Routing and Peering in a Competitive Internet

Ramesh Johari (rjohari@mit.edu)
John Tsitsiklis (jnt@mit.edu)

January 30, 2003

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. In the general case, with nonlinear cost functions, we consider allowing providers to charge each other for flow on their links. We study the efficiency of such a scheme, and arrive at bounds on the worst possible efficiency loss.

Traditional analyses of routing in data networks have assumed that the network is owned by a single operator. Typically, the network operator attempts to achieve some overall performance objective—e.g., low average delay or low packet loss rates. Such analyses persisted throughout the growth of the Internet (see, for example, Bertsekas and Gallager [2]), since the network's development was largely maintained and directed by the government and academia: it could be reasonably assumed that improving performance of the network was a goal in the best interests of all parties.

However, since the National Science Foundation ended its involvement in shaping the Internet in 1995, the Internet has taken on a very different shape. Today, it 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—we can no longer expect that engineering a network properly can be independent of the economic realities of the implementation of that network.

To understand the economic incentives driving the actions of network providers, we must first understand the structure of the interconnections they form with each other. Most 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 such an agreement, provider A accepts the responsibility of carrying any traffic entering from B across their interconnection link.

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. This follows logically from the situations which give rise to peering agreements. 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 Marcus [13], Gao and Rexford [4], Norton [16], and Griffin and Wilfong [6].)

To understand the performance impact of peering relationships, we must first understand how these relationships affect the providers. 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: in order to meet the quantity of traffic entering through a peering link, a provider may need to increase its network capacity. A network provider may also see a cost for entering traffic on a faster timescale; when the amount of incoming traffic increases, congestion on the network increases, and this leads to increased operating and network management costs. 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.

An important driving force behind the network providers' interconnection agreements is the value gained by end users through that interconnection. A home cable modem user, for example, gains value from the fact that the entire Internet is visible to him, not just the segment of the network to which he is connected. A complete characterization of the incentives behind interconnection should therefore include not only network costs, but also the benefits of being able to provide a more attractive service to end users. This approach to modeling interconnection agreements has been pursued extensively in the economics literature; see, e.g., Laffont et al. [12], Mason and Valletti [15], and Hermalin and Katz [9]. In this paper, however, we will work from models where two providers have *already agreed* to connect to each other, and we will assume that the value to the end users is implicitly captured by this agreement. The resulting models we consider will only include the network costs (whether provisioning or congestion) experienced by each network provider in a peering relationship.

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 Marcus [13]). 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 1, determining which placements are most preferred by the sender and receiver is, in general, computationally intractable. Nevertheless, several special cases (in particular, when 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 general, however, we will not address the economic question of how the preferences of the providers translate to an actual placement of peering links; this will typically be the outcome of a bargaining procedure between the two parties, and an interesting future problem would involve formulating a tractable model of this bargaining stage. The results of Section 1 may provide a starting point for this procedure.

In Section 2, 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 [10] and Roughgarden and Tardos [17] 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 4.

In Section 3, we consider an extension of the basic peering model, and suppose that provider R may charge provider S a price per unit flow sent through each link of provider R's network. We might expect that such a modification would prevent the phenomenon of nearest exit routing, and encourage the providers to use network resources efficiently (i.e., in a cost-minimizing fashion). However, using the economic theory of monopoly, we will demonstrate that when both S and R control a single link, the resulting routing of flow is not generally efficient, and give estimates of the loss of efficiency. We reach the important conclusion that the efficiency loss is related to the degree of convexity of the sender's cost function; and that in the worst case, this efficiency loss can be arbitrarily large.

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., [4, 6, 7, 11]. Recently, however, several efforts at understanding the impact of provider economics at network design have also begun, including results by He and Walrand [8], Gopalakrishnan and Hajek [5], and Feigenbaum et al. [3].

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.

1 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. Nevertheless, the differences in their topology have important economic consequences, and accounting for these effects remains open for future research.

In the first two subsections, we assume that both providers share the topology of a linear network. Network cost will be measured by distance travelled across the line; one might interpret this to mean that network cost is incurred linearly in the amount of flow carried and the distance this flow must travel. Under certain assumptions on the symmetry of traffic, we can explicitly compute the peering link locations which would be most desirable to both the sender and receiver. In addition, for more general distributions of traffic between sender and receiver, we give an efficient dynamic programming algorithm to compute the optimal placement for both sender and receiver. In the third section we consider optimal peering point placement when both providers share a tree topology; again, under some simple assumptions on the traffic distribution, we are able to compute an optimal placement 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 will find in Section 1.4. In other words, when the providers try to determine where to place peering links between each other, they face a computationally intractable problem. An open research problem, therefore, involves determining which approximations would be viable in practice, and how these approximations might be applied in the bargaining procedure leading to actual peering link placement.

