context of Internet topology discovery have been created by network researchers, as owners of these datasets, they have largely failed to communicate to the users or consumers of their data the main limitations and issues. A notable exception is the original work by Pansiot and Grad [39], but unfortunately, their diligent efforts listing critical issues with `traceroute` and highlighting important artifacts in the obtained data has been all but ignored and forgotten by the networking community [110]. As illustrated in this survey, this failure to properly educate the networking community and scientific community at large about the pitfalls and drawbacks of using the available data "as is" has led to numerous dead-end research efforts. Importantly, it has in general hampered progress in this important area of Internet research as evidenced by a lack of high-quality maps of the Internet topology at any of the described levels, even to this date.

To improve upon this unfortunate situation, we provide in the following a checklist that is a compilation of the main lessons learned from past work in this area. We encourage every researcher interested in working on Internet topology-related problems to consult this checklist before embarking on their own work of any measurement, analysis or modeling efforts that make use of tools, datasets, or methods that have been precisely addressed in past work in this area. Our checklist consists of a number of increasingly more detailed questions for that any interested researcher should ask upfront:

- What datasets are used or generated for the planned work?
- What techniques have been used to obtain the data?
- What are the (known) limitations of the used techniques and what is known about how these limitations impact the quality of the data?
- How can the known data quality issues impact the results of the planned work?
- If the known data quality issues are claimed to be minor, do the obtained results and findings withstand further scrutiny based on alternative data or available domain knowledge?

In a nutshell, by raising researchers awareness of the limitations of the used measurement techniques and how they may affect the resulting data, researchers will be able to answer for themselves whether or not the used tools, datasets, or methods are of sufficient quality to successfully tackle the particular research problem they are interested in. We view consulting this straightforward checklist as a first step that will hopefully prevent researchers from repeating some of the same or similar mistakes that have been made in the past and that have negatively impacted the progress in this important area of networking research.

7.3 Outlook

If past experience is any indication, progress in terms of discovering more accurate, complete, and internally consistent Internet topologies at different levels will come from renewed effort that gives quality priority over quantity when it comes to Internet measurements. Here, quality refers first and foremost to the choice of the locations of the vantage points used in large-scale measurement campaigns but also includes the diligence necessary to eliminate as many of the known idiosyncrasies inherent in most of the presently-used measurement techniques. Recent examples that demonstrate the promising results that measurement platforms with purposefully-chosen vantage points can produce over conventional measurement infrastructures in the context of Internet topology discovery are seen in [35, 49, 46, 111]. However, these are

early efforts, and the potential for carefully and purposefully-designed next-generation measurement platforms with programmable vantage points in strategic locations indicates an exciting future for research in Internet topology discovery.

Regarding many of the measurement techniques underlying Internet topology discovery, we have shown in this survey that despite gradual and significant progress and achievements in increasing our understanding of the many idiosyncrasies of `traceroute` or `BGP` and how they affect the integrity and quality of the resulting data, the Internet topologies that one can infer are at best inadequate. A main reason for this unfortunate situation is that neither `traceroute` nor `BGP` have been designed for Internet topology discovery but have been "re-purposed" by researchers for that very task. Instead of accepting this situation as a "fait accompli", the time seems ripe to try and do away with the utilization of such "engineering hacks" for the purpose of Internet topology mapping. To this end, we advocate for the pursuit of a "clean slate" design of techniques and/or protocols for the explicit and exclusive purpose of Internet topology discovery at a given level. The objective of such an effort is plain and simple: the design of new measurement techniques that enable researchers to measure what they *want* to measure, and not just what they *can* measure.

