

A network formation model for Internet transit relations*

Aemen Lodhi
Georgia Institute of Technology
aemen.lodhi@gatech.edu

Constantine Dovrolis
Georgia Institute of Technology
dovrolis@cc.gatech.edu

Abstract

Most Autonomous Systems in the Internet need to select one or more transit providers. The provider selection process is complex, influenced by dynamic pricing, contracts, performance, marketing and other factors. We propose a simple dynamic model that captures the salient features of the provider selection process. The model creates a positive feedback effect, where “the bigger a provider is the bigger it gets”. We then study the resulting internetwork formation process, showing that it always leads to a stable, but not unique, internetwork. We also use computational experiments to understand how the convergence delay scales with the size of the network, the factor(s) that affect the number of distinct equilibria, and the impact of three key model parameters.

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 coupled with financial exchanges. Additionally, routing, provider selection and peering policies are often dictated by economic and strategic objectives. Most of the interactions between Autonomous Systems (ASes), or just networks, are local in nature, without coordinated control or regulation. They often have global impact, however, affecting the performance and reliability experienced by users, the financial viability of network and service providers, and to some degree the global economy.

In this paper, we study AS internetwork formation by focusing on *transit relations* between networks, also known as *customer-provider* relations. In a transit relation, one network (the provider) provides global Internet connectivity to another network (the customer) for a transit fee. If the provider cannot reach every other network directly, it has to become the customer of another provider (or to establish peering links with all other top-level, or *Tier-1*, providers).

We propose a simple, yet realistic, provider selection model in which each network selects a provider based on geographic and hierarchical constraints, aiming to minimize its transit costs. The model includes a positive feedback component, creating a “rich get richer” effect for transit providers. We study the dynamics of this provider selection model and show that it always converges to a stable internetwork. That equilibrium may not be unique, however. We also study computationally the convergence delay, the number of distinct equilibria, the topological differences between equilibria, the impact of pricing and of geographic constraints (i.e., locations of each network), x . Finally, we attempt to predict the network that

will become the wealthiest transit provider *before* the internetwork converges to an equilibrium.

Our objective is not to produce an interdomain topology generator, but to study the dynamics of the internetwork formation process, the existence of equilibria in that process, and the economic and topological characteristics of such equilibria. Our model should be viewed as a first step towards more sophisticated models (that will include peering links, multihoming, geographical expansion, etc) through which we aim to understand the microdynamics of the AS internetwork formation process and evolution, quantify the impact of individual parameters, and ask what-if questions.

In the rest of this section we review the most relevant related work. Section II describes the model. Section III presents the stability result (the proof is in the Appendix). Section IV summarizes our computational results. We conclude in Section V.

Related work:

Fabrikant et al. [10] present an internetwork growth model based on the Heuristically Optimized Trade-offs (HOT) framework. Chang et al. [6] build upon the HOT model by incorporating a vector of economic factors in their provider selection scheme. Wang and Loguinov [15] propose that the number of links established by a network is directly proportional to its wealth. Holme et al. [12] consider traffic, economic and geographic constraints. Nodes accumulate an income that is proportional to the traffic they carry, while costs are associated with spatial expansion and capacity increases. Corbo et al. propose a model with link revisions or “rewiring” [8]. In the previous works, the focus has been on network growth and on the topological features, such as power-law degree distributions or strong clustering that have been observed in the actual AS-level topology [7, 16, 17]. Our work is more relevant to earlier network formation models, mostly in the theoretical computer science and economics literature [2, 3, 4, 13]. Those models, however, do not aim to capture the specifics of Internet transit relations, and so they differ significantly with our model in terms of the provider selection process.

2 Model description

We consider a population \mathcal{N} of N Autonomous Systems, or just “networks”. Each network is present in a given set of locations. An internetwork traffic matrix gives the average traffic volume sent from each network to every other network. The model only captures the interconnections between distinct networks - it does not capture the “intradomain” connectivity or traffic flow within the same network. We do not associate a certain business type with any network; they may act as transit providers, content providers, access providers, enterprise networks or any combination of these functions, depending on their generated/consumed traffic, transit pricing, or geographi-

* This material is based upon work supported by the National Science Foundation under Grant No. 1017139 and by a research award from Cisco Systems.

cal presence. The geographic locations and price of each network as well as the traffic matrix remain constant during the timescales of interest in this model, i.e., we do not model growth or other long-term changes in the internetwork characteristics. Instead, we focus on the dynamics of the provider selection process: how does the internetwork change with time as each network attempts to select the best provider, even if nothing else changes? Obviously, this is just a starting point: more complex dynamics (involving traffic matrix variations, peering links, dynamic pricing or geographical expansion, for instance) would take place in larger timescales.

2.1 Locations and geographical presence

Each network is placed in one or more *locations*. We may think of a “location” as a distinct Internet Exchange Point, for instance. The “world” consists of G_M locations. Network x is located at a given set $G(x)$ of locations. $G(x)$ is an input to this model. Two networks *overlap* if they share at least one location. For network x , let $S(x)$ be the set of networks that overlap with x .

