Internet Topology Discovery

Reza Motamedi
University of Oregon
motamedi@cs.uoregon.edu

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 to the vulnerability analysis of Internet infrastructure. The scale of the Internet topology coupled with its distributed and heterogeneous nature makes it very challenging to capture a complete and accurate snapshot of the topology.

In this report, we survey some of the main research studies on the discovery and characterization of the Internet topology during the past 15 years. Toward this end, we classify prior studies based on the "resolution" of the topology that they have considered as follows: interface-level, router-level, PoP-level and AS-level. For studies related to each resolution, we examine techniques and tools for data collection along with their limitations and summarize their key findings. We also discuss modeling efforts and geographic characteristics for studies at certain resolution. Our structured examination of prior research on Internet topology also reveals some exciting research problems in this area that deserve further investigation.

1. INTRODUCTION

Composed of approximately 45,000 networks, the Internet reigns as the ultimate network of networks, each separately owned and managed. These networks, which are referred to as Autonomous Systems (AS), have different coverage, resources, and purposes. For example an AS can be either a Network Service Provider (NSP), an Internet Service Provider (ISP), an education network, or a Content Distribution Network (CDN). The diversity in network type and mission along with their autonomous management indicates that individual ASes are likely to have a different topology, deploy a different intra-domain routing protocol with its own policies, and use devices from different vendors.

Capturing an accurate view of the Internet topology is essential to the network research community as it has many uses including the following areas: *(i)* The topological properties of the Internet affect the performance of network protocols, network applications and services. Having a clear understanding of the Internet topology and its main characteristics enables network researchers to properly design and evaluate network protocols, *(ii)* An accurate map of the Internet is extremely useful for allocating resources (*e.g.*, proxies, replica servers, and data centers), *(iii)* A correct map of Internet topology with certain attributes can inform a wide range of security-related problems and protocols such as backtracking malicious traffic or assessing the vulnerability of the Internet to attacks or blackouts.

Mapping the Internet topology is inherently challenging due to the following reasons: First, the scale of the Internet coupled with its distributed and heterogeneous nature makes it difficult to capture a complete and correct snapshot of the topology. Second, there is no protocol or service whose sole purpose is the discovery of network topology[1, 2]. The measurement tools and data sources that are most often used for topology discovery are merely hacks that researchers proposed to collect information about the Internet topology. In particular, the two most commonly used sources of data for topology discovery, namely `traceroute` measurements and BGP information, have entirely different purposes. More specifically, `Traceroute` is a network debugging tool [3, 4, 5] and BGP is the inter-AS routing protocols that indicates reachability for individual ASes [6]. Despite these challenges, a large body of research has focused on capturing and characterizing the Internet topology.

This report presents a structured survey of some of the main studies on measuring and modeling Internet topology during the last 15 years. Due to its complexity, Internet topology can be viewed at different resolutions, namely interface-level, router-level, Point-of-Presence (PoP) level or AS-level. We classify these studies mainly based on their target resolution of the topology. For each resolution, we further discuss data types, data collection techniques and tools, and topology inference techniques along with their limitations. We present geographic characteristics and proposed topology models that have been presented at certain resolutions. Finally, we summarize the main findings of prior studies.

The rest of this report is organized as follows: Section 2 presents the notion of topology resolutions which motivates our taxonomy. Sections 3, 4, 5, and 6 cover Internet topology at interface-level, router-level, PoP-level, and AS-level, respectively. At each resolution, we introduce the data and techniques used to infer the topology at that level. Finally, we conclude the document in Section 7.

2. TAXONOMY

The Internet's topology is often presented as a graph. However, the term “Internet graph" is used to refer to different structures by different communities. This ranges from the graph structure of the World Wide Web (WWW) and overlay networks to the Internet's infrastructure topology. The focus of this document is the latter, where nodes represent network entities and links represent relations between entities. Even with this definition in place, Internet topology graph could have different meanings to different interested parties.

From the network connectivity stand point, we use the following organization to taxonomize the prior studies. At the very high level, the *resolutions* of the captured Internet topology graph is used to categorize these studies [1, 7, 8, 9]. At each resolution we address two issues: (*i*) The overall classification of data and the techniques employed to collect data in order to discover the topology at that specific resolution. (*ii*) Geographical characteristics of the discovered topology and the extent to which the topology at the target resolution is annotated with geographic attributes.

Internet Graph at Different Resolutions: The Internet topology can be viewed at four different levels. These resolutions are organized as follows from finest to coarsest level.

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 1-1 mapping between nodes and IPs [10, 11]. On the other hand, a link between two nodes shows direct network layer connectivity between the two nodes. This implies that topology at this level ignores devices functioning at OSI layers lower than the network layer (*e.g.*, hubs and switches).

II) Router level: Topology at this level is often the result of grouping interfaces that belong to the same router [12]. At this level, a node represents a network device *e.g.*, a host or a router with multiple interfaces. Two nodes are connected with 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 an AS [13, 14]. ASes commonly impose hierarchical principles through PoP structures. In this context an AS is built from a collection of PoPs [15]. A PoP is used by the AS to provide interconnectivity to PoPs of other ASes or the PoPs of the same AS. In this sense, a node in the PoP level topology represents a PoP that belongs to one AS and a link between two PoPs represents physical connectivity among routers of the two PoPs.

IV) AS level: As opposed to previous views, the AS level topology graph represents a more logical view of the Internet [16, 17]. A node in 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. This business relationship leads to the transfer of data traffic based on a financial agreement [18]. These agreements are certainly the bread and butter of the Internet, since the Internet as a whole is built on the concept of cooperation among networks. These networks without cooperation will downgrade to separate networks without global reachability. Traditionally, these agreements are categorized into three types as follows: *a*) customer-provider (C2P), *b*) peer-peer (P2P), and *c*) sibling relation or peering. As ASes cover an area and often times own multiple PoPs, the actual connectivity between two ASes might happen at multiple locations. Thus, the logical AS relation is an abstraction with multiple physical connectivities between the two ASes [19].

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

Data Types and Data Collection: The nature of the data and the type of data collection techniques is another element that we used to classify prior studies. Regarding its nature, data can be collected from the *control plane* or the *data plane*. In the measurement of the control plane, the collected data reveals information about the routing in the Internet. For instance, `BGP` tables store the AS path to reach different prefixes and they are classic examples of the control plane data. Data plane measurements aim to discover the actual path that packets travel. The simplest measurement of the data plane is `Ping`. It measures the reachability and the Round Trip Time (RTT) delay of a target IP from a source, based on the route that packets take in the Internet. Regarding the collection technique, a measurement can be either *active* or *passive*. In active measurements, probe messages are sent into the network, then successive 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 respectively. The list of common data sources and techniques used for discovering Internet topology at each resolution is summarized in Table 1.

Geographic Attributes of The Topology: Although the main element of a topology is connectivity, *geography* is another element that can be added to the topology to increase its usability. However, the definition of a geographically annotated topology varies for different Internet topology resolutions. Interfaces, routers and PoPs are entities that can be geographically pin-pointed to a location on a map. A geographical Internet map in these three resolutions involves assigning a pair of longitude and latitude to each entity. Therefore, the topology graph consists of points on the map and links that connect those points together. In the case of ASes however, geography translates to the scope of AS. In this case, an AS as a node is shown with a colored area on a map that represents its coverage. The AS relations are represented by connecting the corresponding nodes on the

