*Project Summary:* **SaTC: CORE: Large: Collaborative: Investigating the Susceptibility of the Internet Topology to Country-level Connectivity Disruption and Manipulation**

To apply a military analogy to Internet research, the science of cybersecurity has focused heavily on weapons and tactics, but has largely neglected terrain. *Strategic points in the macroscopic Internet topology constitute key terrain in the cyberspace battlefield*. Adversaries — hackers, terrorists or nation-states — can disrupt, intercept or manipulate the Internet traffic of entire countries or regions by targeting structural weaknesses of the Internet topology. Despite much recent interest and a large body of research on cyber-attack vectors and mechanisms, we lack rigorous tools to reason about how the macroscopic Internet topology of a country or a region exposes its critical communication infrastructure to compromise through targeted attacks. Part of the problem is that collecting and interpreting data about the Internet connectivity, configurations and associated vulnerabilities is challenging. Due to the massive scale and broadly distributed nature of Internet infrastructure and the scarcity of publicly available data, we must resort to complex measurement and inference methodologies that require significant effort in design, implementation, and validation. Despite these obstacles, the Internet measurement community and, in the past two decades, the two research groups leading this research project, have developed a unique set of tools, methodologies and data sets that illuminate various aspects of the Internet infrastructure.

Our proposed research is organized as a collaborative three-phase project. In the first phase, our goal will be to identify important components of the Internet topology of a country/region — Autonomous Systems (ASes), Internet Exchange Points (IXPs), PoPs, colocation facilities, and physical cable systems which represent the "key terrain" in cyberspace. To achieve this goal we will undertake a novel multi-layer mapping effort to discover the key components, relationships between them, and their geographic properties. In the second phase, we will develop methods to identify components that represent *potential topological weaknesses*, i.e., compromising a few such components would allow an attacker to disrupt, intercept or manipulate Internet traffic of that country. Our multi-layer view of the system will enable an assessment of weaknesses, holistically as well as at specific layers, under various assumptions about the capabilities and knowledge of attackers. Geographic annotations will enable us to consider risks related to the geographic distribution of critical components of the communication infrastructure. In the third phase, we will develop a systematic approach to mitigate the impact of observed weaknesses, framing the study as an optimization problem that incorporates socio-economic and political factors as constraints.

**Intellectual merit:** This research will lead to methodologies to *highlight, quantify, and mitigate* macroscopic vulnerabilities of the Internet infrastructure, especially from the perspective of cyber-terrorist attacks and cyber-conflicts between nation-states. The project also promises significant advances in understanding the mapping and relationships between logical topologies at the AS-level and the physical topology of cables and Internet exchanges.

**Broader impact:** Understanding topological weaknesses for countries or regions is of significant interest to not just the research and operational communities, but also national security agencies, policy bodies and in daily life. This research program both enables and benefits from an education and outreach program that will enhance curriculum, foster collaborations, and build community. Findings from this project will influence the development of new course materials that will be used in classes at the undergraduate and graduate level. The students who take these classes as well as those who work directly on the projects will receive guidance and training. All of the tools and data sets that are developed over the course of the grant will be made openly available to the community. Finally, we will disseminate our research results by publishing in highly respected academic conferences and workshops.

**Key Words: Internet topology, macroscopic vulnerabilities, cyber-conflict**

## Contents

***Project Description: SaTC: CORE: Large: Collaborative: Investigating the Susceptibility of the Internet Topology to Country-level Connectivity Disruption and Manipulation***

## 1 Introduction and Motivation

Various types of malicious actors — hackers, terrorists or nation-states — can disrupt, intercept or manipulate the Internet traffic of entire countries or regions by targeting structural weaknesses of the Internet topology. As such, it is crucial to identify critical elements of the Internet topology and to understand how attackers could compromise those elements to affect availability, integrity, or confidentiality of Internet communications. Applying a military analogy to Internet research, we argue that the science of cybersecurity has focused heavily on *weapons* and *tactics*, but has largely neglected *terrain*. Strategic points in the macroscopic Internet topology constitute *key terrain* in the cyberspace battlefield. Despite much recent interest and a large body of research on cyber-attack vectors and mechanisms, ***we lack rigorous tools to quantitatively reason about how the macroscopic Internet topology of a country (or of a region) exposes its critical communication infrastructure to well-targeted attacks***.

Events in North Africa and the Middle East in 2011-2012 demonstrated the ability of authorities in certain countries to disconnect their nation from the rest of the Internet [1, 2, 3, 4]. These events highlighted the lack of diversity in international Internet connectivity in certain countries. In 2012, Dyn Research performed a census *"of all the domestic providers in each country who have direct connections to foreign providers"* and suggested that, surprisingly, only 30 countries were not at risk of being decoupled from the global Internet [5]. Dyn's analysis focused on the specific scenario of a government forcing local operators to shut down their international links. However, depending on an attacker's goal – traffic disruption, eavesdropping, manipulation – and her ability to compromise selected targets, there are many potential attack patterns and a multitude of ways in which a country could be susceptible to compromise of its cyberspace terrain.

State and non-state actors conduct offensive cyber operations to achieve a variety of political, economic, or military objectives and cyber conflict has intensified in the last few years, involving both large and smaller countries [6, 7] as well as terrorist groups [8]. In November 2014, likely in retaliation to the planned release of a satirical film, North Korea conducted a cyberattack against Sony Pictures Entertainment. The US DoD called the attack *"one of the most destructive cyberattacks on a U.S. entity to date"* [9]. In the following weeks North Korea was almost entirely disconnected from the Internet for several days [10]; speculation attributed this event to counterretaliation (*e.g.,* from US-based hacktivists). Denial-of-service (DoS) attacks also played a major role in the cyber-attacks against Estonia, Georgia and Kyrgystan that took place between 2007 and 2009 [11, 12, 13]. These attacks were aimed at paralyzing the communication infrastructure of these states [14]. In the case of Kyrgyzstan, the attack targeted the only four Internet service providers in the country, but caused the majority of the Internet services to collapse. Other notable recent attacks involved cable connections and key Internet exchange points (IXPs). For instance, Ukrainian telecom providers reported disruptions to a key IXP and cable connections during Russian military activity in the Crimean peninsula in 2014 [15]. In 2013, Egypt arrested three divers who were found cutting through a major Internet cable servicing parts of Europe, Africa, the Middle East and Asia [16].

Finally, natural disasters represent a different perspective wherein an attacker does not select specific targets, but rather the event could impact all targets in a certain geographic area. Several recent natural disasters have caused significant disruptions in communications *e.g.,* Hurricane Sandy on the US east coast in 2012 [17, 18], the Tohoku earthquake and tsunami in Japan, and the Christchurch earthquake in New Zealand [19, 20, 21] (both happened in early 2011). These events

highlight the fact that even the most developed countries are not immune to this threat.

In this project we propose a first rigorous attempt at understanding the ways in which the Internet infrastructure of a country or region is susceptible to a range of threats. We propose to develop a modular and extensible framework that will catalyze efforts to define, assess and mitigate these structural topological weaknesses. The proposed research will improve the current state of the art in synthesizing the empirical data that we will use to support our analysis. These efforts will guarantee that our approach to identify and model these weaknesses will be reproducible and repeatable.

Collecting and interpreting data about the Internet infrastructure is challenging due to the distributed nature of the Internet's structure and administration, its scale, and the scarcity of publicly available data. Thus, we must resort to complex large-scale measurement and inference methodologies that require significant effort in design, implementation, and validation. Despite these obstacles, the Internet measurement community and, in the past two decades, the research groups leading this research project, have developed a solid set of tools and methodologies to illuminate various aspects of the Internet infrastructure.