2.2 Traffic matrix and transit traffic

Let T_{xy} be the traffic volume (over a given timescale) generated by node x and consumed by node y . The N-by-N matrix T is referred to as the interdomain *Traffic Matrix (TM)*. If we aggregate across all consumers and producers of traffic from/to x , respectively, we have that x generates traffic $V_G(x)$ and consumes $V_C(x)$. Note that $T_{xx}=0$ - we do not capture the local traffic within a network. The *transit traffic* $V_T(x)$ of x is the traffic volume that is neither generated nor consumed by x - it only passes through x enroute towards its destination. If $V_O(x)$ represents the total *outbound* traffic at x , then the transit traffic of x can be expressed as

$$V_T(x) = V_O(x) - V_G(x) \quad (1)$$

Note that the transit traffic of a network depends on the underlying internetwork topology as well as on the routing algorithm. Even if the TM is constant, the transit traffic of a network may change as networks choose different providers.

2.3 Transit fee and price

A network y charges a transit fee to each network x that connects to y as customer. In that case we say that y is the transit provider of x , or that x is a transit customer of y . For simplicity, we assume that each network can have at most one provider, i.e., *singlehoming*. Even though multihoming is very common in practice, often only one provider is actively used while the other providers act as backups. The transit fee depends on the traffic volume $V_P(x)$ that x sends to and receives from its provider y . The precise formula for computing transit fees does not matter in our context, and so for simplicity we use a linear relation: if $P(y)$ is the transit price of y , the transit fee that x should pay to its provider y is $P(y) \times V_P(x)$. The transit prices are exogenous factors to this model and they are given as inputs.

2.4 Provider selection

Informally, a network x selects a provider y if: a) x overlaps with y , b) y is “larger” than x for an appropriate met-

ric of network size, and c) x is the least expensive among all networks that satisfy the previous two constraints. The first constraint relates to geographical presence which occurs in practice: a network in one country would not select a provider located on a different continent. The third objective is also obvious: choose the cheapest provider, in terms of their transit price, among eligible networks that satisfy the two other constraints. The second constraint is unusual and it requires some explanation.

In practice, a large network such as AT&T or Level3 would not select as transit provider a small regional network. How can we define formally however that *a network x is larger than a network y* ? Several metrics could be proposed: number of locations, number of customers, size of customer tree, total traffic, transit traffic, and others. Based on extensive discussions with network operators we came to the conclusion that the *transit traffic* $V_T(x)$ is the most realistic indicator for the size of a transit provider. We do *not* claim that in practice an AS would reveal its transit volume to its potential customers. Instead, we claim that the previous constraint on the relation between the customer-provider transit volumes acts as a simple but realistic *proxy* for what often happens in practice. In Appendix B, we show experimental evidence for this assumption, based on traffic data from PeeringDB and customer-provider relations that have been inferred from BGP data. Additionally, transit volume is probably positively correlated with most other plausible metrics. For instance, if x is present in many locations it would also have the ability to attract more customers, which would then increase its customer tree size as well as its transit traffic. On the other hand, the generated or consumed traffic of a network does not relate necessarily to that network’s transit capabilities; for instance, content providers (e.g., Facebook or Akamai) generate large volumes of traffic but they do not operate a backbone network.

Formally, the provider of network x is denoted by $R(x)$ and it is computed as

$$R(x) = \arg \min_{y \in S(x)} \{P(y) : V_T(y) > V_T(x)\} \quad (2)$$

Ties between networks with the same price are broken deterministically, based on unique network identifier.

Note that it is possible that $R(x) = \emptyset$. This can happen when x cannot find a network in its set of locations that has larger transit volume than itself. In that case, x is referred to as a *Tier-1* network and it does not have a provider. To ensure that the internetwork remains connected however, a Tier-1 network is connected using *peering links* (as opposed to customer-provider links) with every other Tier-1 network. This is also true for the Internet: Tier-1 networks are interconnected in a clique using peering links to ensure global connectivity. Note that these peering links are formed even between networks that do not overlap in a region.

2.5 Routing

Given that each network can have at most one provider, and Tier-1 networks are interconnected in a clique, there is a unique shortest path between any two networks.

2.6 Network formation process

The network formation process proceeds in discrete time units called *rounds*. We start in round t_0 with an arbitrary initial topology (described later). We route the traffic between every pair of networks, and compute the transit volume of each network. Then, all networks select a provider based on (2) in a synchronous manner. Any resulting Tier-1 networks are interconnected in a clique using peering links. This completes the first round.

The previous steps are repeated in each successive round. If at some point we reach a round in which the resulting internetwork topology is identical to that of the previous round, each network has selected the same provider as in the previous round (or it remained a Tier-1 network). From that point on the internetwork will not change, in terms of topology, provider selection or traffic flow, and we say that it has *converged* to an equilibrium. Whether an equilibrium always exists is the subject of the next section.

Algorithm 1 Network formation process

1. Route each traffic flow along its shortest path.
 2. For each network $x \in \mathcal{N}$, compute transit volume $V_T(x)$.
 3. For each network $x \in \mathcal{N}$, perform provider selection based on Equation 2.
 4. Interconnect all networks that have no provider (Tier-1 networks) with a clique of peering links.
 5. If the internetwork topology has not changed relative to the previous round, stop. Else, proceed to the next round and jump to Step-1.
-

2.7 External inputs and initial topology

The following parameters are exogenous to this model and they are provided as inputs initially: the set of locations for each network G , the price vector P , and the TM.