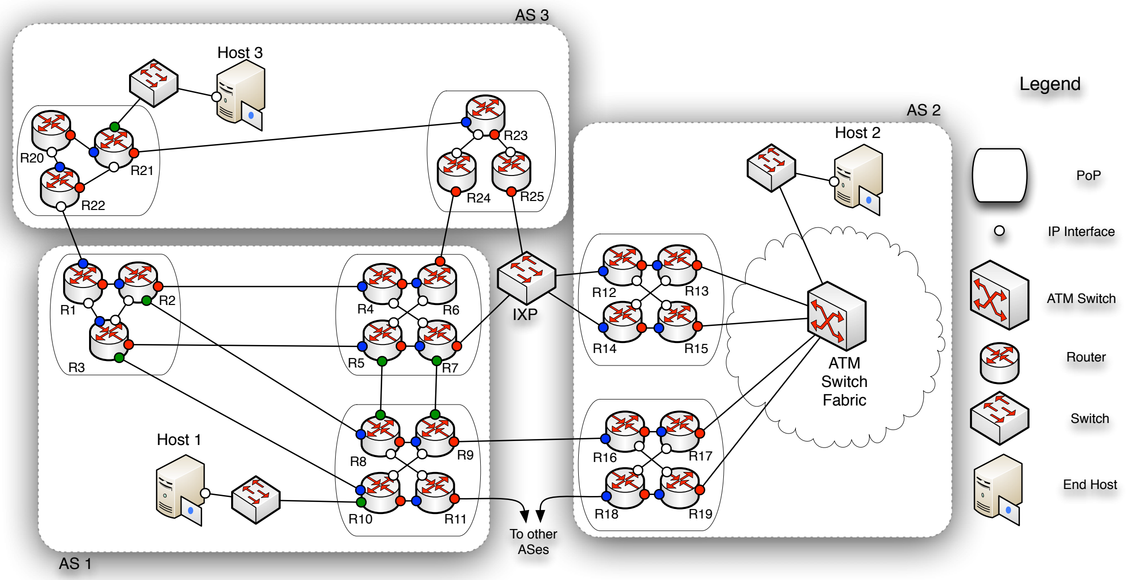


Figure 1: A detailed toy topology representing the Internet topology at different granularities

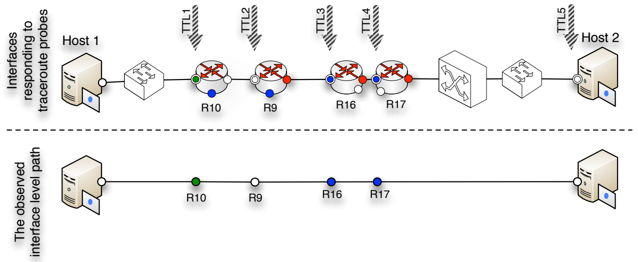


Figure 2: `Traceroute` from $Host1$ towards $Host2$ and the corresponding interface-level path.

graph. The representation can be incorporated with more detail if multiple AS connections between two ASes are represented as individual edges. A link can connect multiple ASes, which is common in Internet Exchange Points (IXP). At an IXP, multiple networks are linked together at one physical location through a mutlipoint connection. As a result the complete geo annotated AS level topology is a hyper-graph [1] where nodes cover areas and links are annotated by the locations of both its ends.

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 multiple interfaces, each router appears as multiple nodes, while normal end-hosts with one interface are presented with one node. The topology is typically simplified by ignoring end-hosts, therefore nodes only represent routers at this simplified abstraction. 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 direct one.

`Traceroute` is the most widely used tool to map the topology of the Internet at this resolution. Based on the nature of data and the collection type, it is an active measurement of the data plane [6, 20]. It uses limited Time To Live (TTL) probes. The `traceroute` from a source to a target successively discovers the IP address of one interface per router along the forward path and at each hop it reports the RRT delay as well. Multiple probe messages with same TTL 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 measure a different delay due to the dynamic network traffic. In the rest of this report we assume that a single probe message is used for each hop discovery. Figure 2 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. In order to discover topology at the interface-level, the outcome of many `traceroutes` 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, vantage points conduct `traceroutes` towards the set of targets. The overall observed topology is

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
	Subnet discovery	Router response inconsistency
Router-level	Alias Resolution	Scalability
		Inaccurate
	`SNMP`	Only applicable to one AS
	`MRINFO`	Only applicable to Ases with multicast-ready routers
PoP-level	Aggregation techniques	DNS name to Geo is not always applicable
		DNS misnaming adds error
		IP to Geo, inaccurate
	Delay based techniques	Sensitive to placement of candidate PoPs
	Online data sources	Obsolete data
AS level	BGP	Reachability announcement protocol with built in information hiding
	`Traceroute`	IP to AS number, not trivial
	Internet Routing Registries	Obsolete data

generated from the union of all the IP paths, each measured by a `traceroute`.

In the subsequent section, we describe `traceroute` as the most common active measurement tool, then discuss its limitations. We then provide an overview of some of the measurement-based studies that use active measurement at interface-level for Internet topology mapping. Finally we cover more recent proposals for collecting interface-level data.

3.1 `Traceroute`

`Traceroute` involves actively sending probes into the network, rather than merely monitoring it. It is the most widely used tool to actively capture the topology of the Internet. Jacobson's `traceroutes` – the first implementation of this tool – uses ICMP packets as probes [21]. However, other versions of `traceroute` use other types of probe messages, for instance UDP and TCP [3].

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

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. ICMP `traceroute` is the default setting for Microsoft Windows.

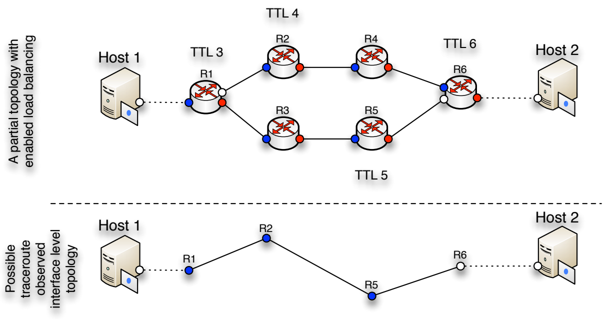

A partial topology with enabled load balancing
TTL 4
Host 1
TTL 3
R1
R2
R4
TTL 6
R6
Host 2
R3
R5
TTL 5
Possible traceroute observed interface level topology
Host 1
R1
R2
R5
R6
Host 2

Figure 3: False links inferred by `traceroute` in the presence of load balanced routers

The main limitation of the UDP and the ICMP is that both UDP messages to high ports and ICMP messages are prone to be filtered by firewalls [22]. To bypass firewalls, TCP `traceroute` 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, especially at the edge of the network.

The comparison of the results of the UDP, ICMP and TCP `traceroute` in topology discovery shows that the ICMP `traceroute` reaches targets more successfully. However, the UDP `traceroute` identifies more IP links, but it is least successful in reaching the targeted IP [23].

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 the inter-domain routing in OSPF [24] and IS-IS [25]. 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 through the same path. Routers use IP headers to identify flows. These header fields include: Source Address, Destination Address, Protocol, Source Port, Destination Port, the IP Type of Service (TOS), the ICMP Code and the Checksum fields. Note that in the traditional `traceroute` the value in some of these fields vary among different probes in order to match the probe and the response. Hence, per flow load balancing may result in the routing of probes of the same `traceroute` measurement through different paths. When measuring a load balanced route, the `traceroute` infers the existence of a false IP segment that does not exist in the topology. Figure 3 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 taken at *R1*. In our example, for TTL 3, 4, and 5 the visited routers are *R1*, *R2*, and *R3*, respectively. As the result, a false link between *R2* and *R5* is inferred. Paris `traceroute` [5] aims 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 on `traceroute` take the same path. Paris `traceroute` resolves the flow based load balancing anomalies in the observed route, but anomalies due to the per packet load balancing are unresolvable.

3.1.1 Limitations & Issues

`Traceroute` is a reachability diagnosis tool and its use for the interface topology discovery is not perfect. Generally, the limitation and issues of `traceroute`-based interface measurements can be categorized into two types: (*i*) The limitations that stem from the nature of measurement method. (*ii*) The issues that arise due to deploying large scale distributed measurement infrastructures. In this section, we summarize the most important limitations and issues when using `traceroute` to actively measure interface-level topology.

