The Internet consists of thousands of autonomous systems (ASes) that voluntarily form bilateral (sometimes conditional) interconnection agreements to provide end-to-end reachability. These interactions between networks are local, without centralized control or regulation, but they often have global impact on the performance and profitability of network and service providers. Despite much recent interest in the economic aspects of the Internet, such as network interconnection (peering), pricing, performance, and the profitability of various network types, two historical developments contribute to a persistent disconnect between economic models and actual operational practices on the Internet. First, the Internet became too complex – in traffic dynamics, topology, and economics – for currently available analytical tools to allow realistic modeling. Second, the data needed to parameterize more realistic models is simply not available. The problem is fundamental, and familiar: simple models are not valid, and complex models cannot be validated. We propose *transformative progress* in both dimensions: creating more powerful, empirically parameterized computational tools, and enabling broader validation than previously possible. We will use measurements of interdomain traffic, topology dynamics, routing policies and peering practices as input to ITER [34], a detailed model of AS interconnection. With ITER, we take an approach that is *computational rather than analytical*. ITER takes as input the interconnection policies of various network types, interdomain traffic demands, routing policies, geographical constraints, and pricing/cost factors, and computes an *equilibrium* – a state where no network has the incentive to change its connectivity. We propose a two-pronged approach to validating ITER. First, we will verify that ITER can reproduce known macroscopic properties of the Internet AS topology. Second, we will use historical, publicly available financial and topological data to verify that ITER can reproduce known trends in the evolution of the Internet. We will then use ITER to study various interconnection practices, the stability and dynamics of interdomain links, and economic properties of the resulting equilibrium.

Intellectual merit: We propose an approach grounded in empirical measurements of macroscopic Internet topology, traffic demand, routing policies, and peering policies. First, we will study the evolution of transit and settlement-free links of ASes, attempting to infer the economic incentives and policies of ASes that underlie topology dynamics. We will extend our previous work on the evolution of transit links over the last 10 years by studying the evolution of settlement-free peering links. Second, we will use flow-level traffic measurements collected at multiple vantage points to infer properties of the interdomain traffic matrix. Third, we will use a combination of active and passive measurements to study the economic implications of routing policies as well as peering policies used by ASes. The data promise to reveal important, and thus far elusive, insights into the economic implications of topology dynamics, interdomain traffic characteristics, and routing policy, but they will also inform the parameterization of ITER, our proposed model of AS interconnection and dynamics. We will validate ITER using publicly available historical financial data, BGP data, and information about the changing business roles and strategies of various networks over time. As an application of ITER, we will simulate several plausible "what-if" scenarios relating to traffic, peering strategies, geography, and peering/cost structures and explore the effects of these factors on macroscopic stability and performance characteristics.

Broader impact: The proposed research will yield not only deeper, empirically grounded interpretation of available data on the most opaque sub-discipline of network research – internetwork economics – but also a broad understanding of how economic forces induce, as well as result from, topology dynamics and architectural evolution. We have structured our computational approach to enable new **transformative research** in complex network modeling, including broader issues such as pricing and policy. The educational side of the project will integrate Internet economics in two Georgia Tech courses, while a PostDoc and a PhD student will graduate as experts in that emerging and rather sparsely populated area. Further, our data and methods will be publicly available and regularly presented to both the research community as well as operator and policy forums, e.g., NANOG, FCC.

Keywords: Internet Economics, Interconnection, Peering, Network Formation, Computational methods, Measurement

Contents

1 Introduction

The Internet consists of thousands of autonomous but interconnected networks of different types and business objectives. In the Internet ecosystem, traffic flow is (sometimes loosely) coupled with financial exchanges, and routing and peering policies are often dictated by economic and strategic objectives. The continual emergence of new applications, pricing schemes, protocol and architectural innovations (to some, "hacks") renders the Internet ecosystem far too intricate for simple models to capture the complexity of network interactions. Most of these interactions are local in nature, without coordinated control or regulation, but they often have global impact affecting the performance and reliability experienced by users, the financial viability of network and service providers, and to some degree the global economy.

The high-level objective of this research is to create a scientific basis for modeling interdomain interconnection and dynamics. Specifically, we aim to understand the structure and dynamics of the Internet ecosystem from an economic perspective, capturing the interactions between network business relations, internetwork topology, routing policies, and the resulting interdomain traffic flow. Though there is a large body of work on economic aspects of the Internet such as pricing and interconnection mechanisms, previous models face two problems. First, the models of a more analytical flavor do not capture most operational realities in the real Internet. Second, and more importantly, few models are parameterized *or validated* using measurements drawn from the real world. There is thus a disconnect between operational realities and economic models of the Internet, because there has been no way to either corroborate or refute the conclusions drawn from previous models. Economically relevant data about interdomain traffic characteristics, routing and peering policies and pricing/cost structures has simply not been available. Empirically parameterized models that capture most real-world intricacies of network interconnection are missing from the literature.

We propose a computational, empirically parameterized model of interdomain connectivity, which allows us to incorporate as much realism as possible. As part of ongoing work, we have developed ITER, a model for interdomain connectivity that includes several components – economics, interdomain traffic flow, geography, routing policies, and provider/peer selection policies of different network types – that play a crucial role in interdomain economics, but have received limited research attention (due to a lack of empirical data to parameterize them). Given an initial topology, interdomain traffic flow, routing policies, and provider/peer selection strategies for each network, ITER iteratively computes an equilibrium – a situation where no node has the incentive to unilaterally change its connectivity. The key difference from previous modeling efforts is that we will parameterize each component of ITER using real-world measurements. To this end, we first propose to characterize a realistic *interdomain traffic matrix*, which determines how much traffic two networks exchange with each other. We propose to use direct measurements of traffic from multiple vantage points to characterize the interdomain traffic matrix. Given an interdomain traffic matrix, ISP routing policies determine how that traffic is actually routed. We propose to use active and passive measurements to characterize ISP routing policies, focusing on the economic implications of routing policy heuristics such as "prefer customer, then prefer peer", or selective advertisement of prefixes by multihomed customer networks. Third, we propose to study the topology dynamics of transit and peering links for a set of ASes, aiming to infer the economic objectives that underlie such dynamics. Finally, we will use resources such as peeringDB [4] to study the peering practices of different network types, in terms of their geographical presence and advertised peering policy. We will validate the empirically parameterized ITER by testing its ability to match known macroscopic properties of the Internet, such as degree distribution, clustering, and hierarchical structure. To explore ITER's predictive power, we will test whether it can reproduce known historical trends in the evolution of the Internet ecosystem, given historical data about the financial performance and topological characteristics of specific networks. Finally, we will investigate several "what-if" scenarios that represent plausible evolution paths in the future of the Internet, studying topological and economic properties of the internetwork at equilibrium.

2 Related Work

There is prior work related to each measurement and modeling task we propose. We present a survey of each thread: traffic characterization, routing/peering policy characterization, topology measurement, and network formation/evolution models.

Despite the importance of understanding inter-AS traffic characteristics, there is little previous work in this area, largely because of a lack of available traffic data. Prior traffic matrix work has relied on coarse estimation [25, 40, 43] with limited ability to validate. We must also rely on estimation to characterize inter-AS traffic, but will base our estimates on *direct measurements* of traffic data (continuous flow statistics rather than short packet traces) collected from multiple vantage points, including edge networks (e.g., Georgia Tech, UCSD) and national backbones that peer with commercial providers, (e.g., Internet2). Attached letters of collaboration express commitments to provide data access.

While studying (and facilitating) operational peering policies of networks, Norton published a series of white papers on the economics of underlying strategies for peering and transit connectivity [73, 74], and evolving trends in the peering ecosystem [75, 77]. recent work has conjectured on aggressive peering by content providers at IXPs [60] and geographical expansion by large content providers [48]. We propose a quantitative study of the peering policies and geographical expansion trends of ASes, using online resources such as peeringDB [4], where networks self-report their business type, peering policy, and presence at IXPs. Prior work on interdomain routing policies studied possible instabilities due to ISP contracts, and designed mechanisms to guarantee stability [41, 42, 46, 49, 50, 51, 45, 64], without much focus on measuring operational routing policies. Six years ago Wang et al. [91] studied ISP routing policies using BGP data, discovering that customers often announce prefixes selectively to their providers. We propose to extend that work by incorporating active traceroute measurements, and explore the implications of ISP routing policies such as "prefer-customer, then prefer peer" and selective prefix advertisement by multihomed customers on interdomain economics.

After the 1999 study by Faloutsos et al. [39], a long sequence of papers investigated both the static topological characteristics of the AS-level Internet topology [29, 24, 28, 31, 53, 70, 80, 97], and the evolution and dynamics of the AS topology [63, 69, 81, 88]. Using a decade of historical BGP data, we recently published the longest evolutionary study of the Internet AS topology; we also classified ASes into different types based on their business function and analyzed trends such as multihoming behavior and provider "attractiveness/repulsiveness" [33]. Our proposed research extends this modest but unprecedented historical analysis, aiming to faithfully capture the more subtle but economically pivotal dynamics of settlement-free peering links.

