

Summary: Mapping Interconnection in the Internet: Colocation, Connectivity and Congestion

As the global Internet expands to satisfy the demands and expectations of an ever-increasing fraction of the world's population, profound changes are occurring in its interconnection structure, traffic dynamics, and the economic and political power of different players in the ecosystem. These changes not only impact network engineering and operations, but also present broader challenges for technology investment, future network design, public policy, and scientific study of the Internet itself. And yet, from both scientific and policy perspectives, the evolving ecosystem is largely uncharted territory.

In particular, two related transformations of the ecosystem motivate our inquiry: the emergence of Internet exchanges (IXes) as anchor points in the mesh of interconnection, and content providers and Content Delivery Networks (CDNs) as major sources of traffic flowing into the Internet. By some accounts over half the traffic volume in North America now comes from just two content distributors (Youtube and Netflix). This shift constitutes the rise of a new kind of hierarchy in the ecosystem, bringing fundamentally new constraints on existing players who need to manage traffic on their networks to minimize congestion. Measurement challenges limit our current capability to describe and understand these dynamics, but evidence of trouble has increased dramatically in the last five years, resulting in tussles among commercial players as well as between the private sector and regulatory bodies, at the expense of users suffering degraded quality of experience.

We propose to characterize the changing nature of the Internet's topology and traffic dynamics, and to describe the implications of these changes on network science, architecture, operations, and public policy. The research is structured as two foundational tasks and a set of research questions that build on those tasks. We will first construct a new type of semantically rich Internet map, which will elucidate the role of IXes in facilitating robust and geographically diverse but complex interdomain connectivity. This map will guide our second task: a measurement study of traffic congestion dynamics induced by evolving peering and traffic management practices of CDNs and ISPs, including methods to detect and localize the congestion to specific points in the network. We will use our own measurement infrastructure as well as measurements from four industry collaborators: Akamai, Netflix, Google, and Comcast. We have designed experimental prototypes in order to test the feasibility of several of our proposed methods, and validated early results with network providers who have provided ground truth on topology and traffic inferences. We will use the results of these two tasks to investigate questions related to infrastructure resiliency, scientific modeling, network economics, and public policy.

Intellectual merit: This project will advance our scientific understanding of the evolving Internet, yielding new methods of analysis and insights into the interaction between CDN dynamics, interconnection at IXes, and network performance. It will also allow us to create and validate a more sophisticated model of the physical Internet topology, as well as provide unique data sets and scenarios by which to evaluate future network architectures.

Broader impacts: We will broadly disseminate the results of this project to interested communities via publications, conferences, relevant committees' meetings and online. We will also organize and host annual international interdisciplinary workshops to close the loop between empirical research results, theoretical modeling activities, and policymaking. By providing unbiased data about the evolving Internet ecosystem, the proposed research has the potential to be of great value to Internet users, network operators, and governments.

Keywords: Internet measurement; performance; traffic engineering and management; economics; policy; future Internet

Project Description: Mapping Interconnection in the Internet: Colocation, Connectivity and Congestion

1 Motivation

What has been called the “flattening” of the Internet [1] is now a recognized phenomenon: peering connections among ISPs (ASes) as primary interconnection paths are increasingly supplanting the model of networks interconnecting hierarchically via a small clique of tier-1 providers [2, 3]. The pervasive growth of two initially independent but synergistic infrastructure sectors have contributed to this shift over the last 15 years: Internet eXchanges (IXes) and Content Distribution Networks (CDNs). IXes facilitate interconnection among networks within a region, allowing traffic to flow along cheaper and lower-latency routes, which CDNs leverage to place (cache) content as close to users as practical. The relationship between IXes and CDNs has grown increasingly symbiotic: both aim to localize traffic near users, optimizing bandwidth efficiency and performance. There are over 400 IXes in the world [4], with diverse architectures, membership, and business/governance models. Most larger IXes are public Internet Exchange Points (IXPs) where participants can interconnect across a shared and sometimes distributed fabric; some also support private peering (colocation) facilities where tenants physically interconnect routers. The diversity in connectivity that IXes enable has not been rigorously quantified, but recent studies by us and others have discovered evidence of more peering connections at IXPs [2, 5] than were previously documented to exist in the entire Internet [6, 7, 8].

Much of today’s Internet traffic originates from large content providers and their (own or partner) CDNs. In 2013, Sandvine [9] reported that about one half of all downstream consumer traffic came from Netflix or Youtube. This transition constitutes the rise of a new kind of hierarchy to the Internet ecosystem, with implications for the balance of power and control among its players. CDNs typically control the source (and thus the path) of data coming into an ISP, so they can increase (or alleviate) loading and congestion on different points of interconnection. At the same time, ISPs can control the capacity of incoming links, which can limit the realistic options for the delivery of high-volume traffic. Strategic behavior by a CDN or ISP may create externalities for other infrastructure operators, and inevitable regulatory interest in the resulting shifts in bargaining, money flows, and potential market power related to interconnection.¹ Although peering disputes over traffic imbalances are not new [11, 12, 13], heated peering disputes between powerful players in the U.S. have increased dramatically in the last four years [14, 15, 16, 17, 18, 19, 20, 21, 22], raising technical questions about appropriate network management as well as concerns about intentional degradation of performance as a business strategy to obtain interconnection fees. For three years we have hosted a Workshop on Internet Economics (WIE), where we have heard ISPs and content providers describe deep conflicts over peering practices. Thus far the research and measurement community has had no scientific framework to study this behavior.

Understanding how these industry sectors interact, and the resulting impact on the rest of the Internet industry, is largely uncharted territory. The shift in Internet traffic dynamics has obvious implications for network engineering and operations, since network operators must change their approaches to traffic engineering, capacity planning, and other techniques to maintain and improve network infrastructure. But these dynamics also pose broader challenges related to the

¹This interest has manifested itself in heated debates on how to update international telecommunications settlements regulations as the focus shifts from telephony to the Internet. The European Telecommunications Network Operators’ (ETNO) proposal [10] (prepared in advance of the ITU plenary meeting in December 2012) and resulting controversy illustrates the tension around the regulation of Internet interconnection.

evolving Internet: technology investment, future network design, public policy, and scientific study of the Internet itself. As the Internet inexorably subsumes the legacy telephony infrastructure, metrics of interest are expanding to include reliability and availability, policymakers and researchers struggle to map network measurements to such metrics.²

The goal of this research is to measure and characterize the changing nature of the topology of the Internet, and to describe the implications of this change on Internet operations, design, science, and policy. The research is structured as two foundational tasks and a set of research questions that build on those tasks. Our first task is to create a new type of Internet map (Section 3.1) to capture the role of IXes in facilitating robust and geographically diverse but complex interdomain connectivity. We will augment traditional BGP sources of AS-level topology data with new measurement and inference techniques to infer massive peering meshes – which have thus far been largely invisible to standard measurement methods – and where they physically occur. We will use this richly annotated topology map to guide discovery of loading and congestion at interconnection points (Section 3.2), which can degrade the experience for many users, and can signal contention in business relationships between ISPs. Our methods include two forms of active probing, combined with data about actual streaming downloads we will receive from Netflix, the dominant provider of content on the Internet, and data from Comcast that can be used to validate the first uses of our test methods.

The development and application of these new methods and data will provide a new lens through which to observe and understand the evolving Internet. We will use the knowledge base and tools created in the first two tasks to explore questions related to infrastructure resilience, network economics, communications policy, and scientific modeling (Section 3.3). Our goals for this inquiry are directly responsive to the NSF’s Networking Technology and Systems (NeTS) program solicitation, and include: advancing our fundamental understanding of how content distribution dynamics affect ISP network management capabilities; developing metrics to quantify the impact of emerging interconnection patterns on the resiliency, efficiency, and market power of modern networks; revisiting longstanding but now questionable topology modeling assumptions; offering new models that are better grounded empirically; and tracking trends over time to establish a baseline against which to evaluate future Internet architecture designs and implementations.

2 Background: State of the Art

The best publicly available data about the global interconnection system that carries most of the world’s communications traffic is incomplete and of unknown accuracy. There is no map of physical link locations, capacity, utilization, or interconnection arrangements. This opacity of the Internet infrastructure hinders research and development efforts to model network behavior and topology, design protocols and or new architectures, much less study real-world properties such as robustness, resilience, and economic sustainability. There are good reasons for the dearth of information: complexity and scale of the infrastructure; information-hiding properties of the routing system; security and commercial sensitivities; costs of storing and processing the data; and lack of incentives to gather or share data in the first place, including cost-effective ways to use it operationally [24]. In this section we review relevant literature in Internet topology and congestion research, and the open questions that motivate our proposed research.

²The FCC’s Chief Scientist proposed a new IETF working group to focus on developing a new architecture for coordinating measurements of broadband network performance [23].

2.1 Inferring and characterizing connectivity

Most research use of Internet topology data today relies on maps at the AS granularity, both because they are useful approximations for a variety of research [25, 26, 7, 27, 28], and because two public repositories of interdomain routing data allow construction of AS-level graphs: RouteViews [29] and RIPE RIS [30]. Although these public data sources provide a reasonable view of customer-provider relationships between networks, they miss a great deal of pervasive regional and non-revenue-based peering activity, because such links tend not to propagate via formal (paid) transit relationships, and thus cannot be observed by remote monitoring instrumentation [31, 32, 33, 34, 35, 36]. In particular, this kind of instrumentation does not typically capture peering at Internet Exchanges (IXes), much of which is not governed by a formal contract [37]. To capture some of these stealthier links in the AS graph, some studies have used other sources of interconnection data, such as routing registries [38, 39] and traceroutes specifically crafted to infer AS links at IXes [40, 6, 41]. Others have claimed (incorrectly) that IXP connectivity is impossible to discover without traffic data from the IXP itself [2].

Researchers have also studied how to meaningfully annotate AS topologies, most commonly trying to infer the type of business relationships between networks [42, 43, 44, 45, 46, 47, 36, 48]. A major limitation of all known AS-relationship algorithms, including our own [49], is that they attach a single type of relationship to an AS link – customer-to-provider (c2p) or peer-to-peer (p2p)³ – while in reality relationships can be more complex, such as region-specific or prefix-specific [50, 51]. These oversimplifications of the AS graph limit its utility for many research questions, including the study of routing policies, path diversity, and resiliency to failures. For example, due to the geographically distributed nature of many network providers, an AS-graph is technically both a multi-graph (two AS nodes connect at multiple locations) and a hypergraph (multiple ASes can be connected by a single “link”, e.g., the shared peering fabric at an IX). Previous efforts to provide Internet maps at finer granularities than the AS-level have focused on either router-level or Points-of-Presence (PoP) level graphs, both of which have daunting research challenges in construction as well as use. Inferring a valid router graph that captures a representative fraction of the core of the Internet requires significant measurement and analysis of massive traceroute data [52], and also results in a hypergraph, but one that can be three orders of magnitude larger than the corresponding AS-level graph, imposing computational costs that inhibit its practical use for some problems. PoP-level maps offer a compromise, depicting geographic locations where networks have infrastructure, but the few examples in the literature [53, 54, 55, 56] are limited in coverage due to lack of measurement infrastructure, suffer inaccuracies in router geolocation, and are extremely difficult to validate.

Although there is wide recognition of the limitations of the existing Internet topology maps, and in particular of the need to move beyond modeling ASes as single nodes connected bilaterally with single links [51, 57], no one has actually attempted this task at scale, much less produced ongoing snapshots of such maps, annotated and validated to the largest extent possible. In the meantime, none of the existing maps can support empirical research on the complex peering ecosystem, its growing synergistic interaction with content industries, and implications of these interactions for the stability and performance of the Internet.