1.1 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 $z \in \mathbb{R}$. For notational convenience we will write our results in terms of the pdfs f and g, and note that the proofs all apply with more general symmetric distributions G.

The second assumption we make is that each origin $y \in S$ 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. This means that several possible interpretations of our model are possible. For example, we may suppose that a total amount of traffic T is being sent from S to R, the origin x is chosen uniformly at random, and then the destination is chosen according to $f(y|x)$. Note further that because we are only interested in optimizing expected cost, we do not need to make any independence assumptions regarding the choice of destination. That is, it may be that the two origins x_1 and x_2 make a correlated choice of destinations y_1 and y_2; nevertheless, this does not change our analysis, as we require only the conditional marginal $f(y|x)$. Alternatively, in a deterministic interpretation, we may assume that given any interval in S of size δx centered at x, an amount of traffic $f(y|x)T\delta x/2l$ is sent to the destination $y \in R$. Both these interpretations, and several other similar variations, give rise to exactly the same analysis in the following development.

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 each provider wishes to minimize their own routing costs, where routing cost is measured by distance traveled (exit and entry 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*. Interestingly, the following theorem shows that there exists a single peering point placement which is optimal for both the receiver and the sender.

Theorem 1 *Let $p_i = -l + (2i-1)l/n$, for $i = 1, \dots, n$; 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 $p_1, \dots, p_n$ is the unique choice which is simultaneously optimal for both the sender and the receiver.*

Proof. We begin by assuming, without loss of generality, that $T = 2l$. Note that formally, we may write the optimization problem facing the receiving provider as follows:

$$\begin{aligned} &\text{minimize} && \sum_{i=1}^{n} \int_R |y - p_i|\, q_i(y|\mathbf{p})\, dy && (1)\\ &\text{subject to} && V_i(\mathbf{p}) = \{x \in S : i = \arg\min_j |x - p_j|\}, && (2)\\ &&& q_i(y|\mathbf{p}) = \int_{V_i(\mathbf{p})} f(y|x)\, dx, && (3)\\ &&& p_1, \dots, p_n \in S. && (4) \end{aligned}$$

In this notation, the vector $\mathbf{p} = (p_1, \dots, p_n)$ represents a possible choice of peering point locations. The set $V_i(\mathbf{p})$ is the Voronoi (minimum distance) region associated with the peering point p_i, given the vector $\mathbf{p}$. Finally, $q_i(y|\mathbf{p})$ gives the density function of the traffic destined for $y \in R$ entering through the peering point p_i, given the vector of peering points $\mathbf{p}$. Notice that q_i may not be a probability density, as it does not have total mass equal to one; in fact, q_i will have total mass equal to T times the fraction of traffic entering through p_i, which is exactly equal to the length $|V_i(\mathbf{p})|$. The objective function in the problem above is the the total cost of transporting traffic experienced by the receiving provider, given the location of the peering points $\mathbf{p}$; the problem is to minimize this cost subject to the assumption that the sending provider uses nearest exit routing. (This is why the destination density q_i is determined by the Voronoi region $V_i(\mathbf{p})$.)

Now consider the problem faced by the sending provider. We claim it is identical to (1)-(4), but with $G(z) = 1$ if $z \geq 0$, and 0 otherwise; this corresponds to a distribution placing unit mass at 0. To see why, note that in this case $q_i(y|\mathbf{p}) = 1$ if $y \in V_i(\mathbf{p})$, and 0 otherwise. In other words, the cost minimization above becomes:

$$\begin{aligned} &\text{minimize} && \sum_{i=1}^{n} \int_{V_i(\mathbf{p})} |y - p_i|\, dy && (5)\\ &\text{subject to} && V_i(\mathbf{p}) = \{x \in S : i = \arg\min_j |x - p_j|\}, && (6)\\ &&& p_1, \dots, p_n \in S. && (7) \end{aligned}$$

This is precisely the problem faced by the sender: because exactly the same (infinitesimal) amount of traffic originates at each point $x \in S$, the sender solves a straightforward *quantization* problem of trying to choose peering points $\mathbf{p}$ which will minimize the total expected distance travelled by traffic exiting the network S. Thus, the problem (5)-(7) is a special case of the problem (1)-(4), with a specific choice of distribution G made above. Since this G is still symmetric, it suffices to show that the peering point locations specified in the theorem statement are optimal for the receiver, i.e., in problem (1)-(4).

