Chapter 6

AN OPTIMIZATION-BASED APPROACH TO MODELING INTERNET TOPOLOGY

David Alderson
California Institute of Technology
alderd@cds.caltech.edu

Walter Willinger
AT&T Labs–Research
walter@research.att.com

Lun Li
California Institute of Technology
lun@cds.caltech.edu

John Doyle
California Institute of Technology
doyle@cds.caltech.edu

Abstract Over the last decade there has been significant interest and attention devoted towards understanding the complex structure of the Internet, particularly its topology and the large-scale properties that can be derived from it. While recent work by empiricists and theoreticians has emphasized certain statistical and mathematical properties of network structure, this article presents an optimization-based perspective that focuses on the objectives, constraints, and other drivers of engineering design. We argue that Internet topology at the router-level can be understood in terms of the tradeoffs between network performance and the technological and economic factors constraining design. Furthermore, we suggest that the formulation of corresponding optimization problems serves as a reasonable starting point for generating "realistic, yet fictitious" network topologies.

Finally, we describe how this optimization-based perspective is being used in the development of a still-nascent theory for the Internet as a whole.

Keywords: Internet topology, network optimization, router constraints, protocol stack, highly optimized tolerance, topology generator.

1. The Importance of Internet Topology

Understanding the large-scale structural properties of the Internet is critical for network managers, software and hardware engineers, and telecommunications policy makers alike. On a practical level, models of network topology factor prominently in the design and evaluation of network protocols, since it is understood that although topology should not affect the *correctness* of a protocol, it can have a dramatic impact on its *performance* [94]. Accordingly, the ability to shape network traffic for the purposes of improved application performance often depends on the location and interconnection of network resources. In addition, a detailed understanding of network topology is fundamental for developing improved resource provisioning, as most network design problems assume a detailed description of existing/available network components.

More broadly, models of network topology may also play an important role in gaining a basic understanding of certain aspects of current large-scale network behavior. For example, the ability to understand, detect, react to, and deal with network attacks such as denial of service (DoS) attacks or network worms/viruses can depend critically on the topology over which those attacks propagate [80]. As the Internet and related communication networks become an increasingly important component of the national economic and social fabric, national security experts and government policy makers seek to understand the reliability and robustness features of this now critical infrastructure [84, 98], and the topological aspects of the Internet are primary to this purpose.

However, understanding the large-scale structural properties of the Internet has proved to be a challenging problem. For a host of technological and economic reasons, the current Internet does not lend itself to direct inspection. Since the Internet is a collection of thousands of smaller networks, each under its own administrative control, there is no single place from which one can obtain a complete picture of its topology. Whereas coordination among the administrative organizations of these separate networks was relatively easy during the Internet's initial days as a research project, the diversity of technologies and organizational entities in the current landscape make this prohibitive. Today, the sheer number of the network components (e.g. nodes, links) in the Internet preclude even the ability to visualize the network in a simple manner [27]. Also, since the decommissioning of the NSFNet in 1995, when administrative control of the Internet was given over to commercial entities, the fear of losing competitive advantage has provided a strong disincentive for network owners and operators to share topology information.

Since the Internet does not lend itself naturally to direct inspection, the task of "discovering" the network has been left to experimentalists who develop more or less sophisticated methods to infer this topology from appropriate network measurements [19, 89, 40, 44, 87, 90]. Because of the elaborate nature of the network protocol suite, there are a multitude of possible measurements that can be made, each having its own strengths, weaknesses, and idiosyncrasies, and each resulting in a distinct, yet fundamentally incomplete, view of the network as a whole.

These factors suggest the need for a theoretical framework that facilitates the modeling of network topology on a large scale and which also provides an understanding of the relationship between network topology and network behavior. This article describes the importance of *an optimization-based perspective* in the development of an explanatory model for Internet topology. Essentially, we argue that Internet topology can be understood in terms of the tradeoffs between network performance and the technological and economic factors constraining design. Furthermore, we suggest that appropriate optimization-based formulations can be used to generate "realistic, yet fictitious" Internet topologies. Finally, we describe how this view of topology is really just a small piece of a much larger picture, where the ultimate goal is to use this optimization-based framework to obtain a fundamental understanding of *the entire Internet protocol stack*. In this manner, recent successes in the use of optimization to capture essential features of network topology and behavior [62] can be viewed as part of an ongoing effort to develop a more comprehensive *mathematical theory for the Internet* [56, 99] and perhaps even a starting point for understanding the "robust, yet fragile" structure of complex engineering systems [33].

2. Previous Work on Internet Topology

Due to the multilayered nature of the Internet protocol stack, there is no one single topology that reflects the structure of the Internet as a whole. Rather, because any two network components (e.g. routers, end hosts) that run the same protocol at the same layer of the architecture can communicate, each protocol induces its own natural graph on the network, representing in turn the connectivity among all such components. For example, the *router-level graph* of the Internet reflects one-hop connectivity between routing devices running the *Internet Protocol (IP)*. Because the router-level graph has received significant attention by the computer networking community, it is sometimes misinterpreted as *the only* Internet graph, but there are many other graphs having very different structural properties and features. For example, the *AS-graph* reflects the "peering relationships" between independent subnetworks, known as *au-*

tonomous systems (ASes). That is, when two independent network providers (e.g. AT&T and Sprint) enter into a business relationship by which they agree to exchange traffic, they connect their router-level infrastructure together at various peering points. Currently, there are over 10,000 ASes in the Internet, and their aggregate peering structure induces an alternate graph in which each AS (composed of hundreds or thousands of router-level components) can be represented as a single node and each peering relationship (again, possibly reflecting many physical peering points) can be represented as a single link between two ASes. At an entirely different level of abstraction, recent interest in the World Wide Web (WWW) has brought attention to the large-scale graph structure among web documents (represented as nodes) that are connected by hyperlinks (represented as links). Thus, the router-level graph reflects a type of physical connectivity, the AS-level graph represents a type of organizational connectivity, and the WWW-graph represents a type of virtual overlay connectivity. However, there is no direct relationship between each of these "Internet graphs", and in general the features of each graph are quite different.

The development of abstract, yet informed, models for network topology generation has followed the work of empiricists. The first popular topology generator to be used for networking simulation was the Waxman model [97], which is a variation of the classical Erdos-Renyi random graph [35]. In this model, nodes are placed at random in a two-dimensional space, and links are added probabilistically between each pair of nodes in a manner that is inversely proportional to their distance. As a representation of the router-level graph, this model was meant to capture the general observation that long-range links are expensive. The use of this type of random graph model was later abandoned in favor of models that explicitly introduce non-random structure, particularly hierarchy and locality, as part of the network design [30, 21, 107]. The argument for this type of approach was based on the fact that an inspection of real router-level networks shows that they are clearly not random but do exhibit certain obvious hierarchical features. This approach further argued that a topology generator should reflect the design principles in common use. For example, in order to achieve desired performance objectives, the network must have certain connectivity and redundancy requirements, properties which are not guaranteed in random network topologies. These principles were integrated into the Georgia Technology Internetwork Topology Models (GT-ITM) simulator.

These *structural topology generators* were the standard models in use until the discovery of scaling or "power law" relationships in the connectivity of both the AS-graph and the router-level graph of the Internet [38, 88] and in the WWW-graph [55, 5, 2]. More specifically, these findings suggest that the distribution of *degree* (i.e. number of connections, denoted here as x) for each node is appropriately represented in the tail by a function $d(x) \propto k_1 x^{-\beta}$, where $0 < \beta < 2$ and k_1 is a positive finite constant. This conjecture/observation has been highly influential in spawning a line of research focused on the identification and explanation of power law distributions in network topology [102, 26, 70, 100, 74], and it has also influenced research on the development of network topology generators. As a result, state-of-the-art generators have recently been evaluated on the basis of whether or not they can reproduce the same types of power law relationships [20]. Since the Transit-Stub and Tiers structural generators in GT-ITM fail to produce power laws in node degree, they have been largely abandoned in favor of new *degree-based* models that explicitly replicate these observed statistics [94]. Examples of these generators include the INET AS-level topology generator [52], BRITE [69], the Power Law Random Graph (PLRG) method [4], the Carnegie Mellon power-law generator [83], as well as general preferential attachment methods [102].

Our belief is that it is possible to capture and represent realistic drivers of Internet deployment and operation in order to create a topology generation framework that is inherently *explanatory* and will perforce be *descriptive* as well, in the sense of [100]. Instead of explicitly fitting certain characteristics of measured Internet topologies, any such agreements between our models and empirical observations would instead be evidence of a successful explanatory modeling effort. For the purposes of router-level topology, this approach naturally focuses on the perspective of the Internet Service Provider (ISP), who acts as the owner and operator of this network infrastructure. As discussed in [9], we believe that an understanding of the key issues facing ISPs will naturally lead to the ability to generate "realistic, but fictitious" ISP topologies and that this understanding in turn will yield insight into the broader Internet. Our starting premise is that the design and deployment decisions of the ISP are to a large degree the result of an (explicit or implicit) optimization that balances the functionality of the network with the inherent *technological* and *economic* constraints resulting from available networking equipment and the need for the ISP to operate as a successful business. The power of this approach to contrast optimization-based models with their degree-based counterparts was recently shown in [62]. It is the purpose of this paper to put this work in a broader context that highlights the role of optimization-based models as a starting point for synthetic topology generators and also suggests the potential for an optimization-based perspective to be a unifying concept for understanding the Internet as a whole. At the same time, we also identify aspects of this story in need of additional work by researchers in optimization and network design.

3. Optimization-Based Topology Models

The use of combinatorial optimization in network design has a long history for applications in telecommunication and computer systems, as well as transportation, scheduling, and logistics planning [67, 73, 3]. In particular, the rapid buildout of telephone infrastructure since the early 1960s led to massive interest in network design problems from the operations research community in capacitated network design problems (see [42] for a comprehensive survey of models and algorithms). Most recently, the prevalence of optical networks have brought significant attention to network design problems at the physical and link layer of the Internet. Recent emphasis has been on problems related to routing and wavelength assignment in wave division multiplexing (WDM) networks [104, 103, 59]; the relationship between equipment at the physical layer and the link layer topology for the purposes of minimizing communication equipment costs [71, 72, 47]; and network survivability in the face of component losses [46, 50, 85, 86, 31, 76, 109, 60].

While an optimization-based framework is natural when faced with important network design decisions having complicated combinatorics, it is not immediately clear how this approach is helpful as a tool for modeling Internet structure. On the one hand, because the Internet was designed and built using a *layered architecture* (more on this in the sequel), there are distinctly different network design problems at different layers of the network. In addition, there can be tremendous differences in model details—such as the arc costs (both installation costs and variable use costs), budget constraints, constraints on traffic patterns, constraints on network configuration, and redundancy/survivability constraints—at each level of network design, and these can have significant effect on the network topologies that result. Finally, while the traditional focus of network optimization has been on obtaining quantifiably "good" problem solutions, there has been little work to explain any possible relationship between good design and empirically observed large-scale network features such as power-laws.

