ABSTRACT

KIM, SANGMIN. The Internet Topology: Illusion and Reality. (Under the direction of Dr. Khaled Harfoush).

Research studies targeted at unveiling the Internet topology are essential for understanding the performance of the Internet and its resilience to failures or distributed attacks, and for generating realistic topologies to simulate Internet protocols and applications. An accurate understanding of the complex Internet structure and behavior, while very rewarding, is very challenging and in fact is a source of controversy in the networking research community. Till now, there is a lack of understanding of the Internet complexity

In this thesis, we make the following contributions. First, we propose an efficient tool, AROMA, to unveil Layer-3 maps of the Internet and use it to reveal ISP maps. AROMA reveals the same number of routers and links as existing tools such as *Rocketfuel* after sending less than 5.1% of the number of probes used by Rocketfuel, and reveals at least 100% more links and routers than Rocketuel while using the same number of probe packets. Second, we study the limitations of existing layer-3 tools such as *traceroute* in unveiling the details of the Internet structure and identify that the power law connectivity observed in the Internet topology is not an illusion as suggested by some researchers. It is mainly manifested due to the blindness of traceroute to layer-2 devices, and this manifestation will persist independent of the nature of the underlying physical topology. Third, we provide a realistic Internet topology model, HINT, which captures the Internet structure and features. HINT is based on economical, performance and security constraints that are typically used to construct networks. Matching HINT topologies to known ISP topologies confirms its superiority to existing Internet topology models.

The Internet Topology: Illusion and Reality

by

Sangmin Kim

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

Computer Science

Raleigh, North Carolina

2007

Approved By:

Dr. George Rouskas	Dr. Carla Savage
Dr. Khaled Harfoush Chair of Advisory Committee	Dr. Arne Nilsson

Dedication

This thesis is dedicated to my wife, two sons, and my parents who have supported me all the way since the beginning of my studies.

Biography

Sangmin Kim was born in Pusan, Korea in 1967. He graduated from Seoul National University in February 1990 with a bachelors degree in Urban Engineering. His Masters degree is from University of North Carolina at Chapel Hill in Urban Planning. Before beginning to work toward the Philosophy of Doctor in Computer Science at the North Carolina State University at Raleigh, he worked as Master of Science degree in Computer Science at the North Carolina State University at Raleigh. His major research interests are Computer Networking, Wireless Sensor Networks, and Computer Programming.

Acknowledgements

Foremost, I would like to thank my advisor Dr. Khaled Harfoush for his thoughtful direction and help for my research. This thesis would not have been materialized without his sincere guidance. I owe Dr. Khaled Harfoush every accomplishments and experiences from my doctoral research at North Carolina State University. I also would like to acknowledge my advisory committee members, Dr. Arne A. Nilsson, Dr. George N. Rouskas and Dr. Carla D. Savage. I sincerely appreciate their comments and constructive suggestions on my work. Last but not least, I want to thank my wife, Misuk Kim, and my sons, Hyeongsup Kim and Hyosup Kim, for their patience, sacrifice and support during my study.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

F Starting from an experimental infrastructure with few hosts and routers, the Internet has evolved to an enormous network with multi-million nodes, and is directly affecting our lives in many ways. Understanding the Internet topology's structural and behavioral characteristics is essential to: (1) simulate new network protocols and evaluate their performance before actual Internet deployment, (2) improve network management capabilities and help better plan for network infrastructures, (3) optimize routing protocols, (4) promote the deployment of network-aware applications which help better utilize network resources, (5) protect the Internet by isolating security bleaches and by identifying vulnerable sections of the Internet that need to be reinforced. An accurate understanding of the complex Internet structure and behavior, while very rewarding, is very challenging and in fact is a source of controversy in the networking research community. Till now, there is a lack of understanding of the Internet complexity [51].

Internet topology models evolved from random topologies in which nodes are connected randomly [33,70], to hierarchical topologies in which nodes are clustered and clusters are connected in a hierarchical manner [33, 74], to scale-free topologies [20, 34, 35] in which few nodes exhibit more connections than others and the degree of nodes exhibit a power law distribution. The transition from one model to the other is justified by Internet measurements that highlight the shortcomings of the previous models in capturing realistic Internet characteristics. However, till now, there is no agreed upon Internet topology model. The power law model, which is backed by the support of a large fraction of the networking community, is questioned by many others [29, 44, 71]. Recent research suggests that (1) the tools used to infer the characteristics of the Internet topology exhibit a sampling bias [44], and that (2) the datasets justifying the claim that the Internet node degree exhibits power law distribution is significantly incomplete [28, 29]. (3) As shown by Lun et al. in [45], different networks may have the same degree distribution. This argues, and we agree with this argument, that the degree distribution, the most popular metric used to characterize the Internet, is inherently inadequate to model the Internet topology. As a result, there is a need to rethink the metrics used to characterize and model the Internet topology. (4) Furthermore, while researchers have focused on identifying the structure of the Internet connectivity, there is hardly any work on characterizing the *features* of this structure like the capacity bandwidth (transmission speed) of the Internet links, the locations of the most congested links, etc.

Also, researchers have tried to identify the economical and technical drivers responsible for the observed large-scale network properties by focusing on the *causal* forces at work rather than only measuring the result [17, 25, 34, 45].[1] The proposals in this direction

[1] Note that these efforts are in line with Barabasi's proposal, which suggests that incremental growth and

suggest that the causal effects of the Internet evolution are in general tradeoffs between yield, cost of resources, and tolerance to risk. This line of research is very rewarding as it does not only allow researchers to model current Internet characteristics but also allows for the possibility to predict the changes in Internet behavior as technology advances and new constraints are introduced, for example on the maximum router connectivity, and as new paradigms, like peer-to-peer systems or end-system multicast, become more widely deployed. The difficulty in this line of research is attributed to the fact that the complex social and economic *forces* which shape the Internet are not well understood [55, 71]. Also, these models need to be validated through actual Internet measurements capable of capturing the accurate Internet structure and feature, which so far do not exist. Till now, there is no topology model that is capable of *closing the loop* by hypothesizing the phenomenas driving the Internet topology's evolution while confirming the hypothesis with accurate Internet measurements. Current models tend to be generic in nature and can lead to wrong conclusions about the causes of the emergent phenomena.

This thesis is dedicated to better understanding the Internet structure and features and to provide more realistic Internet topology models. To this end, we make the following contributions.

- We propose an efficient tool, AROMA, to unveil Layer-3 maps of the Internet and use it to reveal the maps of four major ISPs. AROMA reveals the same number of routers and links as existing tools such as *Rocketfuel* after sending less than 5.1% of the number of probes used by Rocketfuel, and reveals at least 100% more links and routers than Rocketuel while using the same number of probe packets.

preferential attachment are the *causes* of the power law degree distribution in the Internet topology.

- We study the limitations of existing layer-3 tools such as *traceroute* in unveiling the details of the Internet structure and identify that the power law connectivity observed in the Internet topology is not an illusion as suggested by some researchers. It is mainly manifested due to the blindness of traceroute to layer-2 devices, and this manifestation will persist independent of the nature of the underlying physical topology.

- We provide a realistic Internet topology model, HINT, which captures the Internet structure and features. HINT is based on economical, performance and security constraints that are typically used to construct networks. As opposed to existing Internet topology models, HINT networks reveal features such as the traffic load on links and nodes. Matching HINT topologies to known ISP topologies confirms its superiority to existing Internet topology models.

The rest of this thesis is organized as follows. In Chapter 2 we survey research related to Internet measurements. In Chapter 3 we provide the details of the AROMA tool, use it conduct Internet measurements, analyze the results, and compare it existing results. In Chapter 4 we explore the limitations of traceroute-like tools in unveiling the Internet characteristics and resolve the long-standing power-law controversy. In Chapter 5 we provide the details of the HINT model and make a case for its superiority to existing models in capturing the Internet structure. We finally conclude in Chapter 6.

Chapter 2

Related Works

For more than 40 years, scientists have modeled complex networks as being completely random. The roots of this paradigm are in the work of Erdös and Rényi's (ER) [33] who suggested that communication networks and networks observed in life sciences can be modeled by connecting its nodes with randomly placed links. Waxman et al. [70] developed a variant of the classical Erdös-Rényi's (ER) random graph in which the placement of a link between a pair of nodes is based on probabilities that are biased by the Euclidean distance between these nodes. Their model became widely used in protocol simulations for more than a decade.

More recently, Zegura et al. emphasized the fundamental role of *hierarchy* in real network topologies [73]. They observed that paths are primarily affected by the hierarchy imposed by transit and stub domains. This reasoning became widely accepted and motivated a new kind of topology generators, referred as *structure*-based generators. The GT-ITM (Georgia Tech Internetnetwork Topology Models) introduced by Calvert et al. [24, 74] pioneered structure-based generators. Their model generates topologies using three different levels of hierarchy: transit domains, stub domains and LAN's attached to stub nodes. The model constructs the topology in a piecewise fashion, where the pieces correspond to domains at the different levels in the hierarchy. The connectivity between domains is treated differently than the connectivity within a domains and a set of input parameters determines the level of connectivity for each case.

A new line of research was initiated when new evidence on the structure of the Internet topology was made available by the Faloutsos brothers [35, 60]. In their study, they observed the presence of power laws in the degree distributions of the router-maps and the AS-maps of the Internet. These findings became the foundation for a new line of topology generators, namely *degree*-based topology generators. Several studies have found that topologies generated by *structure*-based generators do not lead to power law degree distributions [40, 48, 53]. Also, a comparative study conducted by Tangmunarunkit et al. concluded that *degree*-based topology generators reproduce the structure presented in the Internet more accurately than *structure*-based topology generators [65]. Their main finding is that the Internet structure is *smoother* than the one obtained by rigorous *structure*-based topology generators.

Meanwhile, in several other fields as diverse as Biology and Physics, some studies have attempted to uncover the mechanisms that shape massive graphs such as metabolic networks and protein interactions [11, 19, 20, 68] to understand the phenomena leading to the power law vertex degree distribution observed in router and AS maps. Of these studies, Barabasi and Albert's work [20] has attracted the most attention in the networking community as it proposes a very appealing construction of network topologies. Barabasi and Albert suggested that the power law degree distribution arises in complex networks, such the Internet, from a simple *dynamic* model that combines *incremental growth* with a *preference* for new nodes to connect to existing ones that are already well connected. As a consequence, high-degree nodes are likely to attract more connections, resulting in the power law degree distribution.[1] Influenced by Barabasi's model, several *degree-based* generators have been proposed [10,23,36,40,48,49,53,72]. These generators can be distinguished by the probabilistic method used in their generation process. Among these, BRITE (Boston University Representative Internet topology gEnerator [48,49] developed by Medina et al. incorporates (1) preferential attachment and (2) incremental growth along with (3) distribution of nodes in space and (4) locality of edge connections. Medina et al. also explored the individual relevance of these forces in synthetically generated network topologies and confirmed Barabasi's hypotheses. The generator proposed by Palmer et al. [53] and the GRG (General Model of Random Graphs) tool [30] by Chung et al. generate the topology starting from an expected degree sequence obtained from a degree distribution that follows a power law distribution. In GRG [30] the construction proceeds by first assigning each node its expected degree and then inserting edges between the nodes according to a probability that is proportional to the product of the degrees of the two given endpoints. If the assigned expected node degree sequence follows a power law, the generated graph's node degree distribution will exhibit the same power law. The construction proposed in [53] is very similar to GRG's construction with the exception that the degree sequence is obtained

[1] For a precisely defined model that incorporates the key features of preferential attachment refer to [22] and references therein.

from real AS-maps. In [10], Aiello proposed the PLRG (Power Law Random Generator) topology generator, which also attempts to replicate a given degree sequence, however its construction process is very different to the previous one. The construction in PLRG involves forming a set L of nodes containing as many distinct copies of a given vertex as the degree of that vertex, choosing a random matching of the elements of L, and applying a mapping of a given matching into an appropriate multigraph. It is believed that PLRG [10] and GRG [30] are "asymptotically equivalent, subject to bounding error estimates" [10].

Degree-based Internet topology generators have recently faced criticism. In [28,29], Chang et al. questioned the widespread belief that the Internet's degree distribution follows a power law and suggested that the vertex degree of the Internet, while highly variable which is a common pattern in power law topologies, could be represented by distributions similar to Weibull. The Weibull distribution exhibits a heavy-tail but the rest of the distribution can have any shape. Chen et al. [29] also questioned Faloutsos' results claiming that the BGP routing tables used for the study are inconsistent/incomplete, and suggested that the Internet's AS level topology may have well developed following a different set of growth processes than those proposed by Barabasi [20]. This claim has been supported by Willinger et al. [71] and more recently by Park et al. [56]. Willinger et al. [71] compared the growth rate of a real AS-map and the one produced by Barabasi's model and found that while both produce similar power law degree distributions, they differ *significantly* in their dynamic behavior, which suggests that topology generators based on Barabasi's model are in general not able to provide accurate explanation of the overall network structure. More recently, Park et al. [56] made a similar experiment to the aforementioned and concluded that no existing model –including Barabasi's model– succeeds to follow the dynamic behavior of the Internet. On top of all the previous arguments, Mandelbrot makes the argument that power law type distributions should be expected to arise ubiquitously for purely mathematical and statistical reasons and hence require no special explanation! [45]. Also, as shown by Lun et al. in [45], different networks may have the same degree distribution. This argues, and we agree with this argument, that the degree distribution is inherently inadequate to model the Internet topology. As a result, there is a need to rethink the metrics used to characterize and model the Internet topology.

In an attempt to solve the controversy, various research efforts targeted measuring the router-level Internet topology *directly* using the *traceroute* tool [3,39,61,63]. Typically, these efforts run traceroute from k sources towards m destinations and use all uncovered routers as an approximation, $\hat{G}$, of the Internet router-level topology, G. The degree distribution of the nodes in $\hat{G}$ is used to approximate the degree distribution G. However, the large-scale nature of the Internet results in a pretty loose approximation. Furthermore, Lakhina et al. have shown in [44] that this way of sampling the Internet is highly biased towards power law topologies, that is, using traceroute measurements can lead to power law degree distributions even if the underlying topology does not follow a power law degree distribution.

Also, researchers have tried to identify the economical and technical drivers responsible for the observed large-scale network properties; that is, by focusing on the *causal* forces at work rather than only measuring the result [17,25,34,45]. In [25], Carlson and Doyle suggest that power laws in systems optimized by engineering design are due to tradeoffs between yield, cost of resources, and tolerance to risk. In [34], Fabrikant et al. propose a model in which nodes join uniformly at random within some Euclidean space, and the newly created edges attempt to balance the distance from its new neighbor with the desire to minimize the average number of hops to other nodes. Based on their results, the authors suggest that power laws tend to arise as a result of complex, multi-objective optimization. In [17], Alderson et al. explore a similar heuristic at the ISP level. More recently, Li et al. [45] observed that simple heuristically designed and optimized models that reconcile the tradeoff between different technical constraints in routers such as link costs and user traffic, result in configurations that also exhibit power law in their degree distribution. In [64, 71], the authors question the Barabasi model and suggest that there might be other causes for the highly variable degree distribution of the Internet such as the AS size.

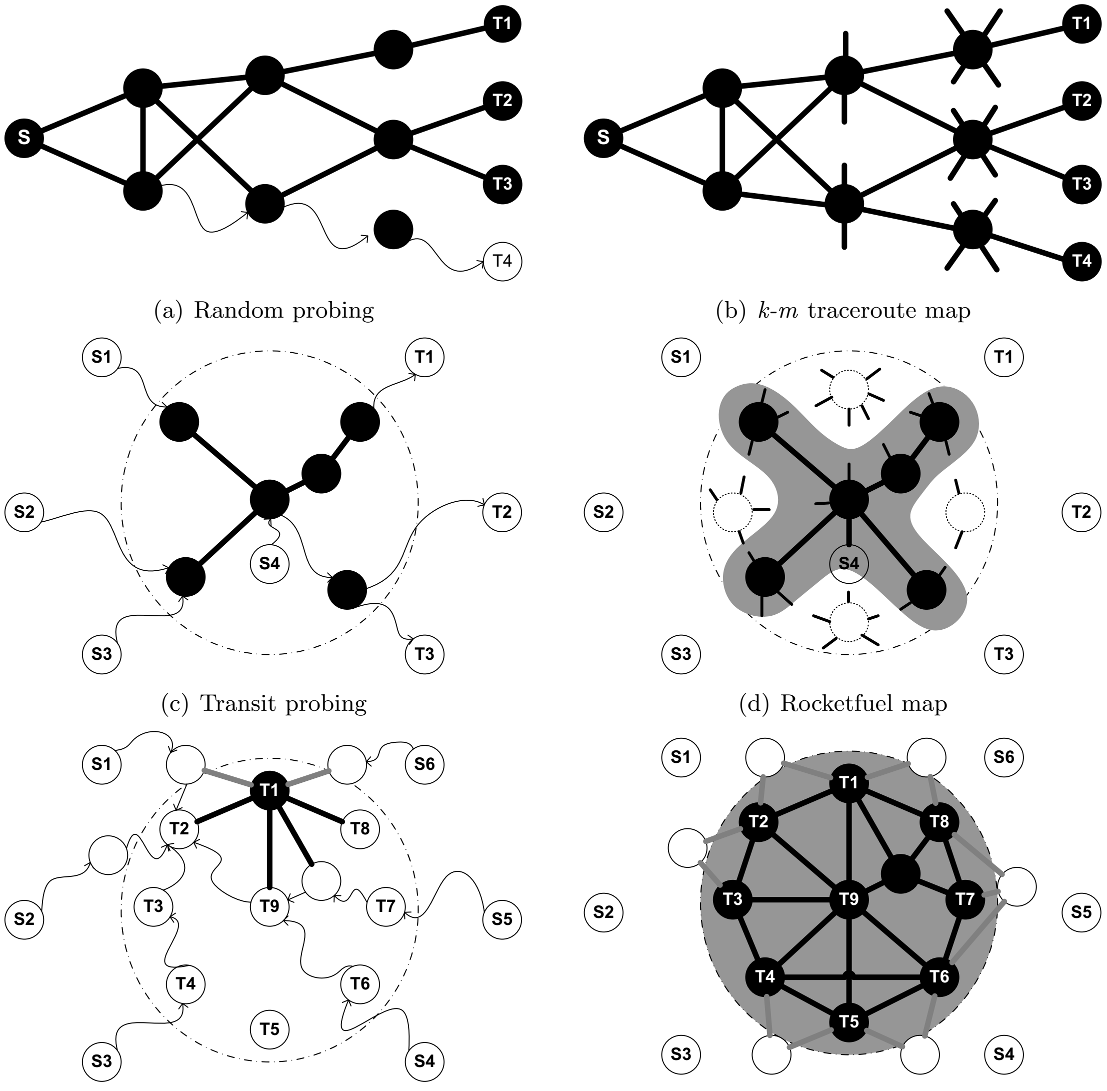
Chapter 3

AROMA: Accurate Router-Level Maps

The huge size and complexity of the Internet, its anarchistic evolution, together with the fact that ISPs do not publish their router-level maps for security reasons contribute to the lack of an accurate Internet Router-Level Maps IRLM. Despite the large body of research targeted at unveiling the IRLM [39, 59, 61], the task remains challenging and the conclusions drawn from these efforts are to-date a source of controversy and debate in the research community.