³In a c2p relationship, the customer pays the provider to transport its traffic to and from ASes the provider can reach, either directly or via its own providers. In a p2p relationship, two ASes gain access to each others’ customers, typically without either AS paying the other (called a *settlement-free* peering relationship)

2.2 Inferring and characterizing congestion

The study of network congestion – adaptive transport-layer congestion control [58], congestion management [59], and detecting its effects [60] – has been a focus of network research, protocol standards development efforts, and network operators for many years. More recently researchers have devoted effort to understanding in-home [61, 62, 63], and broadband access [64, 65, 66, 67, 68] performance issues, and whether these components are really the end-to-end performance bottlenecks for most users [69, 70].

Another related field of measurement of network properties is network tomography [71], which uses end-to-end measurements to discover network-internal characteristics, such as topology [72], loss rates [73, 74, 75], or delays [76, 77, 78, 79]. A specific class of tomography is *binary tomography* [80], which assumes end-to-end measurements and internal link states are in one of two states: “good” or “bad” (or congested vs. uncongested). The binary tomography problem is usually underconstrained – many more links (variables) than end-to-end paths (equations) – and thus is usually simplified to finding the *smallest* set of congested links that explains the end-to-end observations. The problem of identifying the smallest set of likely congested links given a set of end-to-end measurements is NP-hard [81]. Greedy approximations to this problem are known to produce good results; in fact, a greedy algorithm achieves the best possible approximation possible for a polynomial-time algorithm [82]. We have previously utilized a binary tomography approach to identify the set of candidate “broken” links that best explain a set of end-to-end reachability measurements [83]. Researchers have also combined elements of Boolean tomography (estimating good or bad links) with analog tomography to estimate a range of actual performance for bad links [84]. We are not aware of any study that applies tomographic techniques to systematically study broadband networks and their peering and interconnection links.

And yet, growing evidence in the press [11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22] indicates that peering disputes between powerful players are leading to congested interconnection links, which affect the performance of many (sometimes millions of) users simultaneously. These tussles raise technical questions about the need for network management mechanisms [85, 86], and policy concerns about potentially intentional technical degradation of interconnection and content delivery as a consequence of adverse business interests. For example, the four major French ISPs (France Telecom, Free, SFR and Bouygues) are reputed to engage in contentious bargaining with CDNs as part of a business strategy to obtain interconnection fees [87, 88, 89, 90]; some have speculated that part of their business strategy is to run some of their interconnection links “hot”, i.e., underprovisioned and thus congested, so that they are not suited for the delivery of high-volume traffic such as streaming video. There have been heated debates on the existence of such behaviors on operational mailing lists [14], but no published scientific research

Access providers are not the only players with an incentive to degrade performance as a business strategy to protect or increase revenue streams [91]. CDNs also play a growing role in traffic management today, performing complex optimizations, measuring download performance and adjusting traffic flows in near-real-time, including from which nodes they source content [92, 93, 94]. They may also have incentives to induce load onto congested links as a part of a business strategy to negotiate favorable peering agreements with network providers. The resulting potentially contentious interactions between ISP routing and CDN sourcing have implications for network stability and performance [95, 96, 97, 98]. This interconnect issue is increasingly complex, and CDNs and operators have emphasized the need for more analysis in this area (see attached letter of collaboration from Comcast).⁴

⁴While the policy and operational implications of this proposal are important it is not our goal to attribute blame for congested links.

3 Proposed Research

Our proposed research is structured as two foundational tasks and a set of research questions that build on those tasks. We will first construct a new type of semantically rich Internet map, which will guide our second task: a measurement study of traffic congestion dynamics induced by evolving interconnection and traffic management practices of CDNs and ISPs. This new map and these measurement techniques will frame our inquiry into issues relevant to network operators, researchers, policymakers, and users. These inquiries will inform infrastructure resiliency assessments, improve network modeling integrity, and enlighten Internet policy debates.

3.1 Task 1: Create an IX-aware map of the Internet

Our first task is to produce a representation of the inter-AS configuration of the Internet that will elucidate the role of Internet Exchanges (IXes) in facilitating robust and geographically diverse interdomain connectivity. In this proposal, we use the general term *Internet eXchange* (IX) to refer to both IXPs and private peering facilities. We will refine a technique we recently developed to infer *multilateral peering relationships* at known IXes [5]. We will augment this set of discovered links with targeted traceroutes performed at scale to discover bilateral peering links crossing IXes, and explore techniques to identify the presence of IXes not explicitly seen in the path or not documented in public databases. We will then annotate nodes with inferred network types, and annotate links with inferred *regional business relationships*, by extending our AS relationship inference algorithm [49] to accommodate regional differences we observe at IXes. The resulting map will provide detailed information about AS-level connectivity, whether it happens at IXPs or private colocation facilities, and how it differs across regions.

3.1.1 Incorporating IX connectivity into an AS-level Internet map

To augment the baseline topology maps created from public BGP data [29, 30] with additional knowledge about peering and IXes, we will start with *colocation information* obtained from several public data sources of self-reported colocation information that have been largely untapped by the research community.

PeeringDB [4] is the richest source of self-reported colocation and peering offerings; this community web site supports voluntary registration of an ISP’s information about their presence at *public peering points* (at IXes with shared fabrics) and *private peering facilities*, providing hints about likely peering offerings of different networks. PeeringDB serves as a matching service for networks seeking peers; for example, the PeeringDB entry for Comcast contains the statement “We do not offer peering, paid or otherwise, on the shared fabric public switches at any IX” and lists their presence at 17 private peering facilities (and no IXPs). Its registered participants are reasonably representative of the Internet’s transit, content, and access provider populations [99]. Two other repositories [100, 101] provide similar voluntarily reported colocation offerings. Some IXes also publish member lists (e.g., [102]) and peering matrices (e.g., [103]). While not directly providing topology information, all of these data sources can be used to infer colocation at specific IXes [104]. *We propose to build a colocation database that will use these data sources to compile information about which networks are colocated at specific IXes and private peering facilities.*

Many IXes also provide a *route server* so that their members can maximize peering richness using *multilateral peering*, i.e., heavily meshed peering established over an IX’s public peering fabric [105]. In addition to facilitating efficient peering among many participants, route servers provide visibility into peering relationships. IX members control which other members receive their

prefixes (i.e. their peerings) by attaching special BGP community values on routes they announce to the route server (Figure 1). Some IXes publish their route server configuration (e.g., [106]), and an emerging effort is trying to standardize publication of such information in the future [107].

We recently developed and experimentally validated a technique to query public route servers at IXes in order to infer multilateral peering (MLP) meshes that they host [5]. In May 2013 we experimented with manual use of this MLP technique on 13 large European IXes hosting route servers, mining the special-purpose BGP communities, and using the IX looking glasses⁵ to gather routes for specific prefixes. The nature of MLP implies that we may infer hundreds of peering links from a single route. We inferred more than 206K p2p links, 88% of which were not visible in public BGP data [5]. We validated 26K of these links using 70 looking glasses provided by IXP members or their customers, proving that at least 98.4% of the peerings we could test actually existed. For two-thirds of the 1445 ISP members at the 13 IXPs we studied, this method inferred at least an order of magnitude more peering links than BGP or traceroute data revealed (Figure 2 and [5]), and discovered most of the peering connectivity inferred using proprietary traffic data from DE-CIX in 2012 [2].

PeeringDB lists 413 IXPs as of November 2013, suggesting a substantial opportunity to discover additional peering with this technique, but also a substantial scaling challenge. Even the 13 IXPs we studied used a variety of software, conventions, and interfaces to their looking glass and route servers, making the acquisition and processing of this data tedious. To the extent possible, we will automate the extraction of multilateral peering links from IXPs that use route servers, and will publish per-IXP details we find and software we use, to help others reproduce or validate our work.

This MLP approach will not reveal private bilateral peering between ASes, which can also occur at IXes. We must leverage other sources of information to estimate who is likely privately peering where, including

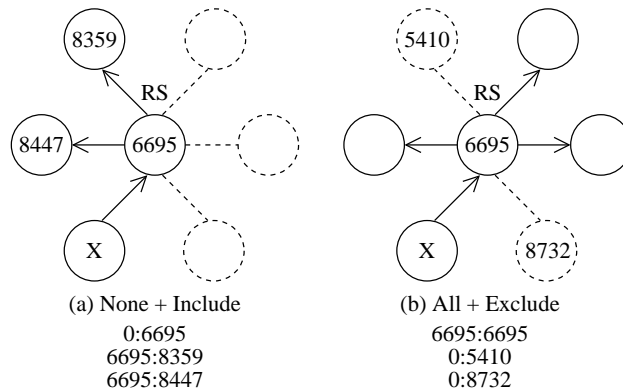


Figure 1: Controlling which members receive a prefix announced to a route server. In (a), X allows ASes 8359 and 8447 to receive the route, and no one else. In (b), X allows all members to receive the route except for 5410 and 8732. Each IXP uses a different ASN and convention for controlling route announcements; 6695 is the ASN for DE-CIX.

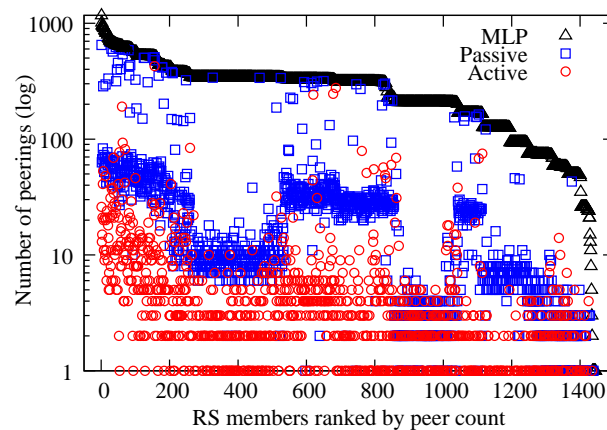


Figure 2: Numbers of MLP links found with our algorithm vs. with passive BGP data (Routeviews, RIPE RIS, PCH) and active traceroute data (Ark, DIMES). Inferred MLP links have little overlap with links observed from current active and passive data sources.

⁵A looking glass is a web-based mechanism some operators support to allow others to safely interact with a privileged or protected network process, such as BGP or traceroute.

traffic data from and personal relationships with IX operators.⁶ We can observe all pairs of ASes exchanging traffic, and subtract those AS pairs we inferred as multi-lateral peers to construct a set of candidate private peering AS pairs. This ground truth data will give us insights into the types of networks likely to engage in private peering, and will inform the final step of our map creation – targeted traceroutes to discover bilateral peering connectivity.

To discover bilateral peering links, we must first know which of these links could exist. Information about network colocation at various IXes and the insights from the ground truth traffic data will guide this process, indicating which links we should look for. We will then perform traceroutes from multiple distributed vantage points, crafted to maximize the possibility of traversing the targeted link. Critical to this technique is the availability of a large number of distributed vantage points, so that we can craft the traceroutes to maximize the possibility of traversing the suspected peering link, ideally with traceroutes in both directions to increase our confidence that the peering exists. In particular, we need vantage points and destinations in the *customer cones* of the targeted ASes, i.e., the set of ASes reachable from that AS by traversing provider-customer links. We will expand on the set of techniques used in [6] in two ways: a more recent and more validated customer cone computation algorithm [49], and an expanded set of vantage points including CAIDA's 77 (as of November 2013) Ark monitors [108], limited use of thousands of RIPE Atlas nodes [109], Akamai nodes, and Mlab servers (see attached letters of collaboration) and more than a thousand public traceroute servers (listed at traceroute.org).