Several topology generation models, based on the "Preferential Attachment (PA)" principle, match observed macroscopic properties of the Internet topology, such as degree distribution or clustering [6, 14, 18, 82, 96, 98]. More recent work extended the PA model to incorporate the economic incentives of networks, most notably the work of Serrano et al. [85], Wang et al. [92] and Shakkotai et al. [86]. The models in this research thread have been *descriptive* in nature, meaning that the models create topologies that match observed graph properties, but they do not explain the dynamics in the Internet topology, how those dynamics are influenced by the economic incentives of ASes, or the economic properties of the resulting networks.

The desire to explain, rather than describe, topology dynamics and evolution led to models that view the Internet topology as the effect of optimization-driven activity by individual ASes [22, 38]. Chang et al. [23, 26, 27] applied these concepts in the context of the Internet's AS-level topology, modeling the peering decisions that an AS must make in its role as a customer and as a peer, and accounting for the effects of geography, interdomain traffic, and economics. Though Chang et al. investigated various "what-if" scenarios with different traffic models and peering strategies, they focused mainly on the topological properties of the resulting network (the presence or absence of power-law degree distributions). In our work, we will investigate a wider range of what-if scenarios arising from the interconnection strategies of various network types, interdomain traffic demands, and changing pricing/cost structures. Further, we will also study the economic properties of the resulting network, such as the financial well-being of individual players, social welfare, and economic efficiency.

Modeling network formation is an emerging area in economics [56]. The basic idea in such games is that the nodes of a network (including social networks) make utility-maximizing decisions about the links they will create. The links may be directed (a node can initiate a link on its own) or undirected (two nodes must reach consensus about the creation of a link) [12, 57]. Bala and Goyal [13] developed a model of non-cooperative network formation where individuals incur costs and benefits from forming and maintaining links with other agents. Recent extensions of that model [16] also consider the choice of behavior in an (anti-)coordination game. Bloch and Jackson [15] presented a network formation game where players bargain by promising or demanding transfer payments when forming links. Anshelevich et al. [8, 9] and Johari et al. [58] used network formation games to study the formation of the AS graph, where ASes incur costs for routing traffic and for a lack of end-to-end connectivity. Other research has also studied the dynamics of the network formation process [10, 93, 94], focusing on the conditions under which the network converges to "good" (in terms of social welfare) equilibria. Although competitive game theory is the most widely used tool in this area, a series of papers by Ma et al. [66, 67, 68] advocate the use of cooperative game theory, in particular, the use of the Shapley value for revenue distribution between ISPs. They show that if profits were shared according to the Shapley value, then the set of desirable "fair" properties inherent in the Shapley solution exist, and the selfish behavior of ISPs leads to globally optimal routing and interconnecting decisions. Shakkottai and Srikant [87] examined price strategies for competing ISPs in the same geographical region, and used a multi-stage game to study the interactions between regional ISPs and global ISPs. He and Walrand [52] proposed a fair revenue-sharing policy that encourages collaboration between ISPs and that can produce higher profits. Laskowski and Chuang [62] proposed network monitoring and contracting systems as a way to avoid commoditization and the resulting price wars between ISPs. Economides [37] and Laffont [61] discussed pricing, competition, and interconnection strategies by backbone providers, focusing mostly on the incentives for providers to engage in anti-competitive behavior. Odlyzko [78, 79] uses backbone traffic growth trends to conjecture about the profitability of service providers, without dealing with issues of interconnection.

Agent-based computational economics is another approach to studying network formation that relies on computational simulations to model economic incentives in decentralized systems that are too complex to model analytically [7, 89, 90, 95]. The complexity and range of Internet peering behavior make this approach particularly appealing to capture the strategic incentives of agents (ASes) in the Internet ecosystem. Holme et al. [54] used the agent-based approach to model networks as individual agents that attempt to maximize their economic fitness. Corbo et al. [32] proposed a model in which new networks connect to existing nodes in such a way as to maximize their economic utility.

Although this wide range of network formation modeling research provides insights into the presence and stability of equilibria in network formation games, the models place little, if any, emphasis on the effects of various real-world interconnection strategies used by networks, changing traffic conditions, pricing/cost structures and geographical constraints. A related but more scientifically problematic issue with these studies is that they have not been validated using real data, other than their ability to reproduce coarse statistical properties of the Internet AS topology such as heavy-tailed degree distributions or clustering. Our proposed model of network formation is far too complex for game-theoretic analysis, but this complexity allows us to capture the operational realities of provider and peer selection in the Internet, parameterize our model based on real-world measurement data, and validate it against measured historical trends using financial and topological data. We believe that our computational, empirically parameterized model will allow us to corroborate or refute at least the qualitative conclusions derived in previous work, and yield additional insights into interdomain economics that cannot be obtained from simpler models.

Finally, there is a rich literature (see [36] and [71] for surveys) in the broader field of network economics, where "network" may refer to transportation, manufacturing, energy, or communication networks. Research in this area mainly studies the economics of systems that show *network externalities*, focusing on price setting and competition in such markets and the effects on social welfare. Though this body of work is useful and broadly applicable, the Internet's complex economic evolution has demonstrated the need for domain-specific models and empirical data to parameterize them.

To summarize, we propose a research agenda that differs in several ways from prior work, most notably that it is *data-driven*, relying on real-world measurement data to parameterize a detailed computational model of network interconnection and dynamics. We will also pursue creative approaches to validating our model, leveraging the use of historical, publicly available topological and financial data.

3 Characterizing Interdomain Traffic

Interdomain traffic characteristics are an important input to our model for interdomain interconnection and dynamics described in § 6. First, an estimate of the interdomain traffic matrix will inform the parameterization of the traffic component of our model. Second, we will use measured traffic characteristics to evaluate different "what-if" scenarios that arise from changes to interdomain traffic characteristics.

Transit pricing in the Internet is typically volume-based, where a transit provider charges a customer network based on the aggregate traffic volume exchanged by the two networks. This aggregate traffic depends on the *interdomain traffic matrix T* (where T_{ij} represents the volume of traffic sourced by network i destined to network j), and interdomain routing policies which determine how this traffic is routed. Research on Internet AS topology dynamics often relies on an estimate of the overall interdomain traffic matrix [25, 27]. Unfortunately, there is little knowledge of the global Internet interdomain traffic matrix, even a snapshot much less its dynamics. But even without precise matrix values, we can extract insights from *qualitative properties* observed from different vantage points. We propose to extract qualitative properties about the Internet interdomain traffic matrix using passive flow-level measurements from participating networks. For example, using netflow data from an edge (non-transit) network E (e.g., UCSD or Georgia Tech), we can estimate one row and one column of the overall traffic matrix (how much traffic E sends to and receives from every other network). We can then study properties such as the presence, stability, and correlations of heavy sources or destinations over time. Further, using AS traffic matrix from a backbone network B such as Internet2, we can determine the traffic exchanged by ASes whose interdomain routes cross B. Finally, by monitoring interdomain traffic statistics over time, we will be able to study the evolution of the interdomain traffic matrix. Attached letters of collaboration from UCSD and Internet2 express commitments to provide access to netflow data.

3.1 Proposed research questions

1. What are the qualitative properties of the interdomain traffic matrix from the perspective of a network N, i.e., the traffic sent (received) by N to (from) each other AS? Is the distribution of inter-AS traffic highly non-uniform per older estimates [40], or has it evolved differently? Do a small number of "heavy-hitter" sources (destinations) account for a large fraction of incoming (outgoing) traffic?
2. How stable are the properties of the traffic matrix from the perspective of network N? For example, does the distribution of incoming (outgoing) traffic show the same qualitative properties over time? Are the heavy hitter sources (destinations) the same when measured over time? Do we see time-of-day or day-of-week effects in the total incoming or outgoing traffic?
3. We will study the properties of the top sources and destinations of traffic from the perspective of a network N. For example, are some of the top sources of traffic also the top destinations? If this is true, it will point to the presence of certain ASes that are heavy hitters with respect to both sent and received traffic. On the

other hand, do we see that the top sources are relatively smaller destinations (and vice versa)? This would indicate an inherent asymmetry in the nature of interdomain traffic.

4. How does the total traffic volume, distribution of incoming (outgoing) traffic per AS, and identity of heavy hitter sources and destinations differ when observed at different vantage points? Do these properties depend on the network type of the vantage point (university, enterprise, or content provider network)?
5. How has the interdomain traffic matrix been evolving over time? Is the heavy-hitter effect in sources and destinations getting stronger (a smaller set of networks accounting for the majority of incoming/outgoing traffic)? A recent study [60] reported that Internet traffic is consolidating, where few large content providers now account for a large fraction of the traffic consumed by most networks. By collecting traffic matrix data over time, we will study the evolution of the heavy hitter sources over time.
6. For a given pair of *neighbor* networks, what is the relation between the 95th percentile traffic volume and the average? How does it differ according to network type? How has this relation changed over time, e.g., as a result of increasing distribution of video [77]?