In general, techniques to uncover IRLM can be classified into two categories. The first relies on a set of k controlled hosts (sources) spread across the Internet sending *traceroute* packets to m arbitrary Internet destinations [3, 9, 37, 39, 47, 63] – Figure 3 (A,B). The topology incorporating *all* the identified routers, together with the links connecting them, is used to represent the IRLM. Donnet et al. have proposed techniques to make this approach more efficient by avoiding probing redundancy over already visited hops [32]. However, the *k-m* traceroute approach has been questioned in [44] and shown to be biased towards routers that are close to the source hosts. That is, links to routers that are close to the source hosts are revealed better than those close to the destinations. Therefore, the resulting graph is deemed unrepresentative of the IRLM. The second category does not traceroute to arbitrary destinations; instead, it focuses on an internet *area*, typically corresponding to an autonomous system (AS). Published BGP tables permit wise selection of the traceroute destination hosts, such that traceroute packets cross the target AS, thus avoiding probing redundancy – Figure 3 (C,D). The fact that it is targeting the complete router-level topologies of ASes permits the study of common factors driving the layout of AS infrastructures, the factors driving AS interconnectedness, and the structure of IRLM in general. The Rocketfuel [61, 66] tool belongs to this category. As we make the case throughout this chapter, Rocketfuel unveils mostly *backbone* routers and links of the targeted AS. This is mainly due to the Border Gateway Protocol (BGP) [57], the mainstream inter-domain routing protocol, which typically leads transit probes to the closest exit router through the targeted AS's backbone links.

Contributions: In this chapter, we propose AROMA, an IRLM mapping tool. AROMA tries to identify all interfaces/links associated with each router in a targeted AS by *directly* probing these interfaces from multiple vantage points – Figure 3 (E,F). As opposed to k-m traceroute, *only* interfaces that have been probed from the vantage points are included in the final map. As opposed to rocketfuel, AROMA does not target address prefixes that transit the targeted AS and thus can deeply penetrate inside an AS; and while Rocketfuel



(a) Random probing

(b) *k-m* traceroute map

(c) Transit probing

(d) Rocketfuel map

(e) Targeted probing

(f) AROMA map

Figure 3.1: Probing philosophies to construct IRLMs: (a) probing random destinations and (b) the resulting *k-m* traceroute- map; (c) probing destinations that lead to probes transiting a particular AS, and (d) the resulting Rocketfuel map; and, (e) targeted probing to a particular AS, and (f) the resulting AROMA map.

relies on *insider* probing servers, if available, to probe *useful* prefixes identified by BGP tables, traced paths still mainly follow backbone paths in the targeted AS towards their intended destinations, and mostly miss the AS details. Address space information of targeted ASes is typically available from authorized registry services like ARIN [1], APNIC [2], and RIPE [7]. The vantage points are filtered-out to avoid probing redundancy and the probed address space is filtered-out to avoid probing IP addresses assigned to end-hosts and IP addresses assigned to nodes in customer ASes, which partly contributes to AROMA's efficiency. Furthermore, AROMA trades off its revealed AS details (*completeness*) for efficiency. These efficiency-aware enhancements result in significant efficiency. Specifically, our results indicate that AROMA reveals the same number of routers and links as Rocketfuel after sending less than 5.1% of the number of probes consumed by Rocketfuel in all investigated ASes, and reveals between 100% and 1700% more links and routers than Rocketuel after consuming the same number of probes as Rocketfuel. We use AROMA to draw the maps of four major ISP networks (SprintLink, Level3, Verio, and Abovenet) and report on the structure of their networks. We also revisit the conclusions drawn by earlier research on the Internet structure and the degree distribution of the Internet routers.[1] Surprisingly, AROMA maps consistently reveal that core routers have a higher degree than edge routers in contrast to the conclusion drawn in [45]. The maps also reveal that routers' degree distribution consistently follows a power-law. This is in contrast to the widely accepted Weibull distribution revealed in the Rocketfuel traces, and in agreement with the highly questioned *k*-*m* traceroute tools. These results highlight the need for more thorough investigations of

[1] Our choice of the degree distribution simply reflects the Internet measurement community's interest in this metric. Finding more interesting metrics that care capable of capturing the Internet structural and behavioral characteristics is an interesting research area on its own.

the different factors driving the Internet structure and behavior.

The rest of the chapter is organized as follows. In Section 3.1 we elaborate on the basic techniques used by AROMA to draw maps. We illustrate the techniques through our four case studies for the Sprintlink, Level3, Verio, and Abovenet networks. In Section 3.2 we highlight tradeoff between completeness and efficiency in the AROMA maps and the edge that it enjoys over current tools. In Section 3.3 we analyze the AROMA maps for the four case studies and and question common perceptions about the Internet router-level topology. Summary is in Section 5.3.

3.1 Detailed AROMA Maps

The Aroma mapping process is sketched in Figure 3.2. Initially, AROMA is fed the IP address space corresponding to the targeted AS from an authorized registry service [1,2,7] and a list of potential probing servers (controlled machines). These lists are refined in Steps 1 and 2 and only k servers and m IP addresses are selected. The process then proceeds recursively in phases. In each phase, targeted interfaces are probed and a new list of target interfaces is identified, which is then fed to the next phase for further processing – steps 3 and 4. The process converges when no new targets are identified. Following convergence, a process for alias resolution is carried before the final map is generated – step 5. We next describe the details of each step.

Table 3.1: Number of IP addresses Targeted by AROMA.

AS	AS #	Addr. Space Size	Addr. with Names	Targeted Addr.
SprintLink	1239	11,615,500	397,920	397,890
Level3	3356	43,506,700	11,623,900	306,500
Verio	2914	6,895,000	788,800	788,800
AboveNet	6461	884,000	466,350	184,630

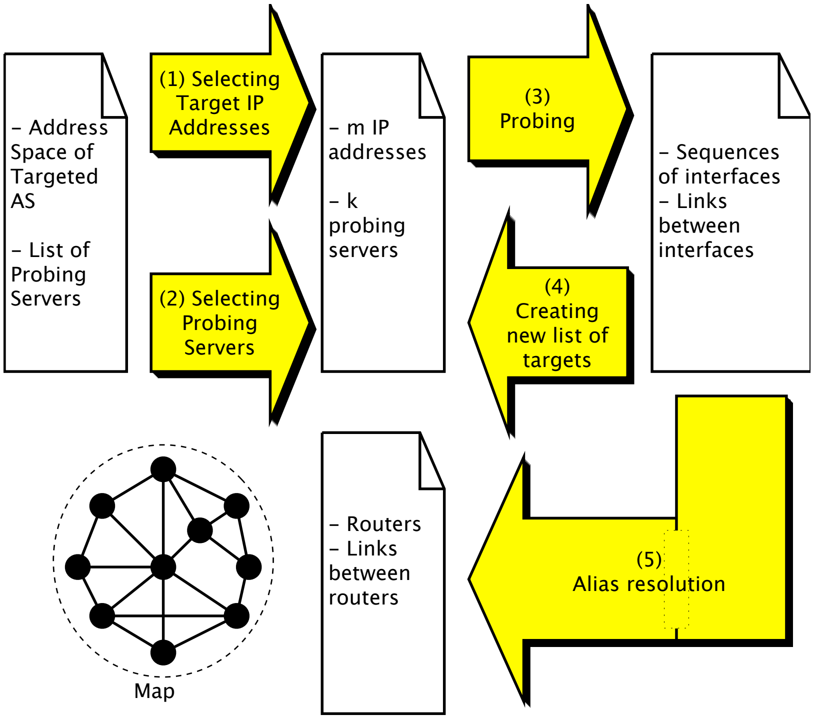

– Address Space of Targeted AS
– List of Probing Servers
(1) Selecting Target IP Addresses
(2) Selecting Probing Servers
– m IP addresses
– k probing servers
(3) Probing
– Sequences of interfaces
– Links between interfaces
(4) Creating new list of targets
(5) Alias resolution
– Routers
– Links between routers
Map

Figure 3.2: AROMA Diagram

3.1.1 Selecting Targeted IP Addresses

Not every IP address in the address space of the targeted AS needs to be probed. Specifically, IP addresses that belong to endpoints and unused IP addresses do not need to be probed since our objective is to draw a router level map. In order to distinguish routers' IP addresses from the rest, we rely on querying the Domain Name Service (DNS) [5]. Typically, ISP administrators assign meaningful names to their routers for management purposes. Naming convention is different from ISP to ISP, but a name typically includes the location and role of the router such as "gw" for gateway routers and "bb" or "bbr" for backbone routers. Reverse DNS is used to convert IP addresses to their names. By selecting only the IP addresses that are assigned names referring to the target AS, then (1) the unused IP addresses and (2) IP addresses belonging to customer ASes of the target AS are left out, which serves our purpose. However, (1) end-hosts in the target AS that have registered DNS names are included and (2) routers in the target AS that are not assigned names are singled out. The former case is resolved by disregarding IP addresses with names including the substrings ”dialup”, ”DSL”, etc. While the latter case is quite uncommon, unnamed IP addresses of routers are revealed through AROMA's recursive process, which we describe in the following sections. Table 3.1 shows the size of the address space corresponding to the four ISPs that we are mapping in this chapter: SprintLink, Level3, Verio, and AboveNet. The table also shows the number of addresses that are assigned names, and the number of selected addresses for probing (neglecting IP addresses with names belonging to customers ASes and those including the ”dialup” or ”DSL” substrings).

3.1.2 Selecting Probing Servers

We have access to almost 280 probing servers that are geographically dispersed all over the globe on the Planetlab facility [6]. However, for each targeted AS, a small fraction of these servers is enough to reveal the routers/links that would be revealed if all servers were used for probing. The reason being that if traceroute probes from two servers to the same destination(s) enter the targeted AS through the same ingress router then these traceroutes will follow the same paths inside the AS and will reveal the exact same information. In this case, only one of these servers should be used for probing. AROMA applies this idea by randomly picking a small set of IP addresses in the targeted AS and probes this set from all 280 probing servers. Then, the largest set of servers for which probe packets enter the targeted AS through different ingress routers are selected for probing. This selection process brings the number of probing servers for the Sprintlink network to 93, and for Level3, Verio, and Abovenet to 92, 105, and 51, respectively.

3.1.3 Probing

In the probing step, each IP address in the target list becomes the destination of traceroute packets (is probed) from multiple probing servers. In order to avoid overwhelming any single router, probing traffic was rate-limited and the order in which IP addresses were probed was randomized. Recall that the traceroute utility relies on a sequence of TTL-limited ICMP packets and a random destination port number. These packets will reveal interfaces of routers along the path to the destination through ICMP_TIME_EXCEEDED messages, and will reveal an interface of the destination router through an ICMP_PORT_

Table 3.2: Percentages of probes that successfully reached their intended destinations and of probes that were blocked.

	Spintlink	Level3	Verio	Abovenet
Successful	32.3%	59.8%	21.6%	35.8%
Blocked	67.7%	40.2%	78.4%	64.2%

Table 3.3: Number of target IPs in each phase.

AS	Phase I	Phase II	Phase III
Sprintlink	397,890	477	5
Level3	306,500	868	8
Verio	788,800	1,171	16
Abovenet	184,630	224	3

UNREACHABLE message. The result of probing is thus sequences of interfaces and links connecting these interfaces. Throughout our mappings, we did not receive complaints from AS administrators. However, many of the probes were *blocked* at some point along the path towards the destination. This is mostly due to firewalls or other security measures. Table 3.2 shows the percentage of probes that successfully made it to their intended destination and the percentage of probes that were blocked.

3.1.4 Creating New List of targets

Any revealed interface from the probing step, that belongs to the targeted AS and that was not included in the target list is included, and steps 3 and 4 are repeated. Identifying whether an interface belongs to the targeted AS is done by checking the interface's IP address against the pool of IP address of the targeted AS, obtained from the registry service. If no new interfaces are revealed, the mapping proceeds to the alias resolution step – step 5. Our results indicate that the number of newly revealed interfaces drops dramatically with each phase, and no new interfaces are revealed after the third phase. Table 3.3 shows the number of IP addresses that are probed in each phase for our four case studies.

3.1.5 Alias Resolution

Alias resolution refers to the process of clustering interfaces (IP addresses) belonging to the same router together. Several approaches have been proposed for alias resolution. They mostly rely on traceroute queries and can be distinguished based on how they process the replies to the traceroute packets. Typically, if a probed router does respond to traceroute queries, then the source IP address field in the header of the reply packets will either correspond to the probed router's (1) default interface, or (2) its outgoing interface towards the probing host. In order to identify whether two IP addresses are aliases for the same router, Mercator [39] compares the source address fields in the reply packets to traceroute queries to the two IP addresses, and if they match, then Mercator concludes that the two addresses are aliases; otherwise, they are not. This approach should work only if routers respond with their default interfaces. A different approach is used by Ally [62], which sends two back-to-back probe packets to the two investigated IP addresses, and inspects the sequence numbers in the reply packets. If the sequence numbers are in order and close enough, then Ally declares that two interfaces are aliases. Mercator and Ally have a couple of weaknesses. First, they only work if routers are responsive to probe packets. Second, their input is a couple of IP addresses. In a network of n interfaces, there are potentially $n!$ aliases to investigate, and exploring all the possibilities is prohibitive. A third popular approach for alias resolution relies on the DNS service. DNS names provide a wealth of useful information for alias resolution. By storing DNS names of routers in a database, aliases of an interface with some domain name are obtained by searching the database for names with common substrings, without the need to communicate with targeted interfaces. As a result, alias resolution can be done even if routers are configured not to respond to probing messages or if they are temporarily unreachable. Our results indicate that almost 30% of the routers are not responsive to probe packets, which highlights the importance of DNS names in resolving aliases. Also, not all $n!$ combinations of interfaces need to be tested, which improves alias resolution efficiency. Since Pansiot [54] introduced the reversed DNS method for alias resolution, virtually most of the Internet topology discovery techniques use the reversed DNS in a certain degree. Research studies have found that about 0.5% of IPs are misnamed [75].

AROMA mainly relies on reverse DNS for alias resolution [54] and complements it using Ally mechanisms by further probing the questioned IP addresses from multiple probing hosts. Interfaces that are suspected to be aliases from their DNS names are verified using Ally and by probing them from multiple PlanetLab servers and the source IP address fields in the reply packets are used to verify aliases. This improves the accuracy as aliases which might not be revealed from one vantage point may be revealed from another. To understand why this is the case, consider a router, which responds to probe packets with the outgoing interfaces towards the probing hosts. Also, consider two interfaces I_1 and I_2 belonging to this router. If I_1 and I_2 are probed from a single probing host, then the replies may be through two different interfaces, which does not indicate that I_1 and I_2 are aliases. However, as we probe from more probing hosts, it becomes more likely that a reply for a probe to I_1 from one of the probing hosts will contain the same outgoing interface as the reply for a probe to I_2 from a different probing host, which signals that I_1 and I_2 are aliases.

The result of the alias resolution step is the set of routers in the target AS map

Table 3.4: Number of Routers and Links Revealed by AROMA and Rocketfuel

	Router		Link	
	AROMA	Rocketfuel	AROMA	Rocketfuel
Sprintlink	35,757	10,332	51,314	25,841
Level3	5,831	1,786	27,144	13,838
Verio	74,114	6,523	176,213	19,289
Abovenet	21,619	654	52,590	2,675

together with links connecting these routers. Table 3.4 shows the number of routers and links revealed by AROMA maps as compared to those revealed by Rocketfuel for our four case studies. The Rocketfuel data can be found in [8]. Clearly, the number of routers/links revealed by AROMA is significantly larger than those revealed by Rocketfuel. However, this comes at the expense of more probing overhead. Furthermore, one can argue that the considered ASes have grown in size since the Rocketfuel experiments were conducted, almost three years ago. We resolve the probing inefficiency problems in Section 3.2, and pinpoint the sources of bias in Rocketfuel in Section 3.3.

Admittedly, the AROMA maps may be less than perfect. AROMA relies on ICMP packets and its accuracy will be offset by routers not responding to traceroute packets, by routers/links that are not revealed in the probing process. While these problems may, more or less, impact the AROMA maps, they persist with much larger magnitude in current state-of-the-art techniques to uncover Internet maps.