Our first goal will be to identify important components of the Internet topology of a country/region — Autonomous Systems (ASes), Internet Exchange Points (IXPs), PoPs, colocation facilities, and physical cable systems which represent the "key terrain" in cyberspace. To achieve this goal we will undertake a novel multi-layer mapping effort, developing Internet measurement and analysis techniques to discover these components and the relationships between them. Our starting point will be the set of ASes active in a country and the logical links (economic contracts) between them. We will then proceed to "peel off the layers" to map the internal structure of those ASes (PoPs), the rich router-level connectivity that constitutes AS links, and the role of IXPs. Finally, at the lowest layer, our map will illuminate the role of physical infrastructure (fiber, cable systems and colocation facilities) in the connectivity of the country. During the course of this mapping process we will annotate topological components with their geographic properties. Next, we will use the multi-layer map and develop methods to identify components that represent *potential topological weaknesses*, *i.e.,* compromising a few such components would allow a potential attacker to disrupt, manipulate, or eavesdrop on national and international communications of that country. Our multi-layer view of the system will enable an assessment of weaknesses holistically as well as at specific layers, under various assumptions about the capabilities and knowledge of attackers. Geographic annotations will enable us to take into consideration risks related to the geographic distribution of critical components of the communication infrastructure.

## 2 Adversaries, Threats and Attacks

In the context of this project, *adversaries* are mainly nation states, terrorist groups or other politically organized groups. Their *objective* is to either disrupt, manipulate or monitor a large fraction of the traffic within the country/region aiming at various possible goals. We enumerate a few examples of *threats* pertaining to our problem. An adversary may deploy traffic monitoring to perform espionage. Both pervasive monitoring and traffic manipulation may aim at conducting information warfare, such as spreading of propaganda, misinformation, and sensitive information, to either demoralize or manipulate the enemy and the public and undermine the quality of opposing force information. Intercepting large fractions of a nation's traffic creates opportunities for an adversary, such as increasing the chance of penetrating sensitive networks normally decoupled from the public Internet (e.g., by taking advantage of a lack of full compliance to security policies in conjunction with the exploitation of undisclosed vulnerabilities), or simultaneously

compromising multiple networks of high strategic relevance (e.g., banks and financial institutions) to maximize the effect of an attack. Connectivity disruption can exacerbate situations of emergency, reduce the ability of the population to be informed, make the government look weak, create strategic military advantage, and significantly affect other critical infrastructure [22]. Finally, we consider the special case of scenarios without the presence of an adversary, in which the impact of disruptive natural events (e.g., severe storms, hurricanes, earthquakes, etc.) on the communication infrastructure undermines public safety.

In this project, we analyze how these threats can be implemented by compromising the integrity, the availability or the confidentiality of the services provided by specific elements of the Internet topology. We focus on these classes of *assets* (described in detail in the following sections) within the communication infrastructure of a region/country, which in our framework represent potential *targets*: Autonomous Systems (AS), Internet eXchange Points (IXP) and other colocation facilities, Cables (including landing stations), and ISP Points of Presence (PoP). In order to affect the vast majority of the communications flowing in and out a country or region, an attacker may consider compromising a proportionate number of assets. However, in practice, these elements play very different roles and may be responsible for servicing largely different amounts of traffic, so that gaining control of (or disrupting) a few of them may yield large results. Our analysis aims at identifying which of these targets are most valuable for the attacker's purposes and when a region or a country is proportionally more exposed than others to the above threats because of a large concentration of strategic value in only few potential targets.

Given the broad set of complex interdependent human, physical and cyber-systems represented by these targets, the range of possible *attack vectors* is vast. Attack vectors include DoS attacks, remote vulnerability exploitation, brute-forcing, hardware/firmware/software backdoors, man-in-the-middle-attacks, viruses, phishing, packet sniffing, route hijacking, cryptographic attacks, social engineering and intimidation, intelligence operations, infiltration, military operations, cable cuts, causing power outages, etc. Similarly, in most scenarios we can assume that the *attacker's capabilities* are highly sophisticated and potentially unknown to the counterpart. Indeed the use of different means of attack can depend on strategic factors such as the disinclination of the adversary to resort to traditional military actions instead of operating exclusively in cyberspace, or the attempt to keep activities covert, or the intent to avoid attribution. It is out of the scope of this investigation to evaluate risks that are a function of the multi-dimensional vulnerabilities of the specific targets we consider, of the capabilities of the adversary, and of political and strategic factors. Our focus is instead to investigate the dimension of the macroscopic topology. Therefore in defining our approach, we are agnostic about the "cost" for a given attacker to compromise each potential target. Nevertheless, in Section 4 we show that our model allows the user to define these and other external variables and take them into account when calculating the risk of different attack patterns from a given adversary.

## 3 Task 1: Constructing a multi-layer topology map at the country/region level

The key aspect of the threats mentioned in Section 2 is the potential for malicious actors to *disrupt, monitor, or manipulate* traffic flowing in and out of specific countries and regions. The potential for such attacks depends on the components of a country's Internet infrastructure — Autonomous Systems (ASes), Internet Exchange Points (IXPs), PoPs, colocation facilities, and physical cable systems. In the first research task we will undertake a novel multi-layer mapping effort to discover these components and the relationships between them. Our starting point will be the set of ASes active in a country and the logical links (*i.e.,* economic contracts) between them. We will then map

the internal structure of those ASes (PoPs), the rich router-level connectivity that constitutes AS links, and the role of IXPs. Finally, at the lowest layer, our map will illuminate the role of physical infrastructure (fiber, cable systems and colocation facilities) in the connectivity of the country. During the course of this mapping process we will annotate topological components with their geographic properties. For simplicity of exposition, we present our approach at the country-level granularity, but emphasize that the approach is general enough that we can apply it at the regional granularity as well.

The mapping task is challenging for three reasons: First, there is no straightforward way to identify the set of networks operating in a country. Second, mapping connectivity between ASes is challenging due to the incompleteness of publicly available data to infer AS-level connectivity, the lack of visibility into backup connectivity that may exist but is not normally visible, and the lack of reliable colocation databases to precisely map logical connectivity to exchanges and private peering facilities. Finally, mapping logical connectivity and traffic flow to the underlying physical infrastructure (cables and fiber) and discovering dependencies at the physical layer is a largely open research problem, but one that offers tantalizing opportunities.

### 3.1 Identify ASes *active* in a country:

It is not straightforward to determine which Autonomous Systems (ASes) are active in a country. WHOIS data only tells us which ASes are registered in a country and is insufficient for identifying ASes that are *active* in a country, since an AS could be active in countries other than the one in which it is registered (*e.g.,* a large multi-national transit provider). Our proposed approach is to fuse data from a variety of sources including WHOIS, BGP, IP geolocation databases and IP address activity censuses [23] to develop a dataset of ASes that are active in a country.

We will start from WHOIS data to identify ASes registered in a country, and use publicly available BGP topology data [24, 25, 26] to derive a mapping between those ASes and the prefixes they originate in the global routing system. We will then use the best available IP geolocation datasets [27] to map those advertised IP prefixes to a set of countries. We will use these constructed AS-prefix and prefix-location mappings to obtain the set countries in which ASes are active. We will investigate the most appropriate methods to merge the WHOIS-derived and BGP-derived sets of ASes active in a country. To do so we will publish our set of inferred AS-country mappings on our webpage or share them informally with ISP contacts to validate our inferences. We note that while available geolocation databases are known to be inaccurate at finer granularities such as city-level, previous work has found them to be mostly accurate at the country-level [28]. Nevertheless, we will devise filtering methods to account for inaccuracies that could skew our results based on the approach taken in prior work by PI-Barford [29].

### 3.2 Infer logical connectivity between ASes active in a country:

The first step in mapping connectivity between ASes active in a country is to infer the logical (AS-level) connectivity between ASes, and annotate those links with business relationships. AS-level connectivity indicates which ASes have established economic contracts, *i.e.,* agreements to exchange traffic. We will proceed by first inferring a baseline set of AS links using publicly available BGP data from Routeviews, RIPE and PCH repositories. In particular, AS paths from BGP vantage points toward prefixes geolocated in a country of interest will reveal the set of ASes and AS links involved in providing international transit to the country. We will use both BGP RIBs and updates to capture backup connectivity that is often only revealed during BGP's route exploration process after a failure [30, 31, 32]. We will augment this dataset with AS-links that are established using

multilateral peering (MLP) with route servers at IXPs, using a technique CAIDA researchers have developed in previous work [33] to discover all peering links established using the router server at an IXP.