The initial topology is constructed as follows. We first select a random number of networks to be Tier-1 nodes and interconnect them with peering links in a clique. We order the rest of the networks randomly and examine each network sequentially. For network x , we assign a provider randomly among the set of networks that a) overlap with x , and b) are *not* in the customer tree of x . At the end of the process, if a network does not have a provider it joins the clique of Tier-1 networks.

3 Convergence to equilibria

In Appendix A, we prove the following result for the previous network formation model.

Theorem-1: *Given a set of networks \mathcal{N} , a set of locations $G(x)$ and a transit price $P(x)$ for each network $x \in \mathcal{N}$, and an internetwork traffic matrix TM, the provider selection process given by (2) converges to an equilibrium in a finite number of rounds.*

This result implies that the provider selection process performed independently by each network cannot go on forever, trapped in an phase space limit cycle. Sooner or later, it will converge to a stable point in which no net-

work has the incentive to change its provider unilaterally, i.e., to a Nash equilibrium. On the other hand, it is easy to show with a simple example of 3-4 networks that we can get different equilibria depending on the initial internetwork topology. Assuming that all other parameters of the model are constant (locations, prices and TM), it is the randomness of the initial topology that can result in different internetworks at the end.

Why is existence of equilibria important in our context? It is well-known that the Internet is in a persistent state of flux. We should notice however that the observed “rewiring” of the Internet topology is accompanied by changes in several factors, such as prices or network geographical presence. If those exogenous factors were constant over a sufficiently long time period, it is reasonable to expect that each network would converge to the best possible provider selection, according to its objective and constraints, instead of switching between providers infinitely. Thus, the model can also be used to study the evolution of an internetwork, computing a new equilibrium every time an exogenous factor changes, *as long as the convergence delay is less than the time between successive changes in those factors*. It is for this reason that we also study, in the next section, the convergence delay of this provider selection model. If the convergence delay is sufficiently short, we can assume that the exogenous parameters of the model (set of existing networks, prices, geographic locations, TM, etc) remain constant while the internetwork moves from one equilibrium to another.

4 Computational results

In this section we investigate several questions about the model: How large is the convergence delay in practice, and how does it scale with the number of networks? What determines the number of distinct equilibria that we can get from different initial topologies and similar questions. It is unlikely that these questions can be answered analytically. Further, we are not interested in worst-case bounds but in the more pragmatic average-case behavior (with appropriate confidence intervals). It is for this reason that we prefer to rely on computational than analytical tools in the following.

First, we describe how we compute the model parameters. Unless if noted otherwise, we simulate $N=500$ networks in a world of $G_M=25$ locations. The number of locations that a network is present at follows a Zipf distribution with exponent 0.8. Even though we do not have data to support the selection of this Zipf distribution, it appears plausible qualitatively that most networks are present in only 1-3 locations, but few networks are present in almost all locations. We refer to $|G(x)|$ as the *expanse* of network x . Unless if noted otherwise, the assignment of a network to $|G(x)|$ locations (i.e., the selection of the elements of $G(x)$) is done randomly without replacement. In some experiments, however, we control the maximum number of networks in a location - this is done as follows. The *density* of a location is the number of networks in that location normalized by N . The *Maximum Density* parameter D_M is the maximum density across all locations. To enforce this parameter, we first place $D_M * N$ randomly selected distinct networks

in a random location. Then, we place each network randomly in the remaining $G_M - 1$ locations, based on its expanse, so that no location has more than $D_M * N$ networks.

The traffic matrix is computed as follows. The traffic $V_G(x)$ that a network x generates also follows a Zipf distribution with exponent 0.8 (see Appendix C). This is consistent with the fact that some networks in the Internet are much larger content producers (e.g., Google, YouTube) than most others. We assume that the traffic flow T_{xy} that network x sends to network y is proportional to the number of locations of y , i.e.,

$$T_{xy} = V_G(x) * \frac{|G(y)|}{\sum_{i=1}^N |G(i)|} \quad (3)$$

The rationale is that a network with large geographic scope would have more access customers (e.g., DSL or cable users), and it is those users that consume most Internet content. Finally, the consumed traffic $V_C(x)$ from a network x is simply

$$V_C(x) = \sum_{i=1, i \neq x}^N T_{ix} \quad (4)$$

Note that $G(x)$ and $V_G(x)$ are independent, while $V_C(x)$ is related to $|G(x)|$ and thus traffic consumption is also long-tailed (few networks are much larger consumers). Regarding transit prices, unless noted otherwise, $P(x)$ is random (the absolute value of the prices does not matter).

Convergence Delay: Convergence delay is the number of rounds it takes for the internetwork to reach equilibrium. The proof of Theorem-1 gives the impression that convergence would take a large number of rounds. In practice, however, we observe that the convergence process takes only few rounds, always less than 5 rounds for up to 1000 networks. Figure 1 shows the mean convergence delay as function of N (with 95th confidence intervals for 100 runs with different initial topologies). Note both the low delay as well as the small variability across different runs.