Lastly, main Internet topology-related modeling studies covered in this survey indicate a clear preference for treating Internet topologies at any level strictly as graph-theoretic constructs and relying mainly graph theory to study their properties and behavior. However, from a networking perspective, such an approach is largely counter-intuitive because the existing Internet topologies at the different layers are highly-engineered systems with pronounced structures and well-defined functionalities. As such, structure trumps randomness when it comes to Internet topology design, and even the latest random graph models (*e.g.*, scale-free networks of the preferential attachment type or the many variants thereof [66]) fail to account for the networks' real-world structures, let alone their functionalities. The HOT models discussed in Section 6 are proof that real-world technological networks such as the Internet's interface-level, router-level, or PoP-level topologies are amenable to mathematical formulations of network design problems that can account for the main underlying engineering-based design criteria and principles and have solutions that are fully consistent with networking reality. Extending this approach to non-technological networks such as the Internet AS-level topology, an inherently economics-driven construct, looms as an exciting open research area, and studies such as [103, 104, 112] are initial attempts in this direction.

8. CONCLUSION

This survey is concerned with the use of "big Internet data" in the form of massive amounts of `traceroute` measurements and `BGP`-derived observations for the main purpose of Internet topology discovery. To this end, we consider the Internet's topology at different resolutions or levels and organize the body of research that has been produced in the past 15-20 years on Internet topology discovery and related topics into studies concerned with the interface-level, router-level, PoP-level, and AS-level topologies of the Internet, respectively.

For each level, we introduce the data used to capture the corresponding topology and classify the data sources based on their type (*i.e.*, data plane vs. control plane measurements) and the techniques used to collect them (*i.e.*, active vs. passive measurement methods). We explain in detail the problems of the different most commonly-used techniques and discuss the limitations and issues that these problems create when using the resulting data for Internet topology discovery at each level. In the process, we show how the main studies in this area have dealt with these known issues of the different data sources and also review the existing literature in this area with an eye on efforts that address geographical properties of the Internet topology and present innovative approaches to Internet topology modeling at different levels.

We conclude with a discussion of some of the main lessons gained from surveying the existing literature on Internet topology discovery. By transforming these lessons into a simple and straightforward checklist, it is our hope that future researchers interested in working on Internet topology-related problems will first consult and reflect upon this checklist, and by doing so will avoid making the same or similar mistakes that have hampered progress in this important area of Internet research. At the same time, we also list a number of challenging new problems as part of an exciting research agenda, and the timely solution of these and similar problems promises the advancement of Internet topology discovery by leaps.

9. ACKNOWLEDGEMENTS

This project is funded by the National Science Foundation (NSF) grant no. CNS 209490. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

10. REFERENCES

[1] X. Cai, J. Heidemann, B. Krishnamurthy, and W. Willinger, "Towards an AS-to-organization Map," in *Proceedings of the 10th ACM SIGCOMM conference on Internet measurement*. ACM, 2010, pp. 199–205.

[2] D. L. Mills and H.-W. Braun, "The NSFNET backbone network," in *ACM SIGCOMM Computer Communication Review*, vol. 17, no. 5. ACM, 1987, pp. 191–196.

[3] M. Roughan, W. Willinger, O. Maennel, D. Perouli, and R. Bush, "10 lessons from 10 years of measuring and modeling the internet's autonomous systems," *Selected Areas in Communications, IEEE Journal on*, vol. 29, no. 9, pp. 1810–1821, 2011.

[4] R. Oliveira, D. Pei, W. Willinger, B. Zhang, and L. Zhang, "The (in) completeness of the observed Internet AS-level structure," *IEEE/ACM Transactions on Networking (ToN)*, vol. 18, no. 1, pp. 109–122, 2010.

[5] M. C. Toren, "tcptraceroute: an implementation of traceroute using TCP SYN packets," man page, 2001, see source code: http://michael.toren.net/code/tcptraceroute/.

[6] Y. Zhang, R. Oliveira, Y. Wang, S. Su, B. Zhang, J. Bi, H. Zhang, and L. Zhang, "A framework to quantify the pitfalls of using traceroute in AS-level topology measurement," *Selected Areas in Communications, IEEE Journal on*, vol. 29, no. 9, pp. 1822–1836, 2011.