The AS-level connectivity inferred from public BGP data and MLP links may still be incomplete, because public repositories of BGP data are known to miss AS links [32, 34, 35, 36, 31, 37, 38], usually peering links lower in the AS-level hierarchy (than the BGP vantage point), or peering links in regions that are not well covered by available BGP vantage points. To augment the AS-level connectivity from BGP route servers, we will design targeted traceroute measurement studies. Specifically, we will use AS customer cone data [39] to identify available traceroute vantage points (using Ark [40], RIPE Atlas [41], and Periscope [42]) located in the customer cone of ASes active in the country of interest. We will then conduct large-scale traceroute measurements from these VPs to destination prefixes geolocated both within and outside the country. Second, we will conduct traceroutes from available traceroute vantage points outside a country toward prefixes geolocated inside the country. To process the traceroutes into AS paths, we will use state of the art IP-to-AS mapping techniques [43]. At the end of this process we will have the best possible map of AS-level connectivity between ASes active in a country of interest.

### 3.3 Develop techniques to map logical connectivity to the router, exchange and facility level:

The next step is to map the logical AS-level connectivity to router-level connectivity, and to infer whether those links are established at Internet Exchange Points (IXPs) and private peering facilities. We will rely on a combination of datasets to infer whether an AS-link is established at a certain IXP. First, if an Ark vantage point is located in an AS of interest, then we can use *bdrmap* [44] a tool we have developed to discover all interdomain links, at the router-level, for the network hosting the VP. In addition to revealing the rich diversity of interconnection at the router-level, *bdrmap* also discovers whether AS links are at IXPs. Second, several IXPs publish colocation data and peering matrices [45] from which we can extract connectivity at the IXP. Second, MLP links obtained from querying an IXP's route server must be established at the IXP. Third, we will use the targeted traceroutes and IP-AS mapping (from the previous step) to infer AS links established at an IXP, using hints in the DNS names of interfaces [46, 47] and a set of known IXP prefixes from peeringDB [48] or PCH.

However, knowing that an AS link is established at an IXP is not enough. Ultimately, the physical infrastructure used to establish connectivity at IXPs resides in *colocation facilities*, which lease customers secure space to locate and install network equipment. An interconnection facility operator may operate multiple facilities in the same city, and connect them, so that networks participating at one facility can access networks at another facility in the same city. IXPs typically partner with colocation facilities located in the same city, and install equipment that enables networks present at a colocation facility to peer with other networks (possibly in different colocation facilities). By deploying infrastructure at a single colocation facility, a network can join several IXPs. It is crucial to uncover the dependency between connectivity at IXPs and the underlying colocation facilities. For instance, a network may be present at 3 different IXPs using equipment hosted at a single colocation facility. This configuration provides redundancy against failures or attacks on the IXP's peering fabric or route server, but does not protect against a failure or attack on the colocation facility or routers at the facility.

To map connectivity at the facility level we will leverage recent results from an ongoing NSF-funded project on mapping connectivity, particularly at Internet Exchanges in the Internet [49]. In that project we have developed techniques to map interdomain links to Internet exchanges and private peering facilities, using active traceroute measurements from RIPE Atlas [41], Ark [40] and

Periscope [50], colocation information from peeringDB [48] and PCH [51] and a *constrained facility search* algorithm [50]. We will extend this facility mapping technique to AS links between all ASes active in a country of interest. The result will be a connectivity map that annotates the AS link between each pair of ASes with the set of IXPs and colocation facilities at which they interconnect. We recognize that connectivity is dynamic and our maps will be generated in a way that captures this behavior longitudinally.

### 3.4 Identifying and mapping *Internet physical infrastructure*

In the final step of our mapping effort, we will identify and map the physical infrastructure in a country. By physical infrastructure, we are referring to the combination of nodes (*e.g.*, colocation, hosting facilities and data centers) and links [1] (*e.g.*, optical fiber conduits) that provide the substrate for connectivity world-wide. While many dynamic aspects of the Internet's topology have been examined in prior work, the underlying physical paths that make up the Internet are, by definition, static [2], and it is this fixed infrastructure which we seek to identify.

Over the past 6 years, the *Internet Atlas project* [52] at the University of Wisconsin has been focused on building a comprehensive, geographically accurate map of the physical Internet. In addition to the map itself the repository includes relevant related data such as details of service providers that are represented. The map repository is accessible through a GIS-based web portal that enables visualization and analysis, and enables consideration of a wide variety of geocoded data that is available from other sources (*e.g.*, country boundaries, geographical features, road and rail infrastructure, census data, weather reports, etc.). Atlas currently contains over 20K node locations and over 25K links for over 1100 service provider networks from around the world (including all tier 1 networks and nearly 200 metro area networks in the US). An example of a portion of the map that is specific to connectivity in the UK can be seen in Figure 1. This map highlights both deployment of infrastructure within the UK and connectivity between the UK and other countries.

To produce a comprehensive map of the physical Internet we use *search* to identify infrastructure maps and other repositories service providers publish online. Unfortunately, the maps we have identified (over 3000 to date) have no consistent format; some are images, others are embedded in Flash applications and all are given with a variety of geographic details. We carefully enter this data into the Atlas repository using a combination of manual and automated processes including consistency checks and methods for geocoding both node and link data. Geographic accuracy is aided by the fact that many service providers list street addresses of the locations of their PoPs and colocation facilities, and listings of the same street addresses

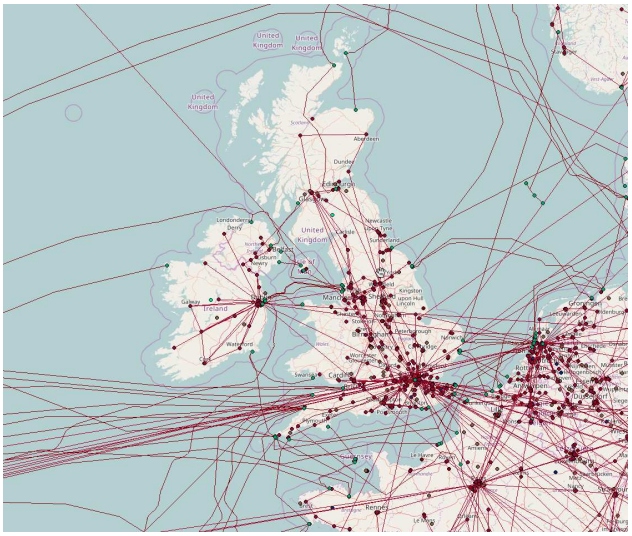


Figure 1: Map of physical infrastructure for the United Kingdom. This representation was extracted from the Internet Atlas repository [52]

[1] In the rest of this proposal, we will use the terms "link" and "conduit" interchangeably — a "tube" or trench specially built to house the fiber of potentially multiple providers.

[2] More precisely, installed conduits rarely become defunct, and deploying new conduits takes considerable time.

from multiple service providers (indicating a third party hosting facility) increases confidence in the overall map.