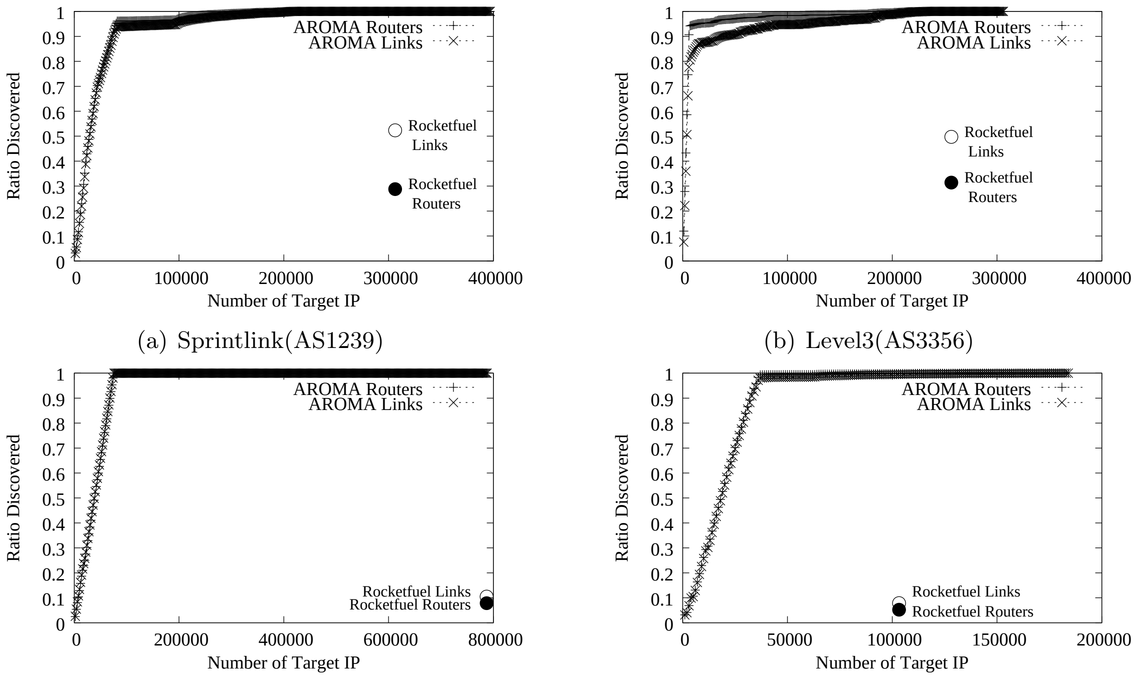
3.2 Efficient AROMA

Large ASes are assigned a large pool of IP addresses and the number of probing servers is potentially large as well. Probing every IP addresses in the targeted AS from every server is intolerable not only because it will put extensive stress on the AS infrastructure but also because it will take months if not years to get a router-level map. The same holds even after the wise selection of servers and destination IP addresses described in Sections 3.1.1 and 3.1.2. In this section we make the case that AROMA (1) can achieve significant efficiency improvement at the cost of a tiny reduction in the details of the resulting map, and (2) can lead to much more detailed maps while being more efficient than existing tools.

We use the number of routers/links revealed in the targeted AS as a measure of **completeness** and the number of traceroute probes as a measure of the **efficiency** (also a measure of **overhead**).[2] We intentionally ignore the overhead introduced by querying the DNS infrastructure. That is not because the overhead is small but because (1) the DNS infrastructure is basically doing what it is intended for, accommodating DNS queries. The DNS system deals with billions of requests per day and accommodating a couple of more million requests – refer to Table 3.1 – is not stressing especially if they are rate-limited and spread over a reasonable period of time. Most importantly, (2) DNS information is used for alias resolution, a major component of any router-level mapping tool. Without using DNS information, the alias resolution part would not scale as explained in Section 3.1.5, and the load on the DNS system would have to shift to routers in targeted ASes in order to accommodate an accurate alias resolution technique like Ally [62]. For this reason, Rocketfuel also relies on the DNS system for alias resolution.

In Figure 3.3, we compare AROMA and Rocketfuel's completeness and efficiency. From one server, we probe every IP address selected as in Section 3.1.1. We then plot the ratio of the number of routers/links revealed to the overall number of routers/links

[2] Note that measuring *accuracy* is difficult due to the lack of a known topology. Furthermore, the sources of inaccuracy in AROMA, outlined at the end of Section 3.1, are similar to those in Rocketfuel. The comparison in this sense, based on completeness, seems fair.



(a) Sprintlink(AS1239)

(b) Level3(AS3356)

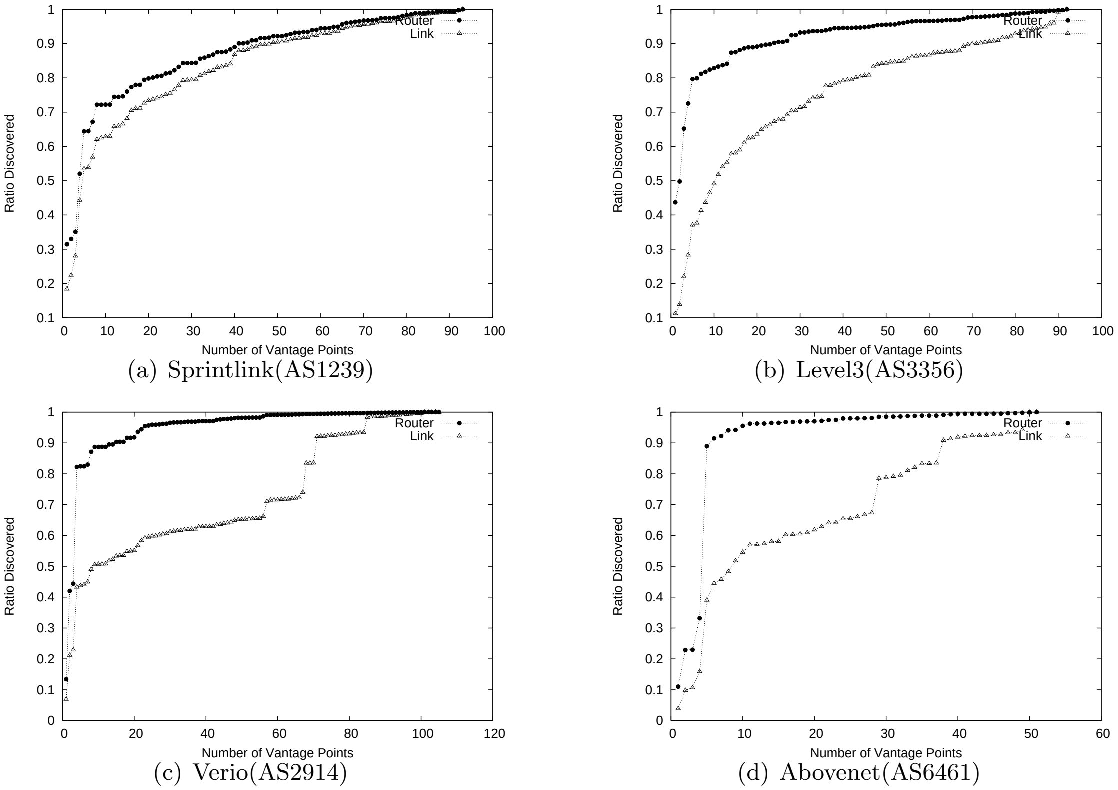
(c) Verio(AS2914)

(d) Abovenet(AS6461)

Figure 3.3: Probing efficiency versus mapping completeness of AROMA using one probing server compared to Rocketfuel.

Table 3.5: Number of Probes Required to Attain Rocketfuel Completeness

	Router			Link		
	AROMA	Rocketfuel	%	AROMA	Rocketfuel	%
Sprintlink	13,300	307,605	4.3%	15,580	307,605	5.1%
Level3	2,235	268,237	0.1%	4,790	268,237	1.8%
Verio	7,090	814,061	0.1%	10,273	814,061	1.3%
Abovenet	680	103,122	0.1%	2,180	103,122	2.1%



(a) Sprintlink(AS1239)

(b) Level3(AS3356)

(c) Verio(AS2914)

(d) Abovenet(AS6461)

Figure 3.4: Mapping completeness as we vary the number of probing servers.

(completeness) as we increase the number of probed IP addresses (efficiency). We also plot the ratio of the number of routers/links revealed by Rocketfuel to map the same ASes and the number of probes used to get these maps. The rocketfuel data is publicly available in [8]. The figure clearly shows that even with one probing server and with much less targeted IP addresses (efficiency) than those selected in Section 3.1.1, the number of routers and links revealed by AROMA (completeness) is much larger than those revealed by Rocketfuel. Based on Figure 3.3, and as shown in Table 3.5, AROMA requires at most 5.1% of the overhead incurred by Rocketfuel to attain similar completeness. Also, based on Figure 3.3, it is clear that it is not necessary to probe all selected IP addresses. A more efficiency-aware implementation would measure the *utility* of probing more IP addresses and would halt probing at any probing server once this utility drops below some threshold. This utility can be expressed as the ratio of the number of revealed routers/links to the number of probed IP addresses during some probing time interval.

In Figure 3.4 we plot the number of revealed routers/links as we increase the number of probing servers. The figure reveals an interesting trend: The rate at which more routers are revealed decreases sharply as we increase the number of probing servers; however, the rate at which more links are revealed does not. In other words, more servers are useful in revealing more links between routers since they access the targeted AS from different ingress points but are not that useful in revealing more routers since the AROMA's initial list includes most routers.

To sum up, AROM can be tuned to be very efficient reducing the number of IP addresses probed from any vantage point with a small degradation in completeness.

However, reducing the number of probing servers will typically come at the cost of hiding the connectivity information between the routers.

3.3 Results and Analysis

In this Section we introduce a simple validation of AROMA by mapping our campus map, and analyze in more detail the maps generated for the SprintLink, Level3, Verio, and AboveNet networks. We also use our results to revisit common perceptions about the structure of the Internet router level topology.

3.3.1 Validation

We used AROMA to map our campus network at NCSU (AS 11442) to validate its completeness and accuracy, and the map is plotted in Figure 3.5. While the network has several different egress points, all traffic from the Planetlab probing servers reach the same ingress router (following the Abilene network). Thus only one probing server was enough to probe our campus network. By targeting 40,785 destination addresses in the campus AS, we found 360 routers and 482 links between them. The map was verified with a campus network administrator, who confirmed that the map is very accurate, missing only a few peering connections.

3.3.2 Topology Structure

In order to get a better understanding of the structure of the revealed maps, we characterize routers based on how close they are from the backbone routers in their AS.

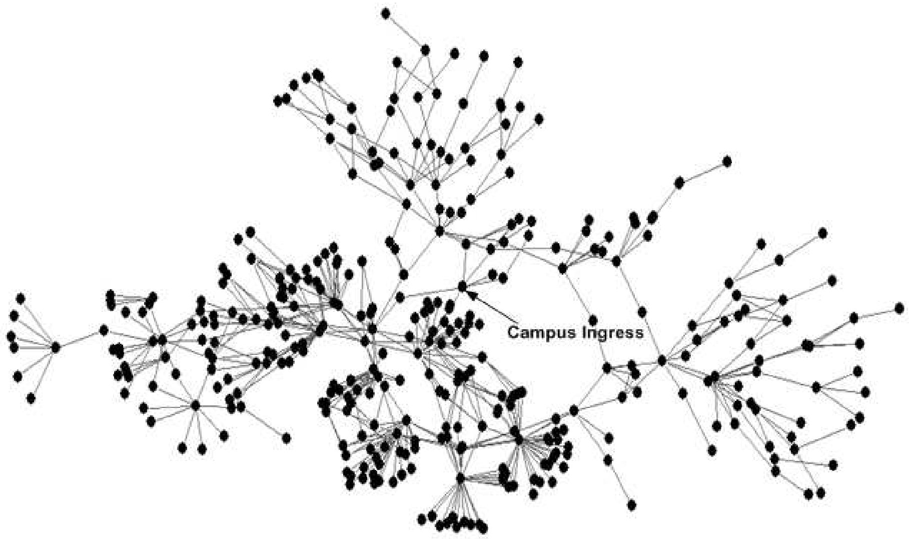
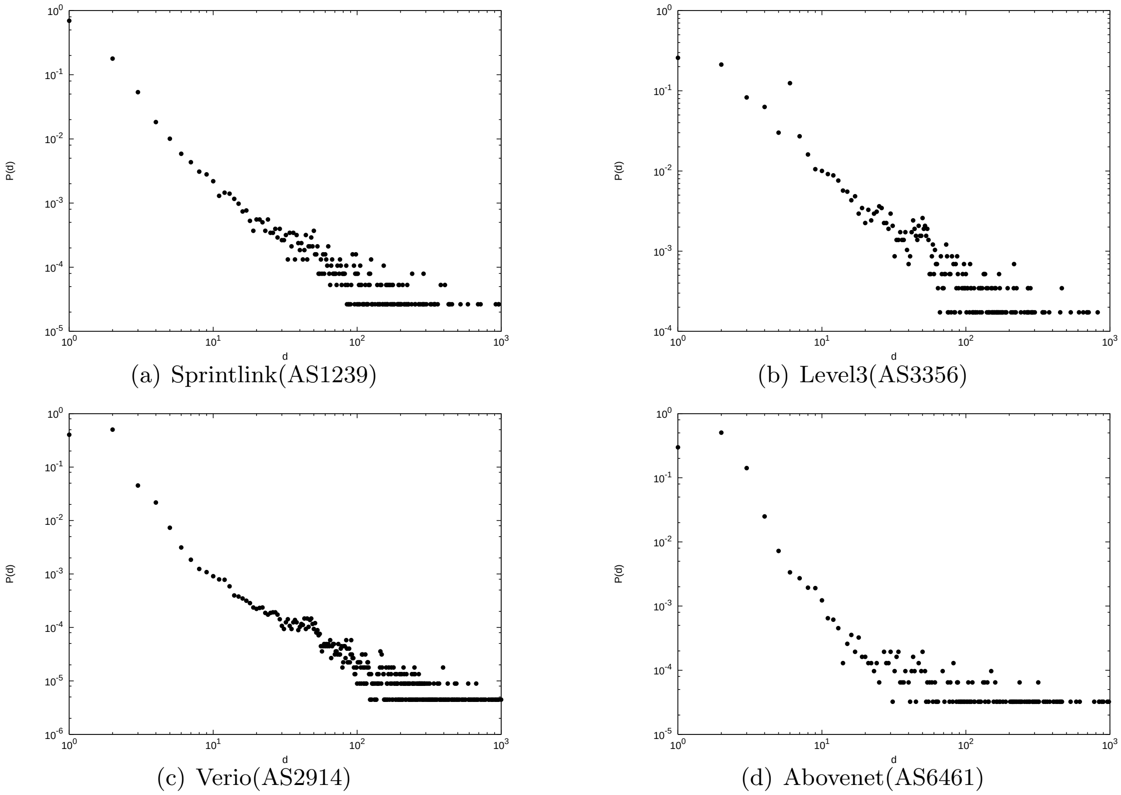


Figure 3.5: AROMA map for NCSU campus network.

Specifically, let L_0 be the set of backbone routers (identified from their DNS names), L_1 be the set of routers that are linked to routers in L_0, L_i be the set of routers that are linked to routers in L_{i-1}, etc. In Table 3.6 we compare the numbers of routers identified in the sets L_0 up-to L_{5+} using both AROMA and Rocketfuel.[3] The numbers reveal the deep hierarchy in these large ASes but also reveal that most of Rocketfuel's revealed routers are at the core of the network, in the sets L_0 through L_3. On the other hand, AROMA finds a considerable number of routers in L_4 and L_{5+}. This confirms our intuition that Rocketfuel is biased towards the backbone of the targeted AS due to the routing of transit traffic through the backbone links. AROMA does not from this structural bias as it probes directly the targeted IPs regardless of where their associated routers are located in the network hierarchy.

[3] In this table, the column labeled L_{5+} refers to the number of routers in the set $L_5 \cup L_6 \cup L_7$. Since the number of routers in L_6 and L_7 is relatively insignificant, we sum the numbers together.



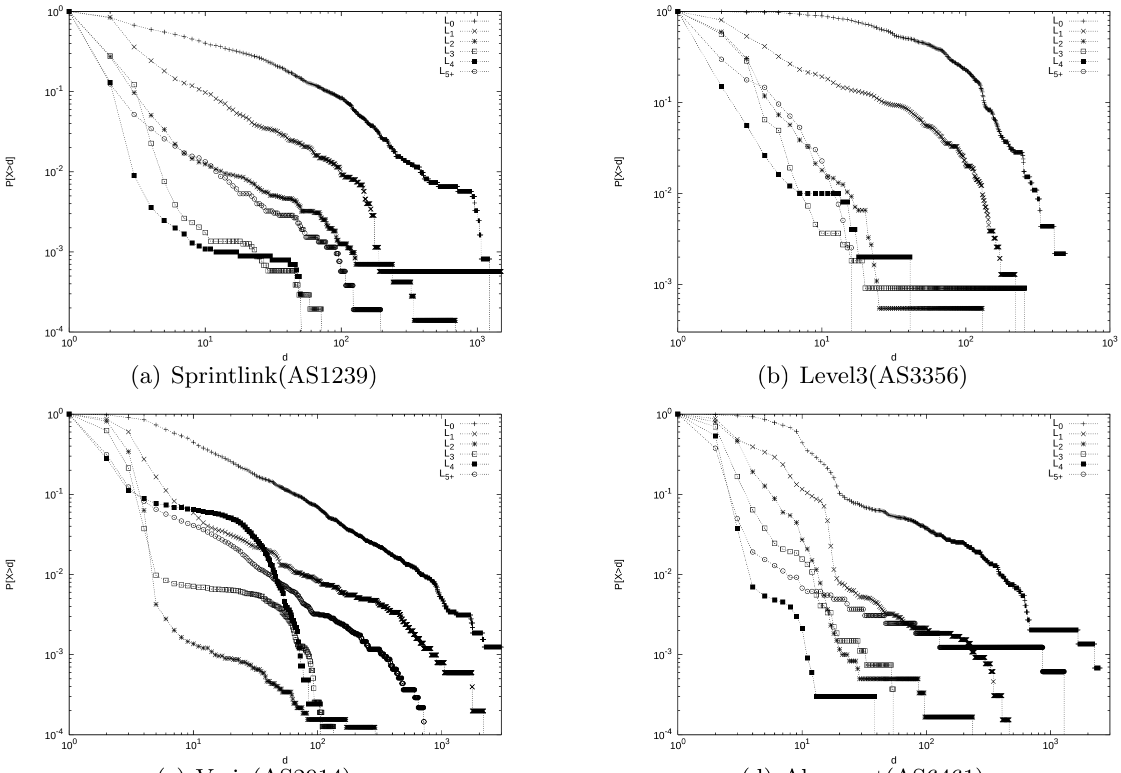
(a) Sprintlink(AS1239)

(b) Level3(AS3356)

(c) Verio(AS2914)

(d) Abovenet(AS6461)

Figure 3.6: Router degree distribution of ISP observed by AROMA.



(a) Sprintlink(AS1239)

(b) Level3(AS3356)

(c) Verio(AS2914)

(d) Abovenet(AS6461)

Figure 3.7: Degree distribution by distance from backbone network: L_i represents the distance from backbone, L_0 are the routers at the backbone while L_3 are routers 3 hops away from the backbone.