Measurement Limitations: First, there is no unique setting a router's response to a TTL zero probe. The router configuration determines its response. Network administrators typically choose one of the five following policies when configuring the router responses. (*i*) *Null interface routers* remain reticent to the probes. For these routers, traceroute detects their existence, but not their address (Anonymous Routers) [26]. In this case, the RTT is not reported as well. (*ii*) *Probed interface routers* respond with the IP address of the probed interface. This configuration is most common when the router is directly probed. (*iii*) *Incoming interface routers* respond with the IP address of the interface from which the probe message was received by the router. This configuration is reported to be the most common setting when the router is probed with indirect TTL-limited messages [27]. (*iv*) *Shortest-path interface routers* respond with the IP address of the interface that is closest to the source. It is worth noting that Internet asymmetry means that incoming interface and shortest-path interface are not necessarily the same. (*v*) *Default interface routers* respond with a designated IP address indifferent to the probed interface. In addition to these router configuration settings, firewalls can also be configured to prevent probed routers from responding. In summary, `traceroute` suggests the existence of one interface per router in the foreword path at best.

Second, the IP address reported 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)configurations sometimes allude to the appearance of private non-routable addresses and carrier-grade NAT (large scale NAT) addresses. These IP addresses can be used by multiple ASes, that could lead to path loops and other anomalies. In addition, these IPs can not be mapped to a single router or an AS and can not be used to pin point the location of the interface due to the one to many relation of the IP and the assigned interfaces.

Third, the RTT delay reported at each hop can not be used to accurately measure the delay to and from the target. `Traceroute` is a foreword route diagnostic tool. A rule of thumb in Internet routing is that routes are not always symmetric. Hence, the path taken by the probe may differ from the path taken by the response. As a matter of fact, the variation in the delay at two consecutive hops could be due to variable queuing delays or the existence of a different backward route.

Forth, layer 2 clouds are opaque to a `traceroute`. These clouds have the explicit purpose of hiding the network infrastructure from the IP layer. 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 topology of AS2 in Figure 1, routers directly connected to the ATM cloud have a mesh like interconnectivity. Multi-Protocol Label Switching (MPLS) is another common layer 2 technology used to manually configure tunnels passing through multiple routers. It has been reported that at least 30% of the paths tested traverse an MPLS tunnel [28, 29]. Routers using MPLS may be configured either to decrement the TTL (MPLS opaque option), as `traceroute` requires, or to ignore the TTL field completely. Typically, the switched path of MPLS is manually configured with the opaque option [29]. Although it might be possible to detect the MPLS tunnels from `traceroute` measurements [28, 29], the inference methods are not guaranteed to be perfect and are very specific to MPLS tunnels.

Large Scale Measurements Issues: First, the distribution of vantage points and targets limits the observable interface topology. The probability of sampling an IP segment is correlated with the placement of the vantage points and the type of IP segment. For instance, back-up inter AS routes are hard to discover. Similarly, IP segments corresponding to inter AS peering relations are among the least discoverable ones [6]. To deal with this bias, two approaches have been proposed. 1) Eriksson *et al.* [30, 31] suggested a mechanism to infer the unseen components of the Internet. Their solution is to map this problem to the statistical 'unseen species problem'. First they estimate the number of unseen components using incomplete observations. Matrix completion techniques are used later 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. 2) Targeted probing is used 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 [6]. For instance, Augustin *et al.* [32] use targeted `traceroute` to discover peering links at Internet Exchange Points (IXPs) where ASes are more likely to peer.

Second, orchestrating a large measurement campaign imposes a high load on the network and the measurement infrastructure. 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 detected as a Denial of Service (DoS) attack by Intrusion Detection Systems (IDS). The redundant measurements are classified into two distinct types [33]. "Intra-monitor redundancy" occurs close to one 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 overhead reduction techniques were proposed to address this issue in the literature. "Far probes" [33] are proposed to address the Intra-monitor redundancy. In this case, when the topology close to the vantage point is fully discovered, a higher TTL value is used instead of using `traceroute` with probes starting with TTL 1. "Stop set" (collaborative probing) [33, 34] 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 follow the same path toward t due to the 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.

3.1.2 Coverage & Completeness

Early studies suggested the utilization of a few vantage points and a large set of targets that were well distributed across the targeted network. The claim was that the gain from adding vantage points increases marginally by adding more vantage points [35]. However, later studies showed that despite the diminishing return of extra vantage points, the observed topology is more complete [36].

In order to produce a more complete picture of the topology, researchers have both increased the number of vantage points and targets [37, 38], and the duration of the measurement [20]. While the former increases the scope of the captured topology, the latter provides a dynamic view of the topology and reveals a more complete one. It is widely accepted that longer measurements observe a more complete view of the topology, since measurement probes may take the rarely used back-up routes.

The large scope of these measurement campaigns could impose a high load on the data plane. Additionally, these measurements may raise red flags in intrusion detection systems [39]. Beverly *et al.* [20] used high frequency measurement with adaptive probing techniques to limit the imposed measurement load, while keeping the discovery rate high. In each cycle, the "interface set cover" algorithm minimizes the `traceroute` load while maintaining a high discovery rate. In order to maximize the gain from each `traceroute`, "subnet centric probing" selects targets to reveal maximum information from the inside of a network.

3.1.3 Measurement Infrastructure

In the past decade, the Internet research community has benefited from many `traceroute`-based Internet topology studies. These studies have either used dedicated instrumentation boxes (*e.g.*, Skitter [40] and Archipelago [41]) and PlanetLab (*e.g.*, iPlane [11], RocketFuel [12] and [34]) or deployed a platform of software agents to collect `traceroute` from a larger number of vantage points. In the latter case the platform's incentive model can be classified into two models: 1) Altruistic Model (*e.g.*, Dimes [10]) where the participation in the platform is just for the good of science. 2) Win-win model (*e.g.*, Ono [37] and Dasu [38]) where the measurement conducted by the software agent is both beneficial to user and the experimenter.

Although using dedicated boxes and PlanetLab are still very common approaches in conducting active measurements, the better coverage of software agent platforms has resulted in the deployment of more *crowd-sourcing measurement campaigns* in the past few years (*e.g.*, Scriptroute [42], Dimes [10], and Bitprobe [43]). However, the large scale of these campaigns often requires extra care in its instrumentation [33] as discussed in section 3.1.1.

The use of public `traceroute`, servers also known as looking glasses, to conduct active measurements has also gained a lot of attention, due to the large coverage in term of the placement of vantage points. However, due to their public nature, these `traceroute` servers impose limits on the rate of the measurements. Therefore their usage is mostly for small scale measurements and validation (*e.g.*, RETRO [44] and [32])

3.2 Other Approaches

Although `traceroute` is the most commonly proposed approach for obtaining interface-level topology, its limitations expedited the proposal of other approaches to collect additional connectivity information. While `traceroutes` with different types of probe messages mainly attempt to penetrate through firewall filters, other active measurement techniques are used to address its other limitations. Due to the diverse nature of these techniques and their limited usages, they are all covered in this section.

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. In order to get 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 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. SR allows the sender to specify the router that packets should go through before reaching the destination. The intermediate router 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 source-routed capable routers is a small fraction of all routers in the Internet (around 8%), Govindan *et al.* [27] show that this number of source route capable routers 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 Internet topology is not random.

The asymmetric nature of Internet routing implies that the discovered routes are only foreword routes from the vantage points to the targets. Reverse `traceroute` [45] uses the "Record Route" (RR) option and "IP Timestamp" to detect the interfaces on the reverse 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.

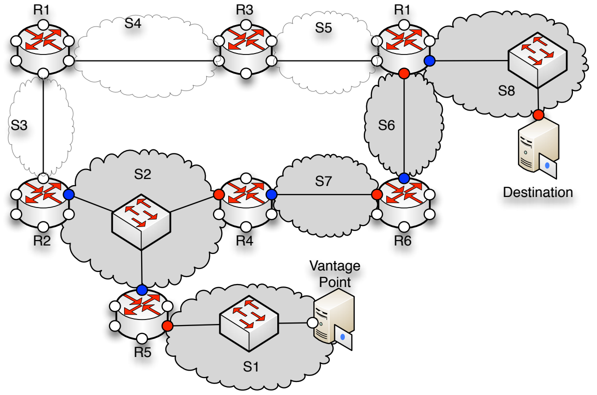


Figure 4: An example topology and corresponding subnets represented by clouds. Subnets identified by `tracenet` are marked grey.

