Routing and Peering in a Competitive Internet

Ramesh Johari
Stanford University
rjohari@stanford.edu

John N. Tsitsiklis
MIT
jnt@mit.edu

***Abstract*— Today's Internet is a loose federation of independent network providers, each acting in their own self interest. In this paper, we consider some implications of this economic reality. Specifically, we consider how the incentives of the providers might determine where they choose to interconnect with each other; we show that for any given provider, determining an optimal placement of interconnection links is generally NP-complete. However, we present simple solutions for some special cases of this placement problem.**

We also consider the phenomenon of nearest-exit, or "hot-potato," routing, where outgoing traffic exits a provider's network as quickly as possible. If each link in a network is assessed a linear cost per unit flow through the link, we show that the total cost of nearest exit routing is no worse than three times the optimal cost.

I. Introduction

The modern Internet is a network owned by a loosely connected federation of independent network providers. Fundamentally, the objectives of each provider are not necessarily aligned with any global performance objective; rather, each network provider will typically be interested in maximizing their own monetary profits. This profit maximizing, self interested behavior has important ramifications for the performance of the network.

Our paper is concerned with connections between network providers. Most economic relationships between two providers may be classfied into one of two types: *transit*, and *peer*. Provider A provides transit service to provider B if B pays A to carry traffic originating within B and destined elsewhere in the Internet (either inside or outside A's network). In this paper, we will instead be primarily interested in peer relationships. In peering agreements, one or more bidirectional links are established between two providers A and B. In contrast to transit service, where traffic is accepted regardless of the destination, in a peering relationship provider B will only accept traffic from A that is destined for points *within* B, and vice versa. Importantly, such agreements are typically negotiated without any transfer of money between the two parties involved, because two providers will only choose to become peers if they are roughly the same size and have similar amounts of traffic to send to each other. Peering agreements are typically seen among the "tier 1" or "backbone" providers at the top level of the Internet hierarchy, who provide national and global connectivity to their customers. (For further details, see [1], [2], [3].)

When two providers form a link connecting their networks (which we shall refer to as a *peering link*), the traffic flowing across that link incurs a cost on the network it enters. Such a cost may be felt at the time of network provisioning, or on a faster timescale as network congestion rises. We will abstract away from making any specific assumptions about the nature of the network costs in our models, with the understanding that these two interpretations are possible.

Consider a situation, then, where providers S and R are peers. Each of these providers will typically have some amount of traffic to send to each other. However, for the purposes of this paper, we will separate the roles of the two providers as sender and receiver; this will allow us to focus on the different incentives that exist in each role. In particular, we suppose that provider S has some amount of traffic to send to destinations in provider R's network. If we assume the only costs incurred are network routing costs, then because the peering relationship includes no transfer of currency, provider S has an incentive to force traffic into provider R as quickly and cheaply as possible. This phenomenon is known as "nearest exit" or "hot potato" routing (see [4]). In practice, for example, traffic travelling from an AT&T subnetwork in Boston to a computer on a Sprint subnetwork in Chicago will enter Sprint's network at a peering point in Boston, then traverse links owned by Sprint until arriving at the destination in Chicago.

We will consider two problems that arise due to the phenomenon of nearest exit routing. First, suppose again that a provider S has agreed to peer with provider R. Given the distribution of traffic flowing from S to R (across all origins in S and destinations in R), both providers assume at the outset that S will use nearest exit routing. We then ask: where would R and S like to establish peering links? This is a question that might be asked, for example, when providers first establish a peering agreement and need to physically construct the links connecting their networks. The decision of where to place these links is, of course, intimately connected to the distribution of the traffic flowing between them. As we will see in Section II, determining which placements are most preferred by the sender and receiver is, in general, computationally intractable. Nevertheless, special cases where both providers have a linear or tree topology can be analyzed, and the link placements most preferred by the sender and receiver can be determined. In particular, we are able to show that when both providers have a linear network, under some symmetry conditions on the traffic, there exists a unique peering point placement which will *simultaneously* satisfy both providers. This leads to the important conclusion that at least in this special case, it is possible to identify the expected outcome of the peering point placement process between the providers.