The general power of an optimization-based approach to understanding power-laws in complex systems has been documented as part of the so-called *HOT* concept, for *Highly Optimized Tolerance* [23] or *Heuristically Optimized Trade-offs* [37]. By emphasizing the importance of design, structure, and optimization, the HOT concept provides a framework in which the commonly-observed highly variable event sizes (i.e., scaling) in systems optimized by engineering design are the results of tradeoffs between yield, cost of resources, and tolerance to risk. *Tolerance* emphasizes that robustness (i.e., the maintenance of some desired system characteristics despite uncertainties in the behavior of its component parts or its environment) in complex systems is a constrained and limited quantity that must be diligently managed; *Highly Optimized* alludes to the fact that this goal is achieved by highly structured, rare, non-generic configurations which—for highly engineered systems—are the result of deliberate design. In turn, the characteristics of HOT systems are high performance, highly structured internal complexity, apparently simple and robust external behavior, with the risk of hopefully rare but potentially catastrophic cascading failures initiated by possibly quite small perturbations [23].

The first explicit attempt to cast topology design, modeling, and generation as a HOT problem was by Fabrikant et al. [37]. They proposed a toy model of incremental access network design that optimizes a tradeoff between connectivity distance and node centrality. More specifically, when adding a new node i, connect it in order to

$$\min_{j<i} \ \alpha \cdot dist(i,j) + h_j,$$

where $dist(i,j)$ is the distance between nodes i and j, and where h_j is some measure of "centrality" (e.g. number of hops). They showed that changes in the relative weights of these two terms in the overall objective function leads to a range of hierarchical structures in resulting topology, from simple star-networks to trees. More specifically, by tuning the relative importance of the two factors, the authors provided analytical proof that the resulting node degree distributions can be either exponential (non-heavy tailed) or of the scaling (heavy-tailed) type. That is, if $d(x)$ equals the number of nodes with degree $\geq x$, then for $\alpha < 1/\sqrt{2}$, resulting topology is a star; for $\alpha = \Omega(\sqrt{n})$, $E[d(x)] < n^2 e^{-k_2 x}$; and for $\alpha \geq 4$ and $\alpha = O(\sqrt{n})$, $E[d(x)] = k_3(\frac{x}{n})^{-\beta}$. Subsequent work on this model has suggested that the resulting degree distribution follows a power law only up to a cutoff [15]. While this work successfully illustrated the power of HOT to generate heavy-tailed distributions in topology generation, their construction was not intended to be a realistic model of router level topology.

3.1 An Engineering-Based Approach

Our approach to modeling the structural topology of the Internet is rooted in two beliefs. First, as key decision makers in the design and operation of their own network topologies, Internet Service Providers (ISPs) play a fundamental role in the ongoing evolution of Internet structure as a whole. Second, an understanding of the layered architecture of the Internet is critical to the appropriate interpretation of key drivers affecting the decisions made by ISPs.

The robustness and user-perceived simplicity of the Internet is the result of a modular architecture that builds complex functionality from a succession of simpler components [29]. These components are organized into vertical *layers* whereby each component relies on the functionality of the layer below it and provides in turn a new set of functionality to the layer above. Each layer can be implemented more or less independently, provided that it adheres to specified rules for interacting with its adjacent layers. In this manner, layering provides modularity and gives rise to the "hourglass" metaphor—a 5-layer suite of protocols (the "TCP/IP protocol stack") where the Internet protocol (IP) constitutes the waist-layer of the hourglass and provides a simple abstraction of a generic but unreliable data delivery service [78] (see Figure 3.1). The physical layer and link layer below the waist deal with the wide variety of existing transmission and link technologies and provide the protocols for running IP over whatever bit-carrying network infrastructure is in place. Above the waist is where the transport layer and application layer provide the protocols that enhance IP (e.g., at the transport layer, TCP ensures reliable transmission) and greatly simplify the process of writing applications (e.g., WWW) through which users actually interact with the Internet. By including multiple layers of feedback control, this architecture provides much more than mere modularization, being largely responsible for the legendary ability of the Internet to perform robustly even in the presence of component losses [99].

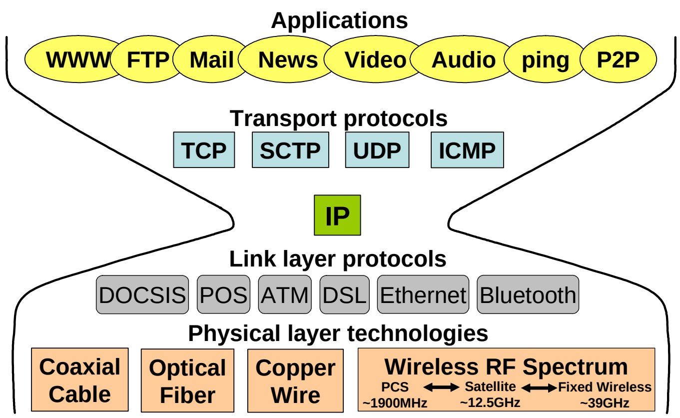


Figure 6.1. The Internet "hourglass". The physical layer and link layer below the waist deal with the wide variety of existing transmission and link technologies and provide the protocols for running IP over whatever bit-carrying network infrastructure is in place ("IP over everything"). Above the waist is where the transport layer and application layer provide the protocols that enhance IP (e.g., at the transport layer, TCP ensures reliable transmission) and greatly simplify the process of Writing applications (e.g., WWW) through which users actually interact with the Internet ("everything over IP").

Internet Service Providers (ISPs) are the owners and operators of the public Internet infrastructure, and as such, the decisions they make in designing and building their networks largely determine the overall structure of the Internet. Modern ISPs face significant challenges in their ongoing operations [49], and network design problems can factor prominently in their ultimate success as a business. Over the last decade, IP-based networking has emerged as the dominant technology in use within the Internet, and IP now functions as a type of "common currency" within the Internet—nearly all applications are designed to run on it, and most physical network infrastructures are designed to support it. Typically, ISPs do not specify which technologies are used at the upper layers, leaving them open instead to the needs of customers (e.g. end users) and their applications. However, an ISP must provide transmission and link technologies as well as the protocols for running IP over whatever physical infrastructure is in place.

3.2 Network Drivers

Our starting premise is that any explanatory framework for router-level Internet topology modeling should incorporate both the *economic* and *technological* factors faced by ISPs. For example, because of the costly nature of procuring, installing, and maintaining the required facilities and equipment, the ISP is economically constrained in the amount of network that it can support. At the same time, ISPs must configure their limited network resources in a manner that satisfies the service requirement of their customers. In designing the topology that best supports its business, the ISP is further constrained by the technologies currently available to it. While a complete review of these issues is beyond the scope of this paper, we argue that these drivers in their simplest form can be understood in terms of *link costs*, *router technology*, *customer requirements* and *service requirements*.

3.2.1 Link Costs. Operation of an ISP at a national scale requires the installation, operation, and maintenance of communication links that span great distances. For the purposes here, we use the term "link" to mean both the physical network cable and the link layer equipment used to send traffic along that cable. At the national level, the cables are usually fiber optic and the equipment consists of transmitter/receivers at the end points and signal repeaters along the way. In addition, we assume in the remainder of this section that there are no significant differences between the connectivity observed at the IP layer and the underlying link connectivity. While this simplifying assumption holds true for some real networks (such as Abilene described below), we describe in Section 5 how higher fidelity optimization models could treats these layers in isolation. While a significant portion of the link cost is often associated with obtaining the "right of way" to install the network cables, there is generally an even greater cost associated with the installation and maintenance of the equipment used to send the traffic across these cables. Both the installation and maintenance costs tend to increase with link distance. Thus, one of the biggest infrastructure costs facing a network provider is the cost associated with the deployment and maintenance of its links.

National ISPs are one type of network provider for which link costs are significant. However, their challenge in providing network connectivity to millions of users spread over large geographic distances is made somewhat easier by the fact that most users tend to be concentrated in metropolitan areas. Thus, there is a natural separation of the connectivity problem into providing connectivity within a metropolitan region and providing connectivity between these regions[1]. In considering the costs associated with providing connectivity between metropolitan regions, the ISP has strong economic incentive to spread the cost of an intercity link over as many customers as possible. This is the basic motivation for *multiplexing*—a fundamental concept in networking by which a link is shared by many individual traffic streams. Multiplexing one of the most basic design principles in networking and has tremendous impact on the types of topologies chosen by network architects. In it simplest form, it states that the only type of design that makes sense from an economic perspective is one that aggregates as much traffic on the fewest number of long distance links. This principle applies at all levels of network design, including the local and regional levels, and not just the national backbone[2].

3.2.2 Router Technology. Another major constraint affecting the types of topologies available to network designers is related to the routing equipment used to control the flow of traffic on the network. Based on the technology used in the cross-connection fabric of the router itself, a router has a maximum number of packets that can processed in any unit of time. This constrains the number of link connections (i.e., node *degree*) and connection speeds (i.e., bandwidth) at each model type, thereby creating an "efficient frontier" of possible bandwidth-degree combinations available for each router. That is, a router can have a few high bandwidth connections or many low bandwidth connections (or some combination in between). In essence, this means that the router must obey a form of *flow conservation* in the traffic that it can handle. While it is always possible to configure the router so that it falls below the efficient frontier (thereby underutilizing the router capacity), it is not possible to exceed this frontier (for example, by having an ever increasing number of high bandwidth connections). For any particular router model, there will be a frontier representing the possible combinations that are available. Router models with greater capacity are generally more expensive.

Consider as an example the Cisco Gigabit Switch Routers (GSRs), which are one of the most widely deployed routers within the Internet backbone[3]. In Figure 6.2(a), we show an example of the technology constraint of the Cisco 12416 GSR. This router has a total of 15 available "slots" for line cards each of which may have one more ports (i.e. physical connections). When the total number of

[1]Within the ISP industry, this distinction often separates service offerings into two lines of business known as "metro service" and "long-haul service" respectively.
[2]The telephone network is subject to the same economics associated with link costs and also exhibits the same type of network design in which local traffic is aggregated along "trunks" which interconnect local regions. Given a history in which the modern data networks grew out of traditional phone networks, the reuse of commonly accepted and successful design principles is not surprising.
[3]As reported in [34], Cisco's share of the worldwide market for service provider edge and core routers was approximately 70% during 2002.