[7] B. Augustin, X. Cuvellier, B. Orgogozo, F. Viger, T. Friedman, M. Latapy, C. Magnien, and R. Teixeira, "Avoiding Traceroute Anomalies with Paris Traceroute," in *Proceedings of the 6th ACM SIGCOMM conference on Internet measurement*. ACM, 2006, pp. 153–158.

[8] R. Bush, O. Maennel, M. Roughan, and S. Uhlig, "Internet optometry: assessing the broken glasses in Internet reachability," in *Proceedings of the 9th ACM SIGCOMM conference on Internet measurement conference*. ACM, 2009, pp. 242–253.

[9] C. M. Bowman, P. B. Danzig, U. Manber, and M. F. Schwartz, "Scalable Internet resource discovery: Research problems and approaches," *Communications of the ACM-Association for Computing Machinery-CACM*, vol. 37, no. 8, pp. 98–107, 1994.

[10] Y. Zhang, H.-L. Zhang, and B.-X. Fang, "A survey on Internet topology modeling," *Journal of Software*, vol. 15, no. 8, pp. 1220–1226, 2004.

[11] B. Donnet and T. Friedman, "Internet topology discovery: a survey," *Communications Surveys & Tutorials, IEEE*, vol. 9, no. 4, pp. 56–69, 2007.

[12] Y. Shavitt and E. Shir, "DIMES: Let the Internet measure itself," *ACM SIGCOMM Computer Communication Review*, vol. 35, no. 5, pp. 71–74, 2005.

[13] H. V. Madhyastha, T. Isdal, M. Piatek, C. Dixon, T. Anderson, A. Krishnamurthy, and A. Venkataramani, "iPlane: An information plane for distributed services," in *Proceedings of the 7th symposium on Operating systems design and implementation*. USENIX Association, 2006, pp. 367–380.

[14] N. Spring, R. Mahajan, and D. Wetherall, "Measuring ISP topologies with Rocketfuel," *ACM SIGCOMM Computer Communication Review*, vol. 32, no. 4, pp. 133–145, 2002.

[15] D. Feldman and Y. Shavitt, "Automatic large scale generation of Internet pop level maps," in *IEEE GLOBECOM 2008*. IEEE, 2008, pp. 1–6.

[16] K. Yoshida, Y. Kikuchi, M. Yamamoto, Y. Fujii, K. Nagami, I. Nakagawa, and H. Esaki, "Inferring PoP-level ISP topology through end-to-end delay measurement," in *Passive and Active Network Measurement*. Springer, 2009, pp. 35–44.

[17] W. Willinger and M. Roughan, "Internet Topology Research Redux," *ACM SIGCOMM eBook: Recent Advances in Networking*, 2013.

[18] Z. M. Mao, L. Qiu, J. Wang, and Y. Zhang, "On AS-level path inference," in *ACM SIGMETRICS Performance Evaluation Review*, vol. 33, no. 1. ACM, 2005, pp. 339–349.

[19] X. A. Dimitropoulos, D. V. Krioukov, and G. F. Riley, "Revisiting Internet AS-level topology discovery," in *Passive and Active Network Measurement*. Springer, 2005, pp. 177–188.

[20] C. Metz, "Interconnecting ISP networks," *Internet Computing, IEEE*, vol. 5, no. 2, pp. 74–80, 2001.

[21] F. Wang and L. Gao, "On inferring and characterizing Internet routing policies," in *Proceedings of the 3rd ACM SIGCOMM conference on Internet measurement*. ACM, 2003, pp. 15–26.

[22] R. Beverly, A. Berger, and G. G. Xie, "Primitives for active Internet topology mapping: Toward high-frequency characterization," in *Proceedings of the 10th ACM SIGCOMM conference on Internet measurement*. ACM, 2010, pp. 165–171.

[23] V. Jacobson, "Traceroute," see source code: ftp://ftp.ee.lbl.gov/traceroute.tar.gz.