Table 3.6: Numbers (and percentages) of routers identified by AROMA compared to Rocketfuel

ISP	Method	L_0	L_1	L_2	L_3	L_4+	Total
Sprintlink	AROMA	1,225 (3.4%)	1,752 (4.9%)	7,132 (19.9%)	10,310 (28.8%)	15,338 (42.9%)	35,757 (100.0%)
	Rocketfuel	700 (6.8%)	6,637 (64.2%)	2,566 (24.8%)	275 (2.7%)	154 (1.5%)	10,332 (100.0%)
Level3	AROMA	459 (7.9%)	1,551 (26.6%)	1,832 (31.4%)	1,096 (18.8%)	1,352 (23.1%)	5,831 (100.0%)
	Rocketfuel	625 (35.0%)	995 (55.7%)	152 (8.5%)	13 (0.7%)	1 (0.1%)	1,786 (100.0%)
Verio	AROMA	3,217 (4.3%)	5,047 (6.8%)	32,231 (43.5%)	15,750 (21.3%)	17,869 (24.1%)	74,114 (100.0%)
	Rocketfuel	1,013 (15.5%)	3,657 (56.0%)	1,269 (19.5%)	268 (4.1%)	316 (4.8%)	6,523 (100.0%)
Abovenet	AROMA	1,472 (6.8%)	6,501 (30.1%)	6,006 (27.8%)	2,696 (12.5%)	4,944 (22.8%)	21,619 (100.0%)
	Rocketfuel	358 (54.7%)	281 (43.0%)	36 (5.5%)	14 (2.3%)	10 (1.5%)	654 (100.0%)

3.3.3 Degree Distribution

The degree distribution of the Internet topology has been a source of controversy [26, 27, 29, 44, 45, 61] since Faloutsos et al. suggested that it follows a power-law distribution [35]. Some researchers advocate the power-law distribution based on their measurements [3, 39], and others question the power-law hypothesis and suggest biased measurements [26,29,44,61].

In Figure 3.6, we use the AROMA maps of the four investigated ASes to plot the probability density function, $P(d)$, showing the fraction of routers with degree d . The figure clearly shows that the distribution follows a power-law, $P(d) \sim d^{-\gamma}$. The power-law exponent, γ, is 2.7, 1.7, 1.94, and 2.3 for Sprintlink, Level3, Verio, and Abovenet, respectively. The estimation of γ is done through least squares regression using the first 5 points of the log-log fitting [41]. This can be justified by that the first 5 points on a log-log scale contain most of the data of a power-law graph, and fitting the graph with all the points, instead of the first 5 points, will distort the results [38]. Previous studies have found power-law exponents of 2.57 with 3,800 routers [35], and of 2.66 with 150,000 routers [39].

The routers' degree distribution revealed by Rocketfuel is quite different from the power-law distribution revealed by AROMA. It has been shown in [61] that the routers' degrees follow a Weibull distribution, $P[X \geq d] \sim e^{-d^c}$, where c is the shape parameter. As opposed to AROMA's power-law distribution, the Weibull distribution is not long-tailed. Rocketfuel's Weibull distribution can be explained by Rocketfuel's bias towards backbone routers. Notice that the edge part of a power-law graph hosts a large fraction of the low degree nodes. By being mostly blind to edge routers, Rocketfuel deflates the *head* of the power-law distribution, where low degree nodes are represented, which inflates the *tail* of the power-law distribution, where high degree nodes are represented. The result is a Weibull-like distribution even if the underlying topology has a strict power-law distribution. In order to prove our hypothesis, in Figure 3.7 we plot the degree distribution for routers belonging to the set L_0 as revealed by AROMA, and the distribution for routers belonging to L_1, etc. Figure 3.7 reveals a consistent pattern in all investigated ASes: The number of low-degree routers is small in the core, while it is relatively large at the edges, which supports our argument that Rocketfuel's Weibull distribution is due to its bias. This observed pattern also contradicts conclusions drawn in [45], in which the authors argue that core routers have a lower degree than edge routers. We intend to investigate this issue in our future work.

3.4 Summary and Future Work

In this chapter, we introduced a probing tool to unveil detailed Internet IP topologies, and made a case for its efficiency compared to existing tools. We used this tool to map four major ISP topologies and revisited earlier conclusions about the Internet IP topology. Our results, contradicting earlier conclusions, raise more questions than answers. While we believe that AROMA provides a positive step towards unveiling Internet IP maps, understanding these maps and the factors affecting their structural and behavioral characteristics is a challenging task. Understanding the reasons behind the contradicting results and Identifying other Internet features, like the location of layer-3 tunnels and generating layer-2 maps will be part of our future research.

Chapter 4

The Causal Factors Behind Internet Power-Law Connectivity

4.1 Power-Law Controversy

The structure of Internet topology can be abstracted as a graph $G = (V, E)$. Vertices, $v \in V$, and edges, $e \in E$, in a graph take different meanings depending on the level of abstraction. A vertex in an *AS-level topology* represents an Autonomous System (AS) and an edge connects two vertices if their corresponding ASes can directly exchange traffic without interim ASes. A vertex in an *IP topology*, also known as *layer-3* or *router-level* topology, represents a router and an edge connects two vertices if their corresponding routers can exchange traffic without interim routers, as enforced by the underlying routing protocol. A vertex in a *physical topology* represents a network device such as a switch or a router and an edge connects two vertices if their corresponding devices are directly connected with a physical cable (network link).[1] While each level of abstraction provides useful information about the Internet, the physical topology is the most difficult to characterize since the most common probing tool, namely *traceroute*, is blind to layer-2 devices. Traceroute relies on the collaboration of networking devices, which have IP addresses and are willing to reveal their IP addresses typically by generating ICMP TIME EXCEEDED messages in response to probe packets with limited TTL values in their IP header. We are not aware of any Internet-scale layer-2 measurement studies. Still, the Internet physical topology is the most detailed and thus the most rewarding in understanding the Internet performance and security features.

One of the most popular and most controversial results in Internet topology studies, is Faloutsos's power law findings [35], which reveal that in AS-level and IP Internet topologies, the distribution of node degrees follows a power law. That is a few nodes are much more connected than the average node's connectivity. The implications are that (1) the Internet enjoys the *small-world* property delivering information very efficiently between nodes, (2) the Internet is resilient to *random failures*, but is vulnerable to *targeted attacks*, since some nodes act as heavily connected *hubs* and bringing these hubs down can be disastrous to the Internet [14]. Since Faloutsos's discovery, many research studies have been conducted whether to support his findings using extensive measurements [60, 65], to explain the causal factors of these findings [17, 25, 34, 45], or to approach differently by using economical and technological constraints to suggest an appropriate arrangement of power law graph nodes in order to deliver Internet-like throughput [18, 45]. To date the most

[1] A Local Area Network (LAN) or any shared medium network may be represented in a physical topology as a vertex, with switches/routers on this network connected with edges to this vertex.

widely accepted Internet topology generators emulate power law topologies similar to those observed by traceroute Internet measurements [10, 16, 23, 40, 49, 53]. Still, some researchers are questioning the power law results, suggesting that they are a side effect of the sampling bias in the measurement methodology and that the observed power law distribution is quite possibly an illusion [29, 44]. Furthermore, the results of these studies are mainly based on measurements by the traceroute tool, and thus are oblivious to layer-2 devices. It is not clear if the drawn conclusions about the Internet structure, performance and security would persist if the Internet physical topology is revealed.

In this chapter, we make the case that our understanding of the Internet properties is mislead by missing layer-2 devices. Specifically, we make the following conclusions: The power law connectivity observed in the Internet IP topology (1) is not an illusion caused by biased data as suggested in [29, 44], (2) It is mainly manifested due to the blindness of traceroute to layer-2 devices, and (3) this manifestation will persist independent of the nature of the underlying physical topology. We also make the case that the Internet physical topology does *not* have a power law connectivity. Non-power law physical topology positively affects the Internet performance and security. These conclusions are not in line with conclusions made by earlier research studies, challenge common wisdom, and highlight the need for more thorough investigations.

The rest of this chapter is organized as follows. In Section 4.2 analyze the impact of layer-2 devices on our perception of the Internet IP topology. In Section 4.3 we investigate the structure of the Internet physical topology and make a case that it is not likely to be a power law. In Section 4.4 we revise earlier conclusions about Internet topology generators, and summarize in Section 4.5.

4.2 Internet IP Topology

In this section we rely on theoretical proof and simulation results to make the case that the perceived power law Internet IP topology is a side effect of layer-2 devices hidden from probing tools like traceroute, but is not simply an illusion caused by sampling bias.

4.2.1 Impact of Layer-2 Devices

Layer-2 devices such as Ethernet or ATM switches or in general forwarding devices which do not have IP addresses, are designed to perform transparent bridging between multiple network segments. They are common in today's Internet, whether at the edge or at the core, given their relatively simple administration and low cost compared to routers, and their resilience to targeted attacks as they do not have IP addresses. As a simple verification of their popularity, packets sent from our networking lab traverse seven layer-2 switches and only two routers to the university gateway router. Figure 4.1 (left) provides an example of a switch connecting a set of routers, and Figure 4.1 (right) shows the corresponding IP graph. The key observation is that in IP graphs, routers connected to switches are perceived as having higher degree than their actual number of physical connections. This simple observation lays the ground for a bolder statement as manifested in Theorem 1.

Theorem 1 *Consider an arbitrary physical topology, G, in which layer-2 and layer-3 devices are spread uniformly at random. The degree distribution of nodes in the corresponding IP topology, G', follows a power law.*

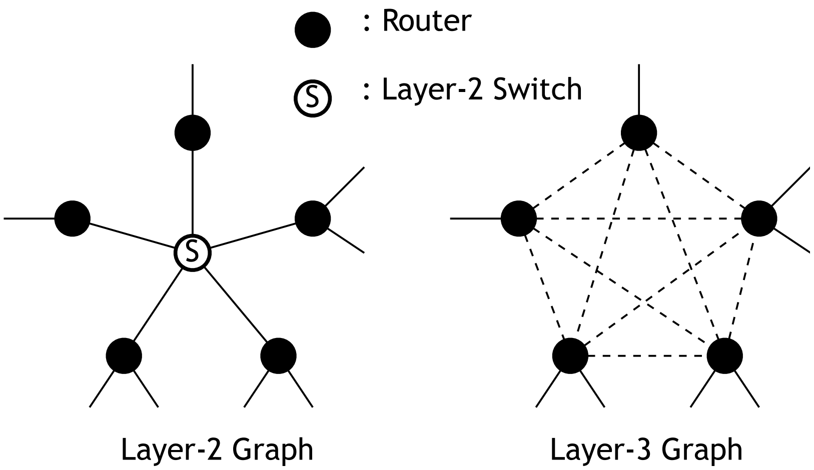


Figure 4.1: A portion of a physical topology and the corresponding IP topology. Solid lines are physical links and dotted lines are connections that would be perceived by a traceroute-like tool.

Proof 1 *Let n be the number of layer-2 devices in G. We convert G into G' in steps, $t = 1 \ldots n$. In step* 1 *we convert G into $G_{(1)}$; in step* 2 *we convert $G_{(1)}$ into $G_{(2)}$; $\ldots$; and in step n we convert $G_{(n-1)}$ into $G_{(n)} \equiv G'$. In each step the physical topology, G, is brought closer to the IP topology, G', by hiding* one *more layer-2 device from G as explained in Figure 4.1. Thus after n steps the resulting graph represents the final IP topology. Consider an arbitrary step, t, and the layer-2 device, s, that is to be considered in this step. After hiding s, the* neighbors *of s in $G_{(t-1)}$ will be perceived as having higher degree in $G_{(t)}$. We next show that if layer-2 devices are placed uniformly at random in G then the resulting G' will be a power law graph. The proof is by mapping the above construction to Barabási and Albert's (1999)* rich gets richer *generative model [20], the most widely accepted model to generate power law graphs. This model was originally developed to reflect the popularity of web pages. New web pages create links to existent ones with a probability distribution which is not uniform, but proportional to the current popularity of web pages, thus popular pages* *become more popular. In our case, we simply need to show that nodes with larger degree are more likely to have even larger degree during the construction of G'. The key idea is that a vertex that is picked uniformly at random in a graph is more likely to be a* neighbor *of a high degree vertex than a neighbor to a low degree vertex. Since layer-2 devices are placed uniformly at random, then at each step, t, the newly considered layer-2 device is likely to be a neighbor of a high degree node in $G_{(t-1)}$, further increasing its degree in the resulting $G_{(t)}$. This concludes the proof.*

The most common probing approach to reveal Internet IP topology is *k-m* traceroute, in which traceroute probes are sent from k controlled probing sources towards m arbitrary uncontrolled destinations. Alias resolution techniques are used to map revealed IP addresses to routers and to reveal links between these routers, generating a topology that is presumed as representative of the Internet. Corollary 1 highlights the impact of layer-2 devices on the perceived IP topology revealed by the *k-m* traceroute approach.

Corollary 1 *Consider an arbitrary physical topology, G, in which layer-2 and layer-3 devices are spread uniformly at random. The degree distribution of nodes in the corresponding IP topology observed by the* k-m *traceroute approach follows a power law.*

Proof 2 *In the case of one probing source, $k = 1$, the graph observed by the source is a* tree *independent of the underlying routing strategy, assuming there are no routing loops. When $k > 1$, the overall observed graph from the k sources is the aggregation of k trees (a* forest*). Theorem 1 applies to arbitrary graphs, and applies naturally to a forest graph. As a result, the IP topology observed by the* k-m *traceroute approach follows a power law. This concludes the proof.*

Theorem 1 and Corollary 1 make the case that the presence of layer-2 devices leads to the power law node degree distribution observed in IP measurement studies. This conclusion is independent of the nature of the underlying physical topology. Note that the assumption that layer-2 devices are spread uniformly at random is reasonable as these devices are quite popular both at the core and the edge of the Internet. It is also common to find layer-2 devices that are physically connected.

4.2.2 Simulation Results

We next verify the results of Theorem 1 and Corollary 1 through simulations. In the simulations, we adopt two well-known random graphs, that were quite popular in representing Internet topologies before the power law models were introduced: (1) Erdös and Rényi (ER) [33], and (2) Waxman [70]. The ER graph has $100,000$ nodes and a link connects two nodes with probability $p = 0.00015$, resulting in an average node degree of 15. The node degree distribution in ER graphs follows a Poisson distribution. The Waxman graph also includes $100,000$ nodes that are added incrementally. Each new node is assigned a coordinate in a plane and a link connects any two nodes, i and j, with a probability, $P(i, j)$, that is proportional to the Euclidean distance between the nodes.

$$P(i, j) = \alpha e^{-d/(\beta L)}$$

where $0 < \alpha$, $\beta \leq 1$, d is the Euclidean distance between nodes i and j, L is the maximum distance between any two nodes. We use $\alpha = 0.15$, $\beta = 0.2$ and the resulting average node degree is about 2. The node degree distribution in the Waxman graph follows an exponential distribution. We pick a fraction of the graph nodes uniformly at random and assume they are layer-2 devices, and all other nodes are assumed to be routers. We also assume shortest path routing between all nodes

For each of the ER and Waxman models, we study four scenarios: (1) Physical, (2) IP, (3) km-physical, and (4) km-IP. In the first scenario we consider a graph including all nodes, whether routers or layer-2 devices, and all links. In the second scenario we consider a graph in which layer-2 devices are hidden as hinted on in Figure 4.1. To do that, we identify the shortest path between each pair of nodes in the graph while hiding layer-2 devices, as if they are not revealed by a tool such as traceroute. We then combine all revealed information from the shortest paths to construct the IP graph. This is the graph that would be revealed using traceroute from each node to all other nodes. In the third scenario we consider the graph that would be revealed by running a *hypothetical* version of traceroute capable of revealing layer-2 devices from k probing sources to m destinations. The sources and the destinations are picked at random. In the fourth scenario we consider the graph that would be revealed by running the traditional traceroute from k probing sources to m destinations, thus missing the layer-2 devices.

In Figure 4.2 we plot the degree distribution of nodes in the four graphs corresponding to the four scenarios in the ER and Waxman models, when the fraction of layer-2 devices is 20%, 40%, and 60%. We plot the curves of scenarios 3 and 4 only for $k = 1$ and $m = 10,000$ since larger values of k lead to similar results. As shown in the figure the presence of layer-2 devices result in power law IP topologies (Scenarios 2 and 4), whether k-m traceroute is used or not, even though the underlying physical topologies are not a

Table 4.1: Characteristic parameters of the ER and Waxman networks by hiding layer-2 vertices: Power law exponent γ, average path length l, and clustering coefficient, C. *Switch %* is the percentage of layer-2 devices in the network.

Switch %	ER			Waxman		
	γ	l	C	γ	l	C
0%	-	4.5	0.00004	-	21.8	0.00004
10%	2.42	4.1	0.00004	3.19	19.3	0.00004
20%	2.41	3.9	0.00005	3.16	17.2	0.00005
30%	2.40	3.8	0.00005	3.14	14.4	0.00005
40%	2.35	3.3	0.00006	3.10	10.1	0.00006
50%	2.40	3.2	0.00007	2.86	9.5	0.00007
60%	2.37	2.9	0.00007	2.85	7.1	0.00007
70%	2.37	2.9	0.00008	2.65	5.5	0.00013

power law. IP topology (scenario 2) shows exponential cut-off at the end of the curve, but *k-m* traceroute IP topology (scenario 4) does not show the exponential cut-off, which is consistent with the result in previous *k-m* traceroute data. The graphs for the scenario 2 and 4 confirm the conclusions from Theorem 1 and Corollary 1. We use least-square fitting to find the power law exponent, γ, in each scenario and those are summarized in Table 4.1. As the fraction of layer-2 devices increases, the exponent tends to be smaller, which means that more nodes are perceived as more highly connected. Also, as expected, the physical curves (Scenario 1) in Figure 4.2 reflect the actual node degree distributions of the ER and the Waxman graphs. What is *not* expected though is that the km-physical curve (Scenario 3) is *not* a power law. As suggested in [44], *k*-m traceroute leads to a power law node degree distribution due the sampling bias in this approach. We clarify this seemingly contradictory result in the Section 4.2.3.