Why does convergence occur so rapidly, despite the fact that networks act without coordination? Figure 2 shows the mean number of networks that have customers (i.e., providers) and the mean highest price among those providers, during convergence. As each node selects the cheapest eligible provider, the number of providers drops significantly in the first couple of rounds. Most networks become “stubs” (i.e., they have no customers), and as shown in the Appendix, a stub does not change its provider selection. The few remaining providers accumulate large transit volumes, and so they attract even more customers - some of them previous providers that now become stubs. This effect further decreases the number of providers, enhancing stability in the internetwork. After investigating systematically all model parameters we found that the geographic overlap between networks, i.e., their placement in locations, has a clear effect on the convergence delay. Figure 3 shows how the convergence delay varies with the maximum density factor D_M .

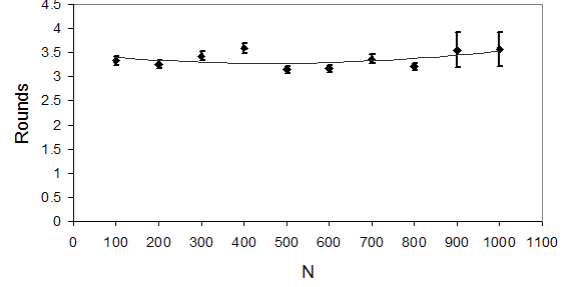


Fig. 1: Mean convergence delay (95% C.I.) versus number of networks N .

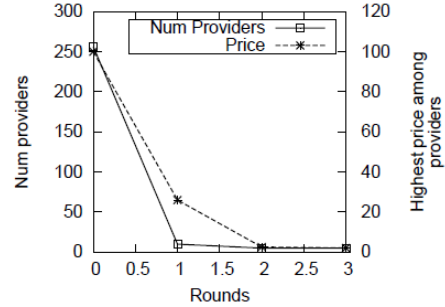


Fig. 2: Number of providers and the highest price among providers during convergence.

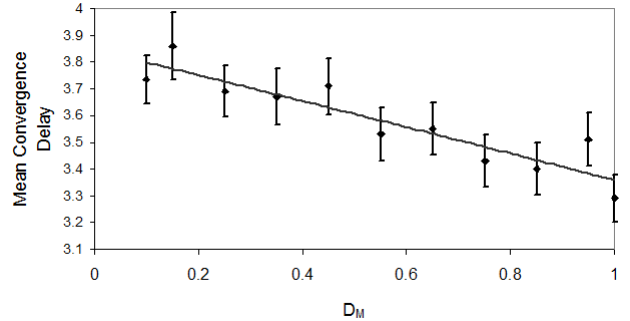


Fig. 3: Mean convergence delay (95% C.I.) versus maximum density.

Although the delay remains below 4 rounds, there is a roughly linear decreasing trend as D_M increases. An increase in D_M causes more networks to overlap, which results in a reduction in the number of providers because more networks can find a cheaper provider in their location(s). The lower the number of providers in the internetwork, the lower the flux in the provider selection process (because more networks become stubs and stubs do not change their provider selection).

Heterogeneity across distinct equilibria: We say that two equilibria are different if there is any topological difference in the corresponding internetworks. We already saw that different initial topologies can result in different equilibria. How common is that however? And how

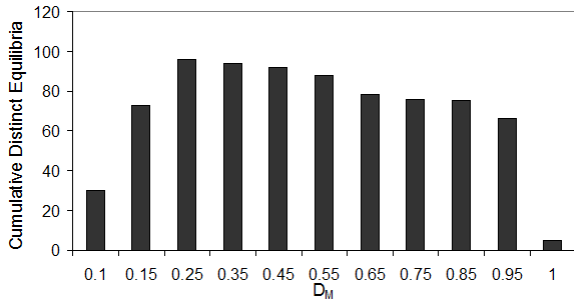


Fig. 4: Number of distinct equilibria (in 100 runs) versus maximum density.

topologically different are the distinct equilibria?

Figure 4 shows the number of distinct equilibria, after 100 runs with different initial topology, as function of D_M . The maximum for any value of D_M is the total number of runs (100). Notice two trends: an increase as D_M increases up to about 0.25, and then a slow decrease as D_M approaches 1. For the extreme value $D_M=1$, the number of equilibria drops suddenly to 5 equilibria. For very low values of D_M , the number of networks in each location is quite low. This decreases the number of provider choices that a small-expense network would have (and the majority of networks have small expense). This creates *oligopoly* in most locations, decreasing the number of distinct topologies that can result from different initial conditions.

For very large values of D_M (close to one), almost all networks are present in a certain location. This means that the cheapest provider is available to (almost) everyone, which creates an effective *monopoly*. This decreases the number of distinct equilibria as it is that cheapest provider that attracts almost all customers. For intermediate values of D_M there are several locations with multiple providers, increasing the sensitivity of the convergence process to the initial conditions.

Historically, the Internet started with low values of D_M (typically a national telecom provider in each country) and it has been moving towards larger values of D_M through the geographic expansion of international transit providers and the establishment of large Internet Exchange Points in few cities (Miami, London, Hong Kong etc) where 1000s of networks are present. If this trend continues, we may see in the future the appearance of effective monopolies in many parts of the world.