[24] S. Savage, "Sting: A TCP-based Network Measurement Tool," in *USENIX Symposium on Internet Technologies and Systems*, vol. 2, 1999, pp. 7–7.

[25] M. Luckie, Y. Hyun, and B. Huffaker, "Traceroute Probe Method and Forward IP Path Inference," in *Proceedings of the 8th ACM SIGCOMM conference on Internet measurement*. ACM, 2008, pp. 311–324.

[26] J. Moy, "OSPF version 2," rfc 2328: see online http://tools.ietf.org/html/rfc2178, 1997.

[27] R. W. Callon, "Use of OSI IS-IS for routing in TCP/IP and dual environments," rfc 1195: see online http://tools.ietf.org/html/rfc1195, 1990.

[28] P. Marchetta, V. Persico, E. Katz-Bassett, and A. Pescapé, "Dont trust traceroute (completely)," in *ACM CoNEXT Student workshop*, 2013.

[29] B. Yao, R. Viswanathan, F. Chang, and D. Waddington, "Topology inference in the presence of anonymous routers," in *In IEEE INFOCOM*, 2003, pp. 353–363.

[30] R. Govindan and H. Tangmunarunkit, "Heuristics for Internet Map Discovery," in *INFOCOM 2000*, vol. 3. IEEE, 2000, pp. 1371–1380.

[31] J. Sommers, P. Barford, and B. Eriksson, "On the prevalence and characteristics of MPLS deployments in the open Internet," in *Proceedings of the 2011 ACM SIGCOMM conference on Internet measurement conference*. ACM, 2011, pp. 445–462.

[32] B. Donnet, M. Luckie, P. Mérindol, and J.-J. Pansiot, "Revealing MPLS tunnels obscured from traceroute," *ACM SIGCOMM Computer Communication Review*, vol. 42, no. 2, pp. 87–93, 2012.

[33] B. Eriksson, P. Barford, J. Sommers, and R. Nowak, "DomainImpute: Inferring unseen components in the Internet," in *INFOCOM, 2011 Proceedings IEEE*. IEEE, 2011, pp. 171–175.

[34] B. Eriksson, P. Barford and J. Sommers and R. Nowak, "Inferring Unseen Components of the Internet Core," *Selected Areas in Communications, IEEE Journal on*, vol. 29, no. 9, pp. 1788–1798, 2011.

[35] B. Augustin, B. Krishnamurthy, and W. Willinger, "IXPs: Mapped?" in *Proceedings of the 9th ACM SIGCOMM conference on Internet measurement conference*. ACM, 2009, pp. 336–349.

[36] R. Sherwood and N. Spring, "Touring the Internet in a TCP sidecar," in *Proceedings of the 6th ACM SIGCOMM conference on Internet measurement*.