Rather than solving this problem directly, we will *relax* the problem, by assuming that the receiving provider may not only choose $p_1,\dots,p_n$, but also adjacent intervals $I_1 = [-l,a_1)$,$I_2 = [a_1,a_2)$,…,$I_n = [a_{n-1},l]$, such that $p_i \in I_i$. We will assume that the interval $I_i \subset S$ sends all its traffic through peering point p_i. In other words, the relaxation of the problem is such that rather than restricting the incoming traffic through peering point p_i to be determined by the Voronoi region $V_i(\mathbf{p})$, we allow the receiving provider to *choose* a surrounding region I_i that sends traffic through p_i. The relaxed problem is therefore as follows:

$$\begin{aligned}
\text{minimize} \quad & \sum_{i=1}^{n} \int_R |y-p_i|\, q_i(y)\, dy && (8)\\
\text{subject to} \quad & q_i(y) = \int_{I_i} f(y|x)\, dx, && (9)\\
& I_i = [a_{i-1},a_i), \;\; p_i \in I_i, \;\; i=1,\dots,n, && (10)\\
& -l = a_0 \le a_1 \le \ldots \le a_{n-1} \le a_n = l. && (11)
\end{aligned}$$

The key simplification yielded by this relaxation is that q_i is now no longer dependent on the choice of the vector $\mathbf{p}$ directly; rather, q_i is determined entirely by the choice of interval endpoints $a_0,\dots,a_n$. Now we note the following: If an optimal solution to the relaxed problem yields peering points $\mathbf{p}$ and intervals I_i such that $I_i = V_i(\mathbf{p})$, then in fact $\mathbf{p}$ must also be an optimal solution to the original receiver problem (1)-(4) (since each candidate solution $\mathbf{p}$ to the problem (1)-(4) yields a candidate solution $(\mathbf{p},I_1,\dots,I_n)$ to the relaxed problem, given by $I_i = V_i(\mathbf{p})$). We now show that a "symmetric" placement of the peering points (i.e., n peering points placed symmetrically about 0, a distance $2l/n$ apart from each other), together with the I_i given by the Voronoi regions around these peering points, yields a solution to the relaxed receiver problem. We need the following lemma.

Lemma 2 *Given the intervals I_i, $i=1,\dots,n$, an optimal choice in the problem* (8)-(11) *is to set p_i equal to the midpoint of the corresponding interval I_i.*

Proof. Because all the intervals I_i are given, our task in this lemma is reduced to the following minimization:

$$\min_{p_i \in I_i} \int_R |y-p_i|\, q_i(y)\, dy \qquad (*)$$

Let m_i denote the midpoint of I_i; i.e., $m_i = (a_i + a_{i-1})/2$. Because g is assumed to be symmetric, it is easily seen that the density $q_i(y)$ must be symmetric about the midpoint m_i. In particular, we find that m_i is a "median" for the density $q_i(y)$, in the following sense:

$$\int_{-\infty}^{m_i} q_i(y)\, dy = \int_{m_i}^{\infty} q_i(y)\, dy$$

But this is precisely the optimality condition for the convex optimization problem $(*)$; consequently, m_i must be an optimal choice of peering point given the interval I_i, as desired. $\square$

The preceding lemma allows us to focus only on choosing the intervals I_i in the problem (8)-(11); given these intervals, we are allowed to choose the peering point p_i to be the midpoint of the interval I_i without any loss of optimality.

To conclude the proof, we consider one final restatement of the original problem. Notice that in (8)-(11), given the symmetry of g, q_i is determined entirely by the midpoint m_i and length l_i of the interval I_i; this is easily seen through a change of variables:

$$q_i(y) = \int_{-l_i/2}^{l_i/2} g(y - m_i - x)\, dx$$

Here it is essential that every origin sees a destination density which is a shift of the same density g; this will allow us to focus only on the length of the interval I_i, and not on the actual location of I_i. In particular, if we choose $p_i = m_i$, we may rewrite the cost contribution of I_i in the expression (8) as:

$$\int_R |y - p_i|\, q_i(y)\, dy = \int_{-\infty}^{\infty} |y| \int_{-l_i/2}^{l_i/2} g(y-x)\, dx\, dy = h(l_i)$$