It is worth mentioning that missing layer-2 devices affect our perception of *topological* metrics other than the node degree distribution, considering, for example, the *clustering coefficient*, which measures the amount of *cliqueness* of a graph [68]. Specifically, the clustering coefficient, C_i, for a vertex, v_i, is the ratio of the number of edges between the neighboring vertices of v_i to the number of edges that could possibly exist between them. In Table 4.1 we show the average clustering coefficient, C, in the ER and Waxman graphs as the fraction of layer-2 devices increases, and thus as the resulting IP topologies tend to power law graphs. It is clear from the numbers that the perceived C is slowly inflated as more layer-2 devices are included, and IP nodes are perceived as more connected.

4.2.3 Sampling bias

In [44], the authors propose that the power law node degree distribution observed in Internet *k-m* traceroute measurements is possibly an illusion resulting from the sampling bias in the measurements. They suggest that nodes that are close to the probing sources reveal their connectivity better than nodes that are further away, and as a result the measurements are distorted leading to power law node degree observations. To make the case for the power law illusion, the authors use the same setup in our ER graph described in Section 4.2.2. The authors do not investigate the impact of layer-2 devices and assume that all nodes are visible to traceroute. Thus one can highlight their conclusions through our km-physical curves in Figure 4.2 (a-c) (Scenario 3), which also assumes that no nodes are hidden to the probing tool. Interestingly, the km-physical curves in Figure 4.2 do not reveal power law trends. The reason for this seemingly contradicting results is that the authors of [44] plot the degree distribution only up-to a degree of 25 (Figure 2 in [44], degree of 1.4 in log scale), while we plot all revealed node degrees. Notice in Figure 4.2 that if we draw the km-physical curves up-to a degree of 25 for the ER setup, we get a power law curve. This does not mean that there is no sampling bias associated with the *k-m* traceroute approach;

it simply means that with thorough probing, whether by increasing k and/or m, the *nature* of the distribution is revealed.[2] In addition, the k-m traceroute itself is *not* able to produce a long heavy-tailed power law graph, because it can *not* inflate the degree that each node actually has. The ER graph in Figure 2 in [44] has maximum degree less than 30. We can hardly tell the k-m traceroute graph in [44] is a power law with that amount of information.

Measurement studies targeting Internet IP topologies relying on k-m traceroute have been quite thorough, revealing node degrees of up-to 1071 in the Skitter data set [4] and 1500 in the Mercator data set [43]. The revealed node degrees certainly exceed the maximum possible physical connectivity of any router. For example, the popular high-end Cisco 7600 series, Cisco's CSR, and Juniper's T640 brands offer routers with a maximum of 128 physical ports. Furthermore, measurement results that are not based on k-m traceroute have revealed also power law traces. Specifically, Accurate Router Level Internet Maps (AROMA) is a project intended at reducing the sampling bias in IP topology measurements by targeting Internet ASes and not arbitrary destinations [42]. Only IP addresses and links belonging to the targeted AS are included in the topology, which limits the bias to nodes that are close to the probing host. Measurement using AROMA have revealed node degrees up-to 1625 in Sprint AS, 492 in Level3, 4557 in Verio, and 4557 in Abovenet. As a result, the sampling bias itself cannot explain the power law prevalent in Internet measurements and the suggestion that the power law results are an illusion seems ill-founded. This is especially true given that layer-2 devices lead to power law results as advocated in Section 4.2.1. We argue that the power law results in the Internet IP topology are not an illusion,

[2] Notice that we used the same k and m as in [44] but in Figure 4.2 we plot all revealed node degrees and not only up-to a node degree of 25.

and are caused by layer-2 devices. The nature of the Internet physical topology is a different story, which we investigate next.

4.3 Internet Physical Topology

Understandably, there are no measurement studies about the Internet physical topology. Layer-2 devices do not have IP addresses and thus cannot respond to IP probe packets, and network administrators do not publish the structure of their networks. Until tools that are capable of revealing layer-2 devices exist, the nature of the layer-2 Internet topology will remain as a mystery. Still, there is standing evidence that the Internet physical topology is *not* a power law.[3] We next make the case for this statement. Some of our arguments are subjective due to the lack of measurement studies, while other arguments are objective and are based on the observed Internet performance and security.

The Internet is not the first large scale network to be built. Other networks with possibly similar characteristics do exist such as the power grid networks and the U.S. highway system. Researchers have invested their effort in classifying large scale networks based on their features. For example, Watt's taxonomy classifies networks into *relational* and *spatial* networks [69]. In relational networks, new connections (links) between vertices are governed by a probability that is a function of existing network connections. Web sites and links between them [13], science papers and citations linking them [67], airline systems connecting cities [19], sexual relationships [46] are all examples of relational networks. In spatial networks, connections between vertices are governed by a probability that is a function of

[3] Recall that, according to Theorem 1, the power-law connectivity observed in the Internet IP topology will persist independent of the structure of the physical topology.

the Euclidean distance between the vertices. The power grid networks [69] and the U.S. highway system [21] are examples of spatial networks. Research studies on spatial networks confirm that they have exponential-like degree distribution [15, 19]. The Internet physical topology can be classified as a spatial network and it is likely that it also has exponential connectivity. While there is no evidence of this conjecture, it is worth highlighting that none of the existing and well-documented spatial networks has a power law connectivity.

Furthermore, there is no doubt that intelligent design is behind the high performance and resilience observed in today's Internet. Network designers typically optimize performance and security subject to cost and technology constraints. Interestingly, cost, technology constraints, performance and security all favor *non*-power law Internet physical topology. We next consider each of these factors separately and show that they all lead to the same conclusion: It is unlikely to have nodes with a degree that is much larger than the average node degree, the characteristic property of power law graphs.

4.3.1 Economical Drivers

Economically speaking, the cost of laying multiple cables from a router to each demand is more expensive than extending a single line to demand side and using switch devices to split lines. Especially the demands are spread over wide area or vertically located, splitting zones with switches and multiplexing them are more economical solution. This cost efficient recursive clustering helps to reduce the degree of the Internet physical graph vertices, making it unlikely that some vertices would have a degree much larger than the average vertex degree. Notice that while recent edge router technology can accommodate thousands of DS0 channels or hundreds of DS1 channels, physical ports on the router are more or less than 100, and thus switch devices should be placed to multiplex a hundred demands before sending them to the router using multiplexing technology.

4.3.2 Internet Performance

We consider two performance metrics: (1) The *maximum throughput*, and (2) the *average path length*. We use the former to make the case that the the maximum throughput that can be achieved by the Internet, subject to technological constraints on router connectivity, is improved with non-power law physical topology. We use the latter to show that the small-world phenomenon observed in the Internet can be explained even if the underlying Internet physical topology has exponential node degree distribution.

Maximum Throughput. Routers and switches have a limited number of physical ports. Utilizing a large number of these ports simultaneously degrades the device's overall throughput, and the throughput of each port [18]. Since network administrators strive for better performance, heavily connected networking devices are unlikely. Furthermore, the typical number of physical ports is not large making it again unlikely that some nodes would have much larger degree than the average node degree.

We define the maximum throughput as the throughput under heavy traffic conditions based on a gravity model [58]. Specifically, we estimate the sum of the maximum throughput of flows on all source-destination pairs of edge routers, assuming shortest path routing. The maximum throughput between a source destination pair is computed under the *router degree-bandwidth constraint* described in [18,45]. The constraint is based on the Cisco 12416 Gigabit Switch Router (GSR) technology constraint, which can be described as follows: If a router is configured to have a degree of 12, then each physical connection can enjoy a maximum throughput of 10 Gbps; while if the router is configured to have larger degree, the throughput per physical connection degrades (Figure 1 in [45]).

In [18, 45], the authors use the above definition of the maximum throughput to argue that the power law observed in the Internet node degree distribution is not enough to describe the Internet topology since networks having the same power law node degree distribution can have different features. They advocate that a power law network should be *disassortative*[4] to obtain the maximum possible throughput while satisfying the router degree-bandwidth constraint [52]. That is high physical connectivity at the router level is expected to be firmly confined at to the network edge, while physical connectivity at the network core is expected to be quite sparse. To make this case, the authors study different topologies with the same power law degree distribution based on the maximum throughput metric and show that their disassortative Heuristically Optimal Topology (HOT) model leads to better throughput than other topologies with the same power law node degree distribution. The HOT model is obtained by rewiring the power law graph obtained by using the Barabási and Albert's rich gets richer model to make it disassortative. Still, the HOT model does not lead to the maximum *possible* throughput as the router degree-bandwidth constraint limits the capacity of highly connected routers at the edge of the HOT model. This itself is unlikely in practice as it is hard to conceive that an administrator would intentionally use a heavily connected router and degrade the potential capacity that would be otherwise available by the use of two less connected routers. A network which does not suffer from the router constraint should thus not have overly connected routers.

[4] In assortative network, a high degree node connects to a high degree node, while in disassortative network, a high degree node is connected to a low degree node. Assortative network usually has high degree nodes at the core, but disassortative network has them at the edge side.

Average Path Length. Researchers have suggested that the Internet has the *small world property* [50], that is a small average path length between nodes. This result can be verified by inspecting traceroute probes showing a relatively small average number of IP hops between Internet hosts. In 1999-2000, the average path length of IP graph was $11 \sim 13$ [12]. The small world property is inherent in power law graphs and Poisson graphs like ER, but *not* in exponential graphs like Waxman. In general, the *small world* property exists is in relational networks, not in spatial networks. In Table 4.1 we show the average path length, l, of the ER and Waxman graphs using the setup described in Section 4.2.2, and as we vary the fraction of layer-2 switches in these graphs. It is clear from the numbers that when no layer-2 devices are included, the ER graph has the small world property but Waxman does not. However, as layer-2 devices are introduced, the observed average path length in the resulting IP topologies decreases dramatically both in the ER and the Waxman networks, since they both converge to power law networks, satisfying the small world property. As a result, the small world phenomenon observed through traceroute measurements is not necessarily persistent in the Internet physical topology. Still, measurement results also reveal a relatively small delay between Internet nodes, typically in the order of milliseconds. This observed efficiency can be attributed by the forwarding efficiency of layer-2 devices compared to routers. In layer-2 devices, packets are forwarded based on MAC addresses at almost full wire hardware speed, which makes it faster than a router. Thus communication between Internet nodes covers a small number of IP hops, each consisting of some number of layer-2 hops, which are quite efficient.

4.3.3 Internet Security

Heavily connected routers represent a security risk. A targeted attack on such routers can affect a significant portion of the Internet. In [14, 16], the authors warn the Internet community about the existence of Internet “Achilles heel”, weak points that are vulnerable to directed attacks. The weak points are the heavily connected routers acting as giant hubs, and handling a large fraction of the Internet traffic. These are based on the perceived power law Internet topology highlighted in [35]. The authors, however, point out that these giant hubs are quite rare making the Internet resilient to random failures. We next revisit these conclusions based on our findings that layer-2 devices are causing the perceived layer-3 Internet power law topology, and that the Internet physical topology does not suffer from high vulnerability.

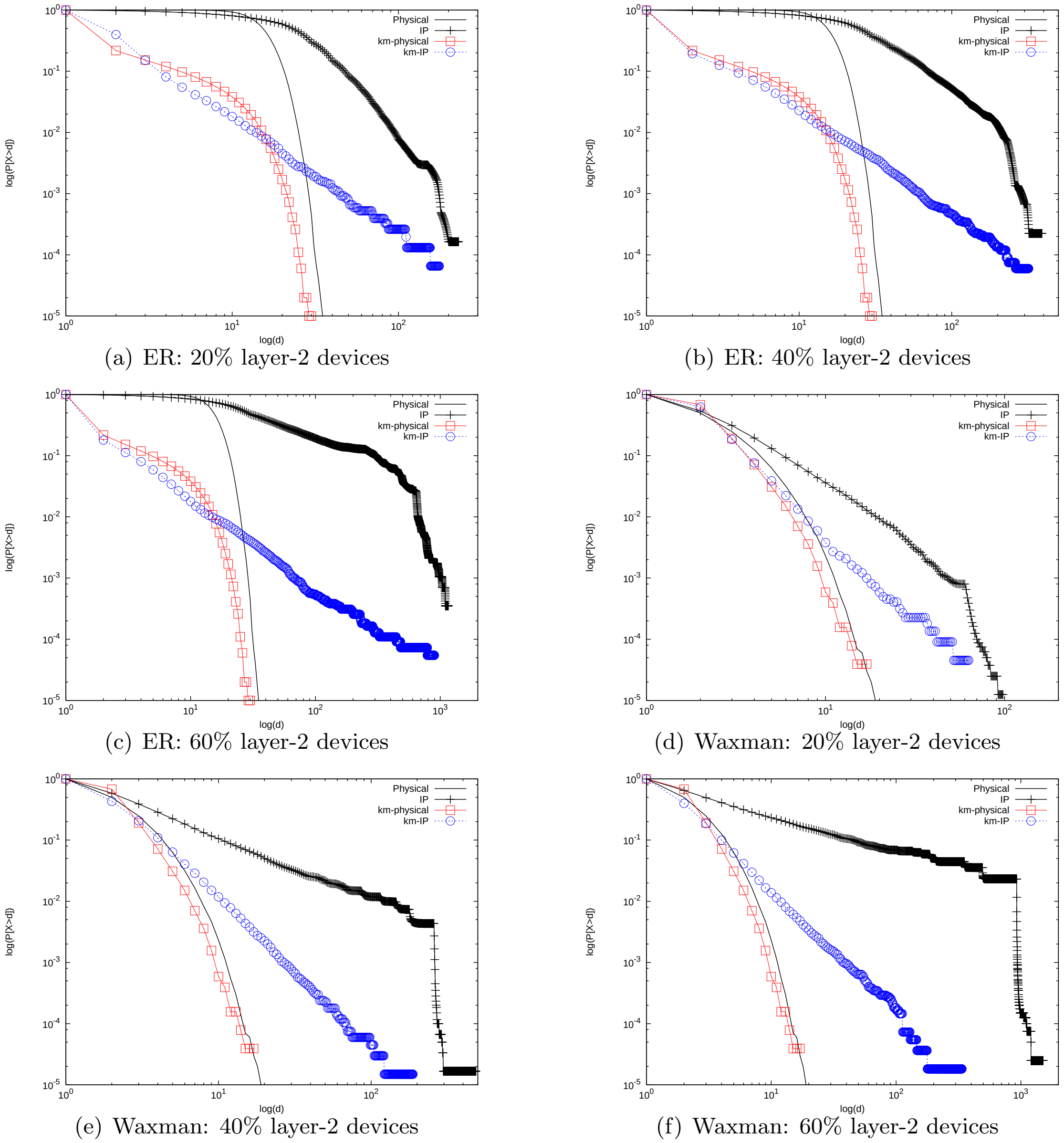
As opposed to a router, a layer-2 switch does not have an IP address, which makes it more resilient to targeted attacks. Furthermore, an Internet physical topology without giant hubs, is also more resilient to failures, targeted and random attacks. Furthermore, switches introduce hidden redundant connectivity that helps for better security to the Internet. For example, in Figure 4.3(b) nodes a and b appear as heavily connected hubs in the perceived IP topology and it seems that attacking and removing these nodes isolates a large portion of the network. However, by observing the underlying physical topology in Figure 4.3(a), removing nodes a and b affects only a smaller fraction of the network. Overall, security favors an Internet physical topology without heavily connected nodes.

4.4 Revisiting Internet Topology Generators

Our conclusions in this chapter about the Internet IP topology and physical topology provide a different perspective on the Internet structure, explain measurement results, and contradicts some earlier conclusions. In this section we revise some conclusions on widely accepted Internet topology generators. Recall from the survey in Chapter 2 that Internet topology generators fall into two main categories: (1) structure-based [24,31] and (2) degree-based [10,30,40,48,49]. Structure-based topology generators model the Internet hierarchical structure known to exist in the Internet, distinguishing *transit* domains from *stub* domains. Degree-based topology generators model power law node degree distributions that exists, or at least perceived to exist, in Internet IP topology measurements. Degree-based topology generators gained more popularity since structure-based generators were shown not to exhibit power law node degree distributions. Furthermore, researchers have made the point that degree-based topology generators are superior to structure-based generators in emulating the Internet structure [65]. The conclusions that we draw in this chapter sheds some doubt into these conclusions. Structure-based topology generators may indeed reflect a power law node degree distribution if layer-2 switches are modeled. Also, the Internet's physical topology is not likely to exhibit a power law node degree distribution because of the reasons we have seen so far. Furthermore, the degree distribution itself may not be enough to represent large-scale properties of the Internet [65]. The hierarchical structure of the Internet is one such large-scale property, but this does not mean that structure-based generators produce more realistic Internet topologies. It is rather that earlier conclusions comparing structure-based and degree-based topology generators are ill-founded, which suggests rethinking Internet topology models.

4.5 Summary

In this chapter, we made the case that (1) layer-2 devices are the reason for the perception of a power law node degree distribution in the Internet's IP topology, not sampling bias or node's tendency of preferential attachment. (2) The Internet physical topology should not be a power law to enjoy better performance and security. Our conclusions help to explain many of the Internet measurement results, to question some of the earlier conclusions, and to call for more thorough investigation of the Internet properties.