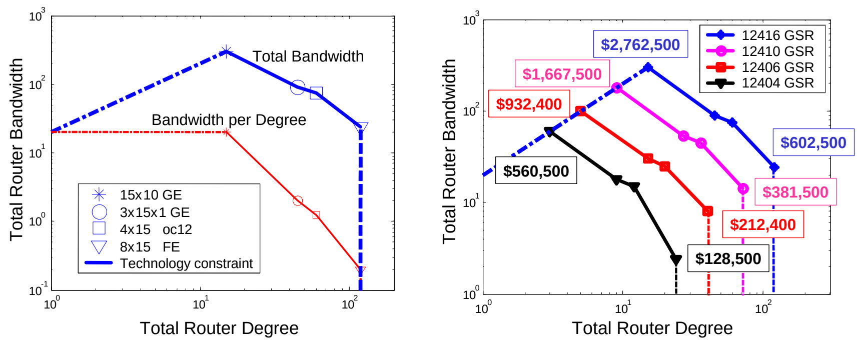


Figure 6.2. (a) Technology constraint for Cisco 12416 Gigabit Switch Router (GSR): Degree vs. maximum throughput. Each point on the plot corresponds to a different combination of ports on each line card. The router can achieve any combination of maximum throughput and degree below the technology constraint line. *(b) Technology constraints for GSR Models 12404, 12406, 12410, 12416.* Each line represents one type of router, and each point on the plot corresponds to routers with different interfaces, with the corresponding price shown in the enclosed inbox.

connections is less than 15, each line card need only support a single port, the throughput of each port is limited by the maximum speed of supported line cards (10 GE), and the router's maximum throughput increases with the number of connections. When the number of connections is greater than 15, the maximum router throughput *decreases* as the total number of ports increases. The reason for this decrease is related to an increased routing overhead for handling traffic over a greater number of ports. Figure 6.2(b) illustrates the efficient frontiers and corresponding prices (associated with different line card configurations) of several Cisco GSR routers taken from a recent product catalog [28]. Although engineers are constantly increasing the frontier with the development of new routing technologies, network architects are faced with tradeoffs between capacity and cost in selecting a router and then must also decide on the quantity and speed of connections in selecting a router configuration.

As noted in Figure 6.2(b), these high capacity core routers can have node degree on the order of only 100 direct connections. As a result, observed IP measurements for nodes having thousands of connections cannot correspond to physical connections between these routers. Until new technology shifts the frontier, the only way to create throughput beyond the frontier is to build networks of routers. In making this claim, we are not arguing that limits in technology fundamentally preclude the possibility of high-degree, high-bandwidth routers, but simply that the product offerings recently available to the marketplace have not supported such configurations[4]. While we expect that companies will continue to innovate and extend the feasible region for router configuration, it remains to be seen whether or not the economics of these products will enable their wide deployment within the Internet.

The current Internet is populated with many different router models, each using potentially different technologies and each having their own technology constraint. However, these technologies are still limited in their overall ability to tradeoff total bandwidth and number of connections. Thus, networking products tend to be specialized to take advantage of one area of an aggregate feasible region, depending on their intended role within the network hierarchy. Figure 6.2 presents an aggregate picture of many different technologies used both in the network core and at the network edge.

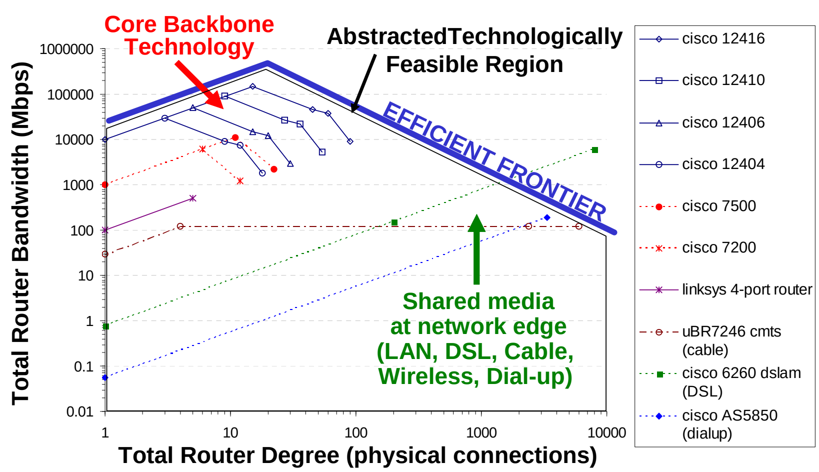


Figure 6.3. Aggregate picture of router technology constraints. In addition to the Cisco 12000 GSR Series, the constraints on the somewhat older Cisco 7000 Series is also shown. The shared access technology for broadband cable provides service comparable to DSL when the total number of users is about 100, but can only provide service equivalent to dialup when the number of users is about 2000. Included also is the Linksys 4-port router, which is a popular LAN technology supporting up to 5 100MB Ethernet connections. Observe that the limits of this less expensive technology are well within the interior of the feasible region for core network routers.

Economic drivers to minimize wiring costs have spawned extreme heterogeneity in the types of technologies that connect at the network edge; for example, dial-up and digital subscriber line (DSL) leverage existing copper telephone lines, broadband cable leverages existing coaxial cable lines, and wireless technology removes the need for wires altogether. These technologies are somewhat

[4]A few companies such as Avici Systems (www.avici.com) have started to offer scalable routing devices built from “stacks” of routers, with some recent success in the marketplace [93].

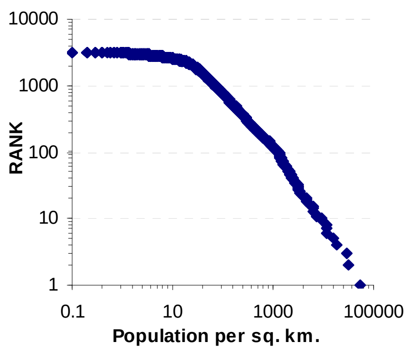
different from core routers in their underlying design, since their intention is to be able to support large numbers of end users at fixed (DSL, dialup) or variable (cable) speeds. They can support a much greater number of connections (upwards of 10,000 for DSL or dialup) but at significantly lower speeds. Collectively, these individual constraints form an overall aggregate feasible region on available topology design.

3.2.3 Customer Constraints. Since the business of the ISP is to provide network service to its customers, there are certain features of ISP network structure that will be driven by the customers it supports. For example, in the current environment there is tremendous variability in the connection speeds used by customers to connect to the Internet. As shown in Table 6.1, an estimated half of all users of the Internet in North America during 2003 still had dial-up connections (generally 56kbps), only about 20% had broadband access (256kbps-6Mbps), and there was only a small number of end users with large (10Gbps) bandwidth requirements [8]. While some of the disparity in consumer choices may be attributed to incomplete deployment of broadband services, it is reasonable to believe that much of this disparity is due to a wide variability in the willingness to pay for network bandwidths.

Table 6.1. Estimated distribution of end host connection types in the United States for 2003. It is important to note that the bandwidth performance seen by an individual user may be less than the total connection speed if the user's network interface card (NIC) is relatively slow. For example, a user on a university campus with a Fast Ethernet (100Mb) card will never achieve more than 100Mbps even if the university has a 10Gbps connection to the Internet.

Type of Edge Connection	*Typical Connection Speed*	*Approx. Connections*	*Relative Frequency*
Campus Users	1.544Mbps(T-1) – 10Gbps(OC-192)	38.1 M	33.6%
Broadband DSL	512kbps – 6Mbps	7.6 M	6.7%
Broadband Cable	300kbps – 30Mbps	13.4 M	11.8%
Dial-Up	56kbps	54.4 M	47.9%
	Total	113.4 M	100%

Another factor facing ISP topology design at the edge is the *location* of its customers. Due to the increased cost of longer links, customers that are located farther away from an ISP's network will be more expensive to service (at least when providing an initial connection). Conversely, regions where potential customers are concentrated over small distances will be more attractive to ISPs. Because population densities themselves range widely by geography (see for example U.S. Census data in Figure 6.4), ISPs that want broad coverage of even the most populated metropolitan regions will need to support wide variability in customer density.



Rank	*Pop. Density (per sq. km)*	*County, State*
1	20239	New York County, NY
2	12594	Kings County, NY
3	11058	Bronx County, NY
4	6888	Queens County, NY
5	5985	San Francisco County, CA
6	4577	Hudson County, NJ
7	4530	Philadelphia County, PA
8	4380	Suffolk County, MA
9	3816	District of Columbia
10	3517	Baltimore City, MD

Figure 6.4. Population density of the United States by county in 1990. (a) Most counties are sparsely populated, but a few counties have extremely high densities. (b) Top ten most densely populated counties. *Source: United States Census Bureau (Released: March 12, 1996. Revised: June 26, 2000.)*

3.2.4 Service Requirements. In addition to the constraints imposed by link costs, router technology limitations, and customer connectivity, it is reasonable to expect that ISPs are driven to satisfy certain service requirements imposed by their customers or the industry at large. For example, most ISPs utilize *service level agreements (SLAs)*, which serve as business contracts with their major customers and their peers. SLAs typically specify terms such as delivered bandwidth and limits on service interruptions, and they often include financial penalties for failure to comply with their terms. While SLAs are often negotiated on an individual basis, competition among ISPs often creates industry norms that lead to standard SLA terms. Conversely, some ISPs use special terms in SLAs as a mechanism for differentiating their services and creating competitive advantage over rival companies.

From the provider's perspective, one simple metric for assessing whether or not a given network topology is "good" is its ability to handle the bandwidth requirements of its edge routers. For the purposes of this paper, we define *network performance* as the maximum proportional throughput on a network under heavy traffic conditions based on a gravity model [108]. That is, starting at the network edge we consider the demand for traffic by an access router to be the aggregate connectivity bandwidth of its end hosts. Then, to determine the flow of traffic across the network core, we consider flows on all source-destination pairs of access routers, such that the amount of flow X_{ij} between source i and destination j is proportional to the product of the traffic demand x_i, x_j at end points i, j,

$$X_{ij} = \alpha x_i x_j,$$

where α is a constant representing the proportional level among all flows. We compute the maximum proportional throughput on the network under the router degree bandwidth constraint,

$$\begin{aligned} \max_{\alpha} \quad & \sum_{ij} X_{ij} \\ s.t \quad & RX \leq B, \end{aligned}$$

where R is the routing matrix (defined such that $R_{kl} = \{0, 1\}$ depending on whether or not flow l passes through router k). We use shortest path routing to get the routing matrix. X is a vector obtained by stacking all the flows X_{ij} and B is a vector consisting of all router bandwidths according to the degree bandwidth constraint (Figure 6.2). Due to lack of publicly available information on traffic demand for each end point, we assume the aggregation of end point traffic demand is proportional to the bandwidth of the higher level router. In this manner, we allow for good bandwidth utilization of the higher level routers[5]. While other performance metrics may be worth considering, we claim that maximum proportional throughput achieved using the gravity model provides a reasonable measure of the network to provide a *fair* allocation of bandwidth.

Another important issue in the design of ISP topologies is related to their reliability in the presence of equipment failure, sometimes known as *survivability*. Generally, network survivability is quantified in terms of the ability of the network to maintain end-to-end paths in the presence of node or link losses. Although survivable network design is not a focus of this article, comprehensive surveys of optimization-based formulations for this type of service requirement are available [46, 50].