While using IP option to provide information that is not available using simple `traceroute`, it increases the chance of processing delay, discard, or alarm at intrusion detection systems.

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 the 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 4 shows the topological structure of a sample network. Corresponding subnets are depicted as clouds.

Subnet level discovery tools such as `XNET` [46] 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 the most one hop distance difference from a vantage point. The problem is that the size of the subnet is unknown. Given IP address t that is n hop 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.

Tracenet [47] 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 4 shows the application of tracenet on a sample topology and identified subnets are greyed. 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, XNETT discovers a bogus subnet on the path. For instance, in Figure 4, if *R1* responds with its green interface, *S5* is discovered instead of *S6* as the fourth subnet on the route.

4. ROUTER-LEVEL

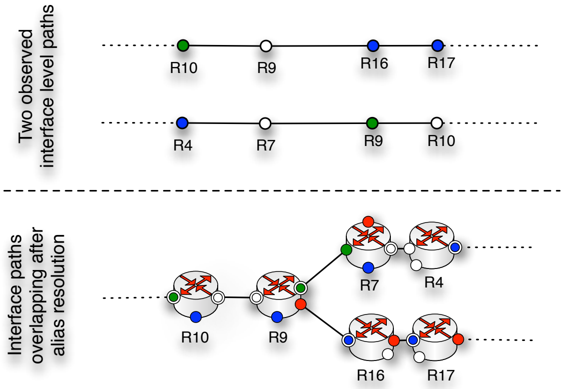
The router-level topology shows the routers and the interconnectivity among their interfaces in the Internet. In this topology, nodes represent end-hosts (with one interface) or routers (with multiple interfaces) and links show layer 3 connectivity between these devices. The topology at this level can be viewed as the outcome of the aggregation of IP interfaces that belong to a single router. When applied to the interface level topology, this aggregation results in the router-level topology. The main techniques to collect router-level connectivity are as follows:

- **Alias resolution:** This approach is the aggregation of traceroute data. The main challenge is to relate different interfaces of a router that were discovered in different traceroute measurements. Alias resolution [27, 48] or router disambiguation [35] is a set of techniques that identify the IP interfaces that belong to the same router.

- **Recursive router discovery:** Another class of techniques rely on a router's capability to be queried for its neighbor on each interface. SNMP and IGMP are two protocols that can be used to discover neighboring routers of a queried router in the scope of 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. As the result of this process true router-level topology is revealed from interface level topology. Figure 5 shows two partial interface paths observed from traceroute measurements in the topology of Figure 1, one from *Host1* to *Host2* and another from *Host3* to *Host1*. The measurements do not have any IP hop in common. However, resolving alias IPs show 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 a false negative case, alias resolution falls short in relating two alias interfaces. Following, we list the most widely used alias resolution methods.



Two observed interface level paths
R10
R9
R16
R17
R4
R7
R9
R10
Interface paths overlapping after alias resolution
R7
R4
R10
R9
R16
R17

Figure 5: Two partial traceroutes with no common hops. Resolving IP aliases shows that the paths overlap.

Common Source Address: This technique was proposed by Pansiot *et al.* [49] and was implemented in Meractor [50]. 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 3.1.1.

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.

Ally's implementation in Rocketfuel [12] uses 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, the addresses are likely to be aliases [12].

Alias resolution based on the ID fingerprint is prone to false negatives due to larger than one ID increment settings on routers. False positives can also occur due to randomly synchronized ID counters of two routers. However it can be mitigated by running more tests after a wait period. The other major drawback of this technique is the overhead of running it on a large set of discovered interfaces, since its complexity is $O(n^2)$. Some heuristics are proposed to improve the efficiency of Ally by restricting the possible alias candidates using delays and TTL [12]. The idea is that alias candidates should have similar TTL from 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 [51] 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 on a large list of IPs. Since routers use a 16-bit counter for the packet ID, counter wraparrounds can occur during measurement. 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 wraparrounds are likely to occur. Although the designers of RadarGun had accounted for the possibility of a single wraparround, the accuracy of the technique diminishes in the presence of multiple wraps.

DNS-Name: The similarities in DNS names associated with router interfaces can also be used to infer aliases [12, 48]. The main limitations of this approach are as follows: i) This technique only works when an AS uses a clear naming convention for assigning DNS names to router interfaces. ii) The complexity of the naming conventions may require human intervention to resolve aliases which limits its scalability. iii) The technique is not very accurate at the AS borders. The interfaces of border routers usually belong to different ASes with different naming conventions, which in turn complicates the alias resolution at the AS borders [27].

Graph-Based Resolution Heuristics: `Traceroute` measurement can offer heuristics on alias inference [48]. Graph-based alias resolution constructs a directed graph by overlaying an individual `traceroute` measurement as demonstrated 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 addressees prior to the common IP should belong to different interfaces of a single router and hence are aliases. Figure 6 shows a partial view of the `tracetoutes` from $Host1$ and $Host2$ toward $Host3$ in our toy example. In this example black interface succeeds the red interface in one `traceroute`, and succeeds the blue interface in another `tracetoute`. The heuristic suggests the blue and the red interfaces are aliases.

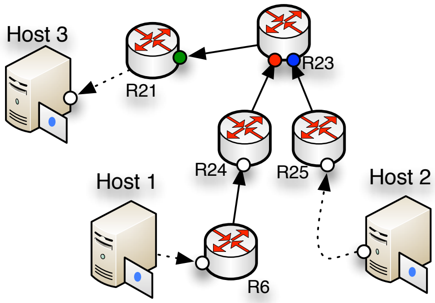


Figure 6: **Graph based alias resolution; The green interface succeeds the blue and the red interface in two `traceroutes` so red & blue are aliases.**

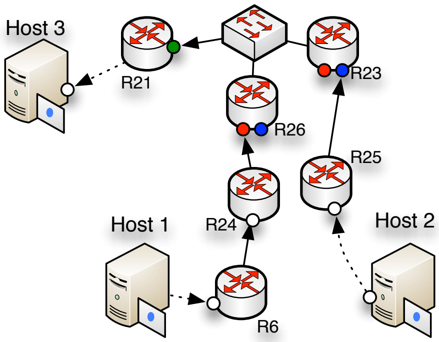


Figure 7: **False positive in graph based alias resolution due to the presence of a layer 2 switch; The green interface succeeds the blue and the red interface in two `traceroutes` so red & blue are inferred to be aliases.**

This heuristic falsely infers aliases in the presence of layer 2 switches or multiple-access clouds. Figure 7 depicts an alternate topology to Figure 6. 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` measured paths, Analytical Alias Resolver (AAR) [52] utilizes 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 link-

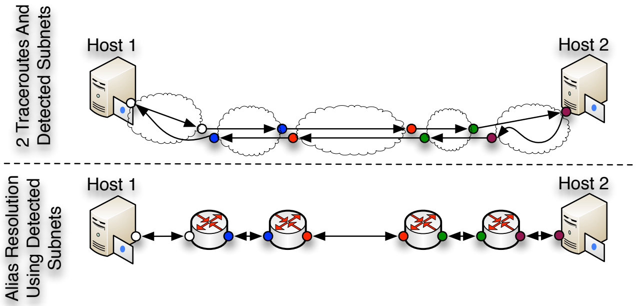

2 Traceroutes And Detected Subnets
Host 1
Host 2
Alias Resolution Using Detected Subnets
Host 1
Host 2

Figure 8: **Analytical Alias Resolution for detecting IP aliases on a symmetric path segment.**

ing 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 `Traceroutes` between *Host1* and *Host2* are shown in Figure 8. The top view shows the two `traceroutes` and the identified subnets. The bottom view depicts how the detected subnets can be used to align the two `traceroutes` and resolve aliases.