ACM, 2006, pp. 339–344.
[37] B. Donnet, P. Raoult, T. Friedman, and M. Crovella, "Deployment of an algorithm for large-scale topology discovery," *Selected Areas in Communications, IEEE Journal on*, vol. 24, no. 12, pp. 2210–2220, 2006.
[38] Y. Tian, R. Dey, Y. Liu, and K. W. Ross, "China's Internet: Topology mapping and geolocating," in *INFOCOM, 2012 Proceedings IEEE*. IEEE, 2012, pp. 2531–2535.
[39] J.-J. Pansiot and D. Grad, "On routes and multicast trees in the Internet," *ACM SIGCOMM Computer Communication Review*, vol. 28, no. 1, pp. 41–50, 1998.
[40] CAIDA, "Macroscopic Topology Measurements Project and the Skitter infrastructure," http://www.caida.org/tools/measurement/skitter/.
[41] CAIDA, "Macroscopic topology measurements project and the archipelago measurement infrastructure," http://www.caida.org/projects/ark/, 2011.
[42] Y. He, G. Siganos, M. Faloutsos, and S. Krishnamurthy, "Lord of the links: a framework for discovering missing links in the Internet topology," *IEEE/ACM Transactions on Networking (ToN)*, vol. 17, no. 2, pp. 391–404, 2009.
[43] N. T. Spring, D. Wetherall, and T. E. Anderson, "Scriptroute: A Public Internet Measurement Facility," in *USENIX Symposium on Internet Technologies and Systems*, 2003.
[44] T. Isdal, M. Piatek, A. Krishnamurthy, and T. Anderson, "Leveraging BitTorrent for end host measurements," in *Passive and Active Network Measurement*. Springer, 2007, pp. 32–41.
[45] K. Chen, D. R. Choffnes, R. Potharaju, Y. Chen, F. E. Bustamante, D. Pei, and Y. Zhao, "Where the sidewalk ends: Extending the Internet AS graph using traceroutes from P2P users," in *Proceedings of the 5th international conference on Emerging networking experiments and technologies*. ACM, 2009, pp. 217–228.
[46] M. A. Sánchez, J. S. Otto, Z. S. Bischof, D. R. Choffnes, F. E. Bustamante, B. Krishnamurthy, and W. Willinger, "Dasu: Pushing experiments to the Internet's edge," in *Proc. of USENIX NSDI*, 2013.
[47] P. Barford, A. Bestavros, J. Byers, and M. Crovella, "On the marginal utility of network topology measurements," in *Proceedings of the 1st ACM SIGCOMM Workshop on Internet Measurement*. ACM, 2001, pp. 5–17.
[48] Y. Shavitt and U. Weinsberg, "Quantifying the importance of vantage points distribution in Internet topology measurements," in *INFOCOM 2009, IEEE*. IEEE, 2009, pp. 792–800.
[49] B. Ager, N. Chatzis, A. Feldmann, N. Sarrar, S. Uhlig, and W. Willinger, "Anatomy of a large european IXP," in *Proceedings of the ACM SIGCOMM 2012 conference on Applications, technologies, architectures, and protocols for computer communication*. ACM, 2012, pp. 163–174.
[50] E. Katz-Bassett, H. V. Madhyastha, V. K. Adhikari, C. Scott, J. Sherry, P. Van Wesep, T. E. Anderson, and A. Krishnamurthy, "Reverse traceroute," in *NSDI*, vol. 10, 2010, pp. 219–234.
[51] W. de Donato, P. Marchetta, and A. Pescapé, "A hands-on look at active probing using the IP prespecified timestamp option," in *Passive and Active Measurement*. Springer, 2012, pp. 189–199.