We will use various properties of the traffic matrix from the perspective of different networks to parameterize traffic parameters in our model, e.g., traffic sourced by or destined to a set of top sources/destinations. These empirical results will also allow us to devise several plausible "what-if" scenarios to investigate with our model, e.g., changing the volume of traffic consumed by edge networks, the fraction of traffic accounted for by heavy hitters, or the relation between 95th percentile and average traffic volume.

3.2 Preliminary Results

We have access to netflow data from the border router at Georgia Tech. We present results using data from each day in the week from Oct 05, 2009 to Oct 11, 2009. We measured the traffic to (from) each destination (source) AS, aggregated over the entire week. Figure 1(a) shows the inbound and outbound traffic per source (destination) AS as a function of the rank. Both curves are approximately linear on a log-log scale, at least for the top 1000 ASes. For inbound traffic from the top 1000 ASes, a Zipf distribution fit with an exponent -1.08 had a correlation coefficient of 0.99. The top 1% of source ASes accounted for close to 90% of incoming traffic; the top 10% of source ASes accounted for more than 99%. For outbound traffic to the top 1000 ASes, a Zipf distribution fit with exponent -0.98 had a correlation coefficient of 0.95. The top 1% of destination ASes accounted for about 85% of outgoing traffic; the top 10% of destination ASes accounted for more than 99%. Both distributions deviated from Zipf on the tail, for ASes that send (receive) small traffic volumes.

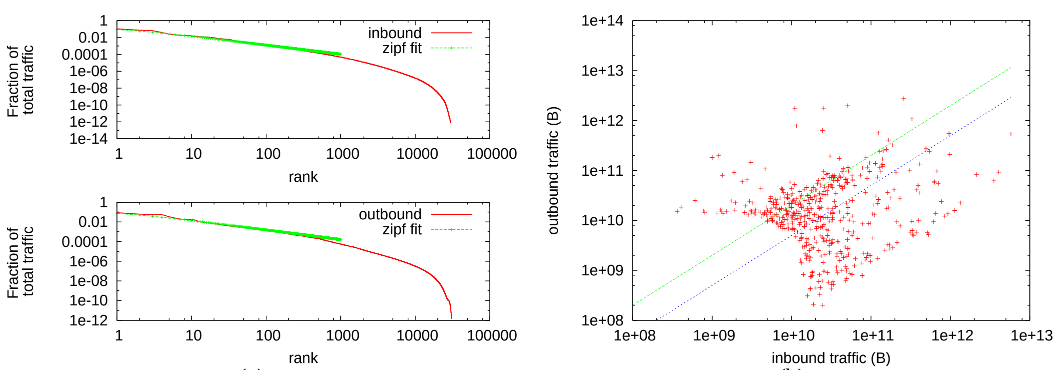


Figure 1: Subfigure (a) shows the inbound and outbound traffic per AS vs rank. Subfigure (b) shows the relation between inbound and outbound traffic for the top 500 ASes.

Figure 1(b) plots the inbound and outbound traffic volumes for the top 500 ASes sending and receiving to/from Georgia Tech. The green and blue diagonal lines distinguish ASes that are “mostly receivers”, i.e., Georgia Tech sends at least twice as much traffic to these networks as it receives from them (above the green line), “mostly senders”, i.e., vice-versa (below the blue line) and “roughly symmetric” (in between the lines). We found that the “mostly receivers” or “mostly senders” categories account for 68% of the top 500 ASes and 82% of total traffic, indicating a strong asymmetry in the sending and receiving behavior of remote ASes. The “mostly senders” category accounts for the largest fraction of ASes (41%), and just over half the total traffic (56%), not surprising for a campus of content-hungry students. The “mostly receivers” category likely consists of ASes with users that download data from Georgia Tech webservers and Linux mirrors, while ASes in the “symmetric” category may exchange a large amount of P2P traffic with Georgia Tech. We will investigate these hypotheses as well as data from other institutions that have recently expressed interest in sharing data under CAIDA’s recently introduced privacy-respecting data-sharing framework [59].

4 Inferring the Economic Incentives Underlying Topology Dynamics

Economic objectives can influence the interconnection policies used by ASes, e.g., some networks may choose the cheapest upstream providers or engage in restrictive peering to force other networks to become their customers. These policies significantly affect the structure and dynamics of the AS-level Internet topology, and the financial well-being of various players in this ecosystem. Significant progress in understanding the structure and evolution of autonomous network interconnection has occurred this decade [6, 14, 18, 85, 86, 92, 96, 98], but there is acknowledgment in the research community that publicly available data mostly captures transit links, missing many of the pervasive peering links [28, 31, 53, 70, 80, 97]. Consequently, previous evolutionary studies have focused on transit links, with only limited conjectures about the dynamics of peering links [81, 33]. Previous work could not study how topology dynamics relates to the economics of AS interconnection, because to fully understand the economics, it is necessary to study the dynamics of both transit and peering (settlement-free and paid peering) links, and data about the latter has been elusive. We propose to study the static and dynamic properties of the *complete interdomain connectivity* of ASes, focusing on the economic implications of the measured topological characteristics. We will extend our previous work on studying the evolution of transit links in the Internet [33], by also studying settlement-free and paid peering links. In addition to illuminating the interactions between economics and topology dynamics, the results of this research will allow us to parameterize and validate our model (§ 6).

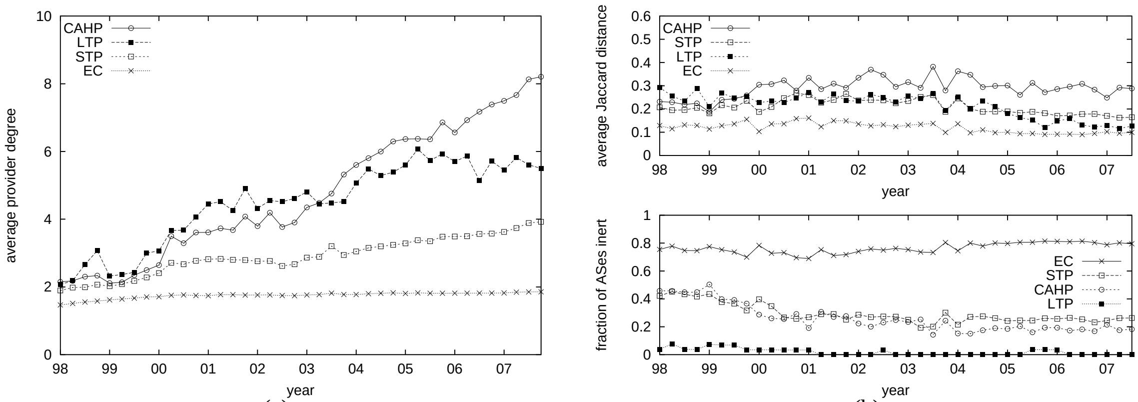
In preliminary work, we have investigated how to detect the complete interdomain connectivity of an AS, including all peering links, as long as that AS acts as a BGP monitor, i.e., it provides a BGP feed to “collectors” from Routeviews [84] and RIPE [83]. We found it possible to observe the complete set of interdomain links (including peering) for contributing ASes that export their full routing tables to Routeviews and RIPE collectors. To determine whether a contributing AS reveals all its links to Routeviews/RIPE (i.e., “a full peer of RouteViews/RIPE”), we are currently using a heuristic based on comparing the number of links of X as seen directly from X with the number of links of X as seen from other Routeviews/RIPE peers, similar to the *semi-global* concept introduced by Broido et al. [17]. To study the dynamics of transit and peering links, we will build on this heuristic and the best known AS relationship inference algorithm [35], in conjunction with additional meta-information about Routeviews/RIPE peers to more accurately classify the business relationships between full peers and their neighbors.

4.1 Proposed research questions

1. How do distinct types of networks differ with respect to topological properties – the number of providers, settlement-free peers and customers?
2. How does the typical duration of peering links differ from transit links? Do peering relationship lifetimes depend on the types of peering networks?

3. How often does a link between the same pair of networks transition between customer-provider and peering? Does the transition frequency depend on the business type of the involved networks?
4. What can we learn about the economic incentives of ASes by studying their topology dynamics? Is an increase in the number of a network's peers accompanied by a decrease in its transit provider count? Does an increase in an AS's customer count incur more aggressive peering?
5. How have topological properties of different types of networks changed over time? Do content or access providers now peer more aggressively, and if so, what are the implications on traffic flow?

We will use the results pertaining to topological properties, frequency of link updates, and link transitions to parameterize our model, in particular configuring the number of providers/peers for different types of networks, the frequency of link evaluation, and whether links can transition between transit and peering. We will use historical trends in peering behavior and dynamics to validate our model's predictive properties, e.g., whether the model can predict increasing peering between content and access providers.