In Section III, we address the second key problem which arises due to nearest exit routing. We will assume that peering links have already been established between the two providers S and R. Given that the sender S is using nearest exit routing, we do not, in general, expect the resulting routing of traffic from S to R to resemble an "optimal" routing, according to some network cost metric chosen *a priori*. Indeed, we will show that if network cost is measured by assessing a cost per unit flow traversing each link, and if we compare nearest exit routing to shortest path routing, then when both sender and receiver share the same topology, we can expect the nearest exit routing cost to be no worse than three times the optimal (shortest path) routing cost. This result follows the spirit of previous work by Koutsoupias and Papadimitriou [5] and Roughgarden and Tardos [6] in bounding the *cost of anarchy*: that is, when selfish agents act in their own interests, what is the resulting shortfall in efficiency relative to some well-defined optimum? In this language, the cost of anarchy in our problem is a factor of three; see Theorem 3.

Our research forms part of a growing body of work on the implications of the current Internet interconnection paradigm. Much of this work has been focused on the protocol level, particularly on the failings of the BGP protocol used for interdomain routing; see, e.g., [1], [3], [7], [8]. Recently, however, several efforts at understanding the impact of provider economics at network design have also begun, including results by [9], [10], [11].

The analysis of these papers suggests that our analytical models may no longer assume that the Internet as a whole acts to optimize some network-wide performance objective. Rather, the actions of the individual network providers will typically lead to quite a different outcome; and quantifying this difference in more general networks remains an important challenge.

II. The Peering Point Placement Problem

In this section, we will investigate the creation of interconnection links between network providers, given that they have already chosen to peer with each other. As discussed in the Introduction, we will assume two network providers S and R, and that S is sending traffic to the receiver R; further we will assume that S is using nearest exit routing. Note that, in general, both S and R will be sending and receiving traffic; however, to isolate the effects of sending and receiving traffic, we will assume only unidirectional traffic flow. We make the further assumption that S and R share the exact same network topology. While this is a strong assumption, it is perhaps founded on the fact that we expect our model to apply to the tier 1, backbone level of the Internet, where most providers control national and international networks. These networks will have many common nodes (major cities, for example), and thus we might reasonably expect some similarity in their topologies.

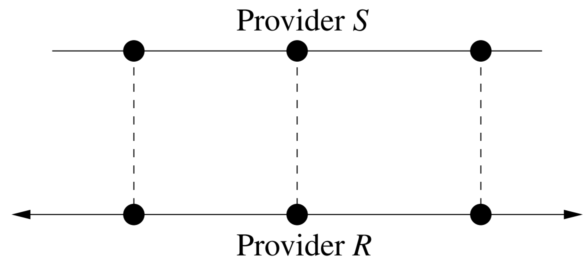

Fig. 1. Two overlapping networks. The vertical lines represent links at peering points between S and R; they are drawn as dashed lines because in this model we assume that traffic experiences no cost travelling across a peering point.

In the first subsection, we assume that both providers share the topology of a linear network; and in the second subsection, we assume both providers share a tree topology. Under certain assumptions, we compute optimal placements for sender and receiver. Nevertheless, for general topologies, computing the peering link locations most preferred by the sender and the receiver is NP-complete, as we briefly describe in Section II-C.

A. Linear Networks

We consider a model consisting of two network providers. The sending provider, S, controls a line segment of length $2l$, while the receiving provider, R, controls a line; we identify S topologically with the interval $[-l,l] \subset \mathbb{R}$, and we identify R with $\mathbb{R}$. The line segment S "overlays" the line R, as depicted in Figure 1.