[52] M. E. Tozal and K. Sarac, "Subnet level network topology mapping," in *Performance Computing and Communications Conference (IPCCC), 2011 IEEE 30th International*. IEEE, 2011, pp. 1–8.
[53] M. Tozal and K. Sarac, "Tracenet: an Internet topology data collector," in *Proceedings of the 10th ACM SIGCOMM conference on Internet measurement*. ACM, 2010, pp. 356–368.
[54] N. Spring, M. Dontcheva, M. Rodrig, and D. Wetherall, "How to resolve IP aliases," *Univ. Michigan, UW CSE Tech. Rep*, pp. 04–05, 2004.
[55] K. Keys, "iffinder, a tool for mapping interfaces to routers," *See http://www. caida. org/tools/measurement/iffinder*.
[56] K. Keys, Y. Hyun, M. Luckie, and K. Claffy, "Internet-scale IPv4 alias resolution with MIDAR," *IEEE/ACM Transactions on Networking (TON)*, vol. 21, no. 2, pp. 383–399, 2013.
[57] W. Chen, Y. Huang, B. F. Ribeiro, K. Suh, H. Zhang, E. d. S. e Silva, J. Kurose, and D. Towsley, "Exploiting the IPID field to infer network path and end-system characteristics," in *Passive and Active Network Measurement*. Springer, 2005, pp. 108–120.
[58] A. Bender, R. Sherwood, and N. Spring, "Fixing ally's growing pains with velocity modeling," in *Proceedings of the 8th ACM SIGCOMM conference on Internet measurement*. ACM, 2008, pp. 337–342.
[59] M. H. Gunes and K. Sarac, "Analytical IP alias resolution," in *Communications, 2006. ICC'06. IEEE International Conference on*, vol. 1. IEEE, 2006, pp. 459–464.
[60] M. Gunes and K. Sarac, "Resolving IP aliases in building traceroute-based Internet maps," *IEEE/ACM Transactions on Networking (ToN)*, vol. 17, no. 6, pp. 1738–1751, 2009.
[61] R. Sherwood, A. Bender, and N. Spring, "Discarte: a disjunctive Internet cartographer," in *ACM SIGCOMM Computer Communication Review*, vol. 38, no. 4. ACM, 2008, pp. 303–314.
[62] R. Siamwalla, R. Sharma, and S. Keshav, "Discovering Internet Iopology," *Unpublished manuscript*, 1998.
[63] J.-J. Pansiot, P. Mérindol, B. Donnet, and O. Bonaventure, "Extracting intra-domain topology from mrinfo probing," in *Passive and Active Measurement*. Springer, 2010, pp. 81–90.
[64] P. Mérindol, V. Van den Schrieck, B. Donnet, O. Bonaventure, and J.-J. Pansiot, "Quantifying ASes multiconnectivity using multicast information," in *Proceedings of the 9th ACM SIGCOMM conference on Internet measurement conference*. ACM, 2009, pp. 370–376.
[65] M. Faloutsos, P. Faloutsos, and C. Faloutsos, "On power-law relationships of the Internet topology," in *ACM SIGCOMM Computer Communication Review*, vol. 29, no. 4. ACM, 1999, pp. 251–262.
[66] R. Albert, H. Jeong, and A.-L. Barabási, "Error and attack tolerance of complex networks," *Nature*, vol. 406, no. 6794, pp. 378–382, 2000.
[67] W. Willinger, D. Alderson, and J. C. Doyle, *Mathematics and the internet: A source of enormous confusion and great potential*. Defense Technical Information Center, 2009.
[68] D. Barnes and B. Sakandar, *Cisco LAN switching fundamentals*. Cisco Press, 2004.