We now explore how topologically different these distinct equilibria are. We carry out two types of experiments. First, we use the same population of networks (with identical G , P and T) but with 100 different initial topologies. Second, we create 100 different population of networks (varying the parameters G , P and T). Figure 5 shows the results for four topological characteristics of the resulting internetwork: number of trees (or number of Tier-1 networks), maximum depth of any tree, average path length between any pair of networks, and number of providers at equilibrium. It is interesting that there is very little variability in all metrics in both types of experiments. In other words, even though we can get

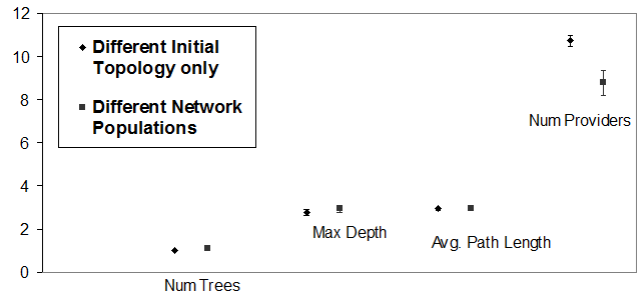


Fig. 5: Four topological characteristics (mean values with 95% C.I.) across distinct equilibria.

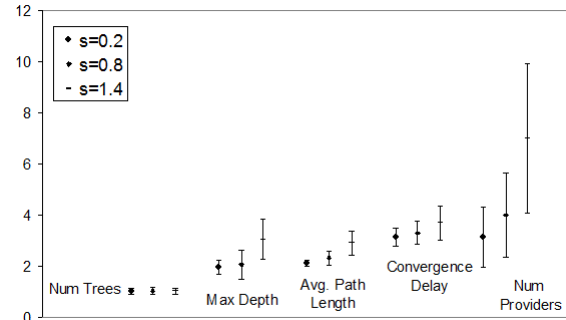


Fig. 6: Topological characteristics and convergence delay (mean values with 95% C.I.) for three exponent values in the expense distribution.

different equilibria, and we cannot predict who will be the provider of each network, certain *macroscopic* topological characteristics of the resulting internetworks are practically identical in all equilibria and thus they are predictable.

Impact of heterogeneity in expense distribution:

What is the role of the heterogeneity in the number of locations that networks are placed at? We vary the exponent of the corresponding Zipf distribution: a larger exponent increases the skewness of the expense distribution. Figure 6 shows the results for three exponents. With a more skewed distribution, the number of networks with large geographic expense is reduced. This increases the likelihood that networks with mid-sized expense are chosen as providers. Thus, the number of providers increases, and the same happens with the maximum depth and the average path length.

Impact of internetwork traffic matrix:

We investigated the role of the traffic matrix by varying the exponent in the Zipf distribution for the generated traffic V_G . Perhaps surprisingly, that parameter has minimal effect in the four topological characteristics we focus on here. The corresponding graph is not shown due to space constraints.

Impact of price-expense correlation:

It used to be the case that that global transit providers were more expensive than regional providers. Even though this may still be true in certain parts of the world, transit prices have dropped dramatically in the last few years in North

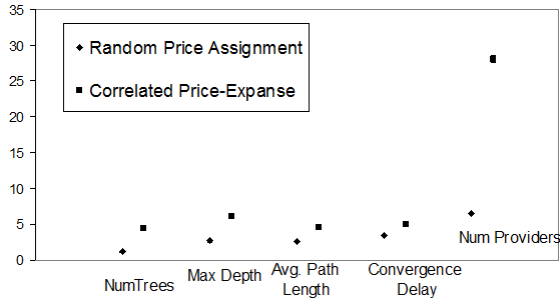


Fig. 7: Topological characteristics and convergence delay (mean values with 95% C.I.) for two transit pricing schemes.

America and Europe, and the price variation across different providers (at those regions) is also quickly diminishing.

What if the transit price vector P was positively correlated with the expanse vector $|G|$? Would the topological characteristics of the resulting internetwork be different compared to a random assignment of prices? Figure 7 shows that the correlated price assignment scheme results in a significantly larger number of providers, as well as in an increase in the number of Tier-1 networks, maximum depth, average path length, and convergence delay. The reason is not hard to see: in the correlated pricing scheme large-expanse networks are more expensive and so stubs would select cheaper networks as providers. Those lower-expanse providers however cannot reach every other location and so they need to connect to larger-expanse providers. This results in more providers and a deeper provider hierarchy. The recent trend towards smaller price differences across providers implies that the Internet is moving towards a lower number of providers and potentially a shallower hierarchy.

Characteristics of wealthiest network in equilibrium:

Who is the wealthiest network, in terms of transit revenues, in the resulting internetwork? Can we predict who that transit provider will be before the convergence process takes place?

The wealth of a network x can be defined as the total transit revenue from all customers of x minus the transit expense of x (paid to its provider $R(x)$)

$$W(x) = \sum_{y \in N: R(y)=x} P(x) V_P(y) - P(R(x)) V_P(x) \quad (5)$$

Recall that $V_P(i)$ is the traffic volume that a network i sends to and receives from its provider.

We measured the percentage of 100 runs in which the wealthiest network (at equilibrium) had certain characteristics. We distinguish between characteristics that we can observe before the convergence process (*a priori* characteristics, such as expanse or price) and characteristics that we can only measure at the equilibrium (*a posteriori*) characteristics, such as whether that network is Tier-1). Figures 8 and 9 show all these characteristics.