3.3 HOT: Heuristically Optimal Topology

Our objective is to develop a simple and minimal, yet plausible model for router-level topology that reflects link costs, conforms to the technology constraints of routers, appropriately addresses the aforementioned issues for high variability in end-user connectivity, and achieves reasonably "good" performance. As noted above, the economic drive to minimize link costs promotes a topology that aggregates traffic as close to the network edge as possible. The use of multiplexing in a variety of routing technologies at the network edge supports this aggregation, and the wide variability in the bandwidth demands and geographies of end user connections suggests that one should expect wide variability in the measured connectivity of nodes at the network edge. Since it is generally accepted that most of the computers in the network are at its edge, it is reasonable to expect that the overall connectivity statistics of the network

[5]We also tried choosing the traffic demand proportional to the product of end points degree as in [43], and a similar result still holds but has different router utilization.

are dominated by those at the edge. Collectively, these constraints suggest that a "good" design is one in which individual links at the edge of the network have are aggregated in a manner such that the link capacities increase as one moves to the network core. In particular, edge routers may connect to a large number of low bandwidth users or a smaller number of high bandwidth users. In contrast, one can expect that backbone links within the network run at high-speeds and that core routers have necessarily fewer links, making the connectivity of the core much more uniform. As can be seen in the example below and in [62], an inspection of real networks reveals a common theme, namely that the topology at the network edge is designed to aggregate traffic within a local region, while the topology within the core of the network is designed to transport aggregated traffic between geographically disparate regions.

3.3.1 Case Study: The Abilene Network. To this point, we have argued that the perspective of an ISP in building a national scale network topology is driven by three factors. First, the need to minimize the long distance link costs means that it is driven to aggregate traffic from its edges to its core. Second, the design of its topology, particularly in the core, must conform to the technology constraints inherent in routers. Third, the network should have good performance, measured in terms of its ability to carry large volumes of traffic in a fair manner. While these are certainly not the only factors affecting design, we claim that these three drivers are a sensible starting point for understanding the relationship between ISP network design and resulting router-level topology. As a preliminary validation of whether or not these factors are reasonable, we seek to compare them to the topology of a national ISP. Given that commercial ISPs are reluctant to share information, we consider the national educational network.

The Abilene Network is the Internet backbone network for higher education, and it is part of the Internet2 initiative [1] (see Figure 6.5). It is comprised of high-speed connections between core routers located in 11 U.S. cities and carries approximately 1% of all traffic in North America[6]. The Abilene backbone is a sparsely connected mesh, with connectivity to regional and local customers provided by some minimal amount of redundancy. Abilene maintains peering connections with other higher educational networks (both domestic and international) but does not connect directly to the commercial Internet. Within Abilene, connectivity from core routers to academic institutions is provided

[6]Of the approximate 80,000 - 140,000 terabytes per month of traffic in 2002 [81], Abilene carried approximately 11,000 terabytes of total traffic for the year [51]. Here, "carried" traffic refers to traffic that traversed an Abilene router. Since Abilene does not peer with commercial ISPs, packets that traverse an Abilene router are unlikely to have traversed any portion of the commercial Internet.

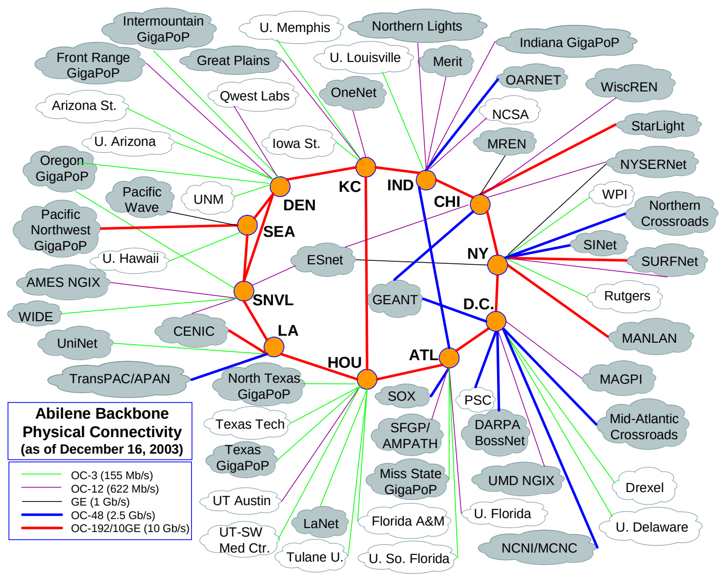


Figure 6.5. *Complete physical connectivity for Abilene.* Each node represents a router, and each link represents a physical connection between Abilene and another network. End user networks are represented in white, while peer networks (other backbones and exchange points) are represented in blue. ESnet is another national backbone network. Each router has only a few high bandwidth connections, however each physical connection can support many virtual connections that give the appearance of greater connectivity to higher levels of the Internet protocol stack.

through local GigaPoPs[7]. The core router in Los Angeles, for example, is connected to core routers in Sunnyvale and Houston, and is also connected to regional networks such as CENIC (the educational backbone for the State of California) and peering networks such as UniNet. In places where no GigaPoP is available, university campuses may be allowed to connect to Abilene directly. For example, the University of Florida is directly connected to an Abilene core router in Atlanta. Within Abilene, there is no difference between the network at the IP layer and the link layer, and the physical connectivity of central core nodes is low (ranging from five to twelve).

We claim that the Abilene backbone is heuristically optimal in its ability to tradeoff performance and link cost, subject to router technology constraints, so we use it as a starting point to construct a toy model of a heuristically optimal topology. Specifically, we replace each of the edge network clouds with a single gateway router whose role is to aggregate the traffic of a certain number of end

[7]The term “PoP” is an abbreviation for “Point of Presence”. A GigaPoP is a point of presence that interconnects many different networks at very high bandwidths.

hosts. Thus, most of the nodes in the network are at its edge, and the high degree nodes are located exclusively at the edge of the network. This flexibility makes it trivial to assign just about any node degree distribution to the network as a whole, and here we assign end hosts to gateway routers in a manner that yields an approximate power law for the overall node degree distribution. Figure 6.6(a) shows the resulting network topology, while Figure 6.6(b) shows the degree distribution for the entire network. This network has total 865 nodes (68 internal routers and 797 end hosts) and 874 links. In approximating Abilene, peering networks ESnet and GEANT are represented as direct connections between Chicago and Sunnyvale, New York and Sunnyvale, and Chicago and Washington D.C. To evaluate the performance of this construction, we assume

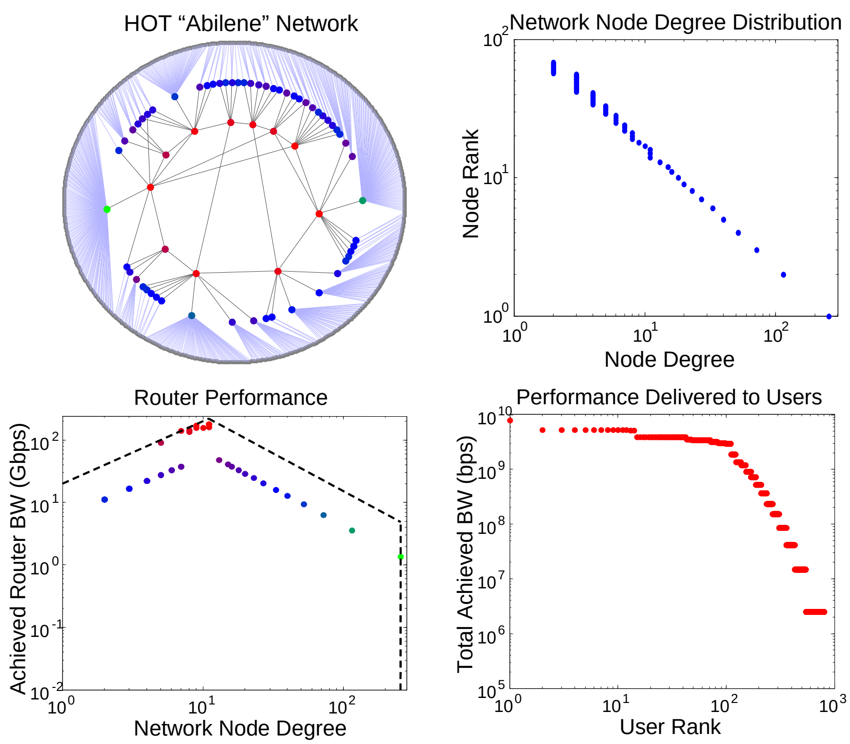


Figure 6.6. Abilene-inspired HOT model of router-level topology. (a) Topology for toy version of Abilene. (b) Network degree distribution. (c) Achieved router performance under maximum flow. (d) Achieved end user bandwidths under maximum flow.

that the network is built using a single router model having an abstracted feasible region, shown in Figure 6.6(c). Then, this network achieves a total performance of 576 Gbps and its routers are used with high efficiency, i.e. they are close to the efficient frontier of the router feasible region. It also provides wide variability in the ultimate bandwidths delivered to its end users, shown in Figure 6.6(d).

This simple toy model is obviously a severe abstraction, and we do not claim that it is an accurate representation of Abilene. However, this toy model illustrates a few important points. First, it is relatively straightforward to use engineering design to construct a network topology that conforms to the constraints of available router technology, has high variability (or even power laws) in overall node degree distribution, and supports a wide variability in end user bandwidths. Second, a network that has low degree, high bandwidth connections in the core and high degree, low bandwidth connections at the edge results in high performance and achieves efficient router utilization (Figure 6.6(c)). From an engineering perspective that explicitly considers the tradeoffs of network design, these points may seem so obvious that it is hard to imagine modeling router-level graphs in any other way. However, there exists a popular alternate approach that considers only the mathematical and graph theoretic aspects of network connectivity.

3.4 Equivalent Degree-Based Models

The starting point for many models of network topology has been to try to replicate the mathematical or statistical properties observed in real networks. With this approach, one usually starts with a sequence of well-understood metrics or observed features of interest—such as hierarchy [107], node-degree distributions [52, 69, 4, 13], clustering coefficients [20], expansion, resilience [94], etc.—and then develops a method that matches these metrics. The result is predictably successful, in the sense that it is always possible to develop models of increasing fidelity in order to tune specific statistics to desired values. Indeed, the common themes of this work are empirical findings that suggest degree based generators provide topologies that are more statistically representative of real networks which are reported to exhibit power law degree distributions [94]. However, this approach suffers from several drawbacks. First, it is hard to choose the "right" metric, since what is right is apt to vary, depending on the intended use of the topology. Second, any generative method that does a good job of matching the chosen metric often does not fit other metrics well. Finally, this approach tends to have little, if any, predictive power, since resulting models tend to be descriptive but not explanatory, in the sense of [100]. Nonetheless, recent attention on power laws in network connectivity has made degree distributions a popular metric for evaluating topology, and degree-based models of Internet topology remain prevalent.