Section 3.2 describes our recent successful attempt to use traceroute data to map Comcast's observable peering connectivity from a given vantage point, all of which is at private peering facilities. But measurement techniques to discover bilateral peering links have never been implemented at the scale we propose, from thousands of distributed vantage points, so collecting, processing, and interpreting this data will involve significant computational and data mining challenges. Our extensive experience in analysis of traceroute data has made us well aware of the pitfalls in inferring AS-level connectivity from traceroute data, requiring careful sanitization to handle artifacts such as third-party addresses and the effects of unresponsive routers [110, 111, 112, 113, 114, 115, 52].

Using traceroute data to infer connectivity at IXes is easier when the IX prefixes are known and appear in the traceroutes, in which case they can be simply elided from the traceroute path and connectivity inferred across them [116, 6]. But inferring peering at unknown IXes is an open research challenge. We will begin by examining identifiable patterns of known IXPs, either host-name conventions, IX prefixes, or topological characteristics, that might yield hints to help us identify undocumented IXes in our traceroute data. We have had success with a similar fingerprinting approach, based on unusual degrees of routers in paths, to infer layer-2 MPLS infrastructure from traceroute data [117]. The final step is to match observed IX signatures to likely geographical locations. For this purpose, we will try to geolocate the suspected IX hops using a variety of techniques: reverse DNS mappings of neighboring IP addresses [118], off-the-shelf geolocation databases [119], or delay and topology-based geolocation [120, 121, 122, 123, 56]. With an estimate of the location of the inferred IX, we can determine if the IX we see in the trace corresponds to IXes for which we know the precise geographical locations (some of which are documented in PeeringDB), or if it corresponds to an undocumented one.

⁶N. Chatzis offered us a code-to-data approach to accessing the European IXP traffic data used in [2].

3.1.2 Inferring complex routing relationships

For current AS-level topology maps, we have developed and validated heuristic techniques to classify nodes into network types [7], and other heuristics to estimate which nodes are part of the same parent organization (“siblings”) [124]. We have also recently improved and heavily validated our algorithm for inferring one of two business relationships between ASes [49], although as described in Section 2, these simplified inference techniques fail to capture much of the complexity in the routing ecosystem. The new data we have gathered will enable us to untangle some of this complexity, improving current inference techniques and enriching the annotations on our map at the same time. To improve node classification, we will include additional inputs to the machine learning classifier in [7] (in addition to customer and peer degrees, we will use customer cone size and advertised address space), as well as a much larger training data set of ground-truth AS classifications from PeeringDB (which has self-reported AS types for more than 4000 ASes). We also regularly receive AS type and sibling ground truth data via our AS-rank interface [125], which we use to continually improve our sibling inferences [124].

To improve link classification we will build on a feature of our AS relationship algorithm [49] that begins to address complex relationships between ASes: the *provider-peer observed customer cone*. In a complex peering relationship, AS X may purchase transit from AS Y in some regions or prefixes, but be peers elsewhere, meaning that Y will not announce customer routes to peers or providers outside of the agreed transit region. Our IX-aware map will allow us to identify and classify some region-specific p2p and p2c relationships that we had no way to classify before. We will explore two methods to infer region-specific relationships. First, for each provider (say Y) of X, we will search for X’s customers that Y announces. If Y only announces some of X’s customers from a particular geographical region to peers and providers, it might indicate a p2c relationship in that region, but a p2p relationship elsewhere. We will use WHOIS and the best available IP geolocation databases to search for geographic trends in observed BGP announcements. Second, our multilateral peering inferences promise insight into region-specific peering behavior. For example, we have found that 69.2% of ASes that self-report a restrictive policy in PeeringDB are actually open in establishing peerings at route servers, though only in certain regions [5]. Such differences indicate the potential to observe complex relationships, such as an AS peering openly in region A but restrictively (e.g., as a provider) in region B.

3.1.3 Validating the IX-aware map

We will pursue at least six techniques to enable us and others to validate the hundreds of thousands of peering links we expect to uncover and annotate. First, we will assign a *confidence* level to each link observed in traceroute data, based on a range of criteria, e.g., peering seen in both directions, whether we could map an IXP address to a specific member AS on one end of a peering link, etc., similar to the methods described in [6]. Second, we will cross-validate inferences from four data sources: BGP tables, multilateral peering inference technique, targeted traceroutes, and peering matrices published directly by the IX (e.g., [103]). Third, we will publish a list of looking glasses hosted in ASes nearby (if not at) IXPs, which can be used to query prefixes for the existence of a peering link; we will automate this querying ourselves to validate the accuracy of our multi-lateral peering inferences over time. Fourth, we will integrate information we learn into our AS-Rank website [125], which publishes our AS relationship inferences and invites corrections from owners of ASes involved in peering. This interactive repository will allow us to solicit feedback directly from network operators regarding our IX map and associated peering inferences, including region-specific ones. Fifth, some BGP community values are currently described

with region-specific annotations [126], which allows us to test our classifications of region-specific relationships. Finally, we will also explore new and creative validation methods by combining information from multiple sources to obtain hints about the likely existence of peering links. For instance, if two networks advertise an open peering policy and claim presence at the same IX, we have high confidence that a peering link we have previously inferred between those ASes is likely to exist.

3.2 Task 2: Inferring congestion at interconnection points

Our goal in this task is to combine the topological data from Task 1 with path performance measurements to detect and localize evidence of performance degradation in the interior of the global Internet. In particular we seek to determine whether and to what extent inter-AS connection links are congested, since they reflect potentially contentious business relationships between ISPs, and associated regulatory concerns about the exercise of market power. We have developed three experimental approaches for detecting evidence of congestion: two based on active measurements of latency and throughput, and one using passive traffic statistics collected by a major content provider. This third method will enable us to overcome a complication with measuring congestion at interconnection points: intelligent coding of video for real-time streaming.

3.2.1 Delay-based detection of congestion

Our first method for detecting congestion involves sending a crafted sequence of pings along the path in question, a method we call *time-sequence ping* or TSP. A single round-trip time delay (ping) measurement does not provide strong evidence of congestion, but a diurnal variation in delay may be an indicator. Delay-based detection of congestion relies on two assumptions: (1) that it is rare for a link to be congested continuously, and (2) that router buffers will fill when an outbound link experiences a threshold of congestion, increasing round trip time (RTT) for packets crossing that link.⁷ The first assumption is consistent with several studies documenting strong diurnal variation in Internet traffic [128, 129, 130, 131, 132], where peak traffic exceeds minimum traffic by a factor of 3 or 4. With these parameters, a continuously congested link would be provisioned to carry no more than 25% to 33% of peak traffic. We believe that links that are so persistently and severely underprovisioned in the core of the Internet would be notorious on operational mailing lists. We therefore assume in most

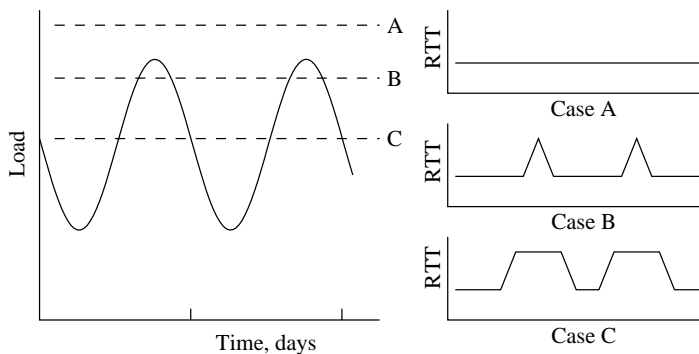


Figure 3: *Idealized representation of load variation over time, and anticipated behavior of RTT for different link capacities. The three horizontal lines represent links with different capacities. Capacity A can handle all peak traffic and we expect no RTT-based signal of congestion. Capacity B becomes congested at the peak, and shows a brief period of elevated RTT. Capacity C is a substantially underprovisioned link, and shows congestion for several hours a day, with a corresponding step function in RTT.*

⁷Different queue management schemes result in different queue behavior, with tail-drop schemes leading to full queues, and active queue management schemes trying to reduce the average queue size during congestion, e.g., RED, CODEL [127].

cases of congestion we will see a diurnal variation in RTT that correspond with traffic loads, as illustrated in Figure 3. Peak loads typically occur in early-mid evening for consumer broadband (see Figure 8 in [132]).

The first goal of this task is to verify that we can detect this kind of variation by measuring delay (round trip time, or RTT) over time, since other factors can contribute to variation in RTT. We ran an experiment using a CAIDA Archipelago (Ark) [108] active monitor at a Comcast residential location in Boston. Each Ark monitor performs continuous traceroutes to all /24 routed IPv4 address blocks, data that we used to determine peering and transit links out of the Comcast network from this location.⁸ We extracted the destinations reached by these background traceroutes, as well as the distance (in hops) from the monitor in Boston to the point of interconnection between Comcast and each of its transit providers and peers. We then sent a time-series of pings (TSP) to measure RTT toward each destination, with the TTL of each packet set to expire at the near end and far side of the interconnection link. If there is no diurnal variation in RTT for the near side of the link, but there is diurnal variation for the far side of the link, then we infer the interconnection link is likely experiencing congestion.

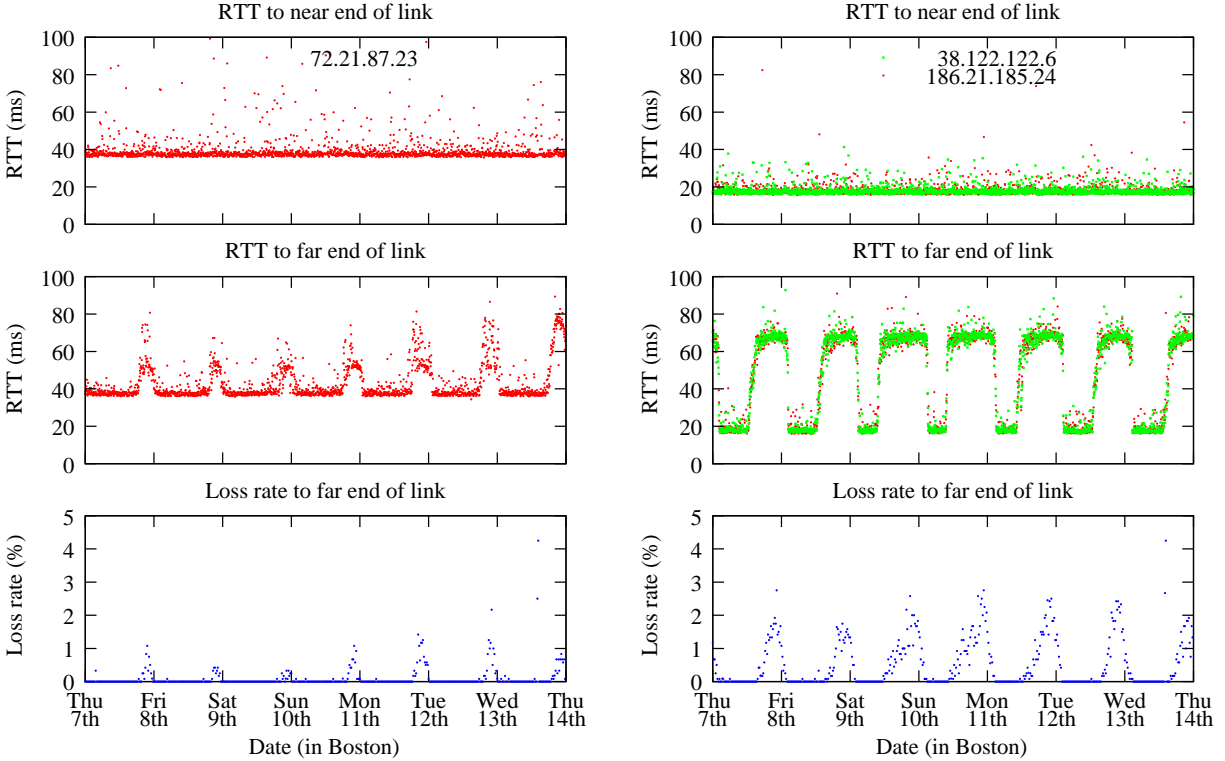
The experiment was encouraging, revealing clear indications of congestion on certain links. We probed over 100 Comcast inter-AS links from Comcast’s Boston region; most links exhibited no indication of congestion, consistent with proper provisioning. Figure 4(a) shows a case with evidence of slight congestion for a brief period, presumably only at peak traffic time. Figure 4(b) shows a link with congestion that persisted for several hours each day; this link connects Comcast to Cogent. Conversation with a Comcast engineer confirmed that this link is heavily congested as a result of a recent decision by Netflix to reroute all their incoming streaming traffic from Level3 to Cogent, overloading links not engineered for this traffic. This situation illustrates that big content providers sometimes have more control than ISPs over how traffic is routed, making sensible traffic engineering difficult for ISPs. In Figure 4(b), one can even estimate the size of the buffer by the change in RTT, and see that the observed congestion lasts longer on weekends.