(a) ER: 20% layer-2 devices

(b) ER: 40% layer-2 devices

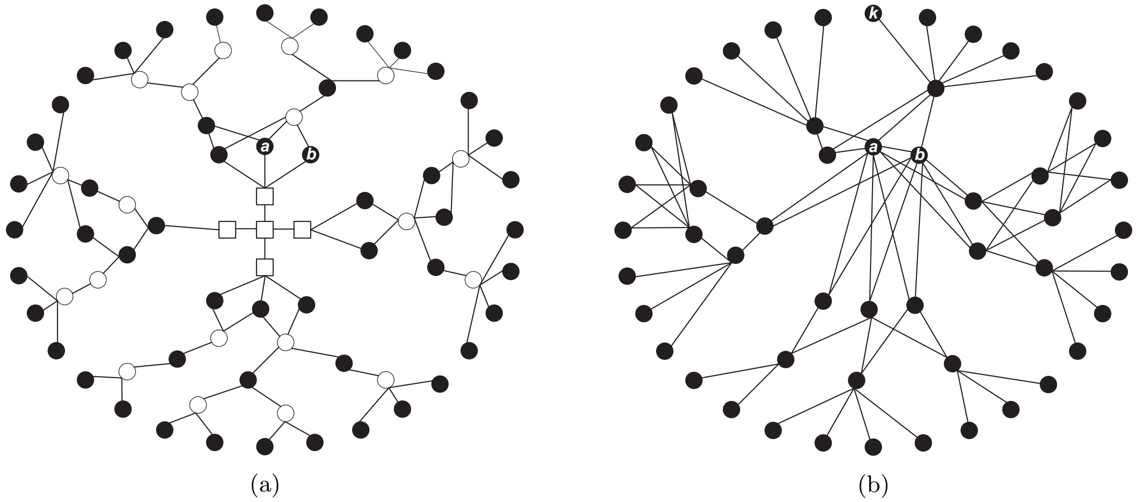
(c) ER: 60% layer-2 devices

(d) Waxman: 20% layer-2 devices

(e) Waxman: 40% layer-2 devices

(f) Waxman: 60% layer-2 devices

Figure 4.2: Node degree distribution of ER and Waxman graphs with 100, 000 nodes, when the fraction of layer-2 devices is 20% (a,d), 40% (b,e), 60% (c,f). Each figure include four curves representing the (1) Physical, (2) IP, (3) km-physical, and (4) km-IP scenarios. For scenarios 3 and 4, we use $k = 1$, and $m = 10,000$.



(a)

(b)

Figure 4.3: (A) Degree distribution for physical connectivity. There is no node highly connected in this level. (B) Degree distribution for IP-level topology based on *k*-*m* traceroute. Traceroute begins from node ‘k’. The result topology of the IP-level is exaggerated by layer-2 devices, and has few highly connected hub nodes. ●: router, ❍: Ethernet switch, ❑: ATM switch, k: traceroute source node

Chapter 5

The Hoops in the Net: Topology based on Cost, Performance, and Security

Existing Internet topology models have progressed from random graphs, to structural based, to degree-based. All of which, however, do not capture realistic Internet structure and features. Popular existing topology generators are degree-based, generating power law networks. Our results in Chapter 4 clarifies the limitations of this approach and brings efforts to model the Internet topology to a fresh start.

In this chapter, we provide a more realistic Internet topology model, HINT, which takes economical, performance and security perspectives into consideration. Furthermore HINT differs from existing models in that it considers traffic loads on each links. The model reveals interesting results that conforms to the published information about some networks.

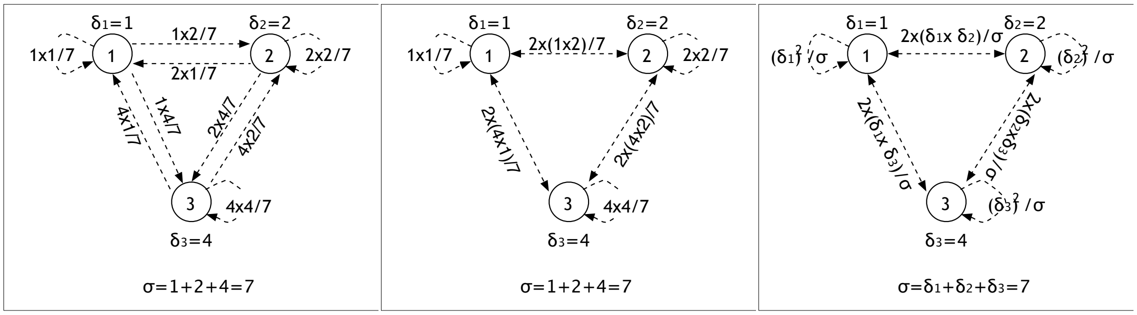


Figure 5.1: (a) Traffic flow between nodes v_1, v_2, and v_3 when $\delta_1 = 1$, $\delta_2 = 2$, and $\delta_3 = 4$. (b) A more compact representation of the traffic flow between nodes. (c) A symbolic representation of flows.

Mainly, that the core of the Internet consists of *cycles* and the *edge* consists mostly of trees.

The rest of the paper is organized as follows. In Section 5.1, we provide the setup, terminology, and explain our model. In Section 5.2 we provide the details of our proposed HINT model and compare the resulting HINT topologies to existing topology models. In Section 5.2.2, we evaluates proposed model with previous ones. We finally conclude in Section 5.3.

5.1 System Model

5.1.1 Basic Model

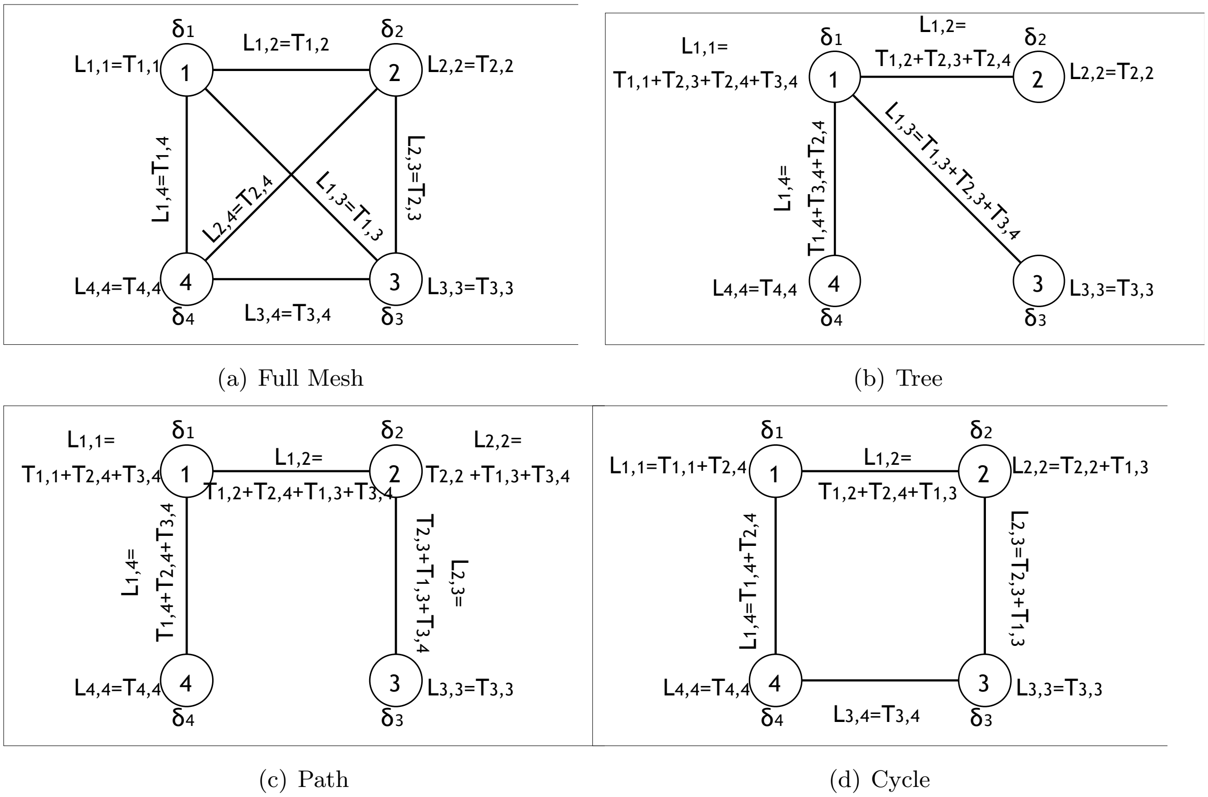
Consider a graph $G = (V, E)$ of $|V|$ nodes (vertices) and $|E|$ edges. An edge $e_{i,j} \in E$ connects vertices v_i,$v_j \in V$. Let $d_{i,j}$ be the length of edge $e_{i,j}$. Each node in G represents a geographic concentration of population reflecting some *demand* for Internet service. One can also think of nodes also as Points of Presence (PoPs). Let δ_i (in units of traffic rate) be the demand at node v_i. δ_i values are normalized such that $\delta_i \geq 1$ and $\delta_{i^*} = 1$ unit, where v_{i^*} is the node with the least demand for Internet service.

We assume that nodes contribute to the Internet as much as they demand from other Internet nodes. This assumption is not necessarily true in practice for all nodes but is reasonable given the lack of other evidence and to simplify our discussions. Cases in which some nodes contribute more that they demand, or demand more than they contribute, is left for future research. Thus node v_i exports δ_i traffic units to all nodes in G and imports/download δ_i units from all nodes in G. The imported δ_i units are downloaded from nodes in G in proportion to their δ_i values. The exported δ_i units are uploaded to nodes in G also in proportion to their δ values. Let $\sigma = \sum_{i=1}^{|V|} \delta_i$. Then a node v_i uploads a fraction $\frac{\delta_j}{\sigma}$ of its δ_i units of traffic to node v_j and downloads the same amount of traffic from node v_j.

Refer to Figure 5.1 for an example on the exchanged traffic units between nodes. In this example we plot three nodes, v_1, v_2, and v_3 for which $\delta_1 = 1$, $\delta_2 = 2$, and $\delta_3 = 4$. In Figure 5.1(left) we plot a directed graph. A directed link from a node v_i to node v_j is annotated with the traffic units that are originating at v_i and destined to v_j. Note the self cycle at each node representing traffic originating and destined to same node. The figure establishes that indeed the sum of the traffic units originating at a node v_i is δ_i and the traffic units intended for v_i is also δ_i. Figure 5.1(middle) provides a more compact representation of the traffic flows between nodes, in which the sum of the traffic units between nodes are summed up instead of distinguishing the traffic units in each direction. Note that the sum of the traffic units over all links in this representation is σ as expected. Figure 5.1(right) provides a symbolic representation of Figure 5.1(right) to help reveal a simple derivation of the traffic units that need to be carried on a *direct* link between any two nodes and the traffic units that not using links connecting nodes (self cycles). One can generalize the following derivations: If physical links $e_{i,j}$ are extended between nodes v_i and v_j, then the *capacity* of these link needs to accommodate a traffic load of $T_{i,j} = 2\delta_i\delta_j/\sigma$ traffic units, when $i \neq j$. Furthermore, a node v_i needs to accommodate a traffic load of $T_{i,i} = \delta_i^2/\sigma + \sum_{\forall j \neq i} T_{i,j}$ traffic units. That is traffic generated through self cycles and traffic communicated with all other nodes. Notice that in the case when a direct link exists between every pair of nodes, nodes do not carry transit traffic originating at other nodes and destined to other nodes. Instead, all traffic crossing a nodes is either originating at that node or destined to it. Thus $T_{i,j}$ values reflect traffic loads on links and $T_{i,i}$ values reflect traffic loads on nodes when the structure of the graph is a full mesh. This case is obviously unrealistic and does not represent the Internet graph. As nodes carry transit traffic for other nodes, other than the full mesh graphs, the traffic load on links and nodes carrying transit traffic will increase. Let $L_{i,j}$ and $L_{i,i}$ be the generalization of the $T_{i,j}$ and $T_{i,i}$ traffic load measures to account for the overall load on links and nodes whether due to transit traffic or not. Thus $L_{i,j} = T_{i,j}$ and $L_{i,i} = T_{i,i}$ in the full mesh graph case (no transit traffic) and $L_{i,j} \geq T_{i,j}$ and $L_{i,i} \geq T_{i,i}$ otherwise to account for transit traffic load on graph links and nodes, respectively. We next explain how to account for the transit traffic.

5.1.2 Accounting for Transit Traffic

A simple example will make accounting for transit traffic load on edges and nodes clear. Consider four nodes v_1, v_2, v_3 and v_4. We next construct four different basic graphs to connect these nodes as illustrated in Figure 5.2: (1) A full mesh, (2) a tree, (3) a path, (4) a cycle.



(a) Full Mesh

(b) Tree

(c) Path

(d) Cycle

Figure 5.2: Basic Network structures: (a) Full mesh. (b) Tree. (c) Path. (d) Cycle

In the *full mesh* case – Figure 5.2(a) –, there is no transit traffic and all traffic going through a node is either originating from it or destined to it. In this case, $L_{i,j} = T_{i,j}$ and $L_{i,i} = T_{i,i}$ as explained in Section 5.1.1.

In *tree* case – Figure 5.2(b) –, basically a star-like topology in this simple setup. In general a tree is a collection of paths without cycles. In this case, three edges have been dropped from the full mesh graph, $e_{2,3}$, $e_{2,4}$, and $e_{3,4}$. The traffic that used to be carried on these links has now to be routed through other nodes and edges. The traffic load $T_{2,3}$ that used to be over $e_{2,3}$ will be routed through node v_1 thus increasing the load on $e_{2,1}$, v_1, and $e_{1,3}$ by $T_{2,3}$; The traffic load $T_{2,4}$ that used to be over $e_{2,4}$ will be routed through node v_1 thus increasing the load on $e_{2,1}$, v_1, and $e_{1,4}$ by $T_{2,4}$; and The traffic load $T_{3,4}$ that used to be over $e_{3,4}$ will be routed through node v_1 thus increasing the load on $e_{3,1}$, v_1, and $e_{1,4}$ by $T_{2,4}$. The end results is that the load on links and the load on the node v_1 will increase compared to the full mesh case.

In the *path* case – Figure 5.2(c) –, three links have been dropped $e_{1,3}$, $e_{2,4}$, and $e_{3,4}$. Again the traffic that used to be carried on these links, $T_{1,3}$, $T_{2,4}$, and $T_{3,4}$ has to be routed through other nodes and edges, adding to the load on these nodes and edges as shown in Figure 5.2(c).

The *cycle* case in Figure 5.2(d) adds only one link, $e_{3,4}$ to the path in Figure 5.2(c) and drops links $e_{1,3}$ and $e_{2,4}$ from to the full mesh case. As a result, traffic load $T_{1,3}$ can be routed between v_1 and v_3 either through node v_2 or through node v_4; similarly, $T_{2,4}$ can be routed between v_2 and v_4 either through node v_1 or through node v_3. This provides some flexibility in balancing the load on the system. Figure 5.2(d) shows the resulting load on nodes and edges as $T_{1,3}$ is routed through v_2, $T_{2,4}$ is routed through v_1. An implicit assumption in our load distribution is that routing follows shortest path to the destination and that routing is symmetric. For example, traffic from v_1 to v_3 follows the same path as traffic from v_3 to v_1.

5.1.3 Network Metrics

We rely on the following metrics to compare different network structures and later on to build and evaluate the proposed HINT topology model.

- Cost. The cost of building the network is a major factor in its structure. This should be obvious as ISPs are mainly investors aiming at maximizing their profit. To do that they are mainly aiming at minimizing the *cost* and maximizing their *profit.* The profit is maximized as more customers join their network and as they carry more transit traffic. The cost of building a network has many facets some of which are *fixed* costs (paid once) such as (1) the cost of digging out to lay out the links (fiber) (2) the cost of the fiber and its installation, and (3) the cost of buying equipment such as routers and switches; and *recurring* costs such as (1) the labor cost running the equipment, (2) the overhead, and (3) the cost of using neighboring ISP's infrastructure to route traffic in transit.

Notice that the digging cost is proportional to the distance between nodes (the length of the link $d_{i,j}$). The fiber cost is proportional to both the length of the link and to the load expected on the link $L_{i,j}$. The cost of equipment (routers, switches, etc), the labor cost, and the overhead are proportional to the traffic load on nodes $L_{i,i}$. While the digging cost is a major component in a network cost, this cost differs depending on whether the links are laid in urban or rural areas. Also, recurring costs such as labor cost constitute a large portion of the overall cost. Typically, investors in network infrastructure do not expect their investment to pay off immediately and they plan for a break-even point, typically within a few years, after which their revenue is expected to exceed the costs.

- Performance. We target an over-provisioned network capable of accommodating the demand of customers. This is typical of the Internet backbone structure. Thus network throughput is not an issue. Instead we use the average delay between nodes as the performance metric. The delay between nodes has two components: (1) the

Table 5.1: Comparison between basic network structures.

	Full Mesh	Cycle	Path	Tree
Cost (Line Digging)	Worst	Bad	Good	Good
Cost (Lines Cost and Installation)	Worst	Good	Bad	Bad
Cost (Equipment, Labor)	Best	Good	Bad	Bad
Performance (Propagation Delay)	Best	Good	Bad	Bad
Node Resilience	Best	Good	Bad	Bad
Link Resilience	Best	Good	Bad	Bad

propagation delay, which is proportional to the length of the links connecting nodes, and (2) the *queuing* delay, which reflects the delay incurred by nodes at intermediate nodes. Again since we are building an over-provisioned network we ignore the queuing delay component and only investigate the average propagation delay between nodes.

- Resilience. Security is an important factor in network design, If the network operation is disrupted by the disruption of a router or a link then investors are likely to lose business. It is desirable to have network sustain at least the disruption of a node or router by rerouting traffic without major disruption of operation.

5.1.4 Comparing Basic Graph Structures

Equipped with metrics discussed in Section 5.1.3 we compare the basic topology structures that we introduced in Section 5.1.2. Our comparison is qualitative and admittedly is ad-hoc but is enough for our purposes.

Table 5.1 compares the cost, performance, and security features of the full mesh, cycle, path, and tree structures. All structure should have the same number of nodes. Note that a tree is basically a collection of paths and their behavior is quite similar.

The full mesh represents an extreme end of the spectrum of network structures, which incurs the highest cost for digging lines and installing links as there is a link between every pair of nodes. However, since nodes are not congested, then cost of buying and managing equipment at nodes is relatively cheaper than other structures. Furthermore, the delay between nodes is the best possible and the resilience to network disconnection due to node or link failures is clearly as best as it can be. The full mesh structure is not practical given the excessive high cost for setting up links between every pair of nodes.

The other three structures are much cheaper than the full mesh and their costs are deemed practical. The cycle structure incurs a cost for establishing an extra link than the path structure and thus incurs an additional digging cost. However, since nodes and links are less congested, then the link capacities in the cycle structure are relatively smaller and thus their cost is relatively less than in the path and tree structures. Furthermore, since nodes are less congested then the equipment at nodes and the labor cost is less. The cycle structure is more resilient than the path and the tree structures to both node and link failures. The average propagation delay between nodes in a cycle is definitely much better than in the path structure. It is also better than the average delay between nodes in a tree.