The drawbacks of using the degree-based approach become clear with a closer look at the methods by which these networks are generated. In general, there are many network generation mechanisms that can yield networks having highly variable (or power-law) degree distributions. However, the aforementioned degree-based topology generators all use one of two methods. The first is *preferential attachment* [13] which says (1) the growth of the network is realized by the sequential addition of new nodes, and (2) each newly added node connects to an existing node preferentially, such that it is more likely to connect with a node that already has many connections. As a consequence, high-degree nodes are likely to get more and more connections resulting in a power law in the distribution of node degree. While, preferential attachment has been a popular mechanism within the complex network literature, its utility as a general network modeling tool is limited to particular power law degree distributions. The second, and more general, generation mechanism is based on graph theoretic methods that yield topologies whose expected degree distribution matches any chosen distribution. An example is the Power-law Random Graph (PLRG) model which constructs a graph by first assigning each node its degree and the randomly inserting edges between the nodes according to a probability that is proportional to the product of the given degrees of two endpoints [4]. If the assigned degree distribution for all nodes follows the power-law, the generated network is expected to reproduce the same power law.

One of the most important features of networks having power law degree distributions that are generated by these two mechanisms is that there are a few centrally located and highly connected "hubs" through which essentially most traffic must flow. For the networks generated by preferential attachment, the central hubs are the earliest nodes, and application of the preferential attachment model to the Internet has suggested that these hubs represent the "Achilles' heel of the Internet" because they make the network highly vulnerable to attacks that target these high-degree hubs [6]. The nodes with high expected degree in PLRG have higher probability to attach to other high degree nodes and these highly connected nodes form a central cluster.

Consider as an example a degree-based model for a network having the degree distribution shown in Figure 6.6. Starting with the PLRG approach and using some additional heuristic tuning, it is possible to obtain a network that matches the degree distribution exactly. The resulting network also has 865 nodes (68 internal routers and 797 end hosts) and 874 links, and it is shown in Figure 6.7(a). While it matches the degree distribution of the HOT model, it achieves an inferior performance of only 4.89 Gbps, more than two orders of magnitude worse. A look at the router performance (Figure 6.7(c)) and the distribution of bandwidths to end users (Figure 6.7(d)) also reveals the functional inferiority of this network. The reason for the poor performance of this degree-based model is exactly the presence of the highly connected "hubs" that create low-bandwidth bottlenecks. In contrast, the HOT model's mesh-like core, like the real Internet, aggregates traffic and disperses it across multiple high-bandwidth routers.

While the graph shown in Figure 6.7(a) is not the only network that could have resulted from a probabilistic degree-based construction, its structure is representative of the features of this genre of graph models. Furthermore, its

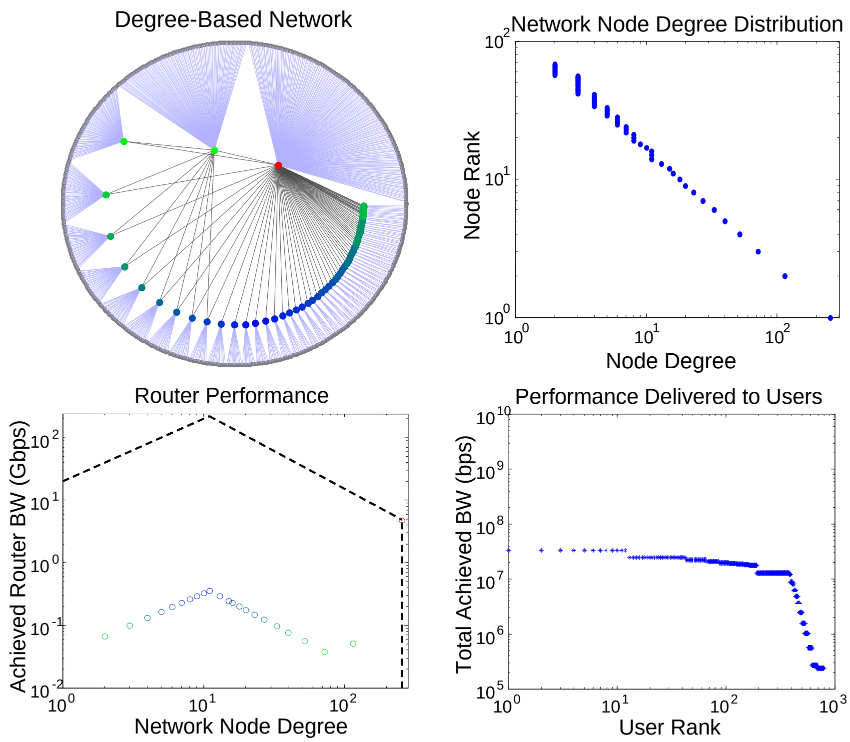


Figure 6.7. Result of degree-based generation for network having the same degree distribution as in Figure 6.6. (a) Network topology. (b) Network degree distribution. (c) Achieved router performance under maximum flow. (d) Achieved end user bandwidths under maximum flow.

comparison with the Abilene-inspired network is fair, in the sense that two different graphs constructed from the same respective processes would compare in a qualitatively similar manner. There are a number of graph theoretic arguments that suggest why probabilistic degree-based networks result in this type of structure (some of which are presented in [62]), but they are beyond the scope of this article.

3.5 Discussion

The objective here has been to uncover "significant" drivers of topology evolution so that we understand the structure of the real Internet and gain insight for generating synthetic topologies. The Abilene-inspired example and degree-based example were selected to highlight how graphs that share certain macroscopic statistical characteristics can appear very different from a perspective that considers issues such as performance, technology constraints, and economic considerations. The HOT framework tries to emphasize optimization, tradeoffs, and robustness. The Abilene-inspired network is the simplest nontrivial example. There may be additional constraints, such as link costs and redundancy, that would be important and needed in the model. However, since HOT views performance and constraints and tradeoffs, we can generalize the HOT model to include any additional objectives and constraints. In this manner, HOT design is an example of a process, whereby additional constraints and tradeoffs can be added in a systematic manner. Degree-based methods are representative of a process too, but the process is a search for macroscopic statistical characterizations to constrain an otherwise random network construction. Although both approaches could be extended, the HOT approach makes engineering sense while the degree-based approach does not. Even so, the results here are really only a "proof of concept" in the development of "fictitious, yet realistic" models of Internet topology.

4. Towards Generative Models

In considering how we might use the aforementioned optimization-based framework to generate synthetic network topologies, we return to the perspective of the national ISP that is looking to build its infrastructure from scratch. While there are many distinct network design problems of concern to ISPs, we distinguish here between the problem of *access* design and *backbone* design. *Network access design* typically occurs at the metropolitan area, where the challenge is to provide connectivity to local customers who are dispersed over some regional geographic area. This connectivity is rooted at the ISP's *point of presence (PoP)*, which serves as the focal point for local traffic aggregation and dissemination between the ISP and its customers. In contrast, *network backbone design* is the problem of providing internal connectivity within the ISP between its different PoPs, which are typically separated by larger geographic distances. The separation of the network design problem into access and backbone design problems and also into different functional layers is important because the resulting optimization formulations often have different objectives and constraints operating at each level of abstraction. In addition, access design problems are inherently local, in the sense that changes to the inputs or incremental growth in the access network have only local affects on the resulting topology or its traffic. In contrast, changes to inputs of the backbone design problem or backbone incremental growth often have global implications on the structure and/or behavior of the network as a whole.

We present here a heuristic "recipe" for generating an annotated router-level graph for a national ISP using this optimization-based approach. The intent here is to provide a minimal level of detail sufficient to illustrate how the various objectives, constraints, and tradeoffs can be synthesized into a reasonable model.

STEP 1: Choose the metropolitan areas to serve (i.e. the PoPs). This decision might be based on the population size of each metropolitan area, or it might be driven by market competition (e.g. the need to have a presence in specific cities).

STEP 2: For each metropolitan area, design an access network, as follows.

- Select the location of the access points (there may be more than one). In modeling specific metropolitan regions, the location of access points may be driven by the presence of *Internet Exchange Points (IXPs)* or *co-location facilities* that serve as central facilities for ISPs to peer with one another.
- Choose an underlying aggregation technology for each access point (e.g. fiber, DSL, broadband cable) or a mixture of them.
- Choose (probabilistic) distributions for customer geography and also for customer bandwidth demands. A reasonable approach would be to choose a mixture of technologies and customer demands that are consistent with market statistics, such as those listed in Table 6.1.
- Formulate an optimization problem (using corresponding technology constraints, link costs, and user demands) in the spirit of [10], and solve to obtain a heuristically optimal local topology.

STEP 3: Design a backbone topology to support traffic between the PoPs.

- Choose a model for traffic demand (e.g a traffic matrix) consistent with the model of local access design (i.e. aggregate PoP demands are derived from the access networks that they serve).
- Decide on the link bandwidths (e.g. OC-48, OC-192) and router models (e.g. Cisco GSR) available for use.
- Select appropriate resource budgets (a simple budget could be the total number of links, or more sophisticated models could use real costs of links and routers) as well as any additional service requirements (e.g. redundancy).
- Formulate the backbone design problem as a constrained optimization problem (in the spirit of [42, 46]). The solution need only be heuristically optimal, but it preferably also has some robustness properties (such as insensitivity to small changes in the inputs or the loss of any single node or link).

STEP 4: Compute theoretical performance of the network as a whole (e.g. throughput, utilization) under initial traffic assumptions.

STEP 5: Consider reverse engineering of real measured ISP topologies (in the spirit of [91]) as a form of heuristic validation, assessment of input assumptions, and examination of engineering details included in the model.

STEP 6: Consider traffic engineering on the generated topology to improve performance, and use improved traffic assumptions as feedback into model and return to STEP 3.

As noted by the final step, this procedure is intended to be a process by which the insights from the resulting topology are used as feedback into the process in order to improve the self-consistency and realism of the model. In the end, the inputs to the model include: distribution of customer demands, distribution of edge technologies, distribution of customer geography (population density), a list of PoPs, a backbone budget constraint, redundancy requirements, and a model of traffic demands. The outputs of the generation process include: an annotated topology with router capacities, link bandwidths, router/link utilization under assumed traffic demand, network performance, and perhaps a measure of robustness to loss of router/links. There are many additional details that could be added, and it's likely that we will need to iterate between modeling, measurement, and analysis to get it right, but this level of modeling is entirely feasible given current understanding of the Internet's most salient features.

There are several advantages of a topology generator that includes this level of detail. First and foremost, the resulting graph models should be consistent with the reality of the engineering systems that they are intended to represent. This alone would represent significant progress in moving beyond the degree-based generators in use today. However, there are a number of additional benefits that would follow. For example, a model at this level of detail provides a natural framework for the study of traffic engineering problems, and moreover it enables one to study of the co-evolving relationship between network design and traffic engineering. In addition, the conceptual framework presented here will enable the exploration of the way in which individual constraints and objectives affect the large-scale features of generated topologies. For example, in comparing the qualitative and quantitative features of real ISP topologies, are the dissimilarities the result of differences in their inputs (customer distributions, geographies) or the result of design decisions (different technologies, different redundancy requirements, budgets)? A better understanding of this "design space" also creates the ability to consider what-if scenarios, for example the potential to predict the consequences of introducing a new technology or the ability to study whether or not there are fundamental limitations in the ability of a network to scale with current technology. Finally, reasonably detailed models of real ISP topologies creates the ability to construct models for the Internet as a whole by interconnecting individual ISPs (similar to what has been suggested for measured ISP topologies [63]). It also enables the systematic investigation of the economics and dynamics of peering relationships (e.g. hot potato routing [95]) which have recently received significant attention by operators and researchers alike.