[69] G. Haviland, "Designing High-Availability Campus Networks," Cisco, 2000.
[70] S. Knight, H. X. Nguyen, N. Falkner, R. Bowden, and M. Roughan, "The Internet topology zoo," *Selected Areas in Communications, IEEE Journal on*, vol. 29, no. 9, pp. 1765–1775, 2011.
[71] R. Durairajan, S. Ghosh, X. Tang, P. Barford, and B. Eriksson, "Internet Atlas: A Geographic Database of the Internet," 2013.
[72] M. Zhang, Y. Ruan, V. S. Pai, and J. Rexford, "How DNS Misnaming Distorts Internet Topology Mapping," in *USENIX Annual Technical Conference, General Track*, 2006, pp. 369–374.
[73] J. Chabarek and P. Barford, "What's in a name?: decoding router interface names," in *Proceedings of the 5th ACM workshop on HotPlanet*. ACM, 2013, pp. 3–8.
[74] A. D. Ferguson, J. Place, and R. Fonseca, "Growth analysis of a large ISP," in *Proceedings of the 2013 conference on Internet measurement conference*. ACM, 2013, pp. 347–352.
[75] I. Poese, S. Uhlig, M. A. Kaafar, B. Donnet, and B. Gueye, "IP geolocation databases: unreliable?" *ACM SIGCOMM Computer Communication Review*, vol. 41, no. 2, pp. 53–56, 2011.
[76] S. Siwpersad, B. Gueye, and S. Uhlig, "Assessing the geographic resolution of exhaustive tabulation for geolocating internet hosts," in *Passive and active network measurement*. Springer, 2008, pp. 11–20.
[77] B. Gueye, S. Uhlig, and S. Fdida, "Investigating the imprecision of IP block-based geolocation," in *Passive and active network measurement*. Springer, 2007, pp. 237–240.
[78] Y. Shavitt and N. Zilberman, "A structural approach for PoP geo-location," in *INFOCOM IEEE Conference on Computer Communications Workshops, 2010*. IEEE, 2010, pp. 1–6.
[79] M. LLC, "GeoIP, 2010," http://www.maxmind.com, 2010.
[80] J. Hawkinson and T. Bates, "Guidelines for creation, selection, and registration of an Autonomous System (AS)," RFC 1930 (Best Current Practice), Internet Engineering Task Force, Mar. 1996, updated by RFC 6996.
[81] W. Mühlbauer, A. Feldmann, O. Maennel, M. Roughan, and S. Uhlig, "Building an AS-topology model that captures route diversity," in *ACM SIGCOMM Computer Communication Review*, vol. 36, no. 4. ACM, 2006, pp. 195–206.
[82] Internet Routing Registry, "Obtaining IRR Data," ftp://ftp.radb.net/radb/dbase, 2013.
[83] Advanced Network Technology Center, "University of Oregon Route Views Project," http://www.routeviews.org, 2013.
[84] "RIPE RIS," https://www.ripe.net/data-tools/stats/ris/routing-information-service, 2011.
[85] B. Zhang, R. Liu, D. Massey, and L. Zhang, "Collecting the Internet AS-level topology," *ACM SIGCOMM Computer Communication Review*, vol. 35, no. 1, pp. 53–61, 2005.
[86] R. Govindan and A. Reddy, "An analysis of Internet inter-domain topology and route stability," in *INFOCOM'97. Sixteenth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings IEEE*, vol. 2. IEEE, 1997, pp. 850–857.
[87] P. Mahadevan, D. Krioukov, M. Fomenkov, X. Dimitropoulos, A. Vahdat *et al.*, "The Internet AS-level topology: three data sources and one definitive metric," *ACM SIGCOMM Computer Communication Review*, vol. 36, no. 1, pp. 17–26, 2006.
[88] Z. M. Mao, R. Bush, T. G. Griffin, and M. Roughan, "BGP beacons," in *Proceedings of the 3rd ACM SIGCOMM conference on Internet measurement*. ACM, 2003, pp. 1–14.
[89] Z. M. Mao, J. Rexford, J. Wang, and R. H. Katz, "Towards an accurate AS-level traceroute tool," in *Proceedings of the 2003 conference on Applications, technologies, architectures, and protocols for computer communications*. ACM, 2003, pp. 365–378.
[90] N. RIPE, "Routing registry consistency check reports," *see http://www.ripe.net/projects/rrcc*, 2009.
[91] G. Huston, "Interconnection, peering, and settlements," in *proc. INET*, vol. 9, 1999.
[92] L. Gao, "On inferring autonomous system relationships in the Internet," *IEEE/ACM Transactions on Networking (ToN)*, vol. 9, no. 6, pp. 733–745, 2001.
[93] G. Di Battista, M. Patrignani, and M. Pizzonia, "Computing the types of the relationships between autonomous systems," in *INFOCOM 2003. Twenty-Second Annual Joint Conference of the IEEE Computer and Communications. IEEE Societies*, vol. 1. IEEE, 2003, pp. 156–165.
[94] J. Xia and L. Gao, "On the evaluation of AS relationship inferences [Internet reachability/traffic flow applications]," in *Global Telecommunications Conference, 2004. GLOBECOM'04. IEEE*, vol. 3. IEEE, 2004, pp. 1373–1377.
[95] R. Oliveira, W. Willinger, B. Zhang *et al.*, "Quantifying the completeness of the observed Internet AS-level structure," 2008.
[96] R. V. Oliveira, D. Pei, W. Willinger, B. Zhang, and L. Zhang, "In search of the elusive ground truth: the internet's as-level connectivity structure," in *ACM SIGMETRICS Performance Evaluation Review*, vol. 36, no. 1. ACM, 2008, pp. 217–228.
[97] PeeringDB, "Exchange Points List," https://www.peeringdb.com/private/participant_list.php, 2013.
[98] A. H. Rasti, N. Magharei, R. Rejaie, and W. Willinger, "Eyeball ASes: from geography to connectivity," in *Proceedings of the 10th ACM SIGCOMM conference on Internet measurement*. ACM, 2010, pp. 192–198.
[99] S. Zhou and R. J. Mondragón, "Accurately modeling the Internet topology," *Physical Review E*, vol. 70, no. 6, p. 066108, 2004.
[100] R. V. Oliveira, B. Zhang, and L. Zhang, "Observing the evolution of Internet AS topology," *ACM SIGCOMM Computer Communication Review*, vol. 37, no. 4, pp. 313–324, 2007.
[101] A. Dhamdhere and C. Dovrolis, "Ten years in the evolution of the Internet ecosystem," in *Proceedings of the 8th ACM SIGCOMM conference on Internet measurement*. ACM, 2008, pp. 183–196.
[102] A. Dhamdhere and C. Dovrolis, "Twelve years in the evolution of the Internet ecosystem," *IEEE/ACM Transactions on Networking (ToN)*, vol. 19, no. 5, pp. 1420–1433, 2011.
[103] H. Chang, S. Jamin, and W. Willinger, "Internet