Analytic and Probe-based Alias Resolver (APAR) [53] consists of analytical and probe-based components. 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 mitigate 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 [54] 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 foreword 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 to traces between 379 sources and 376,408 destinations is reported to be intractable.

4.1.1 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 a router [55]. This technique can recursively be used to discover new routers and the connectivity between them. Although accurate, the usage of this technique is limited within an AS and can only be used by the network administrators with adequate privileges.

More recently `MRINFO` has been used to discover topology at the router-level using `IGMP` messages with a similar incremental method [56, 57]. 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 visibility of this technique is also limited to multicast-enabled routers.

4.2 Modeling

The most cited work on Internet topology modeling is by Faloutsos *et al.* [58]. In their paper, they studied the `traceroute` data collected by Pansiot *et al.* [49] in mid-1995, which showed the actual router-level paths taken by packets in the Internet and the observed router topology. One of their main observations was the scale free structure of the network and the power-law degree distribution of routers. This indicates the existence of a small number of high-degree core routers and a large number of lower degree edge routers. This paper fueled many following modeling studies on router-level topology (eg. [59]) that aimed to simulate the observed scale free structure of Internet topology as a given fact.

Although their observations seem plausible, many domain experts argued that they are indeed erroneous [60]. First, no publicly available Internet topology exhibits the scale-free graph topology. For example, in the public maps of Internet2, there is no evidence of a few highly connected central routers. Second, technology constraints do not allow the formation of the power-law degree distribution. When configuring a router, network administrators are limited by the trade off between traffic vs. degree. In particular, a central router that processes a large volume of traffic on each interface can not 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 a degree distribution opposite to the observed power-law. Third, there is a clear mismatch between the observed scale free topology and the design philosophy of the Internet. An important requirement of original DARPA net design was that "Internet communication must continue despite loss of networks or gateways" [15]. However, In a scale free topology, a failed high degree central router can lead to partitioning of the network as shown by Albert *et al.* [59], a property that became well-known as the Internet's "Achilles' heel". Lastly, it has been shown that the errors in the observed router-level topology can be explained by the following limitations of the measurement tools. *i*) The router degree is directly correlated with the accuracy of alias resolution methods. Since there is no perfect solution to this problem, the router degree is an unreliable property of any inferred router-level topology. *ii*) The inferred high-degree nodes can also be an artifact of `traceroute` inability to observe opaque layer-2 clouds. The observed topology of a group routers at the edge of a layer-2 cloud is a mesh-like (full graph) interconnection among all routers.

Alternatively, Heuristically Optimal Topology (HOT) models are proposed to model the Internet topology by reverse-engineering. These models rely highly on domain knowledge as the alternative resource instead of `traceroute` measurement models. HOT models have three main elements as follows: *i*) The objective of the ISP or the type of ISP, *ii*) the ISP trade-off between cost and efficiency which affects the router topology design. *iii*) the uncertainty in the environment such as the ISP's traffic demands and the traffic matrix. When combining all these ingredients, *constraint optimization* can be used to construct an optimal topology for a given AS's objective and demands. The constriction of the optimal solution can be NP-hard. However, HOT models are not concerned with the construction of the "best" topology, but a "good" performance derived from a heuristically optimal solution is sufficient [15]. The optimization process results in topology that is constant with the constraint. This topology for a single AS has a pronounced backbone, which is fed by tree-like access networks, with additional links added for redundancy and resilience.

5. POP LEVEL

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. Therefore, PoPs are the reflection of hierarchical design in an AS which results in scalability and maintainability. Network designers often apply "cookie cutter" methods to design PoPs. This results in the appearance of PoPs as repeated patterns in the AS network. This modular design strategy simplifies network debugging and management. Figure 9(a) depicts an example cookie cutter design applied in designing a PoP. The design provides connectivity with additional redundancy between customers of the PoP and the rest of the Internet. A node in the PoP-level topology is the PoP of one AS ideally tagged with the PoP's owner and geographical information of the PoP. Inter PoP Links can be categorized into two types. (*i*) *Core-links* or *backbones* connect two PoPs of the same AS. (*ii*) *Peering* links connect PoPs of different ASes. Figure 9(b) shows the PoP-level topology corresponding to the network of 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. Backbone links are represented by lines and dotted lines show in peering links.

PoP-level topology is the ideal resolution to study the connectivity and redundancies of an AS. The topology at this level is also very useful for potential customers since it provides information about the geographical coverage of the AS.

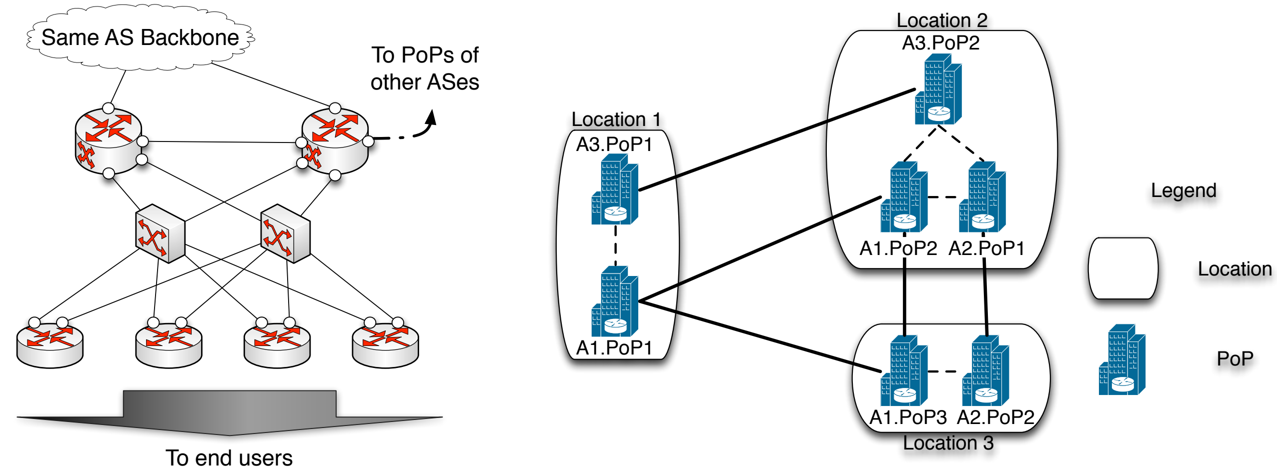
Three approaches have been followed to map PoP-level topology of the Internet. First, the most common approach is to identify PoPs by aggregating data collected from active measurements. This method receives either an interface level topology or a router level topology as its input and groups nodes that belong to one PoP. The related studies are covered in section 5.1.

The second approach per hop information from `traceroute` is replaced by estimation of delay from `ping`. Yoshida *et al.* [14] used this technique to detect the PoPs of four major ISPs in Japan. They argue that an ISP's core network information such as routers and DNS names that are obtained through `traceroute` are unreliable. Hence, they used end-to-end delay measurements, using their infrastructure deployed in all major cities in Japan. Their model relates the end-to-end delay to the sum of the delays between consecutive traversed PoPs. Using many end-to-end measurements, they detected the PoPs that a probe should pass since the total delay should be equal to the sum of the delay between traversed PoPs.

The last approach relies on the resources that are published by ISPs on their websites. Figure 10 shows one example of these maps for Cogent Communications. The map depicts PoP cities and the interconnection among PoPs of the same AS. Topology Zoo [61] is a collection of about 200 topology maps taken from online pages for ISPs. 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 to the out-dated data. These maps only show the connectivity of one AS and do not reveal AS peerings. In the following section, we cover prior studies in the context of 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.

5.1 Aggregation Methods

The first study that focused on the discovery of PoPs was Rocketfuel [12]. It tried to measure the structure of an AS using `traceroute` measurement and used 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



(a) Cookie cutter design used in the PoP of an AS

(b) The PoP-level topology our example

Figure 9: PoP level topology.