5. Optimization: A Broader Framework

In this article, we have used optimization primarily as a tool for the *analysis* of Internet topology. That is, we have assumed that the structural features observed in the router-level Internet are the result of some (implicit) optimization problem faced by network architects, and we have tried to identify the way in which fundamental technological and economic forces shape both the current structure and its ongoing development. Furthermore, we have suggested optimization-based formulations as a means for generating representative models of the router-level Internet. This framework contrasts with the traditional use of optimization for solving difficult decision problems related to network *synthesis*—that is, the design of new or incremental network infrastructure. Unlike the problem here, network synthesis problems typically start with precise mathematical formulations where the variables, objectives, and constraints are clearly defined, and the challenge in solving them typically comes from difficult combinatorics or uncertainty in the decision making. The purpose of this section is to illustrate how these two complimentary perspectives are being used to develop a broader theory of the Internet, and we also identify research areas where additional work is needed.

The modularity resulting from a layered network architecture (see Figure 3.1) provides an appropriate *separation theorem* for the engineering of individual network components. That is, it is often possible to investigate the features, behavior, and performance of an individual layer either in isolation or assuming that the lower-layer functionalities are all specified. It is within each individual layer where one observes a horizontal decomposition of functionality into decentralized components. This horizontal decomposition provides robustness to the loss of individual components, supports scalability of the network, and facilitates fully distributed provisioning, control, and management mechanisms, usually implemented through multiple layers of feedback regulation between different end systems.

This collective picture presents a framework for the analysis, evaluation, and design of network components. For the most part, the aforementioned modularity allows network engineers to focus exclusively on a specific portion of the overall architecture when addressing issues related to performance. However, changes to individual components/protocols have often unanticipated and undesirable consequences on the network as a whole or on the entire TCP/IP protocol stack, and it is for those situations when a comprehensive theory of the Internet is most needed. What follows is a brief summary of some of the progress to date in developing this theory for parts of the overall architecture.

Since each of these areas is a deep research area in its own right, the examples here are meant to be representative and not necessarily comprehensive.

5.1 Link Layer: Connectivity

Most of the discussion in this paper has assumed that the network topology defined by the physical/link layer technologies is the same as the topology at the IP (routing) layer, but this is not true for many real networks. Despite its name, the router-level graph (as inferred from measurement studies based on the *traceroute* program) does not necessarily reflect the structure of its underlying physical layer. Depending on the technologies in use, two routers that are "connected" by a single hop at the IP layer may or may not be physically connected to one another. The use of different link layer technologies, such as Ethernet at the network edge or Asynchronous Transfer Mode (ATM) technology in the network core, can give the illusion of direct connectivity from the perspective of IP, even though the routers in question may be separated at the physical level by many intermediate networking devices or even an entire network potentially spanning hundreds of miles. In some cases, network provisioning at these different layers is handled by different companies, and it's possible that an ISP who sells IP networking services actually subcontracts the provisioning of its optical network to another provider.

The separation of the link layer technologies from IP routing leads to a number of interesting and important network design problems, each of which have their own optimization-based formulation. For example, when designing optical network topologies, the question is which fiber pathways to "light up" to form the basic circuit connectivity and what to do in the event that a particular pathway is interrupted either from accident (e.g. a fiber cut) or failure (equipment malfunction). For example, recent work on the design of WDM networks has focused on cost-based solutions to the placement of optical networking equipment within a mesh-like backbone structure [17, 72]. Other work has focused on the ability to protect and restore fiber circuits at the link layer, thereby hiding such failures from the layers above [85, 86, 103, 71]. Current efforts are focusing on optimization-based approaches to the multi-layer design of both link layer structure and routing policies [47, 59, 60, 104].

5.2 IP Layer: Routing

Given the physical infrastructure, the Internet relies on IP to switch any packet anywhere in the Internet to the "correct" next hop. Addressing and routing are crucial aspects that enable IP to achieve this impressive task. Maintaining sufficient and consistent information within the network for associating the identity of the intended recipient with its location inside the network is achieved by means of *routing protocols*; that is, a set of distributed algorithms that are part of IP and that the routers run among themselves to make appropriate routing decisions. Robustness considerations that play a role in this context include randomly occurring router or link failures and restoration of failed network components or adding new components to the network.

The routing protocols in use in today's Internet are robust to these uncertainties in the network's components, and the detection of and routing around failed components remains largely invisible to the end-to-end application—the Internet sees damage and "works" (i.e., routes) around it. The complexity in protocol design that ensures this remarkable resilience to failures in the physical infrastructure of the Internet is somewhat reduced by a division of the problem into two more manageable pieces, where the division is in accordance with separation of the Internet into *Autonomous Systems (AS)* or *autonomous routing domains*: each AS runs a local internal routing protocol (or *Interior Gateway Protocol (IGP)*; e.g., Open Shortest Path First or OSPF), and between the different ASs, an inter-network routing protocol (or *Exterior Gateway Protocol (EGP)*; e.g., Border Gateway Protocol or BGP) maintains connectivity and is the glue that ties all the ASs together and ensures communication across AS boundaries. However, the de-facto standard hybrid BGP/OSPF routing protocol deployed in today's Internet is largely an engineering "solution" and little (if anything) is known about its optimality, fundamental limits, or inherent tradeoffs with respect to, changing or uncertain traffic demands [11], the design of more sophisticated algorithms for tuning OSPF weights [39], the development of semantically richer BGP routing policies [45], or the gradual deployment of new protocols such as *Multi Protocol Label Switching (MPLS)* [18] or *Multiprotocol Lambda Switching (MPIS)* [105]. Furthermore, recent work has provided a systematic investigation into how routing policies are used in real operational networks and has reported that the traditional view of interior and exterior routing policies is insufficient to capture the intention of traffic engineers [68]. Despite an incomplete understanding of routing-level behavior, current research is proposing changes to the routing infrastructure in order to facilitate the routing, switching, and forwarding of packets through next-generation networks as well as support the demands of novel applications.

5.3 Transport Layer: TCP-AQM

By assuming a given physical network infrastructure and a fixed routing matrix, recent investigation into the network transport layer has brought new understanding to the behavior of both the Transmission Control Protocol (TCP) and Active Queue Management (AQM) schemes for providing optimal allocation of network resources. The main insight from this work [57, 64, 75] is to view the TCP-AQM protocol as a distributed *primal-dual algorithm*, in which TCP source rates are viewed as primary variables, while link congestion measures are viewed as dual variables. Collectively, these two protocols solve an implicit, global utility maximization problem across the Internet. Furthermore, it has been shown that different protocol variants solve the same basic resource allocation problem, but use different utility functions [65]. This theoretical framework suggests that by studying the underlying optimization problem, it is possible to understand the equilibrium properties of a large network under TCP/AQM control. These properties include network throughput, transmission delay, queue lengths, loss probabilities, and fairness. Here, insight from this optimization-based perspective comes despite the fact that TCP and AQM were designed and implemented without regard to utility maximization. These results have subsequently been combined with models from control theory to provide additional insight into the dynamics and stability of these networking protocols, and they have even led to new proposals for transport protocols to replace TCP itself [54, 82, 58, 53].

5.4 Application Layer: Mice and Elephants

At the top of the protocol stack, applications are designed to meet particular performance objectives, and their design typically relies on all of the network resources available from the lower layers of the protocol stack. The technology supporting the Wold Wide Web (WWW) is a prime example. The behavior of the WWW is defined by the interaction of web servers and web browsers that run the HyperText Transfer Protocol (HTTP), web documents that are encoded using the HyperText Markup Language (HTML), and users who navigate via feedback through those interconnected documents. Within this scheme, HTTP relies on TCP to provide reliable packet delivery and robust flow control when transferring files. Sometimes, however, the interaction between layers results in unintended and unexpected poor performance that may be hard to identify and resolve. For example, as discussed in [99], version 1.0 of HTTP interacted badly with TCP, causing problems with the latency perceived by web users as well as problems with web server scalability. These problems were ultimately fixed in HTTP 1.1, when the protocol was tuned to provide better performance.

There has been relatively little theoretical work to formalize the relationship between application behavior and the dynamics of the underlying protocol stack. A noticeable exception is [110] that pursues a HOT-based approach to develop a toy model for Web layout design. This model suggests that the organization and layout of web pages is optimized to minimize the latency experienced by users who search for items of interest. Minimizing user-perceived latency is roughly equivalent to minimizing the average size of downloaded files and is motivated by the limitation on the bandwidth available to both the network and the user. In particular, it is highly desirable for the frequently accessed files that are used for navigational purposes to be small and download quickly (“mice”), as the user's next action awaits this information. At the same time, the large files ("elephants") that tend to represent the endpoints of the search process require in general large average bandwidth for timely delivery, but per packet latency is typically less of an issue.

This type of Web layout design problem has features similar to conventional source coding, but with substantial differences in the constraints and design degrees of freedom (e.g., grouping of objects into files, location of hyperlinks). The toy model produces distributions of file sizes and file accesses that are both heavy-tailed in nature and are in remarkable agreement with measurements of real Web traffic. That these heavy-tailed or highly variable characteristics of Web traffic are likely to be an invariant of much of the future network traffic, regardless of the applications, is one important insight to be gained from this new research direction. The current split of most traffic into mice and elephants is likely to persist. Most files will be mice, which generate little aggregate bandwidth demand, but need low latency. Most of the packets will come from elephants, which demand high average bandwidth, but can tolerate varying per packet latency. After all, much of the human-oriented communication process that involves both active navigation and ultimately the transfer of large objects can naturally be "coded" this way, with important implications for proper protocol design at the transport layer.

5.5 Discussion

These successes in the use of an optimization-based framework to gain deeper understanding of individual components within the overall architecture (either from first principles or via "reverse engineering") suggest a potential benefit for extending this perspective to deal with issues that span multiple layers. In its ultimate form, one hopes for a coherent theory in which the various protocol layers can be viewed as part of one giant optimization problem that the network solves in a completely decentralized manner, and where each protocol can be interpreted as a local algorithm that works to achieve some part of the global objective. Yet, much remains to be done before such a theory is in hand. While recent work at the transport layer suggests that the behavior of TCP-AQM is nearly optimal, is it known that the collective behavior of TCP-AQM and IP is not optimal. One recent attempt to understand the ability of IP and TCP-AQM to simultaneously solve the optimal routing problem in conjunction with the resource allocation problem has shown that utility maximization over both source rates and their routes is NP-hard and hence cannot be solved in general by shortest-path routing within IP [96].