In recent work we used the Atlas repository to create a first-of-its-kind map of the US long-haul fiber-optic infrastructure [53]. We defined a long-haul link as one that spans at least 30 miles, or that connects population centers of at least 100,000 people, or that is shared by at least 2 providers. These numbers are not proscriptive, rather they emerged through an iterative process of refining our base map using the fiber maps in Atlas from tier-1 ISPs and major cable providers which contain explicit geocoded information about long-haul link locations. Importantly, we validated these link locations and inferred whether fiber conduits were shared using a variety of public record documents such as utility right-of-way information. We also added links from publicly available ISP fiber maps (both tier-1 and major providers) which have geographic information about link endpoints, but which do not have explicit information about geographic pathways of fiber links and again employ a variety of public records to infer the geographic locations of this latter set of links added to the map. The result is the map shown in Figure 2. The map contains contains 273 nodes/cities, 2411 links, and 542 conduits (with multiple tenants). Prominent features of the map include *(i)* dense deployments (*e.g.*, the northeast and coastal areas), *(ii)* long-haul hubs (*e.g.*, Denver and Salt Lake City) *(iii)* pronounced absence of infrastructure (*e.g.*, the upper plains and four corners regions), *(iv)* parallel deployments (*e.g.*, Kansas City to Denver) and *(v)* spurs (*e.g.*, along northern routes).

Figure 2: Map of long-haul conduits and PoPs for networks considered in the continental US.

A striking characteristic of the identified US long-haul fiber-optic network is a significant amount of *observed infrastructure sharing*. Such infrastructure sharing is the result of a common practice among many of the existing service providers to deploy their fiber in jointly-used and previously installed conduits and is dictated by simple economics — substantial cost savings as compared to deploying fiber in newly constructed conduits. By considering different metrics for measuring the risks associated with infrastructure sharing, we examined the presence of high-risk links in the long-haul infrastructure, both from a connectivity and usage perspective.

From a connectivity perspective, we simply consider the amount of sharing per conduit. Figure 3 shows the number of conduits (y axis) for which at least k ISPs (x axis) share the conduit. For example, there are 542 distinct conduits in our physical map (Figure 2), thus the bar at x=1 is 542, and 486 conduits are shared by at least 2 ISPs, thus the bar at x=2 is 486. This plot highlights the fact that it is relatively uncommon for conduits not to be shared by more than two providers. We observe that 89.67%, 63.28% and 53.50% of the conduits

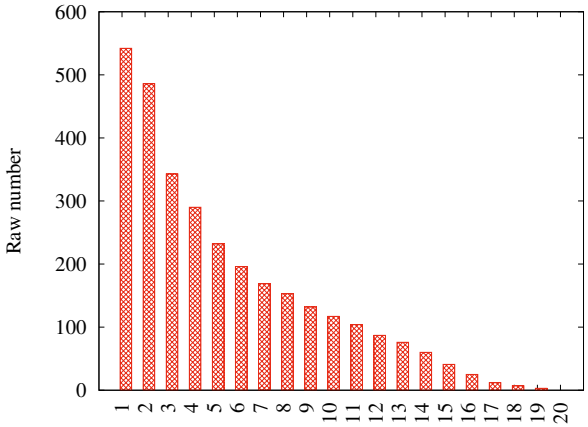


Figure 3: Ranking of US long-haul link sharing by ISPs.

are shared by at least two, three and four major ISPs, respectively. This sharing can be considered a risk in the sense that physical damage to a conduit will affect multiple providers - some of whom may use each other for backup paths without actually knowing they share the same physical link.

An important focus of our proposed work is to develop capability to link maps of physical connectivity to the logical mappings of ASes described above. Natural anchors for such linking include ASes, colocation facilities and IXPs identified in both logical and physical representations. However, inconsistencies are sure to arise and resolving those will require careful consideration in order to produce accurate and complete multi-scale maps that can be used for risk and vulnerability analysis described later in this proposal.

### 3.5 Expanding the perspective of connectivity within a country

While the Atlas repository is a compelling starting point for identifying the Internet's physical infrastructure, one should not assume that the maps represent all nodes and links in any given country. Unfortunately, incomplete representations can lead to incorrect conclusions thus we must take steps to reinforce and enrich the baseline physical maps from Atlas.

In recent work [54], we investigated the hypothesis that physical maps can be used to guide and reinforce the process of collecting layer 3 probe data toward the goal of expanding the scope of physical infrastructure captured in network-layer maps. This conjecture led directly to two key research questions that inform our proposed work: *(i)* how do physical layer maps compare and contrast with network-layer maps? and *(ii)* how can we improve an active probe-based methods used to discovery connectivity at layer 3 to reveal a larger portion of physical infrastructure? We contend that some of the challenges inherent in generating maps from layer 3 probes can be overcome by using the constructive approach of first identifying key infrastructure (POPs, etc.) and then identifying nodes (identified by disambiguating IP addresses or using DNS names) that reside in those locations.

We began by considering physical map data from the Internet Atlas and network-layer map data from Ark [55]. We focused specifically on infrastructure in North America. We resolved the IP addresses from the Ark to DNS names and then used location hints to associate these with physical locations (*e.g.,* cities), which became the basis for comparisons. Several characteristics were immediately evident in the data. Most prominent was the fact that among the 50 networks that were the focus of our comparison study, we observed significant differences between the two data sets. The differences suggested opportunities for reinforcement.

We define the *targeting problem* as identifying source-destination pairs for layer 3 probes that reveal nodes indicated in the physical maps. Probing sources (VPs) are Looking Glasses (LG), RIPE Atlas nodes, Ark nodes or PlanetLab nodes from which we can send probes. We began our target-

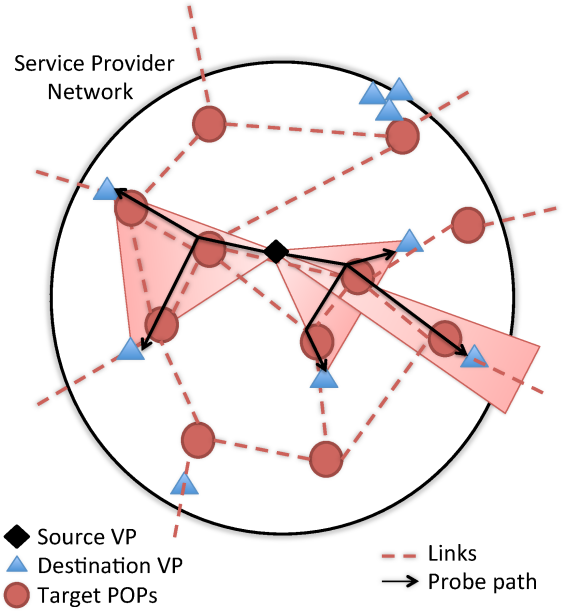


Figure 4: POPsicle targeting process for identifying infrastructure within service provider networks. Vantage Points (VPs) within the ISP that are geographically closest to the target are selected along with destinations that are geographically closest to the target and on the other side of the VP

ing analysis by conducting extensive probe-based measurements using two core ideas: *(i)* source-destination pairs should be proximal to the target POPs (identified in Atlas maps but not present in layer 3 probe data) geographically and in address space, and *(ii)* verification of measurements using multiple sources is required. Our analysis shows that probing between sources and destinations that are both within the same AS as the target(s) reveals the most physical infrastructure. These results motivated a new heuristic algorithm for probe targeting that we call *POPsicle*. Figure 4 shows an overview of the POPsicle approach. Our experiments showed that POPsicle found 2.4 times as many nodes as identified by Ark.

In our proposed work, we plan to expand on the idea of targeted probing that is reinforced with layer-1 information. Specifically, we recognize that outside of the US, the scope and consistency of data that we can collect at either the logical or physical layers is limited. Thus, we will expand the traceroute probing configurations and experiments that seek to map AS connectivity to countries and logical connectivity to routers and facilities to be layer-1 aware.

## 4 Task 2: Identifying and quantifying susceptibility to attacks

Building on the previous mapping task, we will develop an approach to identify the topological assets in a country that are most valuable for the security of its communication infrastructure (*i.e.,* the key terrain) and to evaluate if their number, role, and geographic distribution offer opportunities for an adversary to conquer a significant fraction of that terrain.