Note that in 100% of the runs, *the wealthiest network \hat{x} is the provider that managed to attract as customer the network with the largest generated traffic volume (the largest “content provider”).* With the largest content

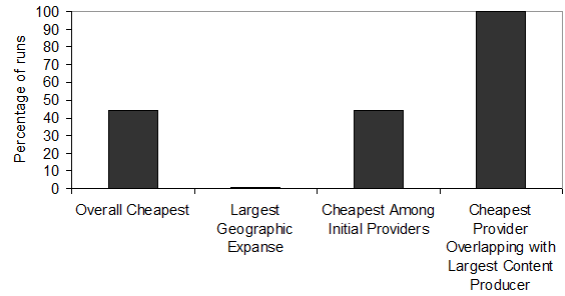


Fig. 8: *A priori* characteristics of wealthiest network.

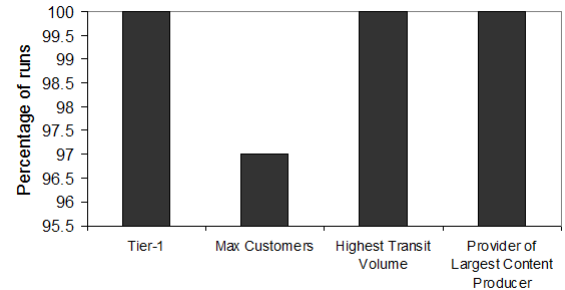


Fig. 9: *A posteriori* characteristics of wealthiest network.

provider as customer, and because V_G follows a skewed Zipf distribution, \hat{x} has enough transit traffic to force other providers to become its customers. It is important to note that \hat{x} is often not the cheapest provider, nor the network with the largest number of customers, and certainly not the network with the largest expanse. On the other hand, it always ends up being a Tier-1 network and it has the largest transit volume at the end.

This result has a significant practical implication: an ambitious transit provider should try to first increase its transit volume by attracting as customers (perhaps using special discounts and favorable contracts) the few major content providers. The transit provider does not need to have the lowest price or the largest expanse (even though its competitive advantage would improve by lowering its price or increasing its expanse).

5 Conclusions and future work

We proposed a simple model that captures the essence of provider selection in the Internet. The model captures the fact that most networks today select a provider based on price, rather than performance or other criteria. Additionally, the model captures two important and realistic constraints: a) collocation between a provider and its customers, and b) a provider should be larger, in terms of transit volume, than any of its customers. This second constraint adds positive feedback in the dynamics of the model. We proved that this provider selection process converges to a stable equilibrium and we investigated the characteristics of the resulting internetworks. In ongoing work, we are extending this model with multihoming and peering. Further, we are investigating the dynamics of network formation when networks can adjust their transit price and expand to new locations.

References

- [1] Peeringdb. <http://www.peeringdb.com>, September 2010.
- [2] ANSHELEVICH, E., DASGUPTA, A., KLEINBERG, J., TARDOS, E., WEXLER, T., AND ROUGHGARDEN, T. The price of stability for network design with fair cost allocation. In *FOCS '04: Proceedings of the 45th Annual IEEE Symposium on Foundations of Computer Science* (2004), pp. 295–304.
- [3] ANSHELEVICH, E., SHEPHERD, B., AND WILFONG, G. Strategic network formation through peering and service agreements. In *FOCS '06: Proceedings of the 47th Annual IEEE Symposium on Foundations of Computer Science* (2006), pp. 77–86.
- [4] BALA, V., AND GOYAL, S. A noncooperative model of network formation. *Econometrica* 68, 5 (2000), 1181–1230.
- [5] CHANG, H., JAMIN, S., MAO, Z., AND WILLINGER, W. An empirical approach to modeling inter-AS traffic matrices. In *Proceedings of the 5th ACM SIGCOMM conference on Internet Measurement* (2005), p. 12.
- [6] CHANG, H., JAMIN, S., AND WILLINGER, W. Internet connectivity at the AS-level: an optimization-driven modeling approach. In *MoMeTools '03: Proceedings of the ACM SIGCOMM workshop on Models, methods and tools for reproducible network research* (2003), pp. 33–46.
- [7] CHANG, H., JAMIN, S., AND WILLINGER, W. To peer or not to peer: Modeling the evolution of the Internet's AS-level topology. In *INFOCOM 2006. 25th IEEE International Conference on Computer Communications. Proceedings* (april 2006), pp. 1–12.
- [8] CORBO, J., JAIN, S., MITZENMACHER, M., AND PARKES, D. C. An economically-principled generative model of AS graph connectivity. In *IEEE INFOCOM Mini-Conference* (2009).
- [9] DHAMDHARE, A., AND DOVROLIS, C. Ten years in the evolution of the Internet ecosystem. In *Proceedings of the 8th ACM SIGCOMM conference on Internet measurement* (2008), pp. 183–196.
- [10] FABRIKANT, A., KOUTSOUPIAS, E., AND PAPADIMITRIOU, C. H. Heuristically optimized trade-offs: A new paradigm for power laws in the Internet. *International Colloquium on Automata, Languages, and Programming, Lecture Notes in Computer Science* (2002), 110–122.
- [11] FELDMANN, A., KAMMENHUBER, N., MAENNEL, O., MAGGS, B., DE PRISCO, R., AND SUNDARAM, R. A methodology for estimating interdomain Web traffic demand. In *Proceedings of the 4th ACM SIGCOMM conference on Internet measurement* (2004), pp. 322–335.
- [12] HOLME, P., KARLIN, J., AND FORREST, S. An integrated model of traffic, geography and economy in the Internet. *SIGCOMM Comput. Commun. Rev.* 38, 3 (2008), 5–16.
- [13] JOHARI, R., MANNOR, S., AND TSITSIKLIS, J. N. A contract-based model for directed network formation. *Games and Economic Behavior* 56, 2 (August 2006), 201–224.
- [14] LABOVITZ, C., IEKEL-JOHNSON, S., MCPHERSON, D., OBERHEIDE, J., AND JAHANIAN, F. Internet inter-domain traffic. *ACM SIGCOMM Computer Communication Review* 40, 4 (2010), 75–86.
- [15] WANG, X., AND LOGUINOV, D. Wealth-based evolution model for the Internet AS-level topology. In *INFOCOM* (2006).
- [16] YOOK, S.-H., JEONG, H., AND BARABASI, A.-L. Modeling the Internet's large-scale topology. *Proc Natl Acad Sci USA* 99, 21 (2002), 13382–6.
- [17] ZHOU, S. Understanding the evolution dynamics of Internet topology. *Physics Review E* 74, 1 (2006), 16–24.