(a) (b)

Figure 2: (a) Evolution of average number of providers for different AS types; (b) peering activity (measured by Jaccard distance) and inactivity (networks that choose no change in peering) for each AS type. The Jaccard Distance quantifies the difference between a set of links E_1 and E_2 as $J(E_1, E_2) = \frac{|(E_1 - E_2) \cup (E_2 - E_1)|}{|E_1 \cup E_2|}$

4.2 Preliminary results

Figure 2 depicts some results from our previous study of AS topology dynamics [33]. We classified Internet ASes into various types – enterprise customers (EC), small and large transit providers (STP and LTP), and content/access/hosting providers (CAHP) – based on their business function, and studied the evolution of the transit connectivity of these networks over the last decade. Figure 2(a) shows the average number of transit providers for different types of networks over time. We found that ECs have not significantly increased their provider degrees over time. On the other hand, transit providers and CAHPs have shown a sharp growth in the number of transit providers over time. We expect CAHPs to have several providers, as CAHPs need reliable and highly available connectivity to the rest of the Internet. We found that STPs and LTPs too have been connecting to a larger number of providers over time. Figure 2(b) shows how active different classes of networks are in changing their upstream connectivity. The top graph shows the average Jaccard distance over every pair of snapshots (a high value of the Jaccard distance implies that most of the links of a network change between two successive snapshots) for each type of network over time. The bottom graph shows the fraction of networks of each type that are "inert", meaning that they make no change to their set of links between two successive snapshots. We find that ECs are the least active and most likely to be inert, while CAHPs have the most active upstream peering dynamics. We propose to extend this analysis, measuring changes in the peers of each network, and links that shift from transit to peering and vice-versa.

5 The Economics of Interdomain Routing and Peering

Most models of interdomain routing assume the "no valley, prefer-customer, then prefer peer" heuristic [19, 44, 47]. The no-valley policy implies that traffic that enters a network from a provider cannot exit through another provider. Also, traffic that enters a network from a peer cannot exit through another peer. Further, an ISP X prefers to route traffic via a customer over peers, and prefers to route through peers over providers (if all these routes are available). This heuristic assumes two conditions about the peering ecosystem: 1) that economics drive an ISP's routing policy, and 2) customer links generate revenue, peering links are free, and transit links incur costs. Whether or not ISPs actually follow this heuristic depends on transit and peering cost structures, and the nature of transit pricing contracts, such as the *direction* of traffic for which an ISP charges its customers. Further, the extent to which a network X has peer routes available depends on the peering policies of X and other networks. For example, networks use peering policies such as "restrictive peering", "peering based on traffic ratios", and "open peering" in practice [27, 73, 74]. We are not aware of a study that characterizes the peering policies of different types of networks, and how they depend on the traffic volume, traffic mix, and business roles of those networks. The traffic volume and traffic mix determine respectively the size of a network and its content/eyeball nature, both of which are important criteria for peering. The business strategy of an AS can also affect its peering policy, e.g., the peering policy of a transit provider should balance its need to minimize (the cost of) its own upstream transit traffic with its need to attract other networks as customers rather than settlement-free peers.

We propose a measurement-based characterization of interdomain routing and peering policies, focusing on the implications of these policies on interdomain economics. First, we will study, using a combination of active and passive measurements, the use of the "prefer-customer, then prefer-peer" rule of thumb by ISPs. Second, we will study the peering policies of networks using online resources such as peeringDB [4], Packet Clearing House [3] and peeringMatrix [5] and Norton's research [75, 77, 72].

Routing policy

We will use both active and passive measurements to characterize routing policies. Using AS relationship datasets from CAIDA [20], we can compute the set of networks $\mathcal{C}_X$ that can be reached from AS X using customer links, and the set of networks $\mathcal{R}_X$ that can be reached from X using peering links. Our goal is to discover how X routes traffic destined to networks in $\mathcal{C}_X$ and $\mathcal{R}_X$. If X provides a BGP feed to Routeviews [84] and RIPE [83], then we can obtain the next hop AS from X toward networks in $\mathcal{C}_X$ and $\mathcal{R}_X$. If AS X provides a public traceroute server, then we will also use active traceroute measurements to probe destinations in $\mathcal{C}_X$ and $\mathcal{R}_X$. From the traceroute path, it is straightforward to obtain the next hop AS used by X. After discovering the next hop from AS X, we use the AS relationship dataset [20] to infer the relation between X and the next hop. Using this approach, we can discover "anomalies", wherein X reaches networks in $\mathcal{C}_X$ via a peer, or networks in $\mathcal{R}_X$ via a provider. If X provides a looking glass server, then we can investigate the cause of the anomalies identified using BGP tables or traceroutes. Consider such an anomalous route where X reaches a network $D \in \mathcal{C}_X$ through a peer P. Detailed routing information from the looking glass of AS X will allow us to distinguish between cases where the customer announces its prefix selectively to P (meaning X never receives the route from D), as opposed to the case where X receives the route from both P and D, but prefers the peer over a customer.

Peering policy

We will use publicly available information from Internet exchange points and online resources such as peeringDB [4], Packet Clearing House [3] and peeringMatrix [5] to characterize the peering behavior of ASes. PeeringDB is an online resource where networks volunteer information about their network type (ISP, enterprise customer, content provider, access provider, hosting provider), traffic volume and mix (mostly outbound traffic, mostly inbound traffic, or balanced), peering policy (open peering, based on traffic ratios, restrictive peering), and the list of IXPs they have presence. Although we can infer the set of peering links at an IXP with active measurements [11], we are more interested in the *potential for networks to peer* based on their business presence at various peering locations and advertised peering policies. We will also use *whois* data to determine the primary geographical region in which a network is present.

5.1 Proposed research questions

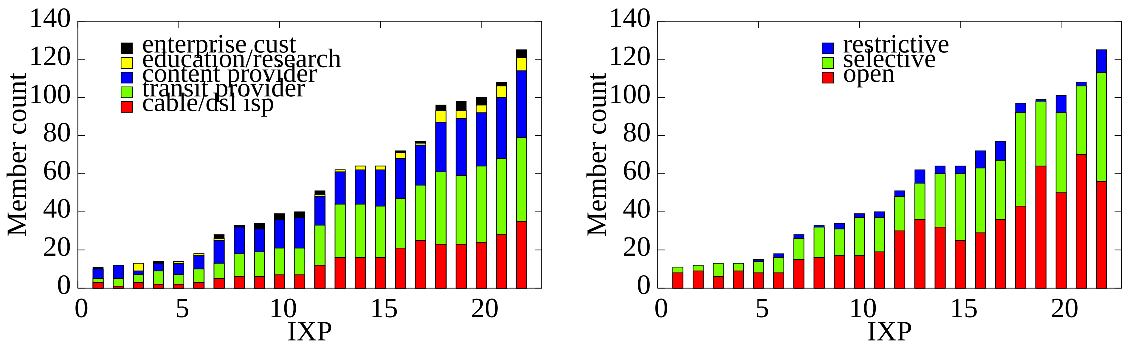
1. How often do we observe potential anomalies – routes that do not agree with the "prefer customer, then prefer peer" routing policy? Does the frequency of potential anomalies depend on network type? How often are these potential anomalies due to customers selectively announcing their prefixes (or announcing more specific prefixes to multiple providers) versus ISPs actually preferring a peer over a customer? Do we infer different behavior using BGP data vs active traceroute measurements?
2. How does the use of "prefer customer, then prefer peer" policy and selective advertisement evolve over time? We will use historical BGP data and also continuously collect traceroutes and looking glass measurements to study the evolution of routing policy.
3. What peering policies do different network types advertise? How many IXPs are different types of networks present at? Are network peering policies and exchange point presence correlated with traffic volume and mix, or geographical (cultural) region? Do networks peer outside their primary geographic region? For example, how many networks primarily based in North America peer at IXPs in Europe (and vice versa)?
4. Do we see trends in geographical expansion, peering policies, traffic volumes and traffic mix that networks self-report? We will use historical data from IXP operators and Internet archives [1], as well as collect snapshots from peeringDB over time to study the evolution of peering policies.

We will use the results of this study to parameterize the routing and peering policies of different network types in our model, e.g., to configure the use of a "prefer customer, then prefer peer" policy by providers, the use of selective prefix advertisement by multihomed customers, and the peering strategies and geographical presence of different network types. We will use the evolutionary trends from these studies as "what-if" scenarios that we will investigate with our model, e.g., what happens if large content providers expand their geographical presence and use an open peering policy?