### 4.1 Graph representations for the multi-layer map

The multi-layer connectivity map we will construct in Task 1 is fundamental to understanding the roles of different components of a country's logical and physical topology, and evaluating their strategic value based on the potential impact of compromising them. We will rely on standard graph representations to represent the multi-layer map, the relationships between the various components, and annotations on nodes (e.g., geographic properties) and links (e.g., AS relationships). For example, the AS-level graph is both a multi-graph (two ASes can have multiple router-level links between them) and a hyper-graph (multiple ASes can be connected by a single "link", e.g., the shared peering fabric at an IX). We can also express the multi-layer graph as a hierarchical graph at the AS, router, PoP, facility and physical level. There are some important considerations about the multi-layer map which will inform and guide our investigation of topological weaknesses.

- Components of this map share attack vectors *e.g.,* they all have an administrative domain (attackable through impersonation, infiltration, social engineering, etc.), they all are mainly based on cyber systems that are connected to the public Internet (thus exposed to different types of cyber attacks), and all have physical facilities (subject to power outages, military operations, sabotage, etc.). Therefore in terms of attacker's capabilities, we must assume that all types of components are a potential target.
- An attacker can select combinations of targets of different types to reach their goal (*e.g.,* remotely compromising an IXP and a major AS).
- The components are interdependent (*e.g.,* compromising an IXP will affect communication between multiple ASes).

### 4.2 Computing the strategic value of topological components

Given our focus on analyzing country-level topological weaknesses, we must identify the key components of a country's topology responsible for connectivity to the rest of the Internet. As described in Section 3, we obtain the set of ASes that are active in a country. Among those ASes, some are *international points of exit* (IPoE), *i.e.,* they operate in multiple countries or are connected to an AS that operates in other countries, and thus represent the points at which traffic flows in and out of the country. From the mapping process we also obtain the paths to/from networks geolocated in the country, and how those paths map to ASes (both IPoE and internal ASes), IXPs, colocation facilities, PoPs and physical cable infrastructure. We can now define the *strategic value* of specific topological components based on a number of factors such as the number of networks/prefixes that depend on a component or estimates of the size of population served. An important step is to refine the notion of strategic value by assigning *weights* to network blocks or ASes geolocated in a country based on factors such as the services they host (*e.g.,* financial services, or government web sites may be more important than others), their type (content, access, enterprise, or transit), or the traffic volume that they send/receive. The ability to apply different weight factors makes our analysis extensible. We provide next some examples of weight factors we will consider, while emphasizing that evaluating and refining methods to assign these weights will be an important direction of this research.

1. The business type of ASes [56], *i.e.,* content, access, enterprise, or transit.
2. The number of subscribers per network, which we will estimate from studies of Internet broadband penetration and market shares [57], and estimates of per-AS end-user population from the APNIC measurement data [58].
3. The utilization of Internet address space in a network, which we will estimate based on recent address space censuses [23, 59] to weigh BGP-announced prefixes. Internet census data will allow us to consider only the fraction of the address space announced by each AS that is inferred as actively used.
4. Another possible weight factor that we will explore, is by taking into account how traffic to/from the most popular content sources in a country flows across the underlying topology. To identify popular content in a country of interest we will use country-specific website rankings provided by Alexa [60] and DNS lookups from VPs (*e.g.,* Ark or RIPE Atlas probes) in the country to resolve hostnames to IP targets. We will then use all available Ark, RIPE Atlas and Periscope VPs in the country to perform traceroute measurements to the top Alexa content destinations. Finally, we will use the multi-layer connectivity map developed in Task 1 to map the discovered paths to ASes, exchanges/facilities, and physical cable infrastructure.

### 4.3 Identifying topological components with high strategic value and quantifying risk

We will develop metrics to characterize, for a given country, the distribution of strategic value among the different components using the previously described weighting approach. Specifically, we can derive the relative importance of ASes, PoPs, IXPs, and IPoE ASes and cables along various dimensions of strategic value (in terms of user population, traffic volume, number of

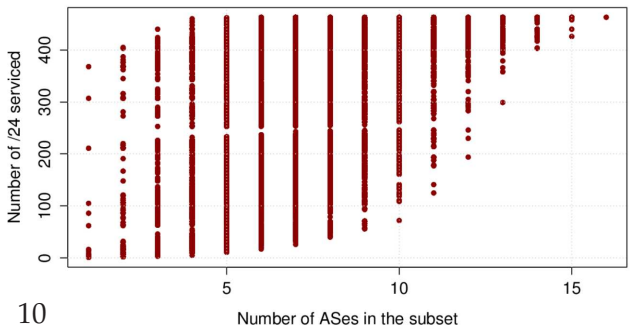


Figure 5: The number of /24 networks geolocated to Bolivia which an attacker could disconnect by compromising a certain number of ASes active in Bolivia (x-axis).

networks, or type of networks). For each of these dimensions, we will then measure the number of targets that an attacker needs to compromise in order to affect significant fractions of overall strategic value.

The simplicity and modularity of this model will enable an assessment of vulnerabilities both holistically, as well as at specific layers. For example, it is possible to evaluate these metrics separately for only certain types of targets (*e.g.,* ASes), in practice simulating that the attacker has only knowledge of a certain layer of the multi-layer map. For example, the diagram in Figure 5 shows the fraction of Bolivian /24 IPv4 address blocks that could be disconnected (y axis) if an attacker was to compromise a certain number of ASes active in Bolivia (x axis). Each point represents a different combination of ASes. If an attacker disabled or gained control of one AS, they could disconnect more than 75% of the Bolivian /24s. With the right combination of 4 ASes the attacker could disconnect 100% of the /24s.

As discussed in Section 2, the risks associated with threats considered in this project are a function of the multi-dimensional vulnerabilities of the specific targets we consider, of the capabilities of the adversary, and of political and strategic factors. It is out of the scope of this project to fully investigate these risks, as our focus is only on the topological dimension. Therefore, in defining our approach, we are agnostic to the *cost* for a given attacker to compromise each potential target. Nevertheless, our model allows the user to define the cost variable and account for it when calculating the risk of different attacks from a given adversary. In other words, we can associate each component of the logical map with *(i)* the multi-dimensional strategic value as previously described and *(ii)* a cost to compromise that component based on a profile of its vulnerabilities considering different attackers. In the proposed research we will conduct case studies in which the attacker's capabilities are limited only to certain classes of attack vectors (*e.g.,* cable cuts, cyber-attacks, military operations) and targets are associated with different costs.

Finally, we will extend this analysis to account for geographic properties. Using the geographical annotations on the multi-layer map, we can evaluate if the geographic distribution of assets in a country potentially exposes its infrastructure to compromise due to natural disasters or military attacks targeting certain areas, and how that risk compares to that of other countries. An example of this kind of analysis is our prior study of outage risks due to natural disasters [61]. In that work, we evaluated risk via the concept of *bit-risk miles*, defined as the geographic distance traveled by the traffic plus the expected outage risk encountered along the specified routing path. Our focus on bit-risk miles allowed for a first-of-its-kind analysis of the tradeoffs of shortest path routing and risk-averse routing. We conducted a detailed risk assessment of service providers by assembling diverse data sets including *(i)* detailed topological maps from Internet Atlas and peering relationships of ASes in the US, and *(ii)* historical information on different types of natural disasters which threaten physical infrastructure. Our analysis highlighted providers that have the highest risk to disaster-based outage events. The notion of bit-risk miles can be extended directly to the geographical threats considered in this proposal.

### 4.4 Evolution

In the previous subsections we described the analysis of snapshots of the multi-layer Internet topology of a country or region. Naturally, the resulting risk analysis applies to the Internet topology as it exists when the snapshot was taken. The measurement and analysis methodology we have presented is modular and repeatable, so we expect that going forward, periodic snapshots

will provide a rich source of data to enable studies of the evolution of country/region-level topological weaknesses.