The points to be made here are: (1) The tree, path and cycle structures are more practical to implement than the full mesh structure. (2) While the cycle structure is relatively more costly in its initial investment (line digging), it requires less costly equipment at nodes, less labor cost, and less costly fiber. (3) The cycle structure offers better delay between nodes than the path and tree structures. (4) Most importantly, the cycle structure is more secure than the path and tree structures as it is more resilient to node and link failures. Bringing one link down from a tree or a path structure will split the network. That is a major concern for path and tree structures.

5.2 HINT Internet Topology Model

The Internet is an associative network and associative networks typically form grids. So should the Internet. We have limited evidence from Level-3 published infrastructure and Abilene infrastructures in Figure 5.3. Whatever the shape of the Internet backbone structure, it is likely that nodes would like to have direct connections to most popular nodes in order to limit the recurring cost they pay to other nodes to redirect their traffic. The Internet construction is driven by supply and demand. An ISP would like the best return on his investment. That goes where the demand is promising the most revenue on his investment with the least possible cost. But performance and security are also important considerations as a degradation in performance or security risks can degrade the demand. Cycles formation provides better performance and security as we have seen in Section 5.1.

The more money spent on the network infrastructure the better the resulting network in terms of performance and security. The premise of the HINT model is to explore this tradeoff and show how the topology should be built as the overall budget shrinks or expands.

5.2.1 HINT Model

Our discussion in Section 5.1.1 motivates the presence of cycles in the Internet but does not explain how cycles are constructed, which nodes are connected, how large/small each cycle is, and does not distinguish between the capacities of the links in the resulting structure. In order to address these issues, we rely on a simple model, which takes into consideration the demand for Internet service and the supply in the form of network infrastructure. More populated nodes (cities) offer/demand more Internet content and ISPs are likely to extend network links between more populated cities than to less populated cities in order to make more profit. Direct links between popular nodes limit the recurring cost that would otherwise need to be paid to other nodes to communicate traffic between these popular nodes. HINT uses the population of cities as a measure of the demand for Internet service at these cities. In Figure 5.5 250 cities in the United States, which have more than 100,000 people within urban boundaries, are selected and the size of dot represents the amount of demand at these cities. On the other hand, the longer the distance between nodes the more the cost of laying out network links between these nodes. This tradeoff can be expressed by adopting Newtonian's Gravitational model, which expresses the connectivity force, $f_{i,j}$, between two nodes i and j as:

$$f_{i,j} = \frac{L_{i,j}}{{d_{i,j}}^2},$$

where $L_{i,j}$ is the traffic load between nodes i and j, $d_{i,j}$ is the distance between these nodes. In other words, the possibility of having a direct connection between cities i and j is proportional to the traffic that is expected to traverse this connection and inversely proportional to the square of the distance between the two cities. Recall from our discussion in Section 5.1 that $L_{i,j} = 2\frac{\delta_i \delta_j}{\sigma}$ when the graph is a full mesh, but $L_{i,j}$ increases when the structure is not a full mesh as links may have to carry transit traffic for other nodes.

The HINT model starts with a full mesh graph between nodes thus initially assuming that cost is not an issue. In this case, $L_{i,j} = 2\frac{\delta_i \delta_j}{\sigma}$. Then least important links, i.e. links with lower $f_{i,j}$ are dropped one at a time until no more links can be dropped without disconnecting the network. As a link $e_{i,j}$ is dropped, its traffic $L_{i,j}$ is redirected through other paths in the structure in order to still connect the $L_{i,j}$ load between nodes i and j – as explained in Section 5.1. HINT thus exploits the set of network structures that are expected to be adopted as cost becomes more and more relevant and converges in the likely condition, when cost cannot be further reduced without disconnecting the network. However, since network resilience to node/link failures is an important issue especially in the Internet backbone, HINT relies on an additional criterion. Specifically, HINT does not drop a link if its removal will lead to a node with a degree of *one*. A structure with a node degree of at least *two* ensure the resilience of the network to the disruption of at least one node/link. Note that this condition leads to the presence of cycles in the Internet backbone. The edge of the Internet is bit different as there is typically little traffic demand between nodes in the in close physical proximity and a tree structure is expected to be the common case. Evidence of the tree structure at the edge can be seen in our own NCSU network. In this case, this degree condition can be removed and the resulting topology will simply be a tree. Figure 5.6 illustrates the HINT process and Algorithm 1 provides the complete HINT network generation algorithm. The efficiency of the algorithm is $O(E(V^2 + E))$ since HINT uses Dijkstra's algorithm with priority queue for the shortest path search, which is $O(VlgV + E)$ and removing and updating each edges requires an $O(E)$ computation.

```
Input: A complete graph, $G$
Output: The Internet backbone network, $G'$
foreach edge $e_{i,j}$ do
    assign traffic load $L_{i,j}$;
    assign connection force, $f_{i,j}$;
end
while $|E|$ in $G'$ is greater than target number of edges do
    select $e_{i,j}$ with least $f_{i,j}$;
    if degree($v_i$) $> 2$ and degree($v_j$) $> 2$ then
        remove $e_{i,j}$;
        find shortest path between $v_i$ and $v_j$;
        add $L_{i,j}$ to each edge on the shortest path
    else
        $f_{i,j} \leftarrow \infty$
    end
end
```

Algorithm 1: The HINT algorithm

5.2.2 HINT Evaluation

Our evaluation of HINT is based on the two networks, which have published structures, Abliene and Level-3. In each case, we use the traffic demands for the cities connected by each network and distances between these cities as the input to the HINT algorithm. Figure 5.7 illustrates the resulting HINT topology connecting the Level-3 cities and Figure 5.8 illustrates the HINT topology connecting the Abilene cities. Even though the Powerlaw and the ER random models do not consider Euclidian distance for model generation, we plot them on the same U.S. map for comparison purposes. Waxman produces a network topology based on proximity of Euclidian distance and thus it produces more realistic network topology for the Internet, but since it does not account for traffic demands, the resulting topology does not match the HINT topology. The figures clearly show that, as opposed to Powerlaw, ER, and Waxman models, the HINT topologies closely match the published Level-3 and Abilene networks. For example, most links in the HINT topology match the Abilene networks except for two links, which connect Houston, TX & New York, NY and Atlanta, GA & New York, NY. The direct connections are included in the HINT topology because the demand factors dominate the distance factor between these cities. The population of New York is about 8 millions and one of Houston is 2 million while Atlanta has only about 400,000 people in the city boundary. This leads to the direct connection from Houston to New York even though Atlanta is much closer to Houston. This limitation in HINT can be remedied by revisiting the removed links after HINT has converged to check whether re-including them in the HINT topology will lead to a reduction in cost. We leave this part for future research.

The total length of links in a network provides a first-order approximation of the network cost. Table 5.2 lists the cost resulting from the HINT, Powerlaw, Waxman, and ER topologies using 250 US cities, which have a population of more than 100,000. HINT leads to one graph but other topologies may differ for different runs due to random parameters used in their topology construction. The numbers in the table are the results of 100 runs. As can be see from the table results, the HINT model is more economical than the others.

To test the resiliency of the network, we remove a node or an edge one at a time

Table 5.2: Total Length To Build Network

HINT	9,025,507km
Powerlaw	11,480,740km ± 824,000km
Waxman	9,775,800km ± 302,000km
ER Random	10,970,680km ± 957,000km

by random and by degree order. Random removing of a node or an edge tests the network in case of device failure, while removing them by order tests the network against targeted attack scenario [16]. To remove a node or an edge by order, we select highest degree node or edge in the graph first. For an edge, we consider the degree of an edge as how many neighboring edges are instant to the given edge. We remove a node or an edge until the largest cluster includes below 30% of the nodes in the network, and measure the percentage of nodes or edges has been removed. We assume a network is broken when it reaches the point that the largest cluster covers only 30% of nodes in the network. The results are summarized in Table 5.3 and Table 5.4. From the result, we see HINT is the most reliable against targeted attack and very resilient to random failure. Against attack, with only 15% of node removing, Powerlaw network is broken down, which is very vulnerable, while HINT requires 31% of node removing to have same effect. Removing an edge does not make big difference. Against random failure, Powerlaw produces the most reliable network, followed by HINT. To break the HINT network, 28% of nodes should be failed, but Powerlaw network is still connected after removing 39% of node failure.

Table 5.3: Resiliency Against Attack

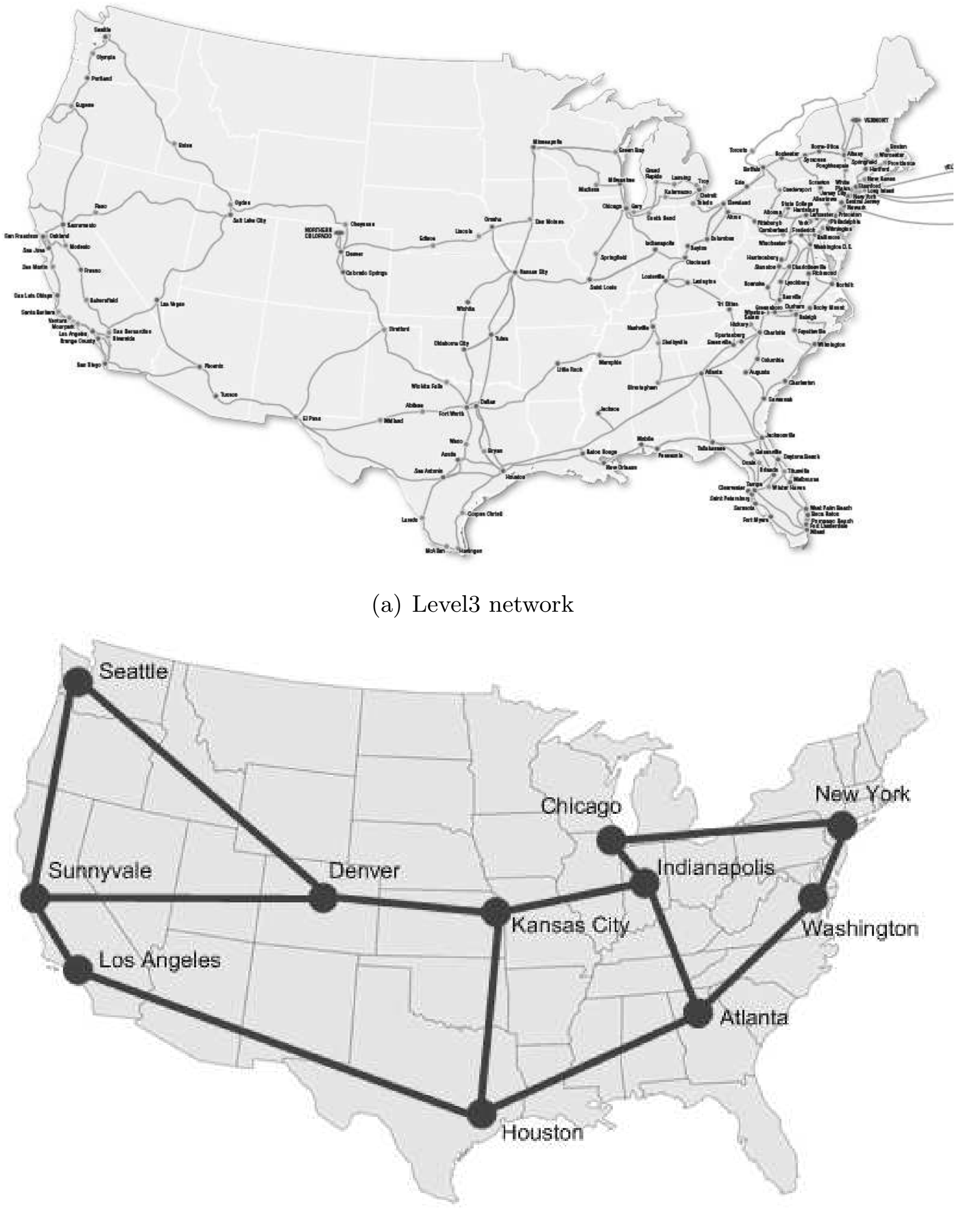
	Node	Edge
HINT	31%	30%
Powerlaw	15%	18%
Waxman	20%	20%
ER Random	19%	18%

Table 5.4: Resiliency Against Failure

	Node	Edge
HINT	28%	27%
Powerlaw	39%	41%
Waxman	20%	21%
ER Random	19%	19%

5.3 Summary

In this chapter, we introduced a new network topology generation model (HINT) which produces cost efficient, better performance and more secure network topology for the backbone of the Internet. Published Internet backbone structures confirm the edge that HINT model enjoys over other existing topology models. Furthermore, HINT labels topology edges with expected traffic load and distance, thus providing more detailed Internet maps than existing topology models.



(a) Level3 network

(b) Abilene network

Figure 5.3: (a) Level3 network. Backbone nodes build a cycle with neighboring nodes and the backbone itself is a collection of cycles. (b) Abilene, Internet2 network. Clear example of cycle backbone.

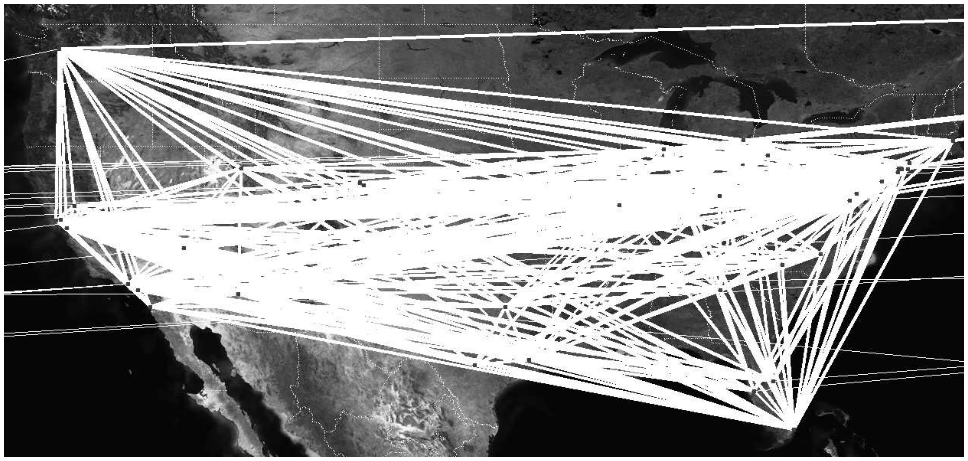


Figure 5.4: Level3 backbone network drawn by Rocketfuel project. Star-like topology directly connecting many remotely apart cities show significant visual difference between projected and actual Level3 backbone network.

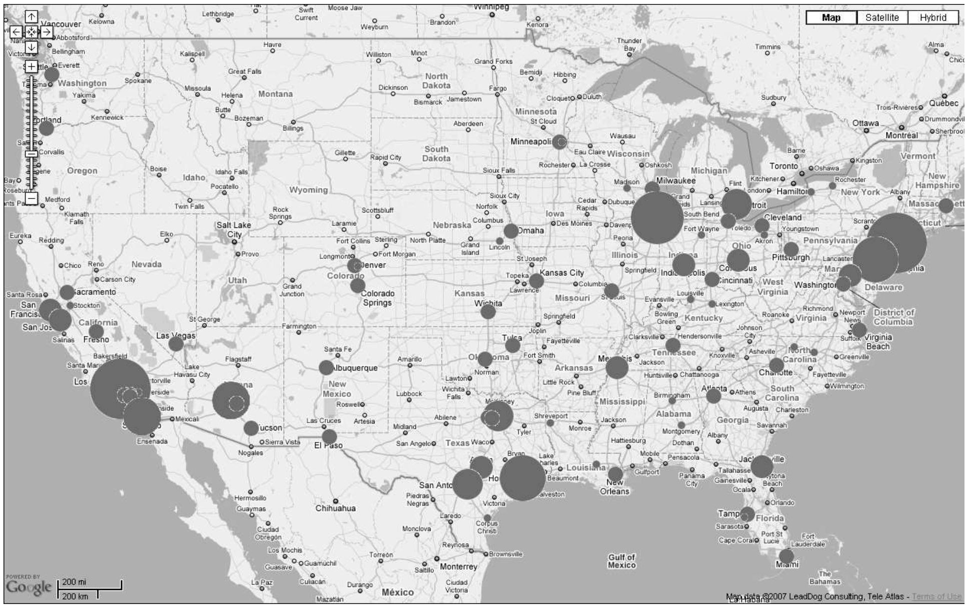
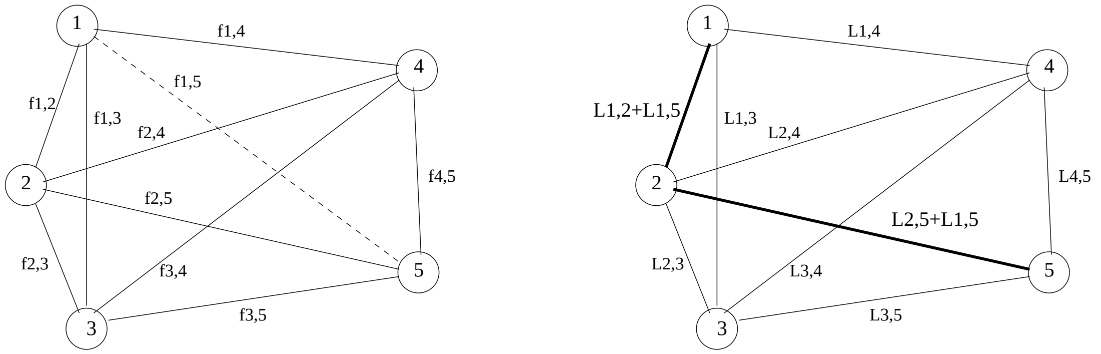

Map
Satellite
Hybrid
200 mi
200 km
Map data ©2007 LeadDog Consulting, Tele Atlas - Terms of Use

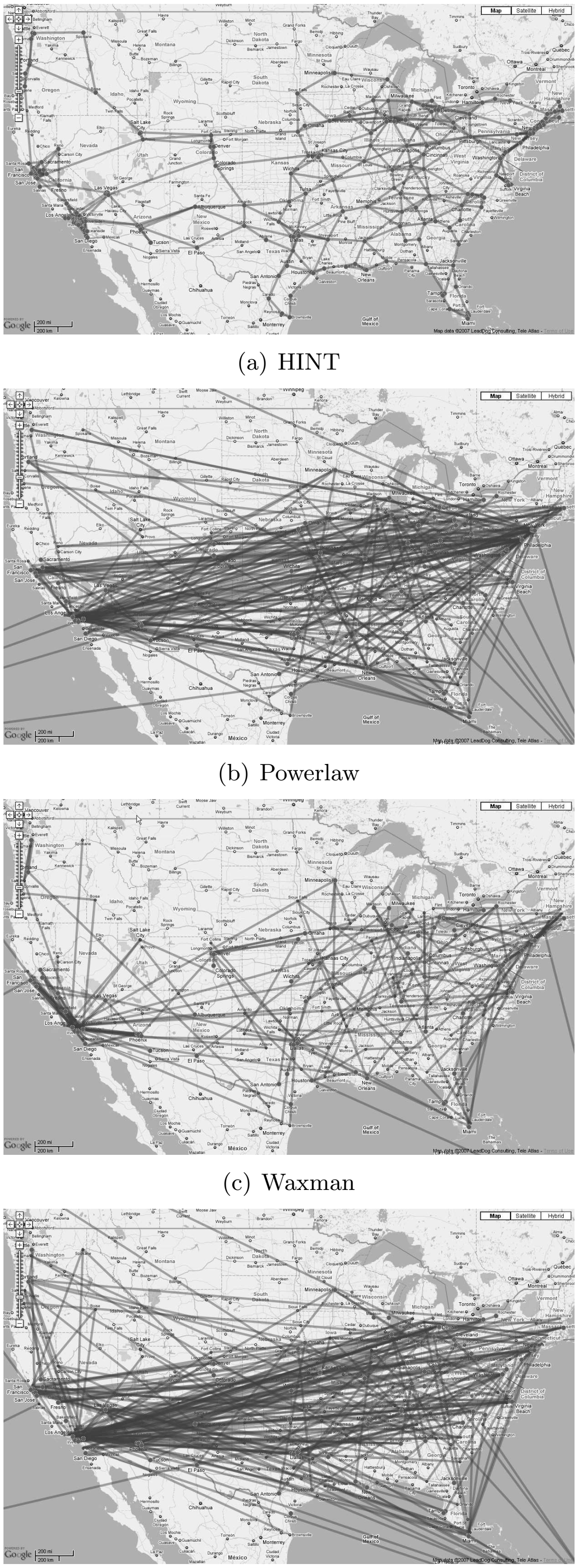
Figure 5.5: Demand model for HINT topology generation



(a) Pick edge

(b) Update

Figure 5.6: HINT process: (a) Select an edge with least connection force. (b) Find shortest path and update traffic load.