5.2 Preliminary results

We report initial results using BGP data from October 10, 2009 to detect routing policy anomalies. We collected BGP tables from Routeviews [84], and extracted the set of contributing peer ASes, which are ASes that provide routing table feeds to the Routeviews collector. For each contributing peer, we found the set of its *direct customers*, using CAIDA's AS relationship dataset [20]. We used this BGP snapshot to find, for each peer AS M, the next hop used by M toward each network in its set of direct customers. If M was using either a peer or a provider to reach a network N that is in M's customer set, then we counted this route as a potential anomaly. We found that in this daily snapshot, several tier-1 providers use their peers to reach directly connected customers, which appears surprising if not necessarily anomalous. AT&T (AS7018) reaches 30% of its customers via its peers, while the corresponding numbers are 18% for Level3 (AS3356), 14% for Cogent (AS174), 22% for NTT (AS2914), and 17% for Global Crossing (AS3549). For each contributing peer M, we measured the fraction of potentially anomalous paths through each peer of M, to determine if some potential anomalies are due to incorrect relationship inference between M and its peer. For AT&T, 81% of the potential anomalies are via peers that are tier-1 networks, which we can assume are settlement-free peers of AT&T. We believe that errors in AS relationship inference did not cause these potential anomalies. Though this observation is surprising, Wang et al. [91] reported similar findings six years ago, and determined that these potential anomalies are either due to selective prefix advertisements by multihomed customers, or hot-potato routing by providers. Our next step will be to use the looking glass servers of Routeviews/RIPE peers (who provide them) to obtain the set of all routes known by a peer M toward a prefix that showed anomalous routes. We can then determine whether the anomaly was caused due to selective advertisement by the customer, or that M indeed preferred a peer route over a customer route.

For our preliminary study of peering policies, we used data from peeringDB [4] in March 2009. Figure 3(a) shows the number of networks listed as present in each IXP in North America (only IXPs that had at least 10 members are shown), split according to the type of network (as self-reported). We found that Content Providers and Cable/DSL ISPs represent a large fraction of the members of every IXP. However, other network types such as transit providers were also present in non-negligible numbers. Figure 3(b) shows the members of each IXP in North America split according to the peering policy advertised by those networks. The largest fraction of IXP members self-reported an open peering policy. We also found a non-negligible fraction of networks that advertise a selective peering policy. Further investigation of peeringDB snapshots will give us insight into the peering behavior of ASes. Also, collecting such snapshots continuously will allow us to parameterize our model to study the evolution of policies over time.



(a) (b)

Figure 3: Number of networks at each IXP in North America, by (a) network type and (b) peering policy, using data self-reported by networks on peeringDB in March 2009. This data will inform our model configuration of geographical presence of different network types and typical peering policies.

6 A first principles model of interdomain connectivity

In this section we describe ITER, a computational model that aims to capture essential aspects of the economic, topological, and traffic interactions in the Internet ecosystem. We follow a bottom-up modeling approach [65]. The main difference with previous modeling approaches is that we will parameterize ITER based on *empirical data* about traffic, routing and pricing/cost structures. In this section, we present the main components of ITER and how we plan to parameterize it from measurement data. The full description of the model is available in a technical report [34].

6.1 Model components

Network types:
We use a first-order network classification that consists of four types of networks: *Enterprise Customers (EC)* are networks at the edge of the Internet, normally mostly either sources of traffic (source-ECs), such as websites and hosting companies, or sinks of traffic (sink-ECs) such as campus, corporate or residential access networks. *Content providers(CP)* mostly sources of traffic, different from source-ECs in that they

also engage in peering, while ECs only act as customers in transit relations. *Small Transit Providers (STP)* mainly provide transit (access) to EC or CP networks, but with limited geographical reach. *Large Transit Providers (LTP)* are similar to STPs but with wider geographical presence. We may consider additional network types, such as Internet Exchange Points (IXPs), but this simple taxonomy is used in presentations to operator groups such as NANOG, and published in databases such as peeringDB [4].

Traffic model:

The traffic model consists of an internetwork traffic matrix, i.e., how much traffic is destined from one autonomous network to another, for all source-destination pairs [25]. We will assign elements of the interdomain traffic matrix in ITER using the interdomain traffic data described in § 3.

Routing model:

Interdomain routing policy determines how traffic is routed by individual networks in the Internet. Previous work assumed most networks use a policy of "no-valley, prefer customers, then prefer peers" [19, 44, 47]. The routing component of ITER is configurable, which allows us to parameterize the model from real-world data and measurements as described in § 5. We will configure how often providers use the "prefer-customer, then prefer peer" rule, and how often multihomed customers selectively advertise their prefixes.

Profit and cost functions:

The revenues of a transit provider consist of payments from its customers, while the provider incurs costs due to transit payments to its own providers, peering maintenance costs, and also local costs to operate and expand its own network. Let $\mathcal{C}_i$ be the set of customers, $\mathcal{P}_i$ the set of providers and $\mathcal{R}_i$ the set of peers of a transit provider i. The profit ("fitness") of i is

$$f_i = \sum_{c \in \mathcal{C}_i} T_i(v_{c,i}) - \sum_{p \in \mathcal{P}_i} T_p(v_{i,p}) - \sum_{r \in \mathcal{R}_i} R_i(v_{i,r}) - L_i(v_i) \quad (1)$$

The function $T_i(v_{c,i})$ is the transit payment made by customer c to provider i when the aggregate traffic exchanged by the two networks is $v_{c,i}$. Similarly, $T_p(v_{i,p})$ is the transit payment made by i to its provider p when the aggregate traffic exchanged by the two networks is $v_{i,p}$. The function $R_i(v_{i,j})$ determines the monetary cost of maintaining the peering link between networks i and j when the aggregate traffic exchanged over the peering link is $v_{i,j}$. The function $L_i(v_i)$ determines the local costs for network i when the aggregate traffic that it handles is v_i.

To parameterize the economic component of our model, we will use survey data about typical transit prices [72, 76], which indicate that transit prices show economies of scale – the per-bit cost decreases as the volume of traffic increases. We will also use data from Internet2 and their survey of participating institutions [55] about typical traffic volumes on interdomain links, as well as transit and operational costs.

Geographical presence:

We capture geographical constraints by assigning a number of points-of-presence (PoP) to each network, using the peering policy measurements described in Section 5.

Provider/peer selection policy:

In our model, each network has a particular policy for choosing its providers and settlement-free peers. ECs may choose the cheapest available providers, while CPs are more likely to choose providers based on some notion of the performance offered by that provider. To parameterize the peer selection policies of different network types, we will use the results of the peering policy measurements from § 5.

Computing the equilibrium network:

Given the initialization and a specific provider and peer selection policy for each network type, we "solve" the previous model, computing the internetwork that results in equilibrium. Equilibrium results when no network has the incentive to change its set of providers, remove a peering link, or add a peering link. The first two conditions correspond to Nash equilibria, while the third condition corresponds to the game theoretic concept of *pairwise stable equilibrium* [57]. We compute equilibria (if they exist) computationally as follows. We pick a network i to "play its move,", which involves first selecting providers (if any providers are needed) unilaterally, based on i's provider selection policy. Network i then attempts to either add one peering link with another network j (requires a bilateral move), or to remove one existing peering link. Both operations are determined based on each network's own peer selection policy. If the move of network i causes the topology to change, we recompute all routing tables and the interdomain traffic flow. If there is no change in the topology, we check the following termination criterion: If every network had the chance to play its move and none of them made any change in their connectivity, then the simulation ends.

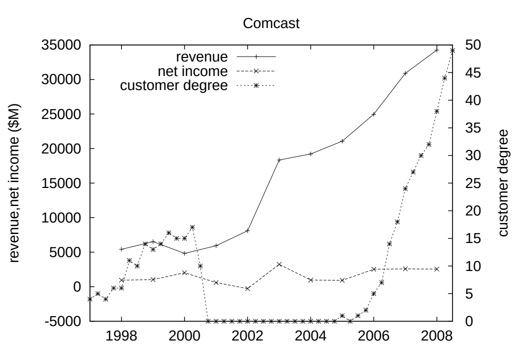
Validating the Model:

We will follow two complementary approaches to validate ITER. First, we will demonstrate that the equilibrium networks produced by the model can reproduce a comprehensive set of well-known macroscopic properties of the Internet. Then we will try to validate ITER's ability to reproduce known historical trends, using historical financial data, information about changes in the business role of ASes, and BGP information.

We have made preliminary progress toward our first validation goal – verifying ITER's ability to reproduce well-known macroscopic properties of the Internet [34]. Based on "best-guess" parameterization of the interdomain traffic matrix, routing policies, peering policies and pricing/cost structures, the equilibrium networks generated by ITER agree with known properties of the AS topology such as heavy-tailed degree distributions, constant average AS path lengths, hierarchical topological structure, profitability of different network types, and link loads. *Two questions remain open: whether ITER can reproduce these properties when parameterized with the explicit measurements we propose, and whether our model can faithfully reproduce historical trends in the evolution of the Internet.*

Comparing with known historical trends

We will validate our model's *predictive power* by comparing with known historical trends. We propose to study the qualitative trends in the evolution of *specific networks in the Internet* over time, attempting to correlate financial performance, documented changes in the business strategy, and observable topological characteristics of these networks. We will extract financial data about different types of networks – transit providers, access providers, content providers – from SEC filings and other online resources. We will use topological data from our recent study [33] to measure changes in the topological properties (number of customers, providers and peers) for these networks. We present two case studies that describe how we plan to use historical financial data and known historical trends as validation scenarios.