To develop more confidence in the RTT signature as a legitimate signal of congestion, for the links in Figure 4(b) we also sent TTL-limited probes once a second to measure loss rate across a larger sample. The bottom figures show that the loss correlates with the RTT-derived congestion signal; when there is no congestion, there is no loss. However, the loss rate increases even when the RTT does not grow, which we hypothesize is due to the presented load increasing. These experiments convinced us of the potential for detection of congestion based on a time-series of delay measurements.

3.2.2 Throughput-based detection of congestion

Our delay-based method to detect congestion can probe any path that has a ping responder at the other end; no special server hardware is needed. But it cannot test for instantaneous congestion, and it cannot distinguish the direction of the congestion, so it may make incorrect inferences, especially if the reverse path is not symmetric. We propose a complementary technique based on measuring TCP throughput, specifically a *rate-limited TCP download probe* (RLTP), which can detect congestion from a single test and should be able to measure each direction separately, although we need to experimentally confirm this last assumption. A tool that reveals congestion in a single test may be more suitable for a broad set of Internet users to run, thus increasing coverage. However,

⁸These background measurements themselves provide a good opportunity to observe signals of congestion for links commonly traversed. But many links are rarely traversed by these measurements, e.g., for Comcast, half of the links we studied (including for some underprovisioned links to content providers) were sampled fewer than ten times across the day.

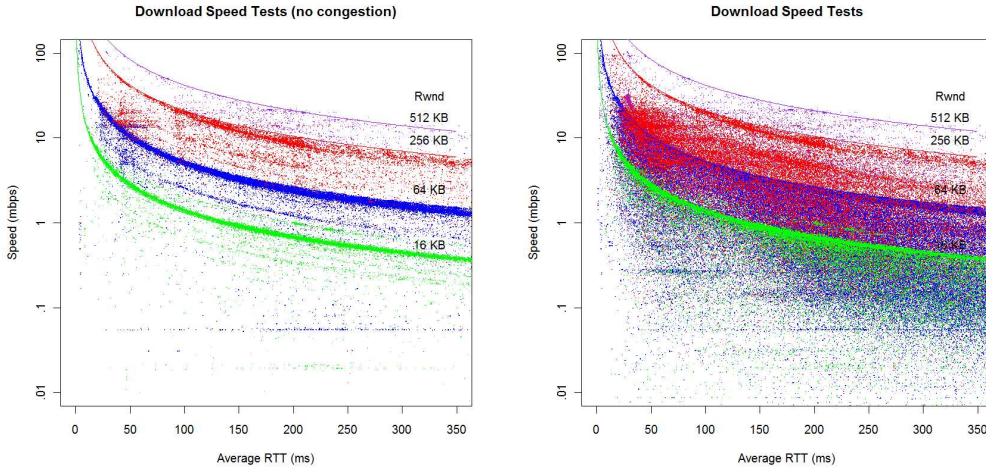


(a) Delay-based method detecting slight congestion between Comcast and EdgeCast (AS15133) in Texas. (b) Delay-based method detecting serious congestion between Comcast and Cogent (AS174) in New York.

Figure 4: Demonstration of feasibility of delay-based method for detecting congestion, using traces for two inter-AS paths leaving the Comcast Boston serving area. Time is EDT, so the peaks correspond to typical peak load times on the east coast in the evening. Top graphs: RTT from the near side of each interconnection link. Middle: RTT from the far side of the same link. Bottom: loss rate from the far side of the same link. The height of the RTT waveforms (middle graphs) corresponds to the size of the buffer, measured in holding time. The height of the loss waveforms (bottom graphs) correspond to the presented load on the buffer, which grows after the link is congested. Weekend days show longer congestion periods than weekdays, except for Monday 11 November (Veterans day).

this test requires a client and a server; it cannot test to different points along the path in order to isolate the point of congestion.

RLTP pre-selects a rate that does not congest the access link to the destination, so that any detected congestion during a TCP download must be elsewhere along the path. Our confidence in this method derives from our earlier experiments with data from the MLab deployment of the Network Diagnostic Test (NDT) [133] tool, which collects hundreds of thousands of daily tests from users trying to evaluate their network connection. Limitations of the resulting NDT data [68, 66] inspired our experimental design. For example, while NDT tests are not intentionally rate-limited, we discovered that many are *unintentionally* rate-limited by insufficiently large TCP receive windows set by many clients' operating systems [66]. These are effectively rate-limited downloads that are suggestive of the rich congestion characteristics one can discover with our proposed test. Figure 5 plots speed vs. round trip time for 380,000 NDT experiments, with colors indicating receive window ($Rwin$) values. Figure 5(a) shows the fraction (one-third) of the tests that ran to completion without receiving any congestion signal; throughput clusters tightly



(a) Experiments with no congestion signals (b) Experiments that encountered congestion

Figure 5: NDT-observed download rate as a function of round trip delay, for tests that did not and did encounter evidence of congestion. For tests that completed without receiving any congestion signal, throughput clusters tightly around the theoretical limit for a given window size and round trip delay. For tests that experienced congestion signals, throughput is dispersed. Without path information we cannot speculate on the correlation of performance degradation with interconnection relationships.

around the theoretical limit for a given window size and round trip delay. In contrast, Figure 5(b) shows the dispersion in throughput when congestion is present. Unfortunately, path information was not being collected for these measurements, so we can only speculate on the relationship between the observed congestion signals and interconnection points along the unrecorded paths. Our proposed active probing approach addresses this limitation.

We will have a client trigger a rate-limited download (e.g., 5 mb/s) from a nearby server. If that download does not show any signals of congestion, we conclude that the path from the client to the server is not congested, and thus we can effectively test other paths. We will then have the client perform an identical set of rate-limited downloads from servers in other ASes, attempting to cross interconnection points on the map we constructed in Task 1. The traceroutes collected in both directions (from the test client to server, and back) during each test, together with the congestion signal obtained from the test, will be an input to a binary tomography algorithm (Section 2.2) that will determine the most likely (topological) location of congestion. Our IX-aware map will further reveal whether those links are internal to an AS, or at IXes.

Our RLTP test is conceptually identical to the *constant window pseudo CBR* (constant bit rate) test proposed by Matt Mathis of Google as part of his broader Model-Based Internet Performance Metrics (MBM) [134] effort. Recognizing this alignment of interests we initiated a collaboration (see attached letter of support from Matt Mathis) to develop and validate this type of test to detect congestion along paths from access ISPs to Mlab servers around the world [135]. Our RLTP experiment requires dedicated client and server side code. We expect to utilize Mlab and our own Ark infrastructure as both clients and servers to test and evaluate this approach; we will share results with collaborator Google to integrate back into their open source MBM tool [136].

3.2.3 Passive traffic-based detection of congestion

Many sources, including Netflix, adapt to congestion by downgrading to a lower resolution coding [137], so the link may not look overloaded as measured by drops, but the quality of experience (QoE) may be substantially degraded. In such cases, both delay-based and throughput-based methods may fail to detect congestion, even while the user experience is degraded. When sources use such adaptive rate coding, the best signal of congestion is that the protocols have downgraded to a lower transmission rate (rather than, e.g., queuing delay or packet drops), most easily observed at the endpoints of the flow, e.g., the CDN or end-user. Download performance on an uncongested path will show a tight clustering of transfer rates around the intrinsic speeds of the different codings; downloads on congested paths will show a dispersion of the achieved speeds around the native encoding speeds (as illustrated in Figure 5).

Our third technique uses exactly such traffic transfer rate data, derived from traffic flow statistics logged by a large CDN. Netflix has agreed (see attached letter) to share per-download data on the transfer rate and time of selected file downloads, along with source and destination address (anonymized to a /24 granularity to protect privacy but allow inference of congestion along paths from their distributed servers to consumers in access networks across the globe). We have already received aggregated data from Netflix that shows clear evidence of diurnal congestion (see figure 6). Our research objective is to see what further structure can be uncovered from the detailed data.

We are not concerned with how current rate adaptation schemes can perform in the presence of congestion [96], but with the sensitivity with which we can detect congestion. The challenge will be to map the data points onto paths, and reason about the location of detected congestion. Although we can again try to use the binary tomography technique discussed in Section 2.2, a simpler approach will be to use this technique in conjunction with the delay-based method (Section 3.2.1), which uses traceroutes and pings to detect the point of congestion along a path. We can compare the congestion signal we get from the delay-based measurements with the passive download performance, and may even be able to find evidence that the passive download data starts to react to congestion even before we see significant queues building up.⁹ We will also use active probing vantage points as available (Ark/RIPE Atlas vantage points and traceroute servers) to determine interconnection paths between the Netflix CDN server and the ISP of the client receiving the video. Once we infer the likely path from Netflix into the access ISP, we can use the congestion indicators provided by Netflix to infer the presence of congestion on interconnection points.

3.2.4 Validation and automation of congestion-related inferences

Initially, we will use these methods to cross-validate each other, and work with ISPs willing to validate our inferences, as with the congested link between Comcast and Cogent. In addition, there are several well-established tools, techniques, and data sets that we can use as validation and

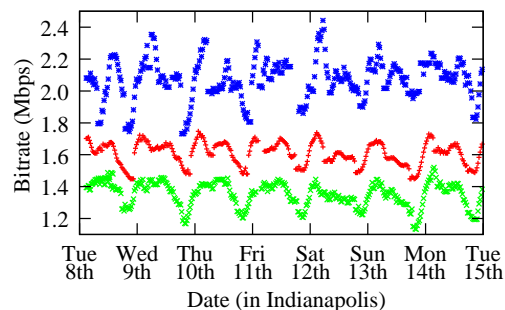


Figure 6: Aggregate bitrate served by Netflix to users of three ISPs in the Indianapolis region. During congested periods, the content bitrate reduces.

⁹Netflix has changed its approach to content distribution several times, as they move from third-party CDNs to their own OpenConnect devices, many of which are currently hosted on two transit networks: Cogent and Tata. In cases where their OpenConnect devices are inside the access network, the data will not cross any interconnection links.

diagnostic aids. Relevant tools include those that test link capacities and identify bottlenecks [138, 139, 140, 141, 142]. Some public network data sets may also be of use; for example, every TCP test hitting the Mlab platform generates Web100 TCP data [143] and a traceroute (recently enabled) to the test client that could provide additional validation of our results.