connectivity at the AS-level: an optimization-driven modeling approach," in *Proceedings of the ACM SIGCOMM workshop on Models, methods and tools for reproducible network research*. ACM, 2003, pp. 33–46.

[104] H. Chang, S. Jamin, and W. Willinger, "To peer or not to peer: Modeling the evolution of the Internet's AS-level topology," *Ann Arbor*, vol. 1001, pp. 48 109–2122, 2006.

[105] A. Lodhi, A. Dhamdhere, and C. Dovrolis, "Genesis: An agent-based model of interdomain network formation, traffic flow and economics," in *INFOCOM, 2012 Proceedings IEEE*. IEEE, 2012, pp. 1197–1205.

[106] W. E. Leland, M. S. Taqqu, W. Willinger, and D. V. Wilson, "On the self-similar nature of Ethernet traffic," in *ACM SIGCOMM Computer Communication Review*, vol. 23, no. 4. ACM, 1993, pp. 183–193.

[107] V. Paxson and S. Floyd, "Wide area traffic: the failure of Poisson modeling," *IEEE/ACM Transactions on Networking (ToN)*, vol. 3, no. 3, pp. 226–244, 1995.

[108] N. Chatzis, G. Smaragdakis, A. Feldmann, and W. Willinger, "There is more to IXPs than meets the eye," *ACM SIGCOMM Computer Communication Review*, vol. 43, no. 5, pp. 19–28, 2013.

[109] N. Chatzis, G. Smaragdakis, J. Böttger, T. Krenc, and A. Feldmann, "On the benefits of using a large IXP as an Internet vantage point," in *Proceedings of the 2013 conference on Internet measurement conference*. ACM, 2013, pp. 333–346.

[110] B. Krishnamurthy, W. Willinger, P. Gill, and M. Arlitt, "A Socratic method for validation of measurement-based networking research," *Computer Communications*, vol. 34, no. 1, pp. 43–53, 2011.

[111] V. Giotsas, S. Zhou, M. Luckie *et al.*, "Inferring multilateral peering," in *Proceedings of the ninth ACM conference on Emerging networking experiments and technologies*. ACM, 2013, pp. 247–258.

[112] A. Lodhi, A. Dhamdhere, and C. Dovrolis, "Analysis of peering strategy adoption by transit providers in the Internet," in *INFOCOM Workshops*, 2012, p. 177.