An interesting question is to what extent we can leverage historical sources of data, such as archives of BGP routing data [24, 25], archived traceroute data [55, 41], and any archived maps of the physical topology to study how the topology has evolved *in the past*? We will investigate whether historical datasets are adequate to perform the types of analysis we have presented here. It may be the case that historical datasets only allow us to measure the topology of a country and analyze weaknesses at certain layers. As a concrete example, BGP data is available historically, but additional measurements required to fully develop the multi-layer map may not be available, or may cover some countries/regions better than others. Nevertheless, we believe that historical data will enable us to create at least a partial picture of the historical evolution of country-level topologies and their weaknesses. In previous work, we demonstrated how archived BGP data [24, 25] can provide surprisingly rich information about evolutionary trends in terms of geographical and economic terms [62].

A longitudinal view spanning the past years, and going forward with our periodic snapshots will enable us to tackle a set of questions about the evolution of the topology (and hence topological weaknesses) of countries/regions over time. How do the topologies of countries or regions, and hence any potential topological weaknesses, evolve over time? Do we see topological weaknesses increasing or diminishing over time?

## 5 Related work

The Internet's basic design [63] makes it robust against failures of its components. However, events such as natural or technological disasters (*e.g.*, [64, 65]), malicious attacks (*e.g.*, [66]) and benign incidents (*e.g.*, [67]) can have localized effects, including the loss of connectivity for varying numbers of Internet users. The main reasons for such localized and temporal Internet outages are typically a lack of (logical and geographic) diversity in connectivity [68, 69] and a tendency for significant physical infrastructure sharing among the affected providers [53].

Dyn Research was among the first to investigate the diversity of international Internet connectivity in countries [5]. They analyzed this problem by performing a census of all the domestic providers in each country who have direct connections to foreign providers. They based their inference on the routing tables they had access to (it is unclear how many and how selected) and classified countries in either resistant, or at low/significant/severe risk of disconnection based on the number of connections to foreign providers. Their approach is pertinent, but lacks details about how the data was collected and – as they state – is simplistic and limited to what is visible on routing tables. Besides covering a larger spectrum of attack patterns, we include in our model many more variables, including the role of different ASes, the traffic and the population that they serve, the physical layer, etc.

Roberts et al. [70] used CAIDA AS relationship data [39] to map the AS-level topology of each country in order to identify which ASes can act as "points of control". They find that in several countries only a few ASes act as points of control, a concept very similar to our IPoE AS. Kang et al. [71] introduced the notion of *routing bottlenecks* – "A routing bottleneck on the routes from S to D is a small set B of IP (layer-3) links such that Bs links are found in a majority of routes whereas the remaining links are found in very few routes." – and show that being a consequence of route-cost minimizations, is a property of Internet design. They show the pervasiveness of routing bottlenecks in 15 countries and 15 cities worldwide and measure their susceptibility to link-flooding attacks. This is one of the many aspects that our multi-layer mapping approach

allows to highlight.

Analyzing the robustness of the physical Internet has been the focus of many prior research efforts. These include studies on its robust yet fragile nature [72, 73], vulnerability [74, 75, 76, 77], survivability [78, 79], resilience analysis [80, 81, 61], reachability [82, 83], security and robustness of components [84], fault detection/localization [85, 86, 87], and the development of resilient routing protocols [88, 89, 90, 91, 92]. A part of the proposed research is based on the availability of high-fidelity maps of long-haul and metro-area fiber-optic routes in the US Internet from the Internet Atlas project [52]. Prior work examined aspects of the US long-haul fiber-optic network (*e.g.,* [68, 93]), but the resulting maps are of uncertain quality, lack important details, and are not reproducible. There have also been prior studies that examine different aspects of the Internet infrastructure and various spatial patterns that have emerged (see for example [94]). Similar to the work by Lakhina *et al.* [95] who use geolocation databases to obtain the approximate link lengths between geolocated routers, our study will consider issues related to router-level granularity.

Finally, another aspect of our work is focused on developing measurement methods that can provide data on static and dynamic properties of the Internet's infrastructure. Examples of prior work that are relevant to our study include inferring and analyzing connectivity at the AS-level [96, 30, 62, 97, 98, 99, 38, 100, 31, 101, 32, 102, 103, 104, 105], annotating AS links with the type of business relationship [106, 107, 108, 109, 110, 34, 32, 111], mapping router-level topology [112, 113, 114], mapping PoP-level topology [112, 115, 116, 117], identifying geographic locations of routers [112, 47] and IP subnet allocations [118], to name a few. Our proposed work will extend and diversify prior mapping efforts with a view of the physical topology.

## References


[1] A. Dainotti, C. Squarcella, E. Aben, K. Claffy, M. Chiesa, M. Russo, and A. Pescap, "Analysis of Country-wide Internet Outages Caused by Censorship," *IEEE/ACM Transactions on Networking*, vol. 22, pp. 1964–1977, Dec 2014.

[2] A. Dainotti, A. King, "CAIDA Blog: Syria disappears from the Internet." http://blog.caida.org/best_available_data/2012/12/05/syria-disappears-from-the-internet/.

[3] J. Cowie, "Egypt Leaves the Internet." http://www.renesys.com/blog/2011/01/egypt-leaves-the-internet.shtml, Jan 27 2011.

[4] J. Cowie, "What Libya Learned from Egypt." http://www.renesys.com/blog/2011/03/what-libya-learned-from-egypt.shtml, March 2011.

[5] Jim Cowie, "Could It Happen In Your Country?." http://research.dyn.com/2012/11/could-it-happen-in-your-countr/.

[6] Center for Strategic and International Studies (CSIS), "Disrupting the Cybersecurity Status Quo." https://www.csis.org/programs/strategic-technologies-program/cybersecurity/disrupting-cybersecurity-status-quo.

[7] Center for Strategic and International Studies (CSIS), "Significant Cyber Incidents Since 2006." https://csis-prod.s3.amazonaws.com/s3fs-public/160824_Significant_Cyber_Events_List.pdf.

[8] M. Ogun, *Terrorist use of Cyberspace and Cyber Terrorism: New Challenges and Responses*. IOS Press, 2015.

[9] US Department of Defense, "The Department of Defense Cyber Strategy." http://www.defense.gov/Portals/1/features/2015/0415_cyber-strategy/Final_2015_DoD_CYBER_STRATEGY_for_web.pdf, 2015.

[10] A. Dainotti, "North Korean Internet outages observed." http://blog.caida.org/best_available_data/2014/12/23/north-korean-internet-outages-observed/, 2014.

[11] Internation Affairs Review, "Denial-of-Service: The Estonian Cyberwar and Its Implications for U.S. National Security." http://www.iar-gwu.org/node/65.

[12] Jose Nazario, "Georgia DDoS Attacks A Quick Summary of Observations." https://www.arbornetworks.com/blog/asert/georgia-ddos-attacks-a-quick-summary-of-observations/.

[13] Secure Works, "Kyrgyzstan Under DDoS Attack From Russia." https://www.secureworks.com/blog/research-20957.

[14] A. Kozlowski, "Comparative analysis of cyberattacks on estonia, georgia and kyrgyzstan," *European Scientific Journal*, 2014.

[15] K. Giles, "Russia's 'New' Tools for Confronting the West: Continuity and Innovation in Moscows Exercise of Power." https://www.chathamhouse.org/sites/files/chathamhouse/publications/research/2016-03-21-russias-new-tools-giles.pdf.

[16] BBC News, "Egypt arrests as undersea internet cable cut off Alexandria." http://www.bbc.com/news/world-middle-east-21963100.

[17] J. Heidemann, "Active Probing of Edge Networks: Outages During Hurricane Sandy." Talk given at NANOG57 as part of panel hosted by James Cowie, February 2013.

[18] A. Dainotti, "Lessons Learned by "Measuring" the Internet During/After the Sandy Storm." Invited talk at FCC Workshop on Network Resiliency, 2013.