We make two assumptions on the nature of the traffic being sent from S to R. First, we assume that given an origin $x \in S$, traffic originating at x is destined for a randomly chosen destination $y \in R$, chosen according to a probability density $f(y|x)$. We make the assumption that $f(y|x) = g(y-x)$, where g is a probability density function such that $g(z) = g(-z)$. Intuitively, each origin $x \in S$ is sending to its mirror image in R, but with some random, symmetric "spread" determined by the density g; notice that *every* origin $x \in S$ sees a spread determined by exactly the same density g. We emphasize here that the only assumption we are making on g is that it be symmetric about the origin. In fact, g may even correspond to a distribution which is symmetric, but does not possess a density; all the results here will continue to hold. In this case, letting G be the distribution corresponding to g, we simply require that $G(-z) = 1 - G(z)$ for all $z \in \mathbb{R}$.

The second assumption we make is that each origin $x \in S$ has exactly the same total amount of traffic to send into R. Formally, we assume that S contains a total amount of traffic T to be sent to the receiver R, and that the total amount of traffic originating in an interval $[a,b] \subset S$ is given by $(b-a)T/2l$; in other words, the origin of any particular unit of traffic is uniformly distributed across the interval $[-l,l]$. In the discussion that follows, we will only be interested in the *expected* cost incurred by traffic flowing from S to R.

We consider the problem of placing at most n *peering points* between providers S and R, i.e., points where traffic exits S and enters R. We assume that n is given, so that the maximum number of points to be placed has been agreed upon *a priori*. We allow the peering points to be located anywhere in the region $[-l,l]$. Each peering point is really two points: an exit point $p \in S$, and an entry point, its mirror image $p \in R$. Note that this is an important restriction; traffic may only enter the provider R at exactly the same point at which it exits provider S.

We will consider two placement problems. First, we will be interested in determining which placement of peering points is most preferred by the sender; next, we consider the same problem for the receiver. Note that entry and exit at peering points is assumed to be costless. Sender S will thus attempt to exit traffic at the peering point nearest to an origin, and receiver R will use a shortest path from the peering point to the destination. The sender wishes to place peering points to minimize the expected distance from any origin to a peering point; the receiver wishes to minimize the expected distance from peering point to destination, *knowing that the sender will use nearest exit routing*. The following theorem shows that there exists a single peering point placement which is optimal for both the receiver and the sender. The proof is omitted; details may be found in [12].

Theorem 1: Let $p_i = -l + (2i-1)l/n$, for $i = 1,\ldots,n$; i.e,. the n peering points are placed symmetrically about 0, a distance $2l/n$ apart from each other. Then, the peering point placement identified by $p_1,\ldots,p_n$ is the unique choice which is simultaneously optimal for both the sender and the receiver.

The theorem demonstrates that, in this special case, the interests of both the receiver and sender are aligned. This highlights the interesting point that if these two providers were in a bargaining procedure to determine the placement of peering points between them, there is a predetermined, easily computed outcome which can be shown to be provably optimal for both providers.

B. Trees

We now consider a model consisting of two network providers, each managing a tree: T_1 will represent the sending provider, and T_2 will represent the receiving provider. Each tree consists of k levels (not including the root node, which by convention is at level 0), with a fan-out of m—i.e., all nodes except the leaves have m children. Thus each tree consists of $N = (m^{k+1}-1)/(m-1)$ nodes. We assume for the moment that the edges of the tree all have unit length; this assumption will be relaxed later.

We first outline our traffic model. We assume that each leaf node in T_1 has 1 unit of traffic to send to a randomly chosen leaf node in tree T_2. In randomly choosing the destination, we fix a parameter p, $0 \le p \le 1$, which determines how "far" the destination is. Note that the distance travelled from origin to destination is determined by the first node i in the tree such that the subtree rooted at i has both origin and destination as a leaf; this is the *common subtree* of the origin and destination. When p is small (resp. large), we will find that this common subtree typically occurs at a very low (resp. high) level in the tree.