(a)

Figure 4: Revenue, income, and the number of customers of Comcast between 1998 and 2008. This historical data will allow us to retroactively test the predictive power of our model.

Case study 1: Comcast

Figure 4(a) shows the revenue, income and number of customers of Comcast over time. We found a sharp jump in revenues and income for Comcast between 2002 and 2003, consistent with Comcast entering the content delivery market by buying AT&T Broadband cable systems, and starting their high-definition TV and video-on-demand services [2]. The number of AS customers of Comcast has been increasing steadily since 2005, consistent with reports of Comcast's aggressive movement into the transit market [60]. We will use the Comcast case study as a validation scenario as follows. In one run of ITER, we will configure a network X as a pure access provider with a presence in a small set of regions, and observe the properties (topological and economic) of X at the resulting equilibrium (eq-A). We will then "replay" Comcast's changes in business strategy by expanding X's geographical presence, making X a large content source, and allowing X to have AS customers. We will then compute a new equilibrium (eq-B). We will investigate whether the change in X's topological and economic properties between eq-A and eq-B are consistent with the changes in Comcast's connectivity and financial performance over time.

Case study 2: Flattening AS topology and change in interdomain traffic flow

A recent study by Arbor Networks [60] reported a flattening of the topology due to content providers that expand their networks and peer aggressively, the rise of "hyper giants" that source a large fraction of total traffic, and traffic bypassing the traditional set of tier-1 providers. We will use these observations as validation scenarios as follows. In one run of ITER, we will assume a uniform distribution for the traffic sourced by CPs, and a small geographical scope for each CP. We will measure the topology and traffic flow in the resulting equilibrium (eq-A), in particular focusing on the traffic share of LTPs and interdomain path lengths. In the next run of ITER, we will simulate the case where a large fraction of traffic is sourced by the top few CPs, and CPs expand their geographical presence, to get the new equilibrium (eq-B). We will investigate whether eq-B does, indeed show a flatter topology and smaller LTP traffic share than eq-A.

6.2 Proposed research questions

Evaluating "what-if" scenarios

We will evaluate several "what-if" scenarios that represent plausible evolution paths for the the Internet, including:

1. What if a few of the largest content providers source a significant fraction of the total traffic (a trend alluded to in the Arbor report [60])?. What changes in the topological and economic properties of the equilibrium network does this modified traffic matrix matrix cause?
2. What if the interdomain traffic matrix is dominated by peer-to-peer file-sharing or video traffic? Increasing amounts of peer-to-peer traffic will induce an interdomain traffic matrix where ECs exchange large traffic volumes among themselves, with topological and economic implications.
3. Internet transit prices have been falling rapidly [72]. What happens to the profitability of transit providers if transit prices further drop by an order of magnitude, as some have predicted [30]?
4. What if the number of IXPs increases, and networks have a presence at a large number of IXPs, thus increasing the potential for networks to peer? Further, what if networks change their peering policies, and most networks use an "open peering" strategy? What happens to interdomain traffic flow? Does "open peering" result in a flatter Internet topology?
5. What if routing policies in the Internet no longer follow the "prefer-customer, then prefer-peer" heuristic, e.g., due to increased use of hot-potato routing? What if multihomed customers frequency advertise prefixes selectively to their providers, due to which providers have to reach their customers via their peers rather than directly? What if a new type of interdomain relationship (and associated routing policy) emerges that allows "peering federations", where a group of transit providers agree to transit traffic for each other? How does such a peering federation affect global traffic flow?
6. Several ISPs are starting to offer "paid peering", where routing decisions are similar to settlement-free peering, but one network pays the other. What happens if paid peering replaces settlement-free peering, and networks exchange a "fair" price, in the sense that both networks see equal monetary value from the peering link? Does paid-peering affect the density of peering links in the resulting equilibrium network?

Which type of networks pay or receive payments when paid-peering is used?

Evaluating provider and peer selection policies

In practice, provider and peer selection by autonomous networks is an ad-hoc process, an art that only few network strategists understand, and ISPs often select their providers and peers using back-of-the-envelope calculations [75]. In this research, we will study the effect of provider/peer selection policies of different network types on the topological and economic properties of the resulting internetwork. We will focus on the profitability of different network types, and topological properties such as path lengths. Studying the effects of provider and peer selection by ISPs is interesting for several reasons. First, individual networks would like to know the policy that would maximize their profit or minimize their cost. Second, we would like to know the effects of these policies on the global Internet, in terms of topological structure, routing dynamics and efficiency, profitability of various network types, and the risk of emerging monopolies or oligopolies. Third, it is important to understand how these provider and peer selection policies perform under different conditions, such as diverse traffic characteristics and application popularity, different pricing structures, or technological advances (e.g., very inexpensive transmission capacity).

Existence, stability and uniqueness of equilibria

Given that ITER's complexity makes it difficult to analytically prove the existence of equilibria, we will use simulations to study the conditions under which equilibria exist. When the simulation does not converge to an equilibrium, we will analyze the trajectory to examine whether the system "oscillates" between a set of states – scenarios under which the topology (and thus economy) is unstable. We will study whether the oscillations are due to an effect that is plausible in the real Internet (as opposed to simulation artifacts) and the extent to which the effects are local or global. When equilibria do exist, we will study their stability by re-starting the simulation after a small perturbation, and checking if the system converges to the same or a different equilibrium. In network formation games, it is common to have multiple equilibria [56], which can depend on the initialization or the playing sequence. If the equilibrium does not depend on the initial conditions or the order of play, then this indicates that certain outcomes are inevitable. On the the other hand, a strong dependence of the equilibrium on the playing sequence indicates that networks making the "right move at the right time" can significantly affect the outcome.

7 Broader Impact, Education and Outreach

Even though the main objective of this research is to contribute in the science of internetwork economics, we firmly expect that a deeper understanding of this area will have a broad and practical impact. First, we expect that the resulting insights and results will be useful in terms of network operations, peering decisions, and pricing. Network operators and strategists need to take short-term and long-term decisions about settlement-free or paid peering, provider selection, pricing, policy-based routing, expansion in new areas, support of new services, etc. All these practical decisions are directly related to the scope of the proposed research, and they can affect not only individual providers but the well-being of the entire Internet.

Second, there is a major debate going on in the US and elsewhere today between ISPs and content providers. Transit and access providers (ISPs) would like to see revenues from the content they deliver, instead of simply "moving the bits" from content providers to users. Content providers, on the other hand, want to make sure that their traffic does not receive degraded quality of service while in transit, and that they do not have to pay extra charges to ISPs. We expect that the results of the proposed research will help to further understand the economic nature of this debate, while the empirically grounded and computationally rigorous evaluation of specific strategies for peering and policy-routing can provide effective solutions to both parties.

In terms of education and mentoring, the PI from Georgia Tech plans to integrate elements of internetwork economics in his existing undergraduate (CS4270) and graduate (CS7260 or CS6250) courses.

Networking students rarely learn about economic constraints, methods and practices. Consequently, there is a major need in the job market today for graduates that can understand both internetworking as well as economic issues and that can design effective solutions that combine technological with economic, pricing or policy mechanisms. Also, the two PIs plan to co-advise the PostDoc and PhD students that will work in the proposed research, expecting that by the end of this research they will have a unique expertise to contribute in the area of internetwork economics.

In terms of outreach, the PIs commit that all data and methods (such as simulators, scripts, or visualizations) that will result from this research will be made publicly available. We will regularly disseminate our results to both research forums (conferences, workshops) and operator forums (such as NANOG or RIPE meetings). Finally, we will communicate our results to related federal policy organizations, such as the FCC, as it may be necessary that certain aspects of the Internet connectivity market are subject to regulation.

8 Collaboration Plan and Timeline

The PIs and senior personnel (SP) have an established history of successful collaboration, and they will participate equally in all research tasks and co-author most resulting papers. Dr. Dovrolis and Dr. Claffy have collaborated on a large DOE project in the past, and students from Georgia Tech regularly do summer internships at CAIDA. Dr. Dovrolis was the PhD advisor of SP Dr. Dhamdhere (now PostDoc at CAIDA). The PIs and SP will work closely with Ph.D. students from Georgia Tech working on this research.

We will proceed with the proposed measurement and modeling tasks in parallel. In the first year, we will focus on characterizing interdomain traffic (§ 3) and topology dynamics (§ 4), and improve the scalability and performance of the ITER simulator. In the second year, we will focus on measurements of routing and peering policies (§ 5) and validate ITER (§ 6). In the third year, we will use a fully parameterized ITER to answer the proposed research questions in § 6. We plan to apply for additional funding to host a Workshop on Internet Economics, following the successful workshop we co-hosted in September 2009 [21].