One open issue we have considered is that our delay-based detection method cannot determine in which direction the congestion is occurring. The congestion could be occurring on the reverse path, and since we do not have traceroute data from destination to source (since we do not control the remote destinations responding to ping), the congestion may be on a different link than the one revealed by the outgoing traceroute. To reduce the probability of an asymmetric path, our delay-based measurement method terminates the RTT probing at the router immediately beyond the interconnection link of interest. Figures 4(a) and 4(b) show nearly identical RTT to both near and far sides of the interconnection during off-peak times, suggesting the response from the far-side of the interconnection is not taking a circuitous route. We believe our approach gives reasonably unambiguous results, but part of this research task will be to validate this assumption. Identification of asymmetric routes is important since it may reveal additional paths not found by the outgoing traceroutes.

We can use at least two approaches to detect and resolve asymmetric routes. The first approach is to use a tool for reverse traceroute [144], which incrementally pieces together the reverse path using measurements from vantage points close to routers on the reverse path. Reverse traceroute builds on two capabilities: (1) the IP record route option, where routers embed one of their IP addresses into packets as they forward them, and (2) the ability to spoof the source address. Reverse traceroute sends echo requests with the record route option to routers on the reverse path, spoofing as our source that detected congestion with TSP so that the echo reply may collect IP addresses on the reverse path. Spoofing packets using a distributed infrastructure is a sensitive issue, so we will first try a much simpler method to get the reverse path. We are primarily interested in congestion that occurs at the interconnection points of networks hosting our vantage points, and thus we are likely to be within the nine hops afforded by the record route option. Therefore, we can simply send an echo request to the far side of the interconnection with the record route option set, in order to observe at least the first few hops on the path back to our source.

The second approach will rely on our planned collaboration with Akamai, which has agreed to give us traceroutes from over 1000 distinct CDN locations to Ark nodes, which means we can obtain paths in both directions. We can also use these nodes as targets, and our Ark nodes as sources, in our delay-based detection experiments. These measurements will give us a set of end-to-end measurements along with paths, which we can feed into a binary tomography algorithm (Section 2.2) to localize observed congestion. The limitation of this approach is that while Akamai has many test sites, there may not be a test site on the other side of every interconnection point of the network hosting our vantage point.

A final issue is scaling our methods to operate on Ark nodes in many access networks, each with many paths to the rest of the Internet.¹⁰ We need automated methods to detect and analyze congestion in huge volumes of raw data. We must also automate as much as possible the parameterization of the experiment: using our map to find routers along the path of interest, and selecting ping targets among them.

¹⁰We are using NSF CRI funding to expand the Ark infrastructure to include additional residential access links.

3.3 Task 3: Exploring implications for network resiliency, policy, and science

The development and application of these new methods and data will provide a new lens through which to observe and understand the evolving Internet. Our results will move the field beyond traditional notions of AS graphs as consisting of nodes and links, toward a more detailed characterization that includes multiple connections between ASes, regional semantics in business relationships, and a characterization of where congestion is likely to occur. Given the knowledge base and tools created in our foundational tasks, we will pursue several research studies that explore infrastructure resilience, inform communications policy, and improve scientific modeling capability. We will also make the bulk of our collected data available to other researchers.

How do IXes influence Internet resiliency?

Modeling an AS as a single node ignores the redundancy obtained from multiple, geographically diverse connections between ASes [51], and such simplified models have led to questionable conclusions about the “robust yet fragile” nature of the Internet infrastructure [145, 146]. Our new map will include information about hundreds of thousands of peering links, including multiple connections between ASes, which will inform our investigation of the nature of Internet resiliency, including the importance of certain ASes and AS-links in maintaining global or local reachability. As heavily aggregated points of connection, IXes are themselves potential vulnerabilities, since the failure of an IX, due to natural disaster or deliberate attack, might severely disrupt connectivity. To mitigate this risk, many large IXes, e.g., AMS-IX, DE-CIX and LINX, are decentralized within a metro-area, i.e., they support multiple physical locations in one virtual distributed exchange, which adds richness and complexity to the nature (and study) of availability. We will investigate whether networks tend to connect at multiple IXes in a city/region (for redundancy) or in different regions (for geo-diversity), and the networks they connect to at various IXes. Our map of inter-AS connections cannot determine the total route diversity of an ISP, because neither traceroutes nor BGP data can reveal paths that exist but are not used. However, our analysis of the tendencies of different sorts of ISPs to seek diverse connections to different IXes will allow us to characterize the degree to which the loss of a major IX might potentially isolate or impair regions of the Internet.

Do regional differences in peering behavior produce choke points or routing inefficiencies?

Many characteristics of the peering ecosystem – multihoming trends, business models [7, 147], transit prices [148], and the prevalence of IXes – vary by geographic region, and manifest themselves in region-specific peering idiosyncrasies. Using our new map, we can compare topological structure of different countries or regions and the role of IXes in improving topological diversity. A few studies have characterized topological structure of specific countries (e.g., China [149] and Germany [150]), but none have had the data for geographic analysis of connectivity, and the role of IXes. The map will also facilitate analysis of the extent to which some countries, regions, or organizations are essential hubs for connecting others to the global Internet [151], such as how the U.S. has served as a hub even for ISPs within the same Asian country [152]. We will use BGP or WHOIS allocation data to determine IP address blocks in specific countries or regions, identify the set of ASes that control these IP address blocks, and also the set of organizations that control ASes in a given country/region. Our IX-aware topology map can augment this view with lower-level information about *where* connectivity between two ASes is established: at an IX or not, and within or outside the region of interest. We will also be able to identify points of regional control of Internet infrastructure (“choke points”), i.e., ASes or IXes that operate most access links into/out of that region. Finally, the map will reveal inefficiencies, such as circuitous routes due to peering issues or unavailability of IXes to facilitate local interconnection.

How can our research inform growing policy concerns such as network transparency, investment incentives, and market power?

The FCC's Measuring Broadband America effort [64] focused almost exclusively on the character of the access links into different ISPs: speed, latency, reliability. Our proposed research will yield a complementary view of the network: the degree and character of the connections between an ISP and the rest of the Internet. Persistent under-provisioning of interconnection links may trigger increased regulatory attention to peering practices. The FCC-sponsored measurements focused on 13 ISPs in the U.S., all of which we intend to study in this project. To compare practices across regions, we will also measure access ISPs in other countries. We expect that the pattern of peering and transit will strongly correlate with the presence of IXes in the region. One research outcome will be a presentation of our data in a form that is accessible and useful to policy-makers. By presenting actual data on interconnection, we can shed light on the question of whether *increased transparency* would avoid the need for more invasive forms of regulation.

There is also concern in business and regulatory circles that ISPs may not have the proper economic incentives to continue to *invest in network infrastructure* [153]. This concern has led to a perception, especially in the developing world, that given the typically higher profitability and leaner capital structure of content providers (and their CDNs) relative to network infrastructure providers, regulatory intervention should seek to mitigate this imbalance, and require payment from CDNs to deliver content into access networks. While highly charged and political, these debates about payment and obligations for interconnection are largely uninformed by the realities of the Internet [154], in particular the ability of content providers to control the origin of content. CDNs can pick paths into the country that have quite different consequences for the cost structure of the domestic network. Our research, properly packaged for regulators, will help avoid unexpected and undesirable policy outcomes. In particular, we intend to extend some of our past modeling work on fair and stable peering settlements [155]. That model quantified peering settlements in terms of the *value* of the link (peering relationship) to each party, where value was defined in purely economic terms. We propose to enrich that model by assigning a performance attribute to the value of a link. Rather than abstract notions of performance such as AS-level path lengths, we will model performance degradations arising due to congested interconnections.

A third controversial policy debate we hope to inform with quantitative metrics derived from our measurements is whether different types of network actors have significant market power. In the past, tier-1 backbone providers were thought to have market power if they were bottleneck paths between many other ISPs. Today it is more common to hear concerns that access providers have market power since they provide exclusive access to their customers by other ISPs (and content providers) that wish to reach those customers. As an indicator of Internet market power, researchers have proposed a simple topological metric such as betweenness (number of paths that go through a given AS node) [156, 157], which is problematic even today, since it does not account for traffic flow and possible alternative paths. But an ecosystem with dominant content providers and access providers will require different metrics for market power, since both types of providers tend to have low betweenness (because they are almost always the first or last hops on an AS path, and customers of access providers generally have no alternate path). Our prior work [158] suggested that if an ISP implemented few interconnection options, and then exhibited signs of congestion on some links, it could be considered evidence of the exercise of that market power. We will examine possible correlations between conventional *metrics of market power* in network industries, such as the number of customers and market share of broadband providers [159], and the fraction of intervals per day when we observed congestion on the interconnection links of those providers.

How valid are modeling assumptions about topological invariants and routing policies?

Measuring and modeling the Internet topology, particularly at the AS-level, has been an active area of research over the last decade, starting with the work of Faloutsos et al. [160], which claimed the existence of power-laws in the Internet topology. Despite subsequent findings that questioned the power-law nature of the Internet topology [161], several topological properties of the AS-level Internet are considered invariants – a heavy-tailed degree distribution that is close to a power-law, strong clustering, assortativity, etc. Much network science research on topology modeling has relied on these properties as part of the validation process [162, 163, 164, 165, 146, 166], though a model’s ability to produce a power-law degree distribution should be viewed only as a sanity check [161, 167, 168]. Similarly, we know that the “valley-free, prefer-customer, then prefer-peer” routing policy [42] is overly simplistic in the real-world, as is the assumption about binary (customer-provider or peer-peer) relationships between ASes [50, 51]. There has never been a comprehensive validation of these assumptions, or even an assessment of how much deviation from these assumptions matters for modeling; the data to do so has not been available.

As we expand our techniques to detect MLP links at as many IXes as possible, we will produce an AS graph that is an order of magnitude larger than those previously analyzed (e.g., [39, 169]), and is both a multi-graph and a hypergraph. The resulting data will require a re-thinking of the definition and significance of a number of topological properties (degree distribution, clustering, assortativity, etc.), with enormous implications for *network science*, which has thus far treated the Internet AS-level topology as just another graph to model. Furthermore, inferring region-specific AS-relationships will illuminate the prevalence of complex relationships in the real world. Using our mechanism for computing AS customer cones, we can quantify the extent to which the “valley free, prefer-customer, prefer-peer” routing policy assumption is violated, and the ASes most likely to produce such violations. These studies will inform our development of more realistic and empirically grounded models for AS business relationships and interdomain routing policies. The data sets that result from this research will also serve as more realistic topologies for simulation and analysis. We (or others) can distill statistical properties from the IX-aware map that are relevant to different modeling activities: colocation and peering behavior of different AS types, distributions of the number of physical locations at which two ASes connect, the scope and nature of region-specific business relationships, etc. We will use these properties to construct *synthetic* maps that are conducive to simulation or model-based analysis, potentially scaled down for usability while retaining the statistical properties of the original [170].

Can we detect trends over time in interconnection and performance?