This behavior is described formally as follows. Given a leaf node i in the tree, let $P(i)$ denote the parent of node i, and $P^l(i)$ denote the $(k-l)$-level parent of node i; i.e., $P^l(i)$ is the node at level $k-l$ in the tree, such that the subtree rooted at $P^l(i)$ contains i as a leaf. Denote the origin node by i_o, and the destination by i_d. With probability $1-p$, $i_d = i_o$. With probability $p(1-p)$, the destination is chosen uniformly at random from among the $m-1$ siblings of the origin, in the subtree rooted at $P(i)$; and in general, for $1 \le l < k$, with probability $p^l(1-p)$, the destination is chosen uniformly at random from among the $m^l - m^{l-1}$ leaf nodes for which $P^l(i)$ is the root of their common subtree with i_o. Finally, with probability p^k, the destination is chosen uniformly at random from among the $m^k - m^{k-1}$ leaf nodes for which the root node of the tree is also the root of the common subtree with i_o.

Given the traffic distribution, we may analyze the optimal placement of peering points for both sender and receiver. For this section, we will assume that the providers may place an *arbitrary* number of peering points. Given this ability, the sending provider would prefer to place m^k peering points at the lowest level—level k—of the tree. Under nearest exit routing, this leads to zero routing cost for the sending provider.

The situation for the receiving provider is more interesting. It is possible to show that there exists an optimal level $l^*(p)$, depending on the parameter p, which minimizes the routing cost; that is, the receiving provider would wish to place $m^{l^*(p)}$ peering points at level $l^*(p)$ of the tree. We omit the details of the derivation, which can be found in [12]. The optimal level $l^*(p)$ is given by:

$$l^*(p) = \begin{cases} k, & p \in [0, 1/2]; \\ k-i, & p \in [(1/2)^{1/i}, (1/2)^{1/(i+1)}],\ 1 \le i \le k-1; \\ 0, & p \in [(1/2)^{1/k}, 1] \end{cases}$$

Note that at the boundary points, there are two possible optimal levels the provider may choose from. Further, the expression for $l^*(p)$ is *independent* of the fan-out m; and the analysis may be extended to the case where not all links have unit length. Generally, the two providers will not agree on where to place peering points in this model: the sender always prefers to place peering points at the lowest level of the tree, whereas the receiver prefers $l^*(p)$, which may or may not be the lowest level.

C. In General

Under some assumptions on the structure of traffic and topology, the previous sections have provided insight into the placements most preferred by sender and receiver. In general, computing these optimal placements is computationally intractable, as we now show.

We assume two providers, and identify each with the same graph: $S = R = (N,A)$. We assume that if $(i,j) \in A$, then $(j,i) \in A$; thus any link from i to j is paired with a return link from j to i. To distinguish the two graphs S and R notationally, we denote sender and receiver by subscripts S and R respectively: thus, N_S represents the set of nodes in the sending network S, etc. The providers are to place a collection of n peering points, labeled by $\mathbf{p} = (p_1, \ldots, p_n)$. Formally, this means the network as a whole will be a graph $G = (N_G, A_G)$ consisting of nodes $N_G = N_S \bigcup N_R$, and arcs $A_G = A_S \bigcup A_R \bigcup \{(p_{1,S}, p_{1,R}), \ldots, (p_{n,S}, p_{n,R})\}$. Each of the last n arcs link from a peering point $p_{i,S} \in N_S$ to a corresponding $p_{i,R} \in N_R$. Traffic may travel from S to R only at these peering points.

S, the sending provider, has some amount of traffic to send to R. The amount of traffic originating at a source $s \in N_S$ and terminating at destination $d \in N_R$ is given by $\mathbf{b} = b_{sd}$; we write $b = (b_{sd})$ for the vector of source-destination flows. Given the peering point locations $\mathbf{p} = (p_1, \ldots, p_n)$, the set of routes available to a source-destination pair (s,d) is given by $P^{\mathbf{p}}(s,d)$; each element $r \in P^{p}(s,d)$ is a path in G consisting of a path from $s \in N_S$ to some $p_{i,S}$, followed by the link $(p_{i,S}, p_{i,R})$, followed in turn by a path from $p_{i,R}$ to d. We let y_r denote the flow sent along route r.