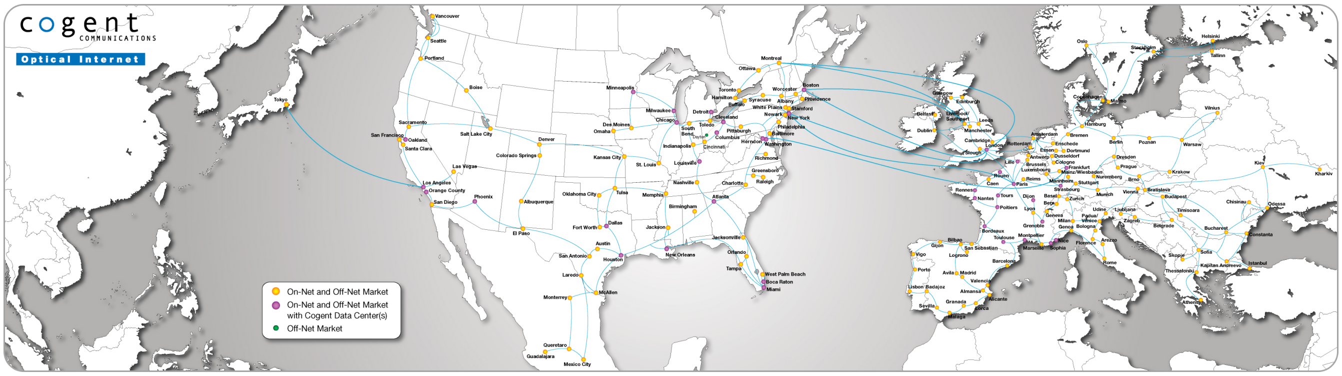


Figure 10: The PoP-level topology of Cogent available online at http://www.cogentco.com/en/network/network-map.

infer the geographical location of an interface. Interfaces in one geographical location are grouped as a PoP.

iPlane [11] extends Rocketfuel. First, a Meractor-like [50] alias resolution is used to identify routers. Additionally, it uses a mate-30 heuristic similar to AAR [52] and identifies subnets to find candidate alias pairs. Packet ID fingerprinting technique is used on the candidate alias pair to infer aliases [12]. Second, DNS names are used to geo-locate routers and group them into PoPs. It is worth noting that the DNS name can be assigned to any of the inferred aliases. However this geo-location is not complete and accurate for three reasons as follows: *i*) For some routers, there is no DNS name assigned to any of their interfaces. *ii*) Extracting geographical information from a DNS name is not a guarantee. *iii*) 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 [62]. Third, routers that are not mapped to a location are assigned a location using a clustering approach. iPlane identifies router clusters including interfaces that are similar from a routing and performance perspective. For this purpose, it probes all interfaces with `ICMP echo` probes from 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 to the clustering problem over these measurements. Interfaces in one cluster are assumed to belong to the same PoP.

Another approach is to use Geo-IP databases to assign a location to an IP address. Tian *et al.* [34] use these databases in conjunction to 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. This annotated graph contains some clusters corresponding to each city. Their heuristic technique re-annotates an interface to a new location if its new annotation results in more coherent groups, where more links are inside a group. Each group is detected as a PoP.

A PoP comprises a set of routers with high interconnectivity and links inside a PoP are usually rather short. These properties were used by Feldman *et al.* [13, 63] to propose a more automatic approach to detect PoPs. In their graph-based approach, network "motifs" are used to detect repeated patterns in `traceroutes` based interface level topologies collected by DIMES [10]. These repeated patterns are used to identify tightly connected interfaces. First, they ignore all links with delay above a certain threshold (5 ms), since these links are likely to be long haul between distant PoPs. This step generates a graph with disconnected components, each of which is a candidate to be detected as a single 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. In order to geolocate the PoP they use several geolocation services including MaxMind GeoIP [64]. Finally, they validated their PoP-level topology map with the DNS name based geo-localization and two geo-IP data bases. They claim that by not using the DNS names in their methodology, this information can be used as a "ground truth" to validate the accuracy of their technique. However, the accuracy of DNS names to infer geographical location of an interface is questionable [62].

6. AS-LEVEL

The topology at the AS-level is typically modeled using a simple graph where a node is an AS identified by an AS number. An Autonomous System (AS) is a collection of IP prefixes under the control of one network operator that presents a common, clearly defined routing policy to the Internet [65]. On the AS graph, links represent logical connectivity between two ASes, and are labeled according to the type of connection; customer-provider, peer-peer, and sibling. The logical connectivity usually represents multiple physical connectivities among PoPs of the two ASes.

This graph representation of the AS topology has some limitations: First, each AS has a geographical footprint that may overlap with the footprint of another AS. This can not be illustrated using a simple node, unless the node is replaced by a plate that covers an area. Second, ASes are widely considered to be coherent entities with a clearly defined routing policy. However, due to their large coverage, some ASes use various policies. For instance, Muhlbauer *et al.* [66] demonstrated that an AS is an atomic structure with respect to its routing policies. Third, two ASes can have multiple inter-AS connections at different locations, which can not be modeled by a simple graph. Fourth, Internet Exchange Points (IXP) also complicate the AS-level topology by providing connectivity between many ASes, most commonly through layer 2 multiple access clouds. As a result, in the most complete AS topology graph, IXPs should be modeled as links that connect more than two ASes. Considering these issues a more detailed structure of the AS topology can be represented using a hyper-graph [1]. However, despite these limitation, the graph representation of the AS-level topology still includes an abundance of important information and has been studied for the past 15 years to a great extent.

6.1 AS Topology Data Sources

Techniques for discovering AS-level topology rely mainly on three data sources: BGP information,`traceroute`, and Internet Routing Registries (IRR) [67]. Next, we introduce each type of data source and its limitations.

BGP Information: BGP is the inter domain routing protocol of the Internet. BGP is a path vector protocol in which routing decisions are made based on reachability via the advertised AS paths and 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 was the first data source used to map the AS-level topology [69]. BGP information has been used in different forms and can be collected from various resources including: *i) BGP archive*: Oregon RouteViews [70] and Rseaux IP Europens's (RIPE) Routing Information Service [71] collect BGP route information through a set of route collectors also known as BGP monitors or vantage points. They provide route table dumps and route update traces. While BGP dumps show the best path to reach other ASes, the back-up links and the dynamic nature of BGP routings are more likely captured by "route updates". Both the BGP dumps and updates are used to capture the AS-level topology [72, 73]. *ii) Route Servers*: A route server is a BGP router that offers interactive login access via `telnet` or `ssh` permitting to run most non-privileged router commands [72]. For example, BGP summary information can be obtained by "`show bgp summary`" command. *iii) Looking Glasses*: A looking glass is a web interface to a BGP router which usually allows BGP data querying and limited use of debugging tools such as ping and traceroute. [72]

Although passive collection of BGP tables and updates have fueled many studies on AS-level topology, there also have been efforts that used the active measurements of BGP. A BGP beacon [74, 6] is a router that advertises and withdraws a prefix. Observing these announcements from the perspective of different route collectors allows an estimation of protocol behavior, *e.g.*, the protocol convergence time and the AS distance an advertisement travels on the control plane. BGP Route poisoning prevents BGP announcements from reaching an AS. Bush *et al.* [6] used this technique to measure the prevalence of default routes in the Internet and explain the difference between the observed topology form control vs. data plane measurements.

Using BGP for collecting AS-level topology has several advantages. First, compared to internet registries, the data collected from BGP shows the actual reachability of the Internet control plane. Hence, the data is not normally prone to being obsolete or incorrect. Second, the BGP update can be used to study the dynamic behavior of internet routing and to discover backup links. Third, engineering solutions can be used on top of BGP to improve our view of the topology. For example BGP beacons and route poisoning are used to detect backup paths and default routing [6].