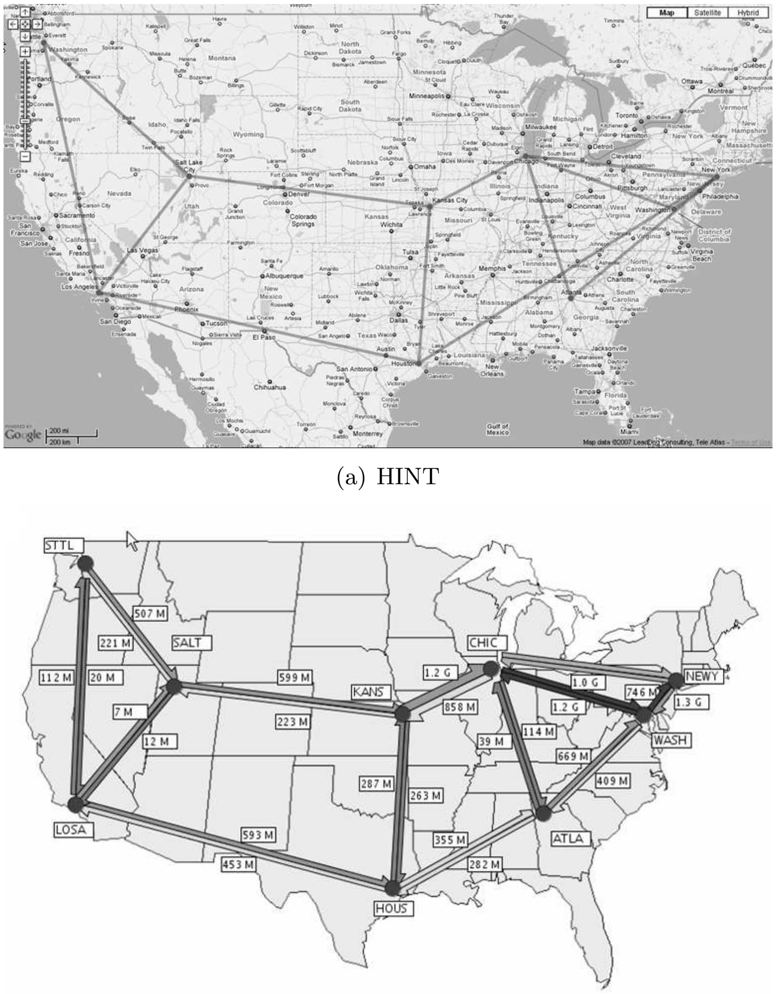
(a) HINT

(b) Powerlaw

(c) Waxman

(d) ER Random

Figure 5.7: (a) The Internet model produced by HINT. (b) Network by powerlaw model. (c) Internet model produced by Waxman. (d) ER Random model for the Internet.



(a) HINT

(b) Abilene

Figure 5.8: (a) A simulated Abilene network produced by HINT. (b) Abilene network.

Chapter 6

Conclusion

This thesis is dedicated to investigating the network property, especially the structure of the Internet. Throughout the thesis, we heavily measure the Internet using servers in Planetlab to produce the map of the Internet. Our level of measurement is router-level topology and our methodology finds more routers and connections than other studies. The mapping result from the measurement is a power-law tendency when it comes to the degree distribution of each router. From the measurement, chapter 2 is discussing why the router level topology shows a power law distribution based on theory and simulation approach. Our claim is that a power law tendency of the router connectivity is because of the hidden layer 2 devices and those devices with limitation of trace routing tools misguides us to wrong direction. Our finding from the study is logical and plausible and thus free us from the power law binding for the Internet structure. Last chapter is introducing a new Internet topology generation model, HINT, which considers demand, supply, cost, performance, and security. Finally product from the model looks very close to the real Internet and thus we claim that HINT is more useful topology generation tool for the Internet.

We investigate the Internet based on the assumption that the network is routing the traffic following shortest path, but it might not necessarily be the case in the real world. The routing is more dynamic to adjust the network environment and finding alternative path for the routing when the shortest path is disconnected. The peak time for the network usage should be different in different time zone and that might be another factor to design the Internet. While people in West coast are sleeping, people in other side of the coast are working. The routing path might be impacted to utilize the paths which is not full functioning in a certain time period. If it is not the case, we can think about how to increase the utilization rate overall areas and make amount of traffic balance. These all factors might be considered when we design the network. As technology grows and changes, so as our researchers.

Bibliography

[1] American registery for internet numbers (arin). http://www.arin.net/.

[2] Asia pacific network information centre (apnic). http://www.apnic.net/.

[3] Cooperative association for internet data analysis(caida), the skitter project. http://www.caida.org/Tools/Skitter.

[4] Cooperative association for internet data analysis(caida), the skitter project. http://www.caida.org/Tools/Measurement/Skitter/router-topology.

[5] Domain name system. http://www.dns.net/dnsrd/.

[6] Planetlab. http://www.planet-lab.org/.

[7] Ripe network coordination centre (ripe). http://www.ripe.net/.

[8] Rocketfuel: An isp topology mapping engine. http://www.cs.washington.edu/research/networking/rocketfuel/.

[9] University of oregon route views project. http://www.routeviews.org/.

[10] W. Aiello, F. Chung, and L. Lu. A random graph model for massive graphs. In *In Proceedings of STOC*, pages 171–180, Portland, Oregon, May 2000.

[11] R. Albert and A.-L. Barabási. Topology of evolving networks: Local events and universality. *Physical Review Letters*, 85(24):5234–5237, December 2000.

[12] R. Albert and A.-L. Barabási. Statistical mechanics of complex networks. *Reviews of Modern Physics*, 74:47–97, January 2002.

[13] R. Albert, H. Jeong, and A.-L. Barabási. Diameter of the world-wide web. *Nature*, 401:130, September 1999.

[14] R. Albert, H. Jeong, and A. Barabsi. The internet's achilles' heel: Error and attack tolerance of complex networks.

[15] Réka Albert, István Albert, and Gary L. Nakarado. Structural vulnerability of the north american power grid. page 025103, 2004.

[16] Reka Albert, Hawoong Jeong, and Albert-Laszlo Barabási. Error and attack tolerance of complex networks. *Nature*, 406:378, 2000.

[17] D. Alderson, J. Doyle, R. Govindan, and W. Willinger. Toward an optimization-driven framework for designing and generating realistic internet topologies. *ACM SIGCOMM Computer Communication Review*, 1(33):41–46, January 2003.

[18] David Alderson, Lun Li, Walter Willinger, and John C. Doyle. Understanding internet topology: Principles, models, and validation. December 2005.

[19] L.A.N. Amaral, A. Scala, M. Barthelemy, and H. E. Stanley. Classes of small-world networks. *Proc Natl Acad Sci U S A*, 97(21):11149–11152, October 2000.

[20] A.-L. Barabási and R. Albert. Emergence of scaling in random networks. *Science*, 286(15):509–512, October 1999.

[21] A.-L. Barabási and E. Bonabeau. Scale-free networks. *Scientific American*, pages 50–59, May 2003.

[22] B. Bollobas and O. Riordan. Robustness and vulnerability of scale-free random graphs., 2003.

[23] T. Bu and D. Towsley. On distinguishing between internet power law topology generators. In *INFOCOM*, New York, June 2002.

[24] K. Calvert, M. Doar, and E. Zagura. Modeling internet topology. In *IEEE Communications Magazine*, number 35, pages 160–163, June.

[25] J. M. Carlson and J. Doyle. Highly optimized tolerance: A mechanism for power laws in designed systems. *Physics Review E*, 2(60):1412–1427, 1999.

[26] Chin-Chen Chang and Kuo-Feng Hwang. Towards the forgery of a group signature without knowing the group center's secret. *Lecture Notes in Computer Science*, 2229, 2001.

[27] H. Chang, S. Jamin, and W. Willinger. Inferring as-level internet topology from router-level path traces. In *In Proceeding of SPIE ITCom*, Denver, CO, August 2001.

[28] Hyunseok Chang, Ramesh Govindan, Sugih Jamin, Scott J. Shenker, and Walter Willinger. Towards capturing representative as-level internet topologies. *Comput. Networks*, 44(6):737–755, 2004.

[29] Q. Chen, H. Chang, R. Govindan, S. Jamin, S. Shenker, and W. Willinger. The origin of power laws in internet topologies revisited.

[30] F. Chung and L. Lu. In *Internet Mathematics*, number 1, pages 91–103.

[31] M. Doar. A better model for generating test networks, 1996.

[32] Benoit Donnet, Philippe Raoult, Timur Friedman, and Mark Crovella. Efficient algorithms for large-scale topology discovery. In *SIGMETRICS '05: Proceedings of the 2005 ACM SIGMETRICS international conference on Measurement and modeling of computer systems*, pages 327–338, 2005.

[33] P. Erdös and A. Réyni. On the evolution of random graphs. *Publ. Math. Inst. Hung. Acad. Sci.*, 1960.

[34] A. Fabrikant, E. Koutsoupias, and C. Papadimitriou. Heuristically optimized trade-offs: A new paradigm for power laws in the internet. *In Proceedings of ICAALP 2002*, 2(60):110–122, July 2002.

[35] Michalis Faloutsos, Petros Faloutsos, and Christos Faloutsos. On power-law relationships of the internet topology. In *SIGCOMM*, pages 251–262, 1999.

[36] M. Fayed, P. Krapivsky, J. Byers, M. Crovella, D. Finkel, and S. Ridner. the emergence of highly variable distributions in the autonomous system topology, 2003.

[37] F. Georgatos, F. Gruber, D. Karrenberg, M. Santcroos, A. Susanj, H. Uijterwaal, and R. Wilhelm. Providing active measurements as a regular service for isps. In *In Proc. PAM 2001*.

[38] Michel L. Goldstein, Steven A. Morris, and Gary G. Yen. Problems with fitting to the power-law distribution. 2004.

[39] Ramesh Govindan and Hongsuda Tangmunarunkit. Heuristics for internet map discovery. In *IEEE INFOCOM 2000*, pages 1371–1380, Tel Aviv, Israel, March 2000.

[40] C. Jin, Q. Chen, and S. Jamin. Internet topology generator, 2000.

[41] James Holland Jones and Mark S. Handcock. An assessment of preferential attachment as a mechanism for human sexual network formation.

[42] Sangmin Kim and Khaled Harfoush. Efficient estimation of more detailed internet ip maps. In *IEEE International Conference on Communications*, pages 377–384, June 2007.

[43] A. Lakhina, J. Byers, M. Crovella, and P. Xie. Sampling Biases in IP Topology Measurements. Technical Report BUCS-TR-2002-021, July 2002.

[44] A. Lakhina, J. Byers, M. Crovella, and P. Xie. Sampling biases in ip topology measurements. In *IEEE INFOCOM*, San Francisco, CA, March 2003.

[45] L. Li, D. Alderson, W. Willinger, and J. Doyle. A first-principles approach to understanding the internet's router-level topology. Portland, OR, 2004. ACM SIGCOMM.

[46] F. Liljeros, C. R. Edling, L. A. N. Amaral, and H. E. Stanley. The web of human sexual contacts. *Nature*, 411:907, June 2001.

[47] Tony McGregor, Hans-Werner Braun, and Jeff Brown. The nlanr: Network analysis infrastructure. May 2000.

[48] A. Medina, A. Lakhina, I. Matta, and J. Byers. Brite: An approach to universal topology generation. In *In Proceedings MASCOTS 2001*, pages 346–356, Cincinnati, Ohio, August 2001.

[49] A. Medina, A. Lakhina, I. Matta, and J. Byers. BRITE: Universal topology generation from a user's perspective. Technical Report 2001-003, Jan 2001.

[50] Stanlay Milgram. The small world problem. *Psychology Today*, 61:60–67, 1967.

[51] M. Newman. The structure and function of complex networks, 2003.

[52] M. E. J. Newman. Assortative mixing in networks. *Physical Review Letters*, 89(20):208701, November 2002.

[53] C. R. Palmer and J. G. Steffan. Generating network topologies that obey power-laws. In *In Proceedings of GLOBECOM 2000*, San Francisco, CA, November 27 - December 1 2000.

[54] Jean-Jacques Pansiot and Dominique Grad. On routes and multicast trees in the internet. *SIGCOMM Comput. Commun. Rev.*, 28(1):41–50, 1998.

[55] Christos H. Papadimitriou. Algorithms, games, and the internet. In *ICALP '01: Proceedings of the 28th International Colloquium on Automata, Languages and Programming,*, pages 1–3, London, UK, 2001. Springer-Verlag.

[56] S. Park, D. Pennock, and C. Lee Giles. Comparing static and dynamic measurements and models of the internet as topology. In *In Proceedings of IEEE INFOCOM 2004*, Hong Kong, March 2004.

[57] Y. Rekhter and T. Li. A border gateway protocol 4 (bgp-4), request for comments: 1771, March 1995.

[58] M. Roughan, A. Greenberg, C. Kalmanek, M. Rumsewicz, J. Yates, and Y. Zhang. Experience in measuring backbone traffic variability: Models, metrics, measurements and meaning. In *International Teletraffic Congress (ITC) 18*, 2003.

[59] R. Siamwalla, R. Sharma, and S. Keshav. Discovering internet topology.

[60] G. Siganos, M. Faloutsos, P. Faloutsos, and C. Faloutsos. Powerlaws and the as-level internet topology, 2003.

[61] N. Spring, R. Mahajan, and D. Wetherall. Measuring isp topologies with rocketfuel, 2002.

[62] Neil Spring, Mira Dontcheva, Maya Rodrig, and David Wetherall. How to resolve ip aliases. Technical Report UW-CSE-TR 04-05-04, May 2004.

[63] Lakshminarayanan Subramanian, Sharad Agarwal, Jennifer Rexford, and Randy H. Katz. Characterizing the internet hierarchy from multiple vantage points. In *Proc. of IEEE INFOCOM 2002, New York, NY*, Jun 2002.

[64] H. Tangmunarunkit, J. Doyle, R. Govindan, S. Jamin, and S. Shenker. Does as size determine degree in as topology. unpublished, 2001.

[65] Hongsuda Tangmunarunkit, Ramesh Govindan, Sugih Jamin, Scott Shenker, and Walter Willinger. Network topology generators: degree-based vs. structural. In *SIGCOMM*

'02: Proceedings of the 2002 conference on Applications, technologies, architectures, and protocols for computer communications, pages 147–159, 2002.

[66] R. Teixeira, K. Marzullo, S. Savage, and G. Voelker. In search of path diversity in isp networks, 2003.

[67] A. Vázquez. Statistics of citation networks. preprint cond-mat/0105031, May 2001.

[68] D. J. Watts and S. H. Strogatz. Collective dynamics of 'small-world' networks. *Nature*, 393:440–442, 1998. Characteristic path length and clustering coefficient.

[69] Duncan Watts. *Small worlds: the dynamics of networks between order and randomness.* Princeton studies in complexity. Princeton University Press, Princeton, N.J., 1999.

[70] B. Waxman. Routing of multipoint connections. *IEEE J. Selec. Areas Communications*, 6(9):1617–1622, December 1988.

[71] W. Willinger, R. Govindan, S. Jamin, V. Paxson, and S. Shenker. In *In Proceedings of PNAS 2002*, pages 2573–2580.

[72] S. Yook, H. Jeong, and A. Barabási. Modeling the internet's large-scale topology. *PNAS*, 99(21):13382–13386, October 2002.

[73] E. Zagura, K. Calvert, and M. Donahoo. A quantitative comparison of graph-based models for internet topology. In *IEEE/ACM Transaction on Networking*, pages 770–783, 1997.

[74] E. W. Zegura. GT-ITM: Georgia Tech Internetwork Topology models (software). http://www.cc.gatech.edu/project, 1996.

[75] Ming Zhang, Yaoping Ruan, Vivek S. Pai, and Jennifer Rexford. How dns misnaming distorts internet topology mapping. In *Proceedings of the 2006 Usenix Annual Technical Conference*, Boston, MA, June 2006.