Despite the inherent challenges for developing such a theory, the ongoing development and deployment of Internet technologies and our increasing reliance upon them demand that we obtain a deeper understanding of the relationship between the underlying architecture and design objectives such as performance, scalability, robustness, and evolvability. More broadly, whenever considering the deployment of new technologies, it would be useful to know where the current architecture stands with regard to optimality. For example, are there other routing protocols that might perform better than IP? What happens if fundamental changes are made at the routing level, such as would be seen with the deployment of new provisioning and routing technologies such as MPLS or lambda switching? What happens when new applications change fundamentally the traffic patterns that the underlying protocol stack needs to manage? What if new technologies fundamentally change the constraints or economics of network provisioning and management? And finally, the envisioned theory should be able to provide insight into the circumstances under which circuit-switching versus packet switching is the optimal thing to do.

6. Conclusion

In this paper, we have described several factors that we suggest are key drivers of the router-level Internet as it is designed, built, and operated by ISPs. While this list of key factors is far from exhaustive, what is striking is how the need to annotate network topologies with even simple domain-specific features shows how graphs that may be sensible from a connectivity-only perspective are no longer viable (e.g., non-realizable or non-sensical) in the real world because of constraints that are imposed by their application domains. In this sense, the models for the router-level Internet that result from optimization-based formulations and degree-based formulations could not be more different. Although degree-based models hold a certain appeal from a macroscopic viewpoint, they are entirely inconsistent with the perspective of network engineering. Networks constructed from degree-based models would be costly to build and would yield poor performance. Optimization-based approaches to modeling router-level topology (and the Internet more generally) hold tremendous promise, but significant work remains before we have the level of clarity that is needed by network operators, corporate managers, and telecommunications policy makers.

Acknowledgments

The ideas presented in this article have benefitted tremendously from conversations with many colleagues, including Steven Low, Ramesh Govindan, Matt Roughan, and Reiko Tanaka. The authors would like to express special thanks to Stanislav Shalunov for help in understanding the physical design of the Abilene backbone. This research was sponsored in part by the Institute for Pure and Applied Mathematics (IPAM) at UCLA, as part of their program on Large-Scale Communication Networks.

References

[1] Abilene Network. Detailed information about the objectives, organization, and development of the Abilene network are available from `http://www.internet2.edu/abilene`.

[2] L.A. Adamic and B.A. Huberman. Power-Law Distribution of the World Wide Web. *Science* 2000; 287(5461):2115.

[3] R.K. Ahuja, T.L. Magnanti, and J.B. Orlin. *Network Flows: Theory, Algorithms, and Applications*. Upper Saddle River, NJ: Prentice-Hall. 1993.

[4] W. Aiello, F. Chung, and L. Lu. A Random Graph Model for Massive Graphs. *Proceedings of the 32nd Annual Symposium in Theory of Computing*, 2000.

[5] R. Albert, H. Jeong, and A.-L. Barabási. Diameter of the World Wide Web. *Nature* 1999; 401:130–131.

[6] R. Albert, H. Jeong, and A.-L. Barabási. Attack and error tolerance of complex networks. *Nature* 2000; 406: 378–382.

[7] R. Albert, and A.-L. Barabási. Statistical Mechanics of Complex Networks. *Reviews of Modern Physics* 2002; 74:47–97.

[8] D. Alderson. Technological and Economic Drivers and Constraints in the Internet's "Last Mile". Technical Report CIT-CDS-04-004, Engineering Division, California Institute of Technology, 2004.

[9] D. Alderson, J. Doyle, R. Govindan, and W. Willinger. Toward an Optimization-Driven Framework for Designing and Generating Realistic Internet Topologies. In *ACM SIGCOMM Computer Communications Review* 2003; 33(1): 41–46.

[10] M. Andrews and L. Zhang. The access network design problem. *Proceedings of the 39th Foundations of Computer Science*, 1998.

[11] D. Applegate and E. Cohen. Making Intra-Domain Routing Robust to Changing and Uncertain Traffic Demands: Understanding Fundamental Tradeoffs. *Proceedings of ACM SIGCOMM* 2003.

[12] A. Balakrishnan, T.L. Magnanti, A. Shulman, and R.T. Wong. Models for planning capacity expansion in local access telecommunication networks. *Annals of Operations Research* 1991; 33: 239–284.

[13] A.-L. Barabási and R. Albert. Emergence of scaling in random networks. *Science* 1999; 286: 509-512.

[14] E. Bender and R. Canfield. The asymptotic number of labeled graphs with given degree sequences. *J. Combinatorial Theory Ser. A* 1978; 24: 296–307.

[15] N. Berger, B. Bollobás, C. Borgs, J. T. Chayes, and O. Riordan. Degree Distribution of the FKP Network Model. *Proceedings of ICALP* 2003; 725–738.

[16] N. Berger, C. Borgs, J.T. Chayes, R.M. D'Souza, and R.D. Kleinberg. Competition-Induced Preferential Attachment. *Proceedings of ICALP* 2004.

[17] G. Birkan, J. Kennington, E. Olinick, A. Ortynski, G. Spiride. Optimization-Based Design Strategies for DWDM Networks: Opaque versus All-Optical Networks. Tech Report 03-EMIS-01, SMU, Dallas, TX; 2001.

[18] U. Black. *MPLS and Label Switching Networks*, 2nd Edition. Prentice Hall PTR, 2002.

[19] A. Broido and K. Claffy. Internet Topology: Connectivity of IP Graphs. *Proceedings of SPIE ITCom WWW Conference* 2001.

[20] T. Bu and D. Towsley. On distinguishing Between Internet Power Law Topology Generators. *Proceedings of IEEE Infocom* 2002.

[21] K.L. Calvert, M. Doar, and E. Zegura. Modeling Internet topology. *IEEE Communications Magazine*, June 1997.

[22] J. M. Carlson and J. C. Doyle. Highly Optimized Tolerance: a mechanism for power laws in designed systems. *Physics Review E* 1999; 60:1412–1428.

[23] J.M. Carlson and J.Doyle. Complexity and Robustness. *Proceedings of the National Academy of Science* 2002; 99(Suppl. 1): 2539-2545.

[24] H. Chang, R. Govindan, S. Jamin, S. Shenker, and W. Willinger. Towards Capturing Representative AS-Level Internet Topologies. *Proceedings of ACM SIGMETRICS* 2002.

[25] H. Chang, S. Jamin, and W. Willinger. Internet Connectivity at the AS-level: An Optimization-Driven Modeling Approach *Proceedings of MoMeTools* 2003 (Extended version, Tech Report UM-CSE-475-03).

[26] Q. Chen, H. Chang, R. Govindan, S. Jamin, S. Shenker, and W. Willinger. The Origin of Power Laws in Internet Topologies Revisited. *Proceedings of IEEE INFOCOM* 2002.

[27] B. Cheswick, H. Burch, and S. Branigan. Mapping and Visualizing the Internet. *Proceedings of Usenix* 2000.

[28] Cisco Catalog. Master Contract for the State of Washington. `http://techmall.dis.wa.gov/master_contracts/intranet/routers_switches.asp`

[29] D. D. Clark. The design philosophy of the DARPA Internet protocols. *Proceedings of the ACM SIGCOMM'88*, in: *ACM Computer Communication Reviews* 1988; 18(4):106–114.

[30] M. B. Doar. A Better Model for Generating Test Networks. *Proceedings of Globecom '96*, Nov. 1996.

[31] J. Doucette and W.D. Grover. Comparison of mesh protection and restoration schemes and the dependency on graph connectivity. *3rd International Workshop on Design of Reliable Communication Networks (DRCN)* 2001; 121–128.

[32] J. C. Doyle and J. M. Carlson. Power laws, Highly Optimized Tolerance and generalized source coding. *Physics Review Letters* 2000; 84(24):5656–5659.

[33] J. Doyle, J. Carlson, S. Low, F. Paganini, G. Vinnicombe, W. Willinger, J. Hickey, P. Parilo, L. Vandenberghe. Robustness and the Internet: Theoretical Foundations, In *Robust design: A repertoire from biology, ecology, and engineering*, E. Jen, Editor, Oxford University Press (to appear).

[34] J. Duffy. "Cisco's loss is Juniper's gain." *Network World Fusion* February 18, 2003. `http://www.nwfusion.com/edge/news/2003/0218mktshare.html`.

[35] P. Erdos and A. Renyi. On random graphs I *Publ. Math. (Debrecen) 9* (1959), 290-297.

[36] P. Erdos and A. Renyi. On the evolution of random graphs. *In Publications of the Mathematical Institute of the Hungarian Academy of Sciences* 1960; 5:17–61.

[37] A. Fabrikant, E. Koutsoupias, and C. Papadimitriou. Heuristically Optimized Trade-offs: A new paradigm for Power- laws in the Internet, *Proceedings of ICALP* 2002; 110–122.

[38] M. Faloutsos, P. Faloutsos, and C. Faloutsos. On Power-Law Relationships of the Internet Topology. *Proceedings of ACM SIGCOMM* 1999.

[39] B. Fortz and M. Thorup. Internet Traffic Engineering by Optimizing OSPF Weights. *Proceedings of IEEE INFOCOM* 2000.

[40] L. Gao. On inferring autonomous system relationships in the Internet. in *Proceedings of IEEE Global Internet Symposium* 2000.

[41] B. Gavish. Topological design of telecommunication networks—local access design methods. *Annals of Operations Research* 1991; 33:17–71.

[42] B. Gendron, T.G. Crainic and A. Frangioni. Multicommodity Capacitated Network Design. In B. Sansó and P. Soriano (eds), Telecommunications Network Planning, pp. 1–29. Kluwer, Norwell, MA. 1998.

[43] C. Gkantsidis, M. Mihail, A. Saberi. Conductance and congestion in power law graphs *Proceedings of ACM Sigmetrics* 2003.

[44] R. Govindan and H. Tangmunarunkit. Heuristics for Internet Map Discovery, *Proceedings of IEEE INFOCOM* 2000.

[45] T.G. Griffin, A.D. Jaggard, and V. Ramachandran. Design Principles of Policy Languages for Path Vector Protocols. *Proceedings of ACM SIGCOMM* 2003.

[46] M. Grötschel, C. Monma, and M. Stoer. Design of Survivable Networks, in *Handbook in Operations Research and Management Science*, Volume on "Networks", 1993.

[47] H. Höller and S. Vo. A Mixed Integer Linear Programming Model for Multi-Layer SDH/WDM Networks. Presented at *Seventh Annual INFORMS Conference on Telecommunications* 2004. Boca Raton, Florida.

[48] H. Höller and S. Vo. Software Tools for a Multilayer Network Design. *Proceedings of the Fourth International Conference on Decision Support for Telecommunications and Information Society (DSITS)*, 2004. Warsaw, Poland.

[49] G. Huston. *ISP Survival Guide: Strategies for Running a Competitive ISP.* New York: John Wiley & Sons, 2000.

[50] *IEEE Communications Magazine, Survivability Issue*, August 1999.