Finally, we recognize the need to develop, maintain, and archive classic data sets to develop some sense of Internet history, which industry players have no incentive to gather and archive. Our ongoing measurement of peering and congestion will allow us to track trends over time in the practices and effects of interconnection. One might expect to see a pattern of gradual emergence of load on a link, followed by a reconfiguration to alleviate the congestion. In our limited observations we have already seen events like the sudden emergence of serious congestion on a link, perhaps as the result of a change in CDN delivery or a failure in some other part of the net shifting load onto this link. Capturing short-term dynamics as well as evolutionary trends in the interaction between the current CDN-driven content delivery and IX industries will not only enlighten studies of today’s Internet, but serve as a use case by which to evaluate or simulate future Internet architectures.

References

- [1] P. Gill, M. Arlitt, Z. Li, and A. Mahanti, "The Flattening Internet Topology: Natural Evolution, Unsightly Barnacles or Contrived Collapse?," in *Proceedings of the 9th International Conference on Passive and Active Network Measurement (PAM)*, 2008.
- [2] B. Ager, N. Chatzis, A. Feldmann, N. Sarrar, S. Uhlig, and W. Willinger, "Anatomy of a Large European IXP," in *Proceedings of ACM SIGCOMM*, 2012.
- [3] C. Labovitz, S. Iekel-Johnson, D. McPherson, J. Oberheide, and F. Jahanian, "Internet Inter-domain Traffic," in *Proceedings of ACM SIGCOMM*, 2010.
- [4] "PeeringDB." <http://www.peeringdb.com>, October 2011.
- [5] V. Giotsas, S. Zhou, M. Luckie, and K. Claffy, "Inferring Multilateral Peering," in *Proceedings of ACM CoNEXT*, 2013.
- [6] B. Augustin, B. Krishnamurthy, and W. Willinger, "IXPs: Mapped?," in *Proceedings of the ACM SIGCOMM Internet Measurement Conference (IMC)*, 2009.
- [7] A. Dhamdhere and C. Dovrolis, "Twelve Years in the Evolution of the Internet Ecosystem," *IEEE/ACM Transactions on Networking*, vol. 19, no. 5, 2011.
- [8] 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 (CoNEXT)*, pp. 217–228, 2009.
- [9] Sandvine, "Global Internet Phenomenon Report, 1H13," 2013.
- [10] Center for Democracy and Technology, *ETNO Proposal Threatens to Impair Access to Open, Global Internet*, 2012. https://www.cdt.org/files/pdfs/CDTAnalysis_ETNO_Proposal.pdf.
- [11] "France Telecom severs all network links to competitor Cogent," 2005. <http://morse.colorado.edu/~epperson/courses/routing-protocols/handouts/cogent-ft.html>.
- [12] S. Cowley, "ISP spat blacks out Net connections," 2005. <http://www.infoworld.com/t/networking/isp-spat-blacks-out-net-connections-492>.
- [13] M. Ricknas, "Sprint-Cogent Dispute Putes Small Rip in Fabric of Internet," 2008. http://www.pcworld.com/article/153123/sprint_cogent_dispute.html.
- [14] Backdoor Santa, "Some Truth About Comcast - WikiLeaks Style," 2010. <http://www.merit.edu/mail.archives/nanog/msg15911.html>.
- [15] J. Engebretson, "Level 3/Comcast dispute revives eyeball vs. content debate," Nov. 2010. <http://www.telecompetitor.com/level-3comcast-dispute-revives-eyeball-vs-content-debate/>.
- [16] J. Engebretson, "Behind the Level 3-Comcast peering settlement," July 2013. <http://www.telecompetitor.com/behind-the-level-3-comcast-peering-settlement/>.
- [17] Verizon, "Unbalanced peering, and the real story behind the verizon/cogent dispute," June 2013. <http://publicpolicy.verizon.com/blog/entry/unbalanced-peering-and-the-real-story-behind-the-verizon-cogent-dispute>.
- [18] J. Brodtkin, "Why YouTube buffers: The secret deals that make and break online video," July 2013. <http://arstechnica.com/information-technology/2013/07/why-youtube-buffers-the-secret-deals-that-make-and-break-online-video/>.
- [19] S. Buckley, "Cogent and Orange France fight over interconnection issues," 2011. <http://www.fiercetelecom.com/story/cogent-and-orange-france-fight-over-interconnection-issues/2011-08-31>.

- [20] "YouTube sucks on French ISP Free, and French regulators want to know why," 2013. <http://gigaom.com/2013/01/02/youtube-sucks-on-french-isp-free-french-regulators-want-to-know-why/>.
- [21] S. Buckley, "France Telecom and Google entangled in peering fight," 2013. <http://www.fiercetelecom.com/story/france-telecom-and-google-entangled-peering-fight/2013-01-07>.
- [22] J. Brodtkin, "Time Warner, net neutrality foes cry foul over Netflix Super HD demands," 2013. <http://arstechnica.com/business/2013/01/timewarner-net-neutrality-foes-cry-foul-netflix-requirements-for-super-hd/>.
- [23] H. Schulzrinne, W. Johnston, and J. Miller, "Large-Scale Measurement of Broadband Performance: Use Cases, Architecture and Protocol Requirements." <https://datatracker.ietf.org/doc/draft-schulzrinne-lmap-requirements/>, September 2012.
- [24] C. Hall, R. Clayton, R. Anderson, and E. Ouzounis, "Inter-X: Resilience of the Internet interconnection ecosystem," Apr 2011.
- [25] G. Siganos, M. Faloutsos, and C. Faloutsos, "The Evolution of the Internet: Topology and Routing," *University of California, Riverside technical report*, 2002.
- [26] R. V. Oliveira, B. Zhang, and L. Zhang, "Observing the Evolution of Internet AS Topology," in *Proceedings of ACM SIGCOMM*, 2007.
- [27] P. Gill, M. Schapira, and S. Goldberg, "Let the market drive deployment: A strategy for transitioning to bgp security," *SIGCOMM Comput. Commun. Rev.*, vol. 41, pp. 14–25, Aug. 2011.
- [28] R. Lychev, S. Goldberg, and M. Schapira, "Bgp security in partial deployment: Is the juice worth the squeeze?," *SIGCOMM Comput. Commun. Rev.*, vol. 43, pp. 171–182, Aug. 2013.
- [29] "University of Oregon Route Views Project." <http://www.routeviews.org/>.
- [30] "RIPE routing information service (RIS)." <http://www.ripe.net/ris/>.
- [31] B. Zhang, R. Liu, D. Massey, and L. Zhang, "Collecting the Internet AS-level Topology," *SIGCOMM Computer Communication Review*, vol. 35, pp. 53–61, Jan 2005.
- [32] H. Chang and W. Willinger, "Difficulties Measuring the Internet's AS-Level Ecosystem," in *Proceedings of the 40th Annual Conference on Information Sciences and Systems*, 2006.
- [33] H. Chang, R. Govindan, S. Jamin, S. Shekher, and W. Willinger, "Towards capturing representative AS-level Internet topologies," *Computer Networks*, vol. 44, no. 6, pp. 737–755, 2004.
- [34] R. Oliveira, D. Pei, W. Willinger, B. Zhang, and L. Zhang, "In Search of the Elusive Ground Truth: The Internet's AS-level Connectivity Structure," in *Proceedings of ACM SIGMETRICS*, 2008.
- [35] M. Roughan, S. J. Tuke, and O. Maennel, "Bigfoot, Sasquatch, the Yeti and Other Missing Links: What We Don't Know About the AS Graph," in *Proceedings of ACM SIGCOMM IMC*, pp. 325–330, 2008.
- [36] 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*, vol. 18, pp. 109–122, Feb. 2010.
- [37] D. Weller and B. Woodcock, "Internet Traffic Exchange: Market Developments and Policy Challenges," *OECD Digital Economy Papers*, no. 207, 2006.
- [38] Y. He, G. Siganos, M. Faloutsos, and S. V. Krishnamurthy, "A Systematic Framework for Unearthing the Missing Links: Measurements and Impact," in *Proceedings of USENIX/SIGCOMM NSDI*, 2007.
- [39] P. Mahadevan, D. Krioukov, M. Fomenkov, B. Huffaker, X. Dimitropoulos, k. claffy, and A. Vahdat, "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.

- [40] 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 Trans. Netw.*, vol. 17, pp. 391–404, Apr. 2009.
- [41] M. A. Sánchez, J. S. Otto, Z. S. Bischof, D. R. Choffnes, F. E. Bustamante, B. Krishnamurthy, and W. Willinger, "Dasu: pushing experiments to the internet's edge," in *Proceedings of the 10th USENIX conference on Networked Systems Design and Implementation (NSDI)*, pp. 487–500, 2013.
- [42] L. Gao, "On inferring autonomous system relationships in the Internet," *IEEE/ACM Transactions on Networking*, vol. 9, pp. 733–745, Dec. 2011.
- [43] L. Subramanian, S. Agarwal, J. Rexford, and R. H. Katz, "Characterizing the internet hierarchy from multiple vantage points," in *IEEE INFOCOM*, (New York, USA), pp. 618–627, June 2002.
- [44] G. D. Battista, T. Erlebach, A. Hall, M. Patrignani, M. Pizzonia, and T. Schank, "Computing the types of the relationships between autonomous systems," *IEEE/ACM Transactions on Networking*, vol. 15, pp. 267–280, Apr. 2007.
- [45] X. Dimitropoulos, D. Krioukov, M. Fomenkov, B. Huffaker, Y. Hyun, and K. Claffy, "AS Relationships: Inference and Validation," *Computer Communication Review*, vol. 37, pp. 29–40, Jan. 2007.
- [46] "Internet topology collection." <http://irl.cs.ucla.edu/topology/>.
- [47] B. Zhang, R. Liu, D. Massey, and L. Zhang, "Collecting the Internet AS-level topology," *ACM SIGCOMM Computer Communication Review*, vol. 35, pp. 53–61, Jan. 2005.
- [48] E. Gregori, A. Improta, L. Lenzi, L. Rossi, and L. Sani, "BGP and Inter-AS Economic Relationships," in *IFIP Networking Proceedings, Part II*, (Valencia, Spain), pp. 54–67, May 2011.
- [49] M. Luckie, B. Huffaker, A. Dhamdhere, V. Giotsas, and k. claffy, "AS Relationships, Customer Cones, and Validation," in *ACM SIGCOMM Internet Measurement Conference (IMC)*, Oct 2013.
- [50] W. Mühlbauer, S. Uhlig, B. Fu, M. Meulle, and O. Maennel, "In search for an appropriate granularity to model routing policies," in *Proceedings of ACM SIGCOMM*, pp. 145–156, 2007.
- [51] 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," *IEEE JSAC, Special Issue on Measurement of Internet Topologies*, vol. 29, pp. 1810–1821, October 2011.
- [52] K. Keys, Y. Hyun, M. Luckie, and k. claffy, "Internet-Scale IPv4 Alias Resolution with MIDAR," *IEEE/ACM Transactions on Networking*, vol. 21, Apr 2013.
- [53] N. Spring, R. Mahajan, D. Wetherall, and T. Anderson, "Measuring isp topologies with rocketfuel," *IEEE/ACM Transactions on Networking*, vol. 12, pp. 2–16, Feb. 2004.
- [54] H. Madhyastha, E. Kataz-Bassett, T. Anderson, A. Krishnamurthy, and A. Venkataramani, "iPlane Nano: Path prediction for peer-to-peer applications," (Boston, MA), pp. 137–152, USENIX, April 2009.
- [55] D. Feldman, Y. Shavitt, and N. Zilderman, "A structural approach for pop geo-location," *Computer Networks*, Decmeber 2012.
- [56] A. Rasti, N. Magharei, R. Rejaie, and W. Willinger, "Eyeball ases: From geography to connectivity," *Internet Measurement Conference (IMC)*, November 2010.
- [57] H. Haddadi and O. Bonaventure, *Recent Advances in Networking*, vol. 1. Aug. 2013. ACM SIGCOMM eBook.
- [58] A. Afanasyev, N. Tilley, P. Reiher, and L. Kleinrock, "Host-to-host congestion control for TCP," *Communications Surveys Tutorials, IEEE*, vol. 12, no. 3, pp. 304–342, 2010.
- [59] S. Bauer, D. D. Clark, and W. Lehr, "The Evolution of Internet Congestion," in *TPRC*, 2009.