[19] BBC News, "New Zealand earthquake: 65 dead in Christchurch." http://www.bbc.com/news/world-asia-pacific-12533291.
[20] BBC News, "Japan earthquake: Tsunami hits north-east." http://www.bbc.com/news/world-asia-pacific-12709598.
[21] A. Dainotti, R. Amman, E. Aben, and K. Claffy, "Extracting Benefit from Harm: Using Malware Pollution to Analyze the Impact of Political and Geophysical Events on the Internet," *Computer Communications Review*, January 2012.
[22] "Presidential policy directive – critical infrastructure security and resilience," Feb 2013. https://www.whitehouse.gov/the-press-office/2013/02/12/presidential-policy-directive-critical-infrastructure-security-and-resil.
[23] A. Dainotti, K. Benson, A. King, B. Huffaker, E. Glatz, X. Dimitropoulos, P. Richter, A. Finamore, and A. Snoeren, "Lost in Space: Improving Inference of IPv4 Address Space Utilization," *IEEE Journal on Selected Areas in Communications (JSAC)*, vol. 34, pp. 1862–1876, Jun 2016.
[24] "University of Oregon Route Views Project." http://www.routeviews.org/.
[25] "RIPE routing information service (RIS)." http://www.ripe.net/ris/.
[26] Packet Clearing House, "Daily Routing Snapshots." https://www.pch.net/resources/Routing_Data/.
[27] Digital Element, "NetAcquity Edge IP Geolocation Service." http://www.digitalelement.com/solutions/netacuity-edge-premium/.
[28] B. Huffaker, M. Fomenkov, and k. claffy, "Geocompare: a comparison of public and commercial geolocation databases - Technical Report ," tech. rep., Cooperative Association for Internet Data Analysis (CAIDA), May 2011.
[29] R. D. M. Syamkumar and P. Barford, "Bigfoot: A Geo-based Visualization Methodology for Detecting BGP Threats," in *Proceedings of IEEE VizSec Conference*, 2016.
[30] R. V. Oliveira, B. Zhang, and L. Zhang, "Observing the Evolution of Internet AS Topology," in *Proceedings of ACM SIGCOMM*, 2007.
[31] 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.
[32] 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.
[33] V. Giotsas, S. Zhou, M. Luckie, and k. claffy, "Inferring Multilateral Peering," in *ACM SIGCOMM Conference on emerging Networking EXperiments and Technologies (CoNEXT)*, pp. 247–258, Dec 2013.
[34] 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.
[35] Y. Zhang, Z. Zhang, Z. M. Mao, C. Hu, and B. M. Maggs, "On the Impact of Route Monitor Selection," in *Proceedings of ACM SIGCOMM Internet Measurmenet Conference (IMC)*, 2007.
[36] 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.
[37] R. Cohen and D. Raz, "The Internet Dark Matter - On the Missing Links in the AS Connectivity Map," in *Proc. IEEE Infocom*, 2006.
[38] 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.

[39] CAIDA, "AS-rank." http://as-rank.caida.org.
[40] CAIDA, "Archipelago (Ark) Measurement Infrastructure." http://www.caida.org/projects/ark/.
[41] "RIPE Atlas." https://atlas.ripe.net/.
[42] V. Giotsas, A. Dhamdhere, and k. claffy, "Periscope: Unifying Looking Glass Querying," in *Passive and Active Network Measurement Workshop (PAM)*, Mar 2016.
[43] 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.
[44] M. Luckie, A. Dhamdhere, D. Clark, B. Huffaker, and K. Claffy, "bdrmap: Inference of Borders Between IP Networks," in *Proceedings of the ACM SIGCOMM Internet Measurement Conference (IMC)*, 2016.
[45] "INEX IPv4 Peering Matrix." https://www.inex.ie/ixp/peering-matrix.
[46] B. Huffaker, M. Fomenkov, and k. claffy, "DRoP:DNS-based Router Positioning," *ACM SIGCOMM Computer Communication Review (CCR)*, vol. 44, pp. 6–13, Jul 2014.
[47] J. Chabarek and P. Barford, "What's in a Name? Decoding Router Interface Names," in *Proceedings of the 5th ACM HotPlanet Workshop*, 2013.
[48] "PeeringDB." http://www.peeringdb.com, October.
[49] Claffy, K., Clark, D., Dhamdhere, A., "Mapping Interconnection in the Internet: Colocation, Connectivity and Congestion ." https://www.caida.org/funding/nets-congestion/.
[50] V. Giotsas, G. Smaragdakis, B. Huffaker, M. Luckie, and k. claffy, "Mapping Peering Interconnections to a Facility," in *ACM SIGCOMM Conference on emerging Networking EXperiments and Technologies (CoNEXT)*, Dec 2015.
[51] Packet Clearing House, "Internet Exchange Directory." https://prefix.pch.net/applications/ixpdir/.
[52] "Internet Atlas, A Comprehensive Repository of the Physical Internet." http://internetatlas.org/index.jsp.
[53] R. Durairajan, P. Barford, J. Sommers, and W. Willinger, "InterTubes: A Study of the US Long-haul Fiber-optic Infrastructure," in *Proceedings of the 2015 ACM Conference on Special Interest Group on Data Communication*, SIGCOMM '15, (New York, NY, USA), pp. 565–578, ACM, 2015.
[54] R. Durairajan, J. Sommers, and P. Barford, "Layer 1-informed internet topology measurement," in *Proceedings of the 2014 Conference on Internet Measurement Conference*, IMC '14, (New York, NY, USA), pp. 381–394, ACM, 2014.
[55] CAIDA, "The IPv4 Routed /24 Topology Dataset." http://www.caida.org/data/active/ipv4_routed_24_topology_dataset.xml.
[56] A. Dhamdhere, "CAIDA's AS Classification Scheme." https://www.caida.org/data/as-classification/.
[57] 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.
[58] APNIC Labs, "Visible ASNs: Customer Populations (Est.)." https://stats.labs.apnic.net/aspop/.
[59] J. Heidemann, Y. Pradkin, R. Govindan, C. Papadopoulos, G. Bartlett, and J. Bannister, "Census and survey of the visible internet," in *Proceedings of the 8th ACM SIGCOMM Conference on Internet Measurement*, IMC '08, (New York, NY, USA), pp. 169–182, ACM, 2008.