Because the sending and receiving networks divide the responsibility of carrying traffic from s to d, we define two new sets of paths. First, let $P_S(s, p_i)$ be the set of all paths available to the sender to route traffic from $s \in S$ to $p_i \in S$; if $r \in P_S(s, p_i)$, and $(i,j) \in r$, we require that $(i,j) \in A_S$. Similarly, we define $P_R(p_i, d)$ as the set of paths available to the receiver to route traffic from $p_i \in R$ to $d \in R$; again, if $r \in P_R(p_i, d)$, and $(i,j) \in r$, we require that $(i,j) \in A_R$.

We will assume that link (i,j) has a length c_{ij}; the cost of sending f_{ij} units of flow on link (i,j) is $c_{ij} f_{ij}$. We assume that distances are symmetric, in the sense that $(i,j) \in A_S$ and $(i,j) \in A_R$ both have length c_{ij}, and we will denote the vector of link lengths by $\mathbf{c} = (c_{ij})$. Also, we assume that given the peering point locations $\mathbf{p}$, all links $(p_{i,S}, p_{i,R})$ have zero length. (Note that if the placement problems are NP-complete with these assumptions, they remain so without the assumptions.) We now define the sender's placement problem:

SenderPlacement$(N, A, \mathbf{b}, \mathbf{c}, n, K)$:

Does there exist a peering point placement $\mathbf{p} = (p_1, \ldots, p_n)$ such that the value of the following optimization problem is less than or equal to K?

$$\text{minimize} \sum_{(i,j) \in A_S} c_{ij} f_{ij} \tag{1}$$

$$\text{subject to} \sum_{p_k} \sum_{r \in P_S(s,p_k)} y_r = \sum_{d \in R} b_{sd}, \ \forall \ s \tag{2}$$

$$\sum_{(s,p_k)} \sum_{r \in P_S(s,p_k):(i,j) \in r} y_r = f_{ij}, \ \forall \ (i,j) \in A_S \tag{3}$$

$$y_r \geq 0. \tag{4}$$

The first constraint ensures all traffic from a fixed source $s \in S$ is routed to a peering point. The second constraint simply identifies the link flow f_{ij} as the sum of flows from routes using that link. According to this formulation the objective of the sender is to use nearest exit routing to send all the flow given by $\mathbf{b}$ out of S into R.

We may similarly define the receiver's placement problem. Let $b'_{p_i,d}$ be the traffic entering at p_i destined for d seen by the receiver R, given that the sender is using nearest exit routing. We note here that the traffic matrix $\mathbf{b}'$ may not be uniquely determined, as there may not be a unique solution to the optimization problem (1)-(4). This technical issue does not play a role in any results presented here, so we may simply assume, for example, that the receiver randomly chooses an optimal solution to the sender's problem (1)-(4). The receiver's placement problem is then:

ReceiverPlacement$(N, A, \mathbf{b}, \mathbf{c}, n, K)$:

Does there exist a peering point placement $\mathbf{p} = (p_1, \ldots, p_n)$ such that the value of the following optimization problem is less than or equal to K?

$$\text{minimize} \sum_{(i,j) \in A_R} c_{ij} f_{ij} \tag{5}$$

$$\text{subject to} \sum_{r \in P_R(p_k,d)} y_r = b'_{p_k,d}, \ \forall \ (p_k, d) \tag{6}$$

$$\sum_{(p_k,d)} \sum_{r \in P_R(p_k,d):(i,j) \in r} y_r = f_{ij}, \ \forall \ (i,j) \in A_R \tag{7}$$

$$y_r \geq 0. \tag{8}$$

The receiver sees the input traffic matrix determined by nearest exit routing at the sender; this traffic is then routed using shortest path routing to the destination.

We prove the following result using a reduction from *VertexCover*; see [12] for details.

Theorem 2: The problems *SenderPlacement* and *ReceiverPlacement* are NP-complete.

We note here that the computational complexity result of Theorem 2 supports an informal claim made by Awduche et al. [13]. In that paper, the authors formulate the optimal peering point location as an integer program, related to the formulation discussed here, and suggest some traditional approximation techniques that might be used by network providers. Our result shows formally that solving the optimal peering point location problem is analytically intractable in general. Nonetheless, our discussion of linear networks and trees shows that for networks with special structure, it is indeed possible to evade the negative conclusion of this theorem.