- [60] F. Dinu and T. S. E. Ng, "Inferring a network congestion map with zero traffic overhead," in *Proceedings of the 2011 19th IEEE International Conference on Network Protocols, ICNP '11*, (Washington, DC, USA), pp. 69–78, IEEE Computer Society, 2011.
- [61] L. DiCioccio, R. Teixeira, M. May, and C. Kreibich, "Probe and Pray: Using UPnP for Home Network Measurements," in *Proceedings of the 13th international conference on Passive and Active Measurement (PAM)*, pp. 96–105, 2012.
- [62] L. DiCioccio, R. Teixeira, and C. Rosenberg, "Impact of Home Networks on End-to-end Performance: Controlled Experiments," in *Proceedings of the 2010 ACM SIGCOMM workshop on Home networks, HomeNets '10*, pp. 7–12, 2010.
- [63] K. L. Calvert, W. K. Edwards, N. Feamster, R. E. Grinter, Y. Deng, and X. Zhou, "Instrumenting home networks," *SIGCOMM Comput. Commun. Rev.*, vol. 41, pp. 84–89, Jan. 2011.
- [64] FCC, "FCC: Measuring Broadband America," <http://www.fcc.gov/measuring-broadband-america>, 2011.
- [65] S. Sundaresan, W. de Donato, N. Feamster, R. Teixeira, S. Crawford, and A. Pescapè, "Broadband Internet Performance: A View From the Gateway," in *Proceedings of ACM SIGCOMM*, pp. 134–145, 2011.
- [66] S. Bauer, D. Clark, and W. Lehr, "Understanding Broadband Speed Measurements," in *TPRC*, 2010.
- [67] C. Kreibich, N. Weaver, B. Nechaev, and V. Paxson, "Netalyzr: Illuminating the Edge Network," in *Proceedings of the ACM SIGCOMM Internet Measurement Conference (IMC)*, pp. 246–259, 2010.
- [68] R. García García, "Understanding the Performance of Broadband Networks Through the Statistical Analysis of Speed Tests," Master's thesis, Massachusetts Institute of Technology, 2011.
- [69] S. Sundaresan, N. Feamster, L. Dicioccio, and R. Teixeira, "Which Factors Affect Access Network Performance?," *Georgia Tech Technical Report*, no. GT-CS-10-04.
- [70] D. Genin and J. Splett, "Where in the Internet is congestion?," *CoRR*, vol. abs/1307.3696, 2013.
- [71] N. Duffield, "Simple network performance tomography," in *Proceedings of the 3rd ACM SIGCOMM Conference on Internet Measurement, IMC '03*, (New York, NY, USA), pp. 210–215, ACM, 2003.
- [72] M. Coates, R. Castro, R. Nowak, M. Gadhiok, R. King, and Y. Tsang, "Maximum likelihood network topology identification from edge-based unicast measurements," in *Proceedings of the 2002 ACM SIGMETRICS international conference on Measurement and modeling of computer systems, SIGMETRICS '02*, (New York, NY, USA), pp. 11–20, ACM, 2002.
- [73] T. Bu, N. Duffield, F. L. Presti, and D. Towsley, "Network tomography on general topologies," in *Proceedings of the 2002 ACM SIGMETRICS international conference on Measurement and modeling of computer systems, SIGMETRICS '02*, (New York, NY, USA), pp. 21–30, ACM, 2002.
- [74] N. Duffield, F. Lo Presti, and V. Paxson, "Network loss tomography using striped unicast probes," *IEEE/ACM Transactions on Networking*, vol. 14, no. 4, 2006.
- [75] M. Coates and R. Nowak, "Network loss inference using unicast end-to-end measurement," in *Proc. ITC Conf. IP Traffic, Modeling and Management*, 2000.
- [76] Y. Tsang, M. Coates, and R. Nowak, "Passive network tomography using em algorithms," in *Proceedings of ICASSP*, 2001.
- [77] R. Castro, M. Coates, G. Liang, R. Nowak, and B. Yu, "Network tomography: Recent developments," *Statistical Science*, vol. 19, no. 3, pp. pp. 499–517, 2004.

- [78] N. Hu, L. E. Li, Z. M. Mao, P. Steenkiste, and J. Wang, "Locating Internet bottlenecks: Algorithms, measurements, and implications," *SIGCOMM Comput. Commun. Rev.*, vol. 34, pp. 41–54, Aug. 2004.
- [79] L. Deng and A. Kuzmanovic, "Pong: Diagnosing spatio-temporal Internet congestion properties," *SIGMETRICS Perform. Eval. Rev.*, vol. 35, pp. 381–382, June 2007.
- [80] H. X. Nguyen and P. Thiran, "The boolean solution to the congested IP link location problem: Theory and practice," in *INFOCOM*, pp. 2117–2125, 2007.
- [81] M. R. Garey and D. S. Johnson, *Computers and Intractability: A Guide to the Theory of NP-Completeness*. New York, NY, USA: W. H. Freeman & Co., 1979.
- [82] U. Feige, "A threshold of $\ln n$ for approximating set cover," *J. ACM*, vol. 45, pp. 634–652, July 1998.
- [83] A. Dhamdhere, R. Teixeira, C. Dovrolis, and C. Diot, "NetDiagnoser: troubleshooting network unreachabilities using end-to-end probes and routing data," in *Proceedings of ACM CoNEXT*, 2007.
- [84] S. Zarifzadeh, M. Gowdagere, and C. Dovrolis, "Range Tomography: Combining the Practicality of Boolean Tomography with the Resolution of Analog Tomography," in *Proceedings of the ACM SIGCOMM conference on Internet measurement (IMC)*, 2012.
- [85] B. Briscoe, "Flow rate fairness: Dismantling a religion," *SIGCOMM Comput. Commun. Rev.*, vol. 37, pp. 63–74, Mar. 2007.
- [86] R. Briscoe, "Congestion Exposure (ConEx), Re-feedback & Re-ECN," 2012. <http://bobbriscoe.net/projects/refb/>.
- [87] EURESCOM, "CDN interconnection," 2010. <http://www.ist-daidalos.org/Public/Projects/P1900-series/P1955/default.asp>.
- [88] Cogent Communications, "Public consultation on the open internet and net neutrality in europe: Contribution by cogent communications," 2010. http://ec.europa.eu/information_society/policy/ecomms/doc/library/public_consult/net_neutrality/comments/01operators_isps/cogent_communications.pdf.
- [89] F. Press, "Response by Free to Public Consultation on the Open Internet and Net Neutrality in Europe," 2010. http://ec.europa.eu/information_society/policy/ecomms/doc/library/public_consult/net_neutrality/comments/01operators_isps/free_iliad.pdf.
- [90] K. Bode, "'free ride' google working with telcos on congestion: France telecom states talks underway about prioritized access," 2011. <http://www.dslreports.com/shownews/Free-Ride-Google-Working-With-Telcos-on-Congestion-114645>.
- [91] B. van Schewick, "Towards an economic framework for network neutrality regulation," *JTHTL*, vol. 5, no. 2, pp. 329–392, 2007.
- [92] V. K. Adhikari, Y. Chen, S. Jain, and Z.-L. Zhang, "Reverse-Engineering the YouTube Video Delivery Cloud," in *Proceedings of the Workshop on Hot Topics in Media Delivery (HotMD)*, Jul 2011.
- [93] V. K. Adhikari, Y. Chen, S. Jain, and Z.-L. Zhang, "Where Do You 'Tube'? Uncovering YouTube Server Selection Strategy," in *Proceedings of IEEE ICCCN*, Aug 2011.
- [94] Y. Chen, S. Jain, V. K. Adhikari, , and Z.-L. Zhang, "Characterizing Roles of Front-end Servers in End-to-End Performance of Dynamic Content Distribution," in *Proceedings of the USENIX/ACM SIGCOMM Internet Measurement Confernece (IMC)*, Nov 2011.
- [95] J. Jiang, V. Sekar, and H. Zhang, "Improving Fairness, Efficiency, and Stability in HTTP-based Adaptive Video Streaming with FESTIVE," in *ACM CoNEXT 2012*, 2012.

- [96] T.-Y. Huang, N. Handigol, B. Heller, N. McKeown, and R. Johari, "Confused, Timid, and Unstable: Picking a Video Streaming Rate is Hard," in *Proceedings of the 2012 ACM conference on Internet measurement conference, IMC '12*, (New York, NY, USA), pp. 225–238, ACM, 2012.
- [97] V. K. Adhikari, Y. Guo, F. Hao, M. Varvello, V. Hilt, M. Steiner, and Z.-L. Zhang, "Unreeling netflix: Understanding and improving multi-CDN movie delivery," in *Proceedings of IEEE INFOCOM*, pp. 1620–1628, 2012.
- [98] V. K. Adhikari, Y. Guo, F. Hao, V. Hilt, and Z.-L. Zhang, "A tale of three CDNs: An active measurement study of Hulu and its CDNs," in *Proceedings of IEEE INFOCOM Workshop*, 2012.
- [99] A. Lodhi, N. Larson, A. Dhamdhare, C. Dovrolis, and C. Kc, "Using PeeringDB to Understand the Internet Peering Ecosystem," in *submitted to ACM SIGCOMM Computer Communications Review (CCR)*, 2013.
- [100] "Euro-IX." <https://www.euro-ix.net>, November 2012.
- [101] "Packet Clearing House." <https://prefix.pch.net/applications/ixpdir/>, November 2012.
- [102] "LINX: Members by ASN/IP." http://www.https://www.linx.net/about/member_ipasn.html, November 2012.
- [103] "SIX: Peering Matrix." http://www.six.sk/aktual_peering.php?lang=en, November 2012.
- [104] N. Chatzis, G. Smaragdakis, A. Feldman, and W. Willinger, "There is more to IXPs than meets the eye," *ACM SIGCOMM CCR*, vol. 43, Oct. 2013.
- [105] J. H. Sowell, "Framing the Value of Internet Exchange Participation," in *Proceedings of Telecommunications Policy Research Conference (TPRC)*, 2013.
- [106] "AMS-IX Route Servers." <https://www.ams-ix.net/technical/specifications-descriptions/ams-ix-route-servers>, November 2013.
- [107] Open-IX, "IXP Technical Requirements." https://docs.google.com/document/d/150nkjX-H7t0p_ZFg-J88njDYvbpGzSLjM5ZjKnuOPA/pub.
- [108] CAIDA, "Archipelago Measurement Infrastructure." <http://www.caida.org/projects/ark/>.
- [109] RIPE Labs, Robert Kisteleki, "RIPE Atlas," February 2011. <https://atlas.ripe.net/doc/udm>.
- [110] M. Luckie, Y. Hyun, and B. Huffaker, "Traceroute Probe Method and Forward IP Path Inference," in *Internet Measurement Conference (IMC)*, (Vouliagmeni, Greece), pp. 311–324, Internet Measurement Conference (IMC), Oct 2008.
- [111] B. Huffaker, A. Dhamdhare, M. Fomenkov, and k. claffy, "Toward Topology Dualism: Improving the Accuracy of AS Annotations for Routers," in *Passive and Active Network Measurement Workshop (PAM)*, (Zurich, Switzerland), PAM 2010, Apr 2010.
- [112] M. Luckie, "Scamper: a scalable and extensible packet prober for active measurement of the Internet," in *ACM SIGCOMM Internet Measurement Conference (IMC)*, 2010.
- [113] M. Luckie, A. Dhamdhare, k. claffy, and D. Murrell, "Measured Impact of Crooked Traceroute," *ACM SIGCOMM Computer Communication Review (CCR)*, vol. 41, pp. 14–21, Jan 2011.
- [114] B. Donnet, M. Luckie, P. Mrindol, and J. Pansiot, "Revealing MPLS tunnels obscured from traceroute," *ACM SIGCOMM Computer Communication Review (CCR)*, vol. 42, pp. 87–93, Apr 2012.
- [115] B. Huffaker, M. Fomenkov, and k claffy, "Internet Topology Data Comparison," *Cooperative Association for Internet Data Analysis (CAIDA) Technical Report*, 2012. <http://www.caida.org/research/topology/topo-comparison>.