References

[1] Internet archive: Wayback machine. `http://www.archive.org/web/web.php`.

[2] Key dates in the history of Comcast Corp. `http://www.google.com/hostednews/ap/article/ALeqM5ir5zIxeURLj5LIMrZDqqK1zIoGcwD9C0T1300`.

[3] Packet Clearing House. `http://www.pch.net/home/index.php`.

[4] Peering Database. `http://www.peeringdb.com`.

[5] Peering Matrix. `https://www.euro-ix.net/member/m/peeringmatrix`.

[6] R. Albert and A. L. Barabasi. Topology of Evolving Networks: Local Events and Universality. *Physical Review Letters 85, 5234*, 2000.

[7] P. Albin and D. K. Foley. Decentralized, Dispersed Exchange Without an Auctioneer : A Simulation Study. *Journal of Economic Behavior & Organization*, 18(1):27–51, June 1992.

[8] E. Anshelevich, A. Dasgupta, E. Tardos, and T. Wexler. Near-optimal Network Design with Selfish Agents. In *STOC '03: Proceedings of the thirty-fifth annual ACM symposium on Theory of computing*, pages 511–520, 2003.

[9] E. Anshelevich, B. Shepherd, and G. Wilfong. Strategic Network Formation through Peering and Service Agreements. In *FOCS '06: Proceedings of the 47th Annual IEEE Symposium on Foundations of Computer Science*, pages 77–86, 2006.

[10] E. Arcaute, R. Johari, and S. Mannor. Network Formation: Bilateral Contracting and Myopic Dynamics. In *Workshop on Internet and Network Economics (WINE)*, 2007.

[11] B. Augustin, B. Krishnamurthy, and W. Willinger. IXPs: Mapped? In *Proceedings of Internet Measurement Conference (IMC)*, Nov. 2009.

[12] R. Aumann and R. Myerson. Endogenous Formation of Links Between Players and Coalitions: An Application of the Shapley Value. *Roth, A. (ed.) The Shapley Value, Cambridge University Press*, 1988.

[13] V. Bala and S. Goyal. A Noncooperative Model of Network Formation. *Econometrica*, 68(5):1181–1229, 2000.

[14] A. L. Barabasi and R. Albert. Emergence of Scaling in Random Networks. *Science 286 509512*, 1999.

[15] F. Bloch and M. O. Jackson. The Formation of Networks with Transfers Among Players. *Journal of Economic Theory*, 133(1):83 – 110, 2007.

[16] Y. Bramouille, D. Lopez-Pintado, S. Goyal, and F. Vega-Redondo. Network Formation and Anticoordination Games. *International Journal of Game Theory*, (33):1–19, 2004.

[17] A. Broido, E. Nemeth, and K. Claffy. Internet Expansion, Refinement, and Churn. *European Transactions on Telecommunications*, 13, 2002.

[18] T. Bu and D. Towsley. On Distinguishing Between Internet Power Law Topology Generators. In *Proceedings of IEEE Infocom*, 2002.

[19] M. Caesar and J. Rexford. BGP Routing Policies in ISP Networks. *IEEE Network*, Nov.-Dec. 2005.

[20] CAIDA. CAIDA's AS Relationships Datasets. `http://www.caida.org/data/active/as-relationships/`.

[21] CAIDA/UCSD and Georgia Tech. The First Workshop on Internet Economics (WIE'09), September 2009.

[22] J. M. Carlson and J. Doyle. Highly Optimized Tolerance: A Mechanism for Power Laws in Designed Systems. *Physical Review E*, 60, 1999.

[23] H. Chang. An Economic-based Empirical Approach to Modeling the Internet's Interdomain Topology and Traffic Matrix. *Ph.D. Thesis*, 2006. Advisor Sugih Jamin.

[24] H. Chang, R. Govindan, S. Jamin, S. J. Shenker, and W. Willinger. Towards Capturing Representative AS-level Internet Topologies. *Computer Networks*, 44(6):737 – 755, 2004.

[25] H. Chang, S. Jamin, Z. M. Mao, and W. Willinger. An Empirical Approach to Modeling Inter-AS Traffic Matrices. In *Proceedings of the Internet Measurement Conference (IMC)*, pages 12–12, 2005.

[26] H. Chang, S. Jamin, and W. Willinger. Internet Connectivity at the AS-level: An Optimization-Driven Modeling Approach. In *Proceedings of ACM SIGCOMM Workshop on MoMeTools*, 2003.

[27] H. Chang, S. Jamin, and W. Willinger. To Peer or Not to Peer: Modeling the Evolution of the Internet's AS-Level Topology. In *Proceedings of IEEE Infocom*, 2006.

[28] H. Chang and W. Willinger. Difficulties Measuring the Internet's AS-Level Ecosystem. In *Proceedings of the 40th Annual Conference on Information Sciences and Systems*, 2006.

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

[30] Cogent Communications. Cogent Communications Group Q3 2009 Earnings Call Transcript. `http://seekingalpha.com/article/172306-cogent-communications-group-q3-2009-earnings-call-transcript?page=9`, Nov 2009.

[31] R. Cohen and D. Raz. The Internet Dark Matter - On the Missing Links in the AS Connectivity Map. In *Proceedings of IEEE Infocom*, 2006.

[32] J. Corbo, S. Jain, M. Mitzenmacher, and D. Parkes. An Economically Principled Generative Model of AS Graph Connectivity. In *Proceedings of the International Joint Workshop on The Economics of Networked Systems and Incentive-Based Computing*, 2007.

[33] A. Dhamdhere and C. Dovrolis. Ten Years in the Evolution of the Internet Ecosystem. In *Proceedings of ACM SIGCOMM Internet Measurement Conference (IMC)*, 2008.

[34] A. Dhamdhere and C. Dovrolis. A Model for Interdomain Network Formation, Economics and Routing. *Georgia Tech Technical Report GT-CS-09-09*, 2009.

[35] X. Dimitropoulos, D. Krioukov, M. Fomenkov, Y. Hyun, K. Claffy, and G. Riley. AS Relationships: Inference and Validation. *ACM SIGCOMM Computer Communication Review (CCR)*, 2007.

[36] N. Economides. The Economics of Networks. *International Journal of Industrial Organization*, 14(6):673 – 699, 1996.

[37] N. Economides. The Economics of the Internet Backbone. *Handbook of Telecommunications Economics Ed. S. Majumdar, I. Vogelsang, M. Cave. Amsterdam: Elsevier Publishers*, 2006.

[38] A. Fabrikant, E. Koutsoupias, and C. H. Papadimitriou. Heuristically Optimized Trade-Offs: A New Paradigm for Power Laws in the Internet. In *Proceedings of ICALP*, 2002.

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

[40] W. Fang and L. Peterson. Inter-AS Traffic Patterns and Their Implications. In *IEEE Global Telecommunications Conference (GLOBECOM)*, Dec 1999.

[41] J. Feigenbaum, V. Ramachandran, and M. Schapira. Incentive-Compatible Interdomain Routing. In *EC '06: Proceedings of the 7th ACM Conference on Electronic Commerce*, pages 130–139, 2006.

[42] J. Feigenbaum, R. Sami, and S. Shenker. Mechanism Design for Policy Routing. In *PODC '04: Proceedings of the twenty-third annual ACM symposium on Principles of distributed computing*, pages 11–20, 2004.

[43] A. Feldmann, N. Kammenhuber, O. Maennel, B. Maggs, R. De Prisco, and R. Sundaram. A Methodology for Estimating Interdomain Web Traffic Demand. In *Proceedings of ACM SIGCOMM Internet Measurement Conference (IMC)*, pages 322–335, 2004.

[44] L. Gao. On Inferring Autonomous System Relationships in the Internet. *IEEE/ACM Transactions on Networking*, 9(6), 2001.

[45] L. Gao, T. Griffin, and J. Rexford. Inherently Safe Backup Routing with BGP. In *Proceedings of IEEE INFOCOM*, pages 547–556, 2001.

[46] L. Gao and J. Rexford. Stable Internet Routing Without Global Coordination. *IEEE/ACM Transactions on Networking*, 9(6):681–692, 2001.

[47] L. Gao and F. Wang. The Extent of AS Path Inflation by Routing Policies. In *Proceedings of IEEE Global Telecommunications Conference (GLOBECOM)*, 2002.

[48] P. Gill, M. Arlit, Z. Li, and A. Mahanti. The Flattening Internet Topology: Natural Evolution, Unsightly Barnacles or Contrived Collapse? In *Proceedings of Passive and Active Measurement Conference (PAM)*, 2008.

[49] T. G. Griffin, F. B. Shepherd, and G. Wilfong. Policy Disputes in Path-Vector Protocols. In *Proceedings of the Seventh Annual International Conference on Network Protocols (ICNP)*, page 21, Washington, DC, USA, 1999. IEEE Computer Society.