III. Nearest Exit Routing vs. Optimal Routing

The previous section considered the problem of where peering points should be placed, given that two providers have decided to peer with each other. In this section, we consider the effects of these peering decisions on routing:

namely, given that two providers have established a set of peering points with each other, how inefficient is the resulting routing of traffic?

We continue to use the notation and model of Section II-C: two network providers S and R share the same topology. Now, however, the peering point vector $\mathbf{p} = (p_1, \ldots, p_n)$ will be assumed *fixed*. Given this set of peering point locations, we will try to investigate the nature of the optimization problems solved by the two providers, given by (1)-(4) for the sender and (5)-(8) for the receiver.

Traditionally, when one network manager controlled the whole network G, routing would be performed according to a global cost minimization problem (see, e.g., [14]):

$$\text{minimize} \sum_{(i,j)\in A_G} c_{ij} f_{ij} \tag{9}$$

$$\text{over} \sum_{r\in P^{\mathbf{p}}(s,d)} y_r = b_{sd}, \quad \forall\ (s,d) \tag{10}$$

$$\sum_{(s,d)} \sum_{r\in P^{\mathbf{p}}(s,d):(i,j)\in r} y_r = f_{ij}, \quad \forall\ (i,j)\in A_G \tag{11}$$

$$y_r \geq 0. \tag{12}$$

Recall that A_G is the global set of arcs, and $P^{\mathbf{p}}(s,d)$ represents the set of paths available from an origin $s \in S$ to a destination $d \in R$, given the set of peering point locations identified by $\mathbf{p}$.

The problem defined by (9)-(12) corresponds to an optimization which minimizes the *sum* of the routing costs experienced by the sender and the receiver. Of course, when sender and receiver act independently (according to the optimization problems (1)-(4) and (5)-(8)), there is no reason to expect them to arrive at the globally optimal solution, and indeed, this is generally not the case. However, we may analytically compare the routing cost of nearest exit routing with globally optimal routing. To emphasize the assumptions, we note here that we have assumed the two networks R and S are identical, and that the two have identical cost functions for their links. We have also assumed a fixed, but arbitrary, placement of n peering points. We then have the following theorem.

Theorem 3: Suppose that $S = R$, and both have identical lengths $c_{ij} \geq 0$ for their links. Then given any placement of n peering points, the cost of nearest exit routing is no more than three times the cost of optimal routing. Further, for all sufficiently small $\varepsilon > 0$, there exist networks such that the cost of nearest exit routing is at least $3 - \varepsilon$ times the cost of optimal routing.

Proof. The proof uses a graphical argument; refer to Figure 2. Recall that because costs are linear, we may treat the link cost coefficient c_{ij} as the length of link (i,j). Suppose that 1 unit of traffic must travel from $s \in S$ to $d \in R$, and the optimal (shortest path) route is the solid black line which passes through p_{OPT}. Let the total distance (and hence the total cost) travelled from s to d along this optimal path be r.

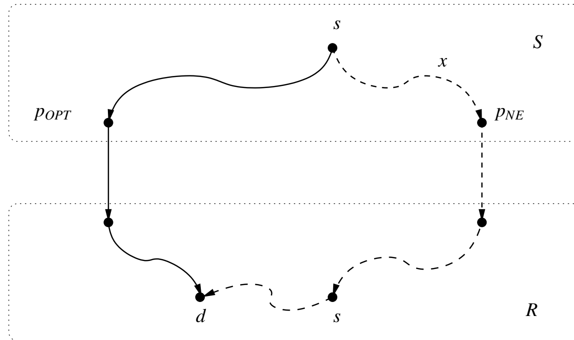

Fig. 2. Proof of Theorem 3: Nearest exit routing cost is at most three times optimal routing cost.