[60] Alexa, "The Top 500 Sites in each Country or Territory." http://www.alexa.com/topsites/countries, November 2012.
[61] B. Eriksson, R. Durairajan, and P. Barford, "RiskRoute: A Framework for Mitigating Network Outage Threats," in *Proceedings of ACM CoNEXT*, 2013.
[62] A. Dhamdhere and C. Dovrolis, "Twelve Years in the Evolution of the Internet Ecosystem," *IEEE/ACM Transactions on Networking*, vol. 19, pp. 1420–1433, Sep 2011.
[63] D. Clark, "The Design Philosophy of the DARPA Internet Protocols," *SIGCOMM CCR*, 1988.
[64] "Quake shakes up the net, Dec. 2006.." http://www.thestar.com.my/story/?file=%2f2006%2f12%2f28%2fnation%2f16426778&sec=nation.
[65] "Impact of the 2003 blackouts on Internet communications (Preliminary report), Nov. 2003.." http://research.dyn.com/content/uploads/2013/05/Renesys_BlackoutReport.pdf.
[66] "The National Research Council of the National Academies, The Internet under crisis conditions: Learning from September 11, 2003.." http://www.nap.edu/catalog/10569/the-internet-under-crisis-conditions-learning-from-september-11.
[67] "The Backhoe: A Real Cyberthreat." http://archive.wired.com/science/discoveries/news/2006/01/70040?currentPage=all.
[68] "A Dissertation So Good It Might Be Classified.." http://archive.wired.com/wired/archive/12.01/start.html?pg=10.
[69] "How to Destroy the Internet.." http://gizmodo.com/5912383/how-to-destroy-the-internet.
[70] H. Roberts, D. Larochelle, R. Faris, and J. Palfrey, "Mapping Local Internet Control," in *Berkman Center for Internet and Society, https://cyber.harvard.edu/netmaps/mlic_20110513.pdf*, 2011.
[71] M. S. Kang and V. D. Gligor, "Routing bottlenecks in the internet: Causes, exploits, and countermeasures," in *Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security*, CCS '14, (New York, NY, USA), pp. 321–333, ACM, 2014.
[72] W. Willinger and J. Doyle, "Robustness and the Internet: Design and evolution," *Robust-Design: A Repertoire of Biological, Ecological, and Engineering Case Studies*, 2002.
[73] J. Doyle, D. Alderson, L. Li, S. Low, M. Roughan, R. Tanaka, and W. Willinger, "The "Robust Yet Fragile" Nature of the Internet," in *PNAS*, 2005.
[74] S. Gorman, L. Schintler, R. Kulkarni, and R. Stough, "The Revenge of Distance: Vulnerability Analysis of Critical Information Infrastructure," *JCCM*, 2004.
[75] S. Gorman, *Networks, Security And Complexity: The Role of Public Policy in Critical Infrastructure Protection*. Edward Elgar, 2005.
[76] L. Zhou, "Vulnerability Analysis of the Physical Part of the Internet," in *IJCI*, 2010.
[77] T. Lewis, *Critical Infrastructure Protection in Homeland Security: Defending a Networked Nation*. John Wiley & Sons, Inc., 2006.
[78] P. Heegaard and K. Trivedi, "Network Survivability Modeling," in *Computer Networks*, 2009.
[79] P. Ho, J. Tapolcai, and H. Mouftah, "On Achieving Optimal Survivable Routing for Shared Protection in Survivable Next-Generation Internet," *IEEE Transactions on Reliability*, 2004.
[80] J. Wu, Y. Zhang, M. Mao, and K. Shin, "Internet Routing Resilience to Failures: Analysis and Implications," in *Proceedings of ACM CoNEXT*, 2007.
[81] P. Agarwal, A. Efrat, S. Ganjugunte, D. Hay, S. Sankararaman, and G. Zussman, "The Resilience of WDM Networks to Probabilistic Geographical Failures," in *Proceedings of IEEE INFOCOM*, 2011.

[82] R. Bush, O. Maennel, M. Roughan, and S. Uhlig, "Internet Optometry: Assessing the Broken Glasses in Internet Reachability," in *Proceedings of ACM Internet Measurement Conference*, 2009.

[83] E. Katz-Bassett, H. Madhyastha, J. John, A. Krishnamurthy, D. Wetherall, and T. Anderson, "Studying Black Holes in the Internet with Hubble," in *Proceedings of USENIX NSDI*, 2008.

[84] K. Kant and C. Deccio, *Security and Robustness in the Internet Infrastructure*. Morgan Kaufmann, 2012.

[85] E. Katz-Bassett, C. Scott, D. R. Choffnes, I. Cunha, V. Valancius, N. Feamster, H. V. Madhyastha, T. E. Anderson, and A. Krishnamurthy, "Lifeguard: practical repair of persistent route failures.," in *SIGCOMM*, pp. 395–406, ACM, 2012.

[86] L.Quan, J. Heidemann, and Y. Pradkin, "Detecting Internet Outages with Precise Active Probing (extended)," in *USF Technical Report*, 2012.

[87] E. Glatz and X. Dimitropoulos, "Classifying Internet One-way Traffic," in *Proceedings of ACM Internet Measurement Conference*, 2012.

[88] H. Wang, Y. Yang, P. Liu, J. Wang, A. Gerber, and A. Greenberg, "Reliability as an Interdomain Service," in *Proceedings of ACM SIGCOMM*, 2007.

[89] D. Andersen, H. Balakrishnan, F. Kaashoek, and R. Morris, "Resilient Overlay Networks," in *Proceedings of ACM SOSP*, 2001.

[90] Y. Zhu, A. Bavier, N. Feamster, S. Rangarajan, and J. Rexford, "UFO: A Resilient Layered Routing Architecture," *SIGCOMM CCR*, 2008.

[91] A. Hansen, A. Kvalbein, T. Cicic, and S. Gjessing, "Resilient Routing Layers for Network Disaster Planning," in *Proceedings of IEEE ICN*, 2005.

[92] K. Gummadi, H. Madhyastha, S. Gribble, H. Levy, and D. Wetherall, "Improving the Reliability of Internet Paths with One-hop Source Routing," in *Proceedings of USENIX OSDI*, 2004.

[93] "GMU Mapping Project..." `http://gembinski.com/interactive/GMU/research.html`.

[94] E. Malecki, "The Economic Geography of the Internet's Infrastructure," *Economic Geography*, 2002.

[95] A. Lakhina, J. Byers, M. Crovella, and I. Matta, "On the Geographic Location of Internet Resources," in *IEEE JSAC*, 2003.

[96] G. Siganos, M. Faloutsos, and C. Faloutsos, "The Evolution of the Internet: Topology and Routing," *University of California, Riverside technical report*, 2002.

[97] 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.

[98] 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.

[99] 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.

[100] H. Chang, R. Govindan, S. Jamin, S. Shekker, and W. Willinger, "Towards capturing representative AS-level Internet topologies," *Computer Networks*, vol. 44, no. 6, pp. 737–755, 2004.

[101] 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.

[102] 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.

[103] B. Augustin, B. Krishnamurthy, and W. Willinger, "IXPs: Mapped?," in *Proceedings of the ACM SIGCOMM Internet Measurement Conference (IMC)*, 2009.
[104] 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.
[105] 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.
[106] L. Gao, "On inferring autonomous system relationships in the Internet," *IEEE/ACM Transactions on Networking (ToN)*, vol. 9, no. 6, pp. 733–745, 2001.
[107] L. Subramanian, S. Agarwal, J. Rexford, and R. H. Katz, "Characterizing the Internet hierarchy from multiple vantage points," in *INFOCOM 2002. Twenty-First Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE*, vol. 2, pp. 618–627, IEEE, 2002.
[108] G. Di Battista, M. Patrignani, and M. Pizzonia, "Computing the types of the relationships between autonomous systems," in *INFOCOM 2003. Twenty-Second Annual Joint Conference of the IEEE Computer and Communications. IEEE Societies*, vol. 1, pp. 156–165, IEEE, 2003.
[109] X. Dimitropoulos, D. Krioukov, M. Fomenkov, B. Huffaker, Y. Hyun, and G. Riley, "AS relationships: Inference and validation," *ACM SIGCOMM Computer Communication Review*, vol. 37, no. 1, pp. 29–40, 2007.
[110] "Internet topology collection." `http://irl.cs.ucla.edu/topology/`.
[111] E. Gregori, A. Improta, L. Lenzini, L. Rossi, and L. Sani, "BGP and Inter-AS Economic Relationships," in *IFIP Networking Proceedings, Part II*, (Valencia, Spain), pp. 54–67, May 2011.
[112] N. Spring, R. Mahajan, D. Wetherall, and T. Anderson, "Measuring ISP topologies with Rocketfuel," *IEEE/ACM Transactions on Networking*, vol. 12, no. 1, 2004.
[113] R. Durairajan, J. Sommers, and P. Barford, "Layer 1-Informed Internet Topology Measurement," in *Proceedings of the ACM Internet Measurement Conference*, 2014.
[114] 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.
[115] H. Madhyastha, E. Katz-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.
[116] D. Feldman, Y. Shavitt, and N. Zilderman, "A structural approach for PoP geo-location," *Computer Networks*, Decmeber 2012.
[117] A. Rasti, N. Magharei, R. Rejaie, and W. l. Willinger, "Eyeball ases: From geography to connectivity," *Internet Measurement Conference (IMC)*, November 2010.
[118] M. Gunes and K. Sarac, "Inferring Subnets in Router-level Topology Collection Studies," in *Proceedings of ACM Internet Measurement Conference*, 2007.