A Proof of Theorem-1

The proof consists of two parts. First, we prove that after a finite number of rounds the population of networks \mathcal{N} is partitioned in a set of stubs, each of them making a stable provider selection, and a set of providers. Second, we prove that after the convergence of stubs, providers

also converge to a stable provider selection in a certain sequence within a finite number of rounds.

It is useful to split the transit volume $V_T(x)$ (at a given round) of a network to two parts: the transit volume $V_T^s(x)$ contributed by customers of x that are stubs, and the transit volume $V_T^p(x)$ contributed by customers of x that are providers. Obviously $V_T(x) = V_T^s(x) + V_T^p(x)$. To treat Tier-1 networks as any other network, we define that each Tier-1 network has a (virtual) provider Φ that does not belong in \mathcal{N} . No other network can select Φ as its provider. So, if x is a Tier-1 network, the statement “ x does not change provider” is equivalent to “ x remains a Tier-1 network”. This convention simplifies the presentation of the proof also because we do not need to consider the peering links between Tier-1 networks; all links are of the customer-provider type.

A.1 Convergence of stub population

By definition, a stub has no customers and so its transit volume is zero. So, based on (2), a stub cannot attract any customers. Once a network becomes a stub it will remain a stub. Thus, if x is a stub at some round t_k , $V_T(x) = 0$ at any round t_m with $m \geq k$.

If a network x is a stub at round t_k (t_k may be the initial round), it selects at t_k as provider the (unique) least-expensive network y that overlaps with x and that has $V_T(y) > 0$ at t_k . This provider selection, $R(x) = y$, will not change after round t_k , because $V_T(x)$ remains zero and there is no other provider than y that overlaps with x and has a lower price. Thus, the provider selection of a stub converges when a network first becomes a stub.

Further, because \mathcal{N} is finite and given that a stub cannot become a provider, after a finite number of rounds the set of stubs converges to a set \mathcal{S} , while the rest of the networks $\mathcal{N} - \mathcal{S}$ remain providers.

Corollary-1: After the set of stubs has converged to \mathcal{S} , the vector V_T^s remains constant.

A.2 Convergence of provider population

Suppose that the set of stubs converges to \mathcal{S} at some round t_k , and that the set of providers after that point consists of p networks. In the rest of the proof we only consider events after t_k . We order the p providers based on their transit volume in a sequence O :

$$O = \{O_i \in \mathcal{N} : V_T(O_i) > 0, V_T(O_1) > \dots > V_T(O_p)\} \quad (6)$$

Ties are broken deterministically based on a unique node identifier. So, O_1 represents the provider with the highest transit volume at any given round, and O_p is the provider with the lowest transit volume at any given round (if we need to clarify that we refer to round t_m we write $O_i(t_m)$).

Suppose that $x = O_i(t_m)$ and that x moves to another position at round t_{m+1} , i.e., $x = O_j(t_{m+1})$ with $i \neq j$. An oscillation will occur at position O_i if x returns to that position at any round after t_{m+1} . We next prove that oscillations cannot occur at any position of O . If this is true, then every provider will converge to a stable provider selection within a finite number of rounds, given that \mathcal{N} (and thus p) is finite.

Base case: An oscillation cannot occur at O_p

At any round after the convergence of stubs the provider

O_p has the following properties:

- $V_T^e(O_p) = 0$: Since the provider O_p has the minimum transit volume among providers, no provider will ever become its customer.
- Its transit volume does not change (follows from last property and from Corollary-1).
- It does not change provider (follows from last property, and because O_p has already selected the least expensive overlapping provider, or Φ).