[51] Internet2 Consortium. Internet2 NetFlow: Weekly Reports. `http://netflow.internet2.edu/weekly/`.

[52] C. Jin, Q. Chen, and S. Jamin. Inet: Internet Topology Generator. Technical Report CSE-TR443-00, Department of EECS, University of Michigan, 2000.

[53] C. Jin, D.X. Wei and S.H. Low. FAST TCP: motivation, architecture, algorithms, performance. *Proceedings of IEEE Infocom* 2004.

[54] D. Katabi, M. Handley, and C. Rohrs. Congestion Control for High Bandwidth-Delay Product Networks. *Proceedings of ACM Sigcomm* 2002.

[55] J. Kleinberg, S. R. Kumar, P. Raghavan, S. Rajagopalan and A. Tomkins. The web as a graph: Measurements, models and methods. *Proceedings of the International Conference on Combinatorics and Computing* July 1999.

[56] F.P. Kelly. Mathematical modelling of the Internet. In *Mathematics Unlimited - 2001 and Beyond*, B. Engquist and W. Schmid (eds). Berlin: Springer-Verlag 2001; 685–702.

[57] F.P. Kelly, A. Maulloo and D. Tan. Rate control in communication networks: shadow prices, proportional fairness and stability. *Journal of the Operational Research Society* 1998; 49: 237–252.

[58] T. Kelly. Scalable TCP: Improving Performance in Highspeed Wide Area Networks. *Computer Communication Review* 2003; 32(2).

[59] Kennington, J., E. Olinick, K. Lewis, A. Ortynski, G. Spiride. Robust solutions for the DWDM routing and provisioning problem: models and algorithms. *Optical Networks Magazine* 2003; 4:74–84.

[60] Kennington, J., E. Olinick, A. Ortynski, G. Spiride. Wavelength routing and assignment in a survivable WDM mesh network. *Operations Research* 2003; 51: 67–79.

[61] A. Lakhina, J.W. Byers, M. Crovella, and P. Xie. Sampling Biases in IP topology Measurements, *Proceedings of IEEE INFOCOM* 2003.

[62] L. Li, D. Alderson, W. Willinger, and J. Doyle. A First Principles Approach to Understanding Router-Level Topology. *Proceedings of ACM SIGCOMM* 2004.

[63] M. Liljenstam, J. Liu, and D.M. Nicol. Development of an Internet Backbone Topology for Large-Scale Network Simulations. *Proceedings of the 2003 Winter Simulation Conference*, S. Chick, P. J. Sánchez, D. Ferrin, and D. J. Morrice (eds).

[64] S. H. Low and D. E. Lapsley. Optimization Flow Control, I: Basic Algorithm and Convergence. *IEEE/ACM Transactions on Networking* 1999; 7(6):861–75.

[65] S. H. Low. A duality model of TCP and queue management algorithms. *IEEE/ACM Transactions on Networking* 2003.

[66] T. Luczak. Sparse random graphs with a given degree sequence. *Random Graphs, vol 2* Poznan, 1989.

[67] T.L. Magnanti and R.T. Wong. Network Design and Transportation Planning: Models and Algorithms. *Transportation Science* 1984; 18(1): 1–55.

[68] D. Maltz, G. Xie, J. Zhan, H. Zhang, G. Hjalmtysson, and A. Greenberg. Routing Design in Operational Networks: A Look from the Inside. *Proceedings of ACM SIGCOMM* 2004.

[69] A. Medina, A. Lakhina, I. Matta, and J. Byers. BRITE: An Approach to Universal Topology Generation. in *Proceedings of MASCOTS*, August 2001.

[70] A. Medina, I. Matta, and J. Byers. On the Origin of Power Laws in Internet Topologies. *ACM SIGCOMM Computer Communications Review* 2000; 30(2).

[71] B. Melián, M. Laguna and J. A. Moreno-Pérez. Capacity Expansion of Fiber Optic Networks with WDM Systems: Problem Formulation and Comparative Analysis. *Computers and Operations Research* 2003; 31(3): 461–472.

[72] B. Melián, M. Laguna and J. A. Moreno-Pérez. Minimizing the Cost of Placing and Sizing Wavelength Division Multiplexing and Optical Cross-Connect Equipment in a Telecommunications Network. July 2004. Submitted for publication. Available from `http://leeds.colorado.edu/faculty/laguna/publications.htm`

[73] M. Minoux. Network Synthesis and Optimum Network Design Problems: Models, Solution Methods and Applications. *Networks* 1989; 19: 313-360.

[74] M. Mitzenmacher. A Brief History of Generative Models for Power Law and Lognormal Distributions. *Internet Mathematics* 2004; 1(2).

[75] J. Mo and J. Walrand. Fair end-to-end window-based congestion control. *IEEE/ACM Transactions on Networking* 2000.

[76] G. Mohan and C.S.R. Murthy. Light-path restoration in WDM optical networks. *IEEE Network*, pp.24-32, November/December 2000.

[77] M. Molloy and B. Reed. A Critical Point For Random Graphs With A Given Degree Sequence, *Random Structures and Algorithms* 1995; 6:161-180.

[78] Computer Science and Telecommunications Board (CSTB), National Research Council. *The Internet's Coming of Age.* National Academy Press, Washington, D.C., 2001.

[79] M.E.J. Newman. Assortative Mixing in Networks. *Phys. Rev. Lett.* 2002; 89(208701).

[80] D. Moore, C. Shannon, G. Voelker, and S. Savage. Internet Quarantine: Requirements for Containing Self-Propagating Code. *Proceedings of IEEE Infocom* 2003.

[81] A.M. Odlyzko. Internet traffic growth: Sources and implications, in *Optical Transmission Systems and Equipment for WDM Networking II*, B. B. Dingel, W. Weiershausen, A. K. Dutta, and K.-I. Sato, eds., Proc. SPIE, 5247:1-15, 2003.

[82] F. Paganini, Z. Wang, S. H. Low and J. C. Doyle. A new TCP/AQM for stable operation in fast networks. *Proceedings of IEEE Infocom* 2003.

[83] C. R. Palmer and J. G. Steffan. Generating network topologies that obey power laws. *Proceedings of GLOBECOM* 2000.

[84] President's Commission on Critical Infrastructure Protection. *Critical Foundations.* Technical report, The White House, 1997.

[85] S. Ramamurthy and B. Mukherjee. Survivable WDM mesh networks, Part I - Protection. *Proceedings of IEEE Infocom* 1999; 744-751.

[86] S. Ramamurthy and B. Mukherjee. Survivable WDM mesh networks, Part II - Restoration. *Proceedings of IEEE International Conference on Communications (ICC)* 1999; 2023–2030.

[87] Route Views. *University of Oregon Route Views Project*, Available at http://www.antc.uoregon.edu/route-views/.

[88] G. Siganos, M. Faloutsos, P. Faloutsos, C. Faloutsos. Power laws and the AS-level internet topology. *IEEE/ACM Transactions on Networking* 2003; 11(4): 514–524.

[89] Cooperative Association for Internet Data Analysis (CAIDA). Skitter. Available at `http://www.caida.org/tools/measurement/skitter/`.

[90] N. Spring, R. Mahajan, and D. Wetherall. Measuring ISP Topologies with Rocketfuel. *Proceedings of ACM SIGCOMM* 2002.

[91] N. Spring, D. Wetherall, and T. Anderson. Reverse-Engineering the Internet. *Proceedings of ACM Workshop on Hot Topics in Networking (HotNets-II)*. November 2003.

[92] L. Subramanian, S. Agarwal, J. Rexford, and R. Katz. Characterizing the Internet Hierarchy from Multiple Vantage Points. *Proceedings of IEEE INFOCOM* 2002.

[93] SURFnet Press Release. "SURFnet Builds Advanced Research Network Based on Solutions, Services from Nortel Networks, Avici Systems, Telindus". Utrecht, 24 March 2004. Available at `http://www.gigaport.nl/publicaties/pers/en_pers240304.html`.

[94] H. Tangmunarunkit, R. Govindan, S. Jamin, S. Shenker, and W. Willinger. Network Topology Generators: Degree-Based vs. Structural. *Proceedings of ACM SIGCOMM* 2002.

[95] R. Teixeira, A. Shaikh, T. Griffin, and G.M. Voelker. Network Sensitivity to Hot-Potato Disruptions. *Proceedings of ACM SIGCOMM* 2004.

[96] J. Wang and L. Li and S. H. Low and J. C. Doyle. Can TCP and shortest-path routing maximize utility? *Proceedings of IEEE Infocom* 2003.

[97] B.M. Waxman. Routing of multipoint connections. *IEEE Journal of Selected Areas in Communication*, 1988; 6(9).

[98] The White House. *The National Strategy to Secure Cyberspace.* February 2003.

[99] W. Willinger and J. C. Doyle. Robustness and the Internet: Design and Evolution. In *Robust design: A Repertoire of Biological, Ecological, and Engineering Case Studies*, E. Jen, Editor, Oxford University Press (to appear).

[100] W. Willinger, R. Govindan, S. Jamin, V. Paxson and S. Shenker. Scaling Phenomena in the Internet: Critically examining Criticality *Proceedings of the National Academy of Science* 2002; 99(1):2573–2580.

[101] K.Wu and A. Liu. The Rearrangement Inequality. `http://matholymp.com/TUTORIALS/Rear.pdf`

[102] S.-H. Yook, H. Jeong, and A.-L. Barabási. Modeling the Internet's large-scale topology, *PNAS Proceedings of the National Academy of Science* 2002; 99:13382-13386.

[103] H. Zang. WDM Mesh Networks: Management and Survivability. Kluwer Publishing, December 2002.

[104] H. Zang, J. Jue, and B. Mukherjee. A review of routing and wavelength assignment approaches for wavelength-routed optical WDM networks. *Optical Networks Magazine* 2000; 1:47-60.

[105] H. Zang, J.P. Jue, L. Sahasrabuddhe, R. Ramamurthy, and B. Mukherjee. Dynamic lightpath establishment in wavelength routed networks. *IEEE Communications Magazine* 2001; 39(9):100–108.

[106] E. W. Zegura, K. L. Calvert, and S. Bhattacharjee. How to model an internetwork. *Proceedings of INFOCOM* 1996.

[107] E. Zegura, K.L. Calvert, and M.J. Donahoo, A quantitative comparison of graph-based models for Internet topology. *IEEE/ACM Transactions on Networking* 1997; 5(6).

[108] Y. Zhang, M. Roughan, C. Lund and D. Donoho. An information-theoretic approach to traffic matrix estimation, Proc. of ACM Sigcomm, Karlsruhe, Germany, 2003. In *Computer Communication Review* 2003; 33(4).

[109] D. Zhou and S. Subramaniam. Survivability in Optical Networks. *IEEE Network*, pp.16-23, November/December 2000.

[110] X. Zhu, J. Yu and J.C. Doyle. Heavy Tails, Generalized Coding, and Optimal Web Layout. *Proceedings of IEEE Infocom* 2001.