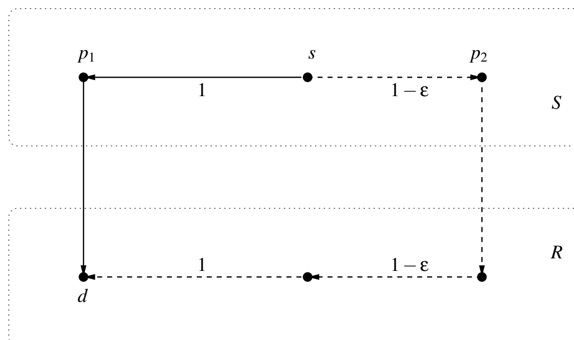

Fig. 3. Proof of Theorem 3: Example where the upper bound is tight.

Now consider the route depicted by the dashed line, passing through p_{NE}. The sender has determined that the nearest peering point to s is p_{NE}; the total distance between them is denoted x. Note that $x \leq r$, since by definition p_{OPT} must be further from s than p_{NE}.

Once traffic destined for d enters R at p_{NE}, the receiver would route it to d at minimum cost. However, consider the following route available to the receiver: first send traffic from p_{NE} back to s, using the same (outgoing) route as used by S to send traffic from s to p_{NE}. This distance is $x \leq r$. After traffic reaches $s \in R$, use a shortest path within R from s to d. This distance must be less than r, since r is the length of shortest path from s to d with the additional constraint that the route must include a peering point. So if the receiver uses this route, from p_{NE} to s to d, then the total cost incurred by the receiver is less than or equal to $2r$. Since the sender incurs a cost no higher than r in sending traffic to p_{NE}, and peering links have zero cost, the total cost incurred in sending traffic from s to d is less than or equal to $3r$. By linearity of the cost functions, this bound may be extended to any arbitrary traffic matrix (b_{sd}), since for each source destination pair (s,d) this bound holds.

We finally show that this bound is tight. Consider the network depicted in Figure 3. The sender has one unit of traffic to send from s to $d = p_1$. Since the distance to peering point p_2 is $1 - \varepsilon$, the sender chooses this exit; the receiver then incurs a cost of $2 - \varepsilon$ in routing the traffic to d from p_2. The total cost, therefore, of nearest exit routing is $3-2\varepsilon$; on the other hand, note that the optimal choice is to send traffic from s to p_1, incurring a cost of 1. Thus the nearest exit cost may be made arbitrarily close to 3 times the optimal cost by a sufficiently small choice of ε. □

Notice that the cost experienced on link (i,j) is linear in the flow f_{ij}, and given by $c_{ij}f_{ij}$. In general, we may define a cost function $C_{ij}(f_{ij})$, which we assume to be convex and increasing, but not necessarily linear; such a framework is discussed by Bertsekas and Gallager [14]. However, note the essential importance of linearity in the current setting, allowing us to decouple individual source-destination pairs from each other; in a general network where costs are nonlinear, any analysis must consider the interaction of flows sharing the same link. In fact, relaxing any of the assumptions in the theorem cause the conclusion to fail; counterexamples exist not only for the constant multiple 3, but for any constant multiple of optimal cost. One may easily construct such cases when the networks are not symmetric (i.e., $S \neq R$), when they do not share common cost functions, or when link costs are allowed to be nonlinear.

Notice that because we have assumed link costs to be linear in flow, the analysis continues to apply even if both providers are sending traffic to each other and receiving traffic from each other. In fact, the result continues to apply even if there are multiple network providers, all peering with each other, and sharing the same topology and link costs. The analysis is done on a route-by-route basis, so these extensions do not affect the final result.

One way to refine our model is to assume, for example, that the link cost c^R_{ij} of link (i,j) in provider R's network and the link cost c^S_{ij} of link (i,j) in provider S's network satisfy $c^R_{ij} \le \beta c^S_{ij}$, for some $\beta > 0$ which does not depend on the link (i,j). In this case, the proof of the theorem above would show that nearest exit routing cost is no worse than $1+2\beta$ times the optimal routing cost; in the setting of our theorem, $\beta = 1$. Thus, through a simple change, we may take into account some degree of heterogeneity in the link costs of the various backbone providers, and relate this to the efficiency loss relative to optimal routing.