Suppose now that $x = O_p(t_m)$ at some round t_m . x can move to $O_j \neq O_p$ at the next round t_{m+1} with $j < p$ if one of the following is true:

1. $V_T(x)$ increases at t_{m+1} so that $V_T(O_p) > V_T(O_j)$ with $j < p$. This is impossible because we showed that $V_T(O_p)$ remains constant.
2. $V_T(y)$, with $y = O_j(t_m)$ ($j < p$), decreases at t_{m+1} so that $V_T(y) < V_T(x)$.

So, we only need to consider the second case. At round t_{m+1} we have that

$$\begin{aligned} V_T(y) &= V_T^s(y) + V_T^e(y) \\ &< V_T(x) \\ &= V_T^s(x) + V_T^e(x) \\ &= V_T^s(x) \end{aligned} \quad (7)$$

and thus $V_T^s(y) < V_T^s(x)$. Thus, a provider at O_p can only be replaced by another provider that has lower transit volume due to stubs.

Now, for network x to return to O_p and replace y at a subsequent round, we must have that $V_T^s(x) < V_T^s(y)$. This is impossible however because the vector V_T^s remains constant after the convergence of stubs. Thus, no oscillation can occur at O_p . Within a finite number of rounds a certain provider will converge to that position, remaining the provider with the lowest transit volume from that point on.

Inductive step: An oscillation cannot occur at O_i $i < p$
Suppose that the providers $O_{i+1}, O_{i+2}, \dots, O_p$ have already converged at a round t_l for any i with $1 \leq i < p$. We now show that there can be no oscillation at O_i . In the rest of the proof we only consider events after t_l .

At any round after t_l the provider O_i has the following properties:

- Its transit volume does not change (any provider O_j with $j < i$ will not select O_i , and any provider O_j with $j > i$, as well as the stubs, have already converged to a provider).
- It does not change provider (follows from last property, and because O_i has already selected the least expensive overlapping provider, or Φ).
- Suppose that the provider O_i at a round t_m is a network x and let $V_T(x, t_m)$ be its current transit volume. At any subsequent round, even if x becomes O_j with $j < i$, the transit volume of x can only be larger or equal than $V_T(x, t_m)$ as every network in

its customer tree has converged - it can only have more customers in the future, not less. We refer to that transit volume as the *minimum transit volume* of x , denoted by $V_T^{min}(x)$.

Suppose now that $x = O_i(t_m)$. x can move to $O_j \neq O_i$ at the next round t_{m+1} with $j < i$ if one of the following is true:

1. $V_T(x)$ increases at t_{m+1} so that $V_T(O_i) > V_T(O_j)$ with $j < i$. This is impossible because we showed that $V_T(O_i)$ remains constant. Thus $V_T(x)$ remains at $V_T^{min}(x)$.
2. $V_T(y)$, with $y = O_j(t_m)$ ($j < i$), decreases at t_{m+1} so that $V_T(y) < V_T(x)$.

So, we only need to consider the second case. At round t_{m+1} we have that

$$V_T(y) = V_T^{min}(y) < V_T(x) = V_T^{min}(x) \quad (8)$$

Thus, a provider at O_i can only be replaced by another provider that has a lower minimum transit volume.

Now, for network x to return to O_i and replace y at a subsequent round, we must have that $V_T^{min}(x) < V_T^{min}(y)$. This is impossible however because the vector of minimum transit volumes remains constant after t_l . Thus, no oscillation can occur at O_i . Within a finite number of rounds a certain provider will converge to that position. We have proved that within a finite number of rounds both stubs and the providers have converged to a provider selection (potentially Φ), and thus the topology of the internetwork has also converged.

B Customer-Provider traffic volume comparison

This section focuses on the question: *is it realistic to assume that the transit volume of a provider is larger than the transit volume of its customers?* We rely on two sources of data to examine this question. First, traffic volumes reported by few thousands of ASes at PeeringDB [1]. Unfortunately these traffic volumes represent total traffic volume ($V_T + V_G + V_C$), not only transit traffic volume (V_T), and they are often reported as (wide) ranges - we use the upper bound of each range. Second, we rely on the reported customer-provider relations that have been inferred by Dhamdhere from BGP routing tables [9]. With this information, we identify 2500 customer-provider pairs that have reported their traffic volumes to PeeringDB. In 90% of these pairs, the customer's traffic volume is significantly less (typically less than half) of its provider's traffic volume. In 9.5% of these pairs, the customer's traffic volume is larger than the provider's but by less than 20%. Finally, in the remaining 0.05% the customer traffic volume is much higher than the provider's. In such exceptional cases we suspect a large difference between total traffic volume and transit traffic volume. For instance, the China Cache Network reports its traffic as 300-500Gbps while its provider TransitRail reports 100Gbps. It is likely that most of the traffic in the former network is locally generated or consumed, not transit. We should also note that multihoming can further decrease the effective traffic volume of a customer relative to its provider.

C Traffic matrix parameterization

Our computational experiments use the Zipf distribution (with exponent 0.8) to model the traffic volume generated by each AS. With this distribution, 0.1% of the ASes generate nearly 28% of the total traffic. This behavior is consistent with the findings of [5] and [11] which report that traffic produced by high-ranking ISPs and content providers follows a Zipf distribution. Additionally, it has been recently reported that the top-30 ASes (out of nearly 30,000, approximately 0.1%) produce 30% of the Internet traffic [14].