Despite all its advantages, using BGP information to infer the AS-level topology is not without limitations. BGP is merely an information hiding protocol and only indicates reachability and not connectivity. More specifically, BGP has the following limitation: First, the AS path announcements are primarily used for loop detection. Adding an AS in the announcement is not uncommon for traffic engineering. ASes also may announce an AS path that does not correspond to the real path [66]. Second, being a path vector protocol, BGP does not announce information on every path. As a result, back-up paths might never appear in the BGP dumps. Third, since BGP only announces the best path, 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 topology is biased to be more complete for these regions. Lastly, 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 [44, 6]. This introduces a measurement bias in BGP based AS topologies, since P2P links are only used for traffic originating from the customers of any of the peering ASes. Hence P2P AS relations are not discovered relatively easily[44]. In fact, the majority of the missing AS links in the topology inferred from BGP data are known to be P2P links [2].

`Traceroute` Measurement: Another approach to discover the 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 [75]. Consecutive IPs that belong to two different ASes reveal the connectivity between the ASes.

This technique has the advantage of revealing a more detailed view of the AS-level topology. We recall that ASes can be connected at multiple locations. The `traceroute` based measurement allows us to distinguish between multiple inter-AS connections between two ASes. In addition, `traceroute` measurements often use more vantage points, since deploying a `traceroute` vantage point is much easier than a BGP route collector. As a result, the AS-level topology generated by large-scale `traceroute` measurement is considered to be more complete than those collected from BGP information [10, 37, 38].

Apart from the limitations of `traceroute` that we discussed in Section 3.1.1, active measurement on the data plane has other limitations for mapping the AS-level topology. First, IP to AS mapping is not a trivial task. Prefix registries are often incomplete and using BGP for mapping IPs to ASs number is not accurate due to its information hiding characteristics. Second, discovering false inter AS connection is likely due to inconsistencies in router responses [37, 4]. 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[4].

Finally, it is worth mentioning that when measuring the AS topology using BGP and `traceroute` measurement, the Internet control plane and its data plane are in fact being measured, respectively. 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. These issues broadly stem from i) the limitation of data that is used to infer the topology, ii) and the lack of knowledge about the effects of these limitations on the observed topology [6]. For instance, “default routing" limits the view of the passive BGP measurements while the active measurement observes the route. The general consensus is that the AS-level topology collected from the measurements on the data plane results in a more accurate and complete view of the Internet [6, 15, 1].

Internet Routing Registries: The Routing Arbiter Database (RADb) maintained by IRR is a group of lookup databases maintained by several organizations. they are designed to provide fundamental information about routing in the networks. Documented routing policies, regulations, and peering information is found amongst the abundance of information kept on these databases.

The main advantage of using IRR is its simplicity. All the information is accessible via `WHOIS` command and can be obtained through `FTP` servers. This resource also does not exhibit the limitations of data obtained through measurements, since it is based on the data provided by the ISPs themselves. However, when using this resources extreme care is a necessity for the following reasons: First, since these registries are voluntarily provided, data may be incomplete due to confidentiality and the overhead of updating an external data store. Second, they may not portray the latest up-to-date state of the network. For instance, reports checking the accuracy of RIPE show inconsistencies in registry overlaps among different databases [68].

6.2 AS Relationship and AS Tiers

Although the logical AS topology is interesting in itself, in order to be more useful in practice, the inter-AS routing policies should also be inferred. The business relation between connected ASes are normally approximated by three categories [76]: (1) Customer-Provider (C2P), (2) Peer-Peer (P2P), and (3) Sibling relations. From the financial perspective, in a C2P relation the customer is billed for the connectivity by the provider. The other two types of relationship are settlement free. P2P relation helps two small ASes with high inter-AS traffic profiles reduce their cost by directly exchanging traffic, hence reducing the traffic sent towards the providers. Sibling relations mostly occur when business

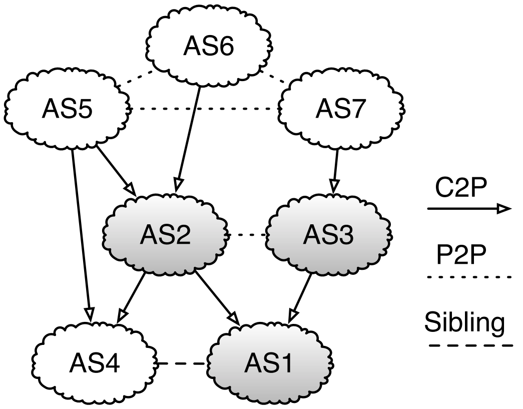


Figure 11: AS graph annotation with AS relations

mergers happen or when an AS is acquired by anther AS.

Primitive approaches to infer 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 customer does and the size of an AS is typically proportional to its degree in the AS graph [77].

The commonly used approach to infer inter-AS relationship is to use the observed routing paths and assume the generality of the "valley-free property" of the Internet [78, 77, 79]. 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 sequence of +1, followed by at most one 0, followed by sequence of -1. The type of relationship assignment can be formulated as an optimization problem. Given an undirected graph representation of AS topology and a set of AS-level paths, they aim to assign policy labels to the links in such a way 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.

The alternative approach is to check the consistency of the inferred relations with other measurements [16]. For instance, Muhlbauer *et al.* [66] used `traceroute` to estimate the accuracy of the inference by comparing the inferred route and the real routes. In their approach, they use multiple quasi-routers to capture route diversity within the ASes.

Traditionally, AS-level topology is widely accepted to be a hierarchical structure, where ASes are categorized into different tiers [78, 79]. Tier-1 ASes are defined as those that don't buy transit from any other AS. These tier-1 ASes form a full mesh connectivity at the highest tier. Tier-2 providers are customers of the tier-1 ASes using them for their transit service. Additionally, tier-2 ASes use peer-peer relations with other tier-2 ASes in order to decrease the transit cost. This hierarchy structure can be extended to more levels. However, this perception is changing in the research community. First, many new ASes (*e.g.*, content providers and Content Distribution Networks (CDN) are not transit ASes but have many connections at various locations. These new types of ASes do not fit in any tier on the hierarchy. In addition, new studies explain this perception using the abundance of missing links and the limited observability of P2P connections [15]. Although the existence of large transit ASes at the highest tier is valid, the tier-based hierarchical view is replaced by a flat but modular view. Figure 11 shows an example of an annotated AS graph. *AS1*, *AS2*, and *AS3* are the ASes in our previous examples. *AS5*, *AS6*, and *AS7* are tier1 ASes form the full mesh at the highest tier. However, the hierarchical structure does not exist beyond tier1.

6.3 Coverage & Completeness

As of 2011, the discovered AS-level topology consists of approximately 40,000 ASes and 115,000 to 135,000 edges where 80,000 to 90,000 are C2P links and the rest are P2P links [15]. While this topology seems to be complete with respect to its nodes, its edges are more prone to exclusion from measurements.

A great deal of research has been dedicated to asses the completeness of the AS-level topology. Lord of the Links study [44] compares BGP routing tables, Internet Routing Registries, and `traceroute` and cross validates the topology captured from various sources and captures a more complete view of the AS topology. It also extracts a significant amount of new information from the Internet Exchange Points (IXPs) and uses this information in the cross validation process.

The incompleteness of the Internet AS map has also been studied (*e.g.*, [2, 80]). Oliveira et al. [81] use the ground truth to validate the accuracy of their derived AS map 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. [2] categorized the missing links into hidden and invisible links. Invisible links are missing due to the limitations imposed by the placement of vantage points. But hidden links 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. [36]. 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.

Quite recently, the AS-level map received a major update using the ground truth data from one of the largest IXPs in Europe with 400 AS members [82]. The main finding was that in this single IXP, there are 50,000 P2P links, which is more than the total number of P2P links already discovered. This suggests the total number of P2P links can be larger than 200,000. These observations show that the discovered

AS-level Internet topology is far from complete and there is still room for improvement.

6.4 Geolocation

Apart from prior studies on the geographic location of the 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). Internet registries and directories such as PeeringBD [83] 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 resource these directories are easy to use but can be out of date and incomplete.

The geographical footprint of eyeball ASes (ISP that serve normal costumers) has been studied in [84]. Using large scale measurement from Peer-to-Peer applications, authors identify a large set of end-host IPs. First, these IPs are mapped to ASes. Then, 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 while reducing the error of each database. 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.