IV. Conclusion

This paper has discussed two important issues which arise in today's Internet between competing network providers: First, where to place interconnection links; and second, the performance of the resulting traffic routing. For both problems, we start from the assumption that the network providers act in their own self interest. This selfish behavior impacts our analysis in two different ways. When placing peering points between each other, the key problem is that providers must agree simultaneously on the placement. This poses an important practical challenge: our computational complexity result shows that aligning the interests of the providers with each other will require tractable approximations which still capture their incentives accurately. The second impact of selfish behavior is in a loss of efficiency, or "cost of anarchy," as discussed in the Introduction. We examine this cost in the context of interdomain routing.

V. Acknowledgements

This research was carried out while the authors were with the Laboratory for Information and Decision Systems, MIT, Cambridge, MA. This work was partially supported by a National Science Foundation Graduate Research Fellowship and by the Defense Advanced Research Projects Agency under the Next Generation Internet Initiative. The authors would like to acknowledge helpful discussions with Tim Griffin, Frank Kelly, and Jennifer Rexford, as well as with seminar participants at the Institute for Pure and Applied Mathematics Workshop on Large Scale Communication Networks.

References

[1] L. Gao and J. Rexford, "Stable Internet routing without global coordination," *IEEE/ACM Transactions on Networking*, vol. 9, no. 6, pp. 681–692, 2001.

[2] W. Norton, "Internet service providers and peering," 2000, available on request from: http://www.equinix.com/press/whtppr.htm.

[3] T. G. Griffin and G. T. Wilfong, "An analysis of BGP convergence properties," *ACM SIGCOMM Computer Communication Review*, vol. 29, no. 4, pp. 277–288, 1999.

[4] J. S. Marcus, *Designing Wide Area Networks and Internetworks: A Practical Guide*. Addison Wesley, 1999.

[5] E. Koutsoupias and C. Papadimitriou, "Worst-case equilibria," in *Proceedings of the 16th Annual Symposium on Theoretical Aspects of Computer Science*, 1999, pp. 404–413.

[6] T. Roughgarden and E. Tardos, "How bad is selfish routing?" *Journal of the ACM*, vol. 49, no. 2, pp. 236–259, 2002.

[7] T. G. Griffin, F. B. Shepherd, and G. T. Wilfong, "The stable paths problem and interdomain routing," *IEEE/ACM Transactions on Networking*, vol. 10, no. 2, pp. 232–243, 2002.

[8] C. Labovitz, A. Ahuja, A. Bose, and F. Jahanian, "Delayed Internet routing convergence," *IEEE/ACM Transactions on Networking*, vol. 9, no. 3, pp. 293–306, 2001.

[9] L. He and J. Walrand, "Dynamic provisioning of service level agreements between interconnected networks," in *Proceedings of the 40th Allerton Conference on Communication, Control, and Computing*, 2002.

[10] B. Hajek and G. Gopalakrishnan, "Do greedy autonomous systems make for a sensible Internet?" 2002, presented at the Conference on Stochastic Networks, Stanford University.

[11] J. Feigenbaum, C. Papadimitriou, R. Sami, and S. Shenker, "A BGP-based mechanism for lowest-cost routing," in *Proceedings of the 21st Symposium on Principles of Distributed Computing*, 2002, pp. 173–182.

[12] R. Johari and J. N. Tsitsiklis, "Routing and peering in a competitive Internet," MIT Laboratory for Information and Decision Systems, Publication 2570, 2003.

[13] D. Awduche, J. Agogbua, and J. McManus, "An approach to optimal peering between autonomous systems in the Internet," in *Proceedings of the 7th International Conference on Computer Communication and Networks*, 1998, pp. 346–351.

[14] D. P. Bertsekas and R. G. Gallager, *Data Networks*, 2nd ed. Upper Saddle River, New Jersey: Prentice Hall, 1992.