- [116] B. Huffaker, M. Fomenkov, M. Luckie, and kc claffy, "Visualizing IPv4 and IPv6 Internet Topology at a Macroscopic Scale," 2013. http://www.caida.org/research/topology/as_core_network/.
- [117] "High degree router analysis." http://www.caida.org/~amogh/rtr_graph/high_deg.html, November 2013.
- [118] B. Huffaker, M. Fomenkov, and kc claffy, "Automating Inference of Router Locations," 2013. in submission, email brad@caida.org.
- [119] MaxMind, "MaxMind GeoLite Country: Open Source IP Address to Country Database." <http://www.maxmind.com/app/geolitecountry>.
- [120] Y. Wang, D. Burgener, M. Flores, A. Kuzmanovic, and C. Huang, "Towards street-level client-independent ip geolocation," in *In Proceedings of USENIX NSDI*, March 2011.
- [121] M. Arif, S. Karunasekera, S. Kulkarni, A. Gunatilaka, and B. Ristic, "Internet Host Geolocation Using Maximum Likelihood Estimation Technique," in *AINA '10: Proceedings of the 2010 24th IEEE International Conference on Advanced Information Networking and Applications*, (Washington, DC, USA), pp. 422–429, IEEE Computer Society, 2010.
- [122] E. Katz-Bassett, J. John, A. Krishnamurthy, D. Wetherall, T. Anderson, and Y. Chawathe, "Towards IP geolocation using delay and topology measurements," in *IMC '06: Proceedings of the 6th ACM SIGCOMM Conference on Internet Measurement*, (New York, NY, USA), pp. 71–84, ACM, 2006.
- [123] B. Eriksson, P. Barford, B. Maggs, and R. Nowak, "Posit: a lightweight approach for ip geolocation," *SIGMETRICS Perform. Eval. Rev.*, vol. 40, pp. 2–11, Oct. 2012.
- [124] B. Huffaker, K. Keys, M. Fomenkov, and K. Claffy, "AS-to-Organization Dataset." <http://www.caida.org/research/topology/as2org>.
- [125] CAIDA, "AS-rank." <http://as-rank.caida.org>.
- [126] V. Giotsas and S. Zhou, "Detecting and assessing the hybrid IPv4/IPv6 AS relationships," in *SIGCOMM posters*, pp. 424–425, Aug. 2011.
- [127] K. Nichols and V. Jacobson, "Controlling queue delay," *ACM Queue*, 2012.
- [128] K. Thompson, G. Miller, and R. Wilder, "Wide-area internet traffic patterns and characteristics," *Network, IEEE*, vol. 11, no. 6, pp. 10–23, 1997.
- [129] M. Roughan, A. Greenberg, C. Kalmanek, M. Rumsewicz, J. Yates, and Y. Zhang, "Experience in measuring internet backbone traffic variability: Models, metrics, measurements and meaning," 2003.
- [130] K. Cho, K. Fukuda, H. Esaki, and A. Kato, "The impact and implications of the growth in residential user-to-user traffic," in *in Proc. ACM SIGCOMM*, pp. 207–218, ACM Press, 2006.
- [131] M. Dischinger, K. P. Gummadi, A. Haeberlen, and S. Saroiu, "Characterizing residential broadband networks," in *Proc. of ACM IMC*, 2007.
- [132] S. Bauer, D. D. Clark, and W. Lehr, "A data driven exploration of broadband traffic issues: Growth, management, and policy," TPRC, 2012. Available at SSRN: <http://ssrn.com/abstract=2029058> or <http://dx.doi.org/10.2139/ssrn.2029058>.
- [133] Rich Carlson, "Network Diagnostic Tool (NDT)." [nowhostedathttp://www.internet2.edu/performance/ndt/](http://www.internet2.edu/performance/ndt/).
- [134] M. Mathis, "Model Based Bulk Performance Metrics." <http://datatracker.ietf.org/doc/draft-ietf-ippm-model-based-metrics/>, Oct 2013.
- [135] "Mlab infrastructure." <http://www.measurementlab.net/infrastructure>, November 2013.
- [136] "Mlab Github." <https://github.com/m-lab/mbm>, November 2013.
- [137] S. Akhshabi, A. C. Begen, and C. Dovrolis, "An Experimental Evaluation of Rate-adaptation Algorithms in Adaptive Streaming over HTTP," in *Proceedings of the second annual ACM*

- conference on Multimedia systems, MMSys '11, (New York, NY, USA), pp. 157–168, ACM, 2011.
- [138] A. Akella, S. Seshan, and A. Shaikh, “An Empirical Evaluation of Wide-area Internet Bottlenecks,” in *Proceedings of the 3rd ACM SIGCOMM conference on Internet Measurement (IMC)*, pp. 101–114, 2003.
- [139] S. Katti, D. Katabi, C. Blake, E. Kohler, and J. Strauss, “MultiQ: Automated Detection of Multiple Bottleneck Capacities Along a Path,” in *Proceedings of the 4th ACM SIGCOMM conference on Internet measurement (IMC)*, pp. 245–250, 2004.
- [140] W. Wei, B. Wang, D. Towsley, and J. Kurose, “Model-based Identification of Dominant Congested Links,” in *Proceedings of the 3rd ACM SIGCOMM conference on Internet measurement (IMC)*, pp. 115–128, 2003.
- [141] M. Jain and C. Dovrolis, “End-to-end available bandwidth: Measurement methodology, dynamics, and relation with tcp throughput,” *IEEE/ACM Trans. Netw.*, vol. 11, pp. 537–549, Aug. 2003.
- [142] C. Dovrolis, P. Ramanathan, and D. Moore, “Packet-dispersion techniques and a capacity-estimation methodology,” *IEEE/ACM Transactions on Networking*, vol. 12, pp. 963–977, Dec. 2004.
- [143] R. A. Carlson, “Developing the web100 based network diagnostic tool (ndt),” in *In: Proc. Passive and Active Measurement Workshop (PAM)*, 2003.
- [144] 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, pp. 219–234, 2010.
- [145] R. Albert and A. L. Barabasi, “Topology of Evolving Networks: Local Events and Universality,” *Physical Review Letters* 85, 5234, 2000.
- [146] S. H. Yook, H. Jeong, and A. L. Barabasi, “Modeling the Internet’s Large-scale Topology,” *National Academy of Sciences*, 2002.
- [147] A. Dhamdhere, M. Luckie, B. Huffaker, K. Claffy, A. Elmokashfi, and E. Aben, “Measuring the Deployment of IPv6: Topology, Routing, and Performance,” in *Proceedings of ACM SIGCOMM Internet Measurement Conference (IMC)*, 2012.
- [148] Telegeography, “IP Transit Pricing Service.” <http://www.telegeography.com/research-services/ip-transit-pricing-service/index.html>, October 2011.
- [149] S. Zhou, G. Q. Zhang, and G. Q. Zhang, “Chinese Internet AS-level topology,” in *IEEE Communications*, 2007.
- [150] M. Wahlisch, T. Schmidt, M. de Brun, and T. Haberen, “Exposing a nation-centric view on the German internet - a change in perspective on the AS-level,” in *in Proceedings of PAM*, 2012.
- [151] J. Karlin, S. Forrest, and J. Rexford, “Nation-state routing: Censorship, wiretapping, and bgp,” *CoRR*, vol. abs/0903.3218, 2009.
- [152] Telecom Asia, “Localizing the Internet in the Philippines.” <http://www.telecomasia.net/blog/content/localizing-internet-philippines>.
- [153] E. . Young, “Top 10 risks in telecommunications 2012,” tech. rep., 2012.
- [154] D. D. Clark, W. Lehr, and S. Bauer, “Interconnection in the Internet: The Policy Challenge,” in *TPRC*, 2011.
- [155] A. Dhamdhere, C. Dovrolis, and P. Franois, “A value-based framework for internet peering agreements,” in *Proceedings of International Teletraffic Congress (ITC)*, 2010.
- [156] A. D’Ignazio and E. Giovannetti, ““Unfair” Discrimination in Two-sided Peering? Evidence from LINX,” *Cambridge Working Papers in Economics*, vol. 0621, 2006.

- [157] A. D'Ignazio and E. Giovannetti, "Asymmetry and Discrimination in Internet Peering. Evidence from the LINX," *International Journal Of Industrial Organization*, vol. 27, pp. 441–448, 2009.
- [158] E. Maida, "The Regulation of Internet Interconnection: Assessing Network Market Power," *Massachusetts Institute of Technology Masters Thesis*, 2012.
- [159] Y. Benkler, "Next Generation Connectivity: A Review of Broadband Internet Transitions and Policy from Around the World," *The Berkman Center for Internet and Society Technical Report*, 2010.
- [160] M. Faloutsos, P. Faloutsos, and C. Faloutsos, "On Power-law Relationships of the Internet Topology," in *Proceedings of ACM SIGCOMM*, 1999.
- [161] Q. Chen, H. Chang, R. Govindan, S. Jamin, S. Shenker, and W. Willinger, "The Origin of Power-Laws in Internet Topologies Revisited," in *Proceedings of IEEE Infocom*, 2002.
- [162] A. L. Barabasi and R. Albert, "Emergence of Scaling in Random Networks," *Science* 286 509-512, 1999.
- [163] S. Park, D. M. Pennock, and C. L. Giles, "Comparing Static and Dynamic Measurements and Models of the Internet's AS Topology," in *Proc. IEEE Infocom*, 2004.
- [164] M. A. Serrano, M. Boguna, and A. D. Guilera, "Modeling the Internet," *The European Physics Journal B*, 2006.
- [165] X. Wang and D. Loguinov, "Wealth-Based Evolution Model for the Internet AS-Level Topology," in *Proc. IEEE Infocom*, 2006.
- [166] Shi Zhou, "Understanding the Evolution Dynamics of Internet Topology," *Physical Review E*, vol. 74, 2006.
- [167] A. Dhamdhere and C. Dovrolis, "The Internet is Flat: Modeling the Transition From a Transit Hierarchy to a Peering Mesh," in *Proceedings of ACM CoNEXT*, 2010.
- [168] H. Chang, S. Jamin, and W. Willinger, "To Peer or Not to Peer: Modeling the Evolution of the Internet's AS-Level Topology," in *Proc. IEEE Infocom*, 2006.
- [169] B. Edwards, S. A. Hofmeyr, G. Stelle, and S. Forrest, "Internet topology over time," *CoRR*, vol. abs/1202.3993, 2012.
- [170] "The dual router+as internet topology generator." <http://www.caida.org/research/topology/generator/>.