6.5 Modeling

The presumed AS topology of the Internet has been examined from a graph theoretic stand point in several studies. However, there is no consensus on which observation is more complete and accurate due to the incompleteness of the measured topology. Zhou *et al.* [85] proposes a growth model with Positive-Feedback-Preference which reproduces many topological properties of the AS-level topology. Their model, however, uses the Skitter [40] `traceroutes` dataset to reveal the target AS-level topology which suffers from known limitations of `traceroute`-based mapping. For instance, the observed power law degree distribution of AS topology is known to be due to the bias in the measurement techniques [1, 15]. Mahadevan *et al.* [73] used the inferred AS topology from multiple data sources including: BGP, `treacroute` and `WHOIS`. They compared the graphs from the graph analysis perspective. They reported that the “joint degree distribution" can be used to characterize the Internet AS graph. They also showed how the data collection peculiarities explain differences in the resulting graph analysis metrics.

The evolution of the Internet AS map has also been investigated. The main challenge with respect to the evolution of the topology over a long term is to distinguish the changes due to the topology change vs. the routing dynamics. Oliveira et al. [86] 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. [87, 88] have 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.

Chan et al. [89] use a policy based graph model, where policies are implemented in a simulated environment that effect ASes decision in creating new AS relations. Similar to HOT models for router-level topology, this model uses a reverse-engineering approach. In the decision process, they consider the gain from P2P links and C2P links, using simulated traffic demands. Using different profiles for ASes with different objectives they can model the behavior of these ASes and model the Internet using an evolutionary framework. In order to validate the model, they use measurement based observations to match their model with observation from reality, however their model parameters can be tuned using more accurate data.

7. CONCLUSION

Internet topology discovery has been on of the most studied man made structures. This is due to the influence of the Internet topology on its functionality. Designing new protocols, managing and debugging of the network, and implementing security measures can all benefit from an accurate Internet topology map.

Being a complex decentralized system, Internet can be view at different resolutions. IP level, router level, PoP level, and AS level. We used the topology resolution to organize the research conducted in the past 15 years on Internet topology discovery. At each level, we introduced the data used to capture the topology at that level. We classified these data sources based on their type of data (information on data plane vs. control plane) and their measurement technique (active vs. passive measurement). We discussed the pros and cons of data sources and topology discovery techniques. We point out that the captured topology is still incomplete and explain this incompleteness using the limitations of the data and techniques. When possible, we also presented the geographical properties of the captured topology. Finally, we covered the topology modeling research at each level.

Despite the large amount of studies performed by the research community, our incomplete view of the Internet topology begs for more work. The research community has come to the conclusion that more measurements do not always compensate for the limitations of the measurement tools. However, knowing the measurement tools, researchers can recalibrate the expectations and revisit their assumptions. Controlled experiments that are based on domain knowledge and validation using ground truth information, and the use of new data sources have lead to great achievements in the Internet topology discovery community.

8. REFERENCES

[1] 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.

[2] 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.

[3] M. C. Toren, "tcptraceroute: an implementation of traceroute using tcp syn packets."

[4] 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.

[5] 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.

[6] 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.

[7] 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.

[8] 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.

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

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

[11] 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.

[12] 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.

[13] D. Feldman and Y. Shavitt, "Automatic large scale generation of internet pop level maps," in *Global Telecommunications Conference, 2008. IEEE GLOBECOM 2008. IEEE.* IEEE, 2008, pp. 1–6.

[14] 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.

[15] W. Willinger and M. Roughan, "Internet topology research redux," *ACM SIGCOMM eBook: Recent Advances in Networking*, 2013.

[16] 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.

[17] 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.

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

[19] 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.

[20] 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.

[21] V. Jacobson, "traceroute," ftp://ftp.ee.lbl.gov/traceroute.tar.gz.

[22] S. Savage, "Sting: A tcp-based network measurement tool." in *USENIX Symposium on Internet Technologies and Systems*, vol. 2, 1999, pp. 7–7.

[23] 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.

[24] J. Moy, "Ospf version 2," 1997.

[25] R. W. Callon, "Use of osi is-is for routing in tcp/ip and dual environments," 1990.

[26] B. Yao, R. Viswanathan, F. Chang, and D. Waddington, "Topology inference in the presence of anonymous routers," in *INFOCOM 2003. Twenty-Second Annual Joint Conference of the IEEE Computer and Communications. IEEE Societies*, vol. 1. IEEE, 2003, pp. 353–363.

[27] R. Govindan and H. Tangmunarunkit, "Heuristics for internet map discovery," in *INFOCOM 2000. Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE*, vol. 3. IEEE, 2000, pp. 1371–1380.

[28] 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.

[29] 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.

[30] 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.

[31] B. Eriksson and 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.

[32] 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.

[33] 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.

[34] 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.

[35] 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.

[36] 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.

[37] 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.

[38] 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 internets edge,” in *Proc. of USENIX NSDI*, 2013.

[39] 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.

[40] CAIDA, “Macroscopic topology measurements project and the skitter infrastructure,” http://www.caida.org/tools/measurement/skitter/.

[41] CIDA, “Macroscopic topology measurements project and the archipelago measurement infrastructure,” http://www.caida.org/projects/ark/, 2011.

[42] N. T. Spring, D. Wetherall, and T. E. Anderson, “Scriptroute: A public internet measurement facility.” in *USENIX Symposium on Internet Technologies and Systems*, 2003.

[43] 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.

[44] 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.

[45] 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.

[46] 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.

[47] 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.

[48] N. Spring, M. Dontcheva, M. Rodrig, and D. Wetherall, “How to resolve ip aliases,” *Univ. Michigan, UW CSE Tech. Rep*, pp. 04–05, 2004.

[49] 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.

[50] K. Keys, “iffinder, a tool for mapping interfaces to routers,” *See http://www. caida. org/tools/measurement/iffinder.*

[51] 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.

[52] 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.

[53] 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.

[54] 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.

[55] R. Siamwalla, R. Sharma, and S. Keshav, “Discovering internet topology,” *Unpublished manuscript*, 1998.

[56] 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.

[57] 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.

[58] 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.

[59] R. Albert, H. Jeong, and A.-L. Barabási, “Error and attack tolerance of complex networks,” *Nature*, vol. 406, no. 6794, pp. 378–382, 2000.

[60] 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.

[61] 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.

[62] 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.

[63] 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.

[64] M. LLC, “Geoip, 2010,” http://www.maxmind.com, 2010.

[65] 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. [Online]. Available: http://www.ietf.org/rfc/rfc1930.txt

[66] 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.

[67] I. R. Registry, "Obtaining irr data," ftp://ftp.radb.net/radb/dbase, 2013.

[68] N. RIPE, "Routing registry consistency check reports," *see http://www.ripe.net/projects/rrcc*, 2009.

[69] 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.

[70] A. N. T. Center, "University of oregon route views project," http://www.routeviews.org, 2013.

[71] "Ripe ris," https://www.ripe.net/data-tools/stats/ris/routing-information-service, 2011.

[72] 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.

[73] 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.

[74] 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.

[75] 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.

[76] G. Huston, "Interconnection, peering, and settlements," in *proc. INET*, vol. 9, 1999.

[77] L. Gao, "On inferring autonomous system relationships in the internet," *IEEE/ACM Transactions on Networking (ToN)*, vol. 9, no. 6, pp. 733–745, 2001.

[78] 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.

[79] 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.

[80] R. Oliveira, W. Willinger, B. Zhang *et al.*, "Quantifying the completeness of the observed internet as-level structure," 2008.

[81] 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.

[82] 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.

[83] PeeringDB, "Exchange points list," https://www.peeringdb.com/private/participant_list.php, 2013.

[84] 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.

[85] S. Zhou and R. J. Mondragón, "Accurately modeling the internet topology," *Physical Review E*, vol. 70, no. 6, p. 066108, 2004.

[86] 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.

[87] 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.

[88] 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.

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