[50] T. G. Griffin, F. B. Shepherd, and G. Wilfong. The Stable Paths Problem and Interdomain Routing. *IEEE/ACM Transactions on Networking*, 10(2):232–243, 2002.

[51] T. G. Griffin and G. Wilfong. A Safe Path Vector Protocol. In *Proceedings of IEEE INFOCOM*, volume 2, pages 490–499, 2000.

[52] L. He and J. Walrand. Pricing and Revenue Sharing Strategies for Internet Service Providers. *IEEE Journal on Selected Areas in Communications*, May 2006.

[53] Y. He, G. Siganos, M. Faloutsos, and S. V. Krishnamurthy. A Systematic Framework for Unearthing the Missing Links: Measurements and Impact. In *Proceedings of 4th USENIX/SIGCOMM NSDI*, 2007.

[54] P. Holme, J. Karlin, and S. Forrest. An Integrated Model of Traffic, Geography and Economy in the Internet. *ACM SIGCOMM Computer Communication Review (CCR)*, 38(3):5–16, 2008.

[55] Internet2. Internet2. `http://www.internet2.edu/`.

[56] M. O. Jackson. A Survey of Network Models of Network Formation: Stability and Efficiency. *Group Formation in Economics: Networks, Clubs and Coalitions, edited by Demange, G. and Wooders, M.*, 2003.

[57] M. O. Jackson and A. Wolinsky. A Strategic Model of Social and Economic Networks. *Journal of Economic Theory, 71:4474*, 1996.

[58] R. Johari, S. Mannor, and J. Tsitsiklis. A Contract-based Model for Directed Network Formation. *Games and Economic Behavior*, 56(2):201–224, August 2006.

[59] E. Kenneally and K. Claffy. Dialing Privacy and Utility: A Proposed Data-sharing Framework to Advance Internet Research. Technical report, 2009. in submission to IEEE Security and Privacy (S&P) special issue on Privacy-Preserving Sharing of Sensitive Information, `http://www.caida.org/publications/papers/2009/dialing_privacy_utility/`.

[60] C. Labovitz, S. lekel Johnson, D. McPherson, J. Oberheide, and F. Jahanian. 2009 Internet Observatory Report. *Arbor Networks White Paper*, 2009.

[61] J. J. Laffont, S. Marcus, P. Rey, and J. Tirole. Internet Interconnection and the Off-Net-Cost Pricing Principle. *The RAND Journal of Economics*, 34(2):370–390, 2003.

[62] P. Laskowski and J. Chuang. Network Monitors and Contracting Systems: Competition and Innovation. *ACM SIGCOMM Computer Communication Review*, 36(4):183–194, 2006.

[63] J. Leskovec, J. Kleinberg, and C. Faloutsos. Graph Evolution: Densification and Shrinking Diameters. *ACM Transactions on Knowledge Discovery from Data (ACM TKDD)*, 2007.

[64] H. Levin, M. Schapira, and A. Zohar. Interdomain Routing and Games. In *STOC '08: Proceedings of the 40th annual ACM symposium on Theory of computing*, pages 57–66, New York, NY, USA, 2008. ACM.

[65] L. Li, D. Alderson, W. Willinger, and J. Doyle. A First-principles Approach to Understanding the Internet's Router-level Topology. In *SIGCOMM '04: Proceedings of the 2004 conference on Applications, technologies, architectures, and protocols for computer communications*, pages 3–14, New York, NY, USA, 2004. ACM.

[66] R. T. Ma, D. Chiu, J. C. Lui, V. Misra, and D. Rubenstein. Interconnecting Eyeballs to Content: A Shapley Value Perspective on ISP Peering and Settlement. In *Proceedings of the Workshop on Economics of Networked Systems (NetEcon)*, 2008.

[67] R. T. Ma, D. Chiu, J. C. Lui, V. Misra, and D. Rubenstein. On Cooperative Settlement Between Content, Transit and Eyeball Internet Service Providers. In *Proceedings of the ACM Conference on Emerging network experiment and technology (CoNEXT)*, 2008.

[68] R. T. B. Ma, D. Chiu, J. C. S. Lui, V. Misra, and D. Rubenstein. Internet Economics: The use of Shapley Value for ISP Settlement. In *Proceedings of the ACM Conference on Emerging network experiment and technology (CoNEXT)*, 2007.

[69] D. Magoni and J. J. Pansiot. Analysis of the Autonomous System Network Topology. *ACM SIGCOMM Computer Communication Review (CCR)*, 2001.

[70] P. Mahadevan, D. Krioukov, M. Fomenkov, B. Huffaker, X. Dimitropoulos, K. Claffy, and A. Vahdat. The Internet AS-Level Topology: Three Data Sources and One Definitive Metric. *ACM SIGCOMM Computer Communication Review (CCR)*, 2005.

[71] A. Nagurney. *Network Economics: A Variational Inequality Approach*. Kluwer Academic Publishers, 1999.

[72] W. B. Norton. Peering vs Transit: The Business Case for Peering. *DrPeering.net white paper*.

[73] W. B. Norton. A Business Case for ISP Peering. *Equinix white papers*, 2002.

[74] W. B. Norton. The Art of Peering: The Peering Playbook. *Equinix white papers*, 2002.

[75] W. B. Norton. The Evolution of the U.S. Internet Peering Ecosystem. *Equinix white papers*, 2004.

[76] W. B. Norton. Transit Cost Survey. `www.nanog.org/mtg-0606/pdf/bill.norton.2.pdf`, Jul 2006.

[77] W. B. Norton. Video Internet: The Next Wave of Massive Disruption to the U.S. Peering Ecosystem. *Equinix white papers*, 2007.

[78] A. M. Odlyzko. Internet Traffic Growth: Sources and Implications. In *Proceedings of SPIE*, volume 5247, pages pp. 1–15, 2003.

[79] A. M. Odlyzko. Threats to the Internet: Too Much or too Little Growth? *Internet Evolution*, Feb 2008.

[80] R. Oliveira, D. Pei, W. Willinger, B. Zhang, and L. Zhang. In Search of the Elusive Ground Truth: The Internet's AS-level Connectivity Structure. In *Proceedings of ACM SIGMETRICS*, 2008.

[81] R. V. Oliveira, B. Zhang, and L. Zhang. Observing the Evolution of Internet AS Topology. In *Proceedings of ACM SIGCOMM*, 2007.

[82] S. Park, D. M. Pennock, and C. L. Giles. Comparing Static and Dynamic Measurements and Models of the Internet's AS Topology. In *Proceedings of IEEE Infocom*, 2004.

[83] RIPE. RIPE Network Coordination Centre. `http://www.ripe.net`.

[84] Routeviews. University of Oregon Route Views Project. `http://www.routeviews.org`.

[85] M. A. Serrano, M. Boguna, and A. D. Guilera. Modeling the Internet. *The European Physics Journal B*, 2006.

[86] S. Shakkottai, M. Fomenkov, R. Koga, D. Krioukov, and K. Claffy. Evolution of the Internet AS-Level Ecosystem. In *First International Conference on Complex Sciences: Theory and Applications (COMPLEX)*, February 2009.

[87] S. Shakkottai and R. Srikant. Economics of Network Pricing with Multiple ISPs. *IEEE/ACM Trans. Netw.*, 14(6):1233–1245, 2006.

[88] G. Siganos, M. Faloutsos, and C. Faloutsos. The Evolution of the Internet: Topology and Routing. *University of California, Riverside Technical Report*, 2002.

[89] L. Tesfatsion. Agent-Based Computational Economics: Growing Economies From the Bottom Up. *Artif. Life*, 8(1):55–82, 2002.

[90] N. J. Vriend. Self-Organization of Markets: An Example of a Computational Approach. *Computational Economics*, 8(3):205–31, August 1995.

[91] F. Wang and L. Gao. On Inferring and Characterizing Internet Routing Policies. In *Procedings of the Internet Measurement Conference (IMC)*, 2003.

[92] X. Wang and D. Loguinov. Wealth-Based Evolution Model for the Internet AS-Level Topology. In *Proceedings of IEEE Infocom*, 2006.
[93] A. Watts and M. Jackson. The Existence of Pairwise Stable Networks. *Seoul Journal of Economics*, 14(3):299–321, 2001.
[94] A. Watts and M. Jackson. The Evolution of Social And Economic Networks. *Journal of Economic Theory*, 106(2):265–295, 2002.
[95] A. Wilhite. Bilateral Trade and ”Small-World” Networks. *Computational Economics*, 18(1):49–64, 2001.
[96] S. H. Yook, H. Jeong, and A. L. Barabasi. Modeling the Internet's Large-scale Topology. *Proceedings of the National Academy of Sciences*, 2002.
[97] B. Zhang, R. Liu, D. Massey, and L. Zhang. Collecting the Internet AS-level Topology. *ACM SIGCOMM Computer Communication Review (CCR)*, 2005.
[98] S. Zhou. Understanding the Evolution Dynamics of Internet Topology. *Physical Review E, vol. 74*, 2006.