This is an expression which depends *only* on l_i, the length of I_i. This observation, together with Lemma 2, yields the following reduction of the problem (8)-(11). First, suppose that we choose the lengths l_i of our intervals, so that $a_0 = -l$, $a_i = a_{i-1} + l_i$ (where we require that $\sum_{i=1}^n l_i = 2l$). Now given these intervals, we simply choose peering point p_i to be the midpoint of the corresponding interval I_i; this choice is made without a loss of optimality. This leads to the following problem:

$$\begin{aligned}
\text{minimize} \quad & \sum_{i=1}^n h(l_i) && (12)\\
\text{subject to} \quad & h(d) = \int_{-\infty}^{\infty} |y| \int_{-d/2}^{d/2} g(y-x)\, dx\, dy, && (13)\\
& \sum_{i=1}^n l_i = 2l && (14)\\
& l_i \ge 0, \;\; i = 1, \ldots, n. && (15)
\end{aligned}$$

Any solution $(l_1, \ldots, l_n)$ to the previous problem will yield a solution to the relaxed problem (8)-(11), through the identification $a_i = -l + \sum_{k=1}^i l_k$, and $p_i = (a_i + a_{i-1})/2$. We will now show that the function h is convex. In this event, it follows that minimization of the objective function (12) is achieved when all l_i are set equal to each other; thus $l_i = 2l/n$ for all i would be an optimal solution, as desired.

We now show that h has nondecreasing derivative, which will insure convexity. Given $\alpha \ge 0$, define the distribution $G_\alpha(z)$ as follows:

$$G_\alpha(z) = \begin{cases} 0, & z < -\alpha, \\ \frac{1}{2}, & -\alpha \le z < \alpha \\ 1, & \alpha \le z \end{cases}$$

In other words, G_α is a distribution that places mass of $1/2$ of each of the points α and $-\alpha$. Any general symmetric pdf g may be written as an integral against distributions of the form G_α; it suffices to show, therefore, that if $f(y|x)$ is determined by G_α, then h has nondecreasing derivative.

There are two cases: either $\alpha \geq d/2$, or $\alpha \leq d/2$. When $\alpha \geq d/2$, then we find that:

$$\int_{-d/2}^{d/2} g(y-x)\,dx\,dy = \begin{cases} \frac{1}{2}, & -d/2-\alpha \leq y \leq d/2-\alpha, \text{ or } -d/2+\alpha \leq y \leq d/2+\alpha, \\ 0, & \text{otherwise} \end{cases}$$

In this case, $h(d) = d\alpha$, so that $h'(d) = \alpha$. On the other hand, if $\alpha \leq d/2$, then:

$$\int_{-d/2}^{d/2} g(y-x)\,dx\,dy = \begin{cases} \frac{1}{2}, & -d/2-\alpha \leq y \leq -d/2+\alpha, \text{ or } d/2-\alpha \leq y \leq d/2+\alpha, \\ 1, & -d/2+\alpha \leq y \leq d/2-\alpha, \\ 0, & \text{otherwise} \end{cases}$$

This yields $h(d) = d^2/4+\alpha^2$, so that $h'(d) = d/2$. In particular, notice that over the entire range of $d \geq 0$, h has nondecreasing derivative; thus, given α, h is convex, so that for any arbitrary symmetric distribution G, h is convex.

In turn, this allows us to conclude that an optimal solution to (12)-(15) is to set all l_i equal to each other; by Lemma 2, this then must also be an optimal solution to (8)-(11). Finally, notice that setting all the l_i equal to each other makes $p_i = -l+(2i-1)l/n$, as stated in the theorem; and further, in this case I_i is exactly the Voronoi region corresponding to p_i. As a result, this yields an optimal solution to the original receiver problem (1)-(4), and in turn to the sender's optimization problem (5)-(7).

Recall that the sender's problem is a special case of the receiver's problem where the target distribution G is given by G_0. In this case, not only is the median of each density q_i unique in the proof of Lemma 2, but we also find that h will be strictly convex. Thus, the solution to (8)-(11) corresponding to $l_i = 2l/n$ for all i is *unique*, implying that the sender's problem (5)-(7) has a unique solution as well. Thus, the only placement of peering points at which both providers will simultaneously achieve their optimal cost is the placement described in the statement of the theorem. □

We note several key features of the proof above. First, the sender and receiver problems are essentially the same, with the former a restricted version of the latter. This allows us to solve both problems by solving the more general case. Second, the solution to the receiver problem we have identified is not guaranteed to be unique; as an example, if we choose the distribution G to place mass $1/2$ at $\pm\alpha$, and 0 elsewhere, with $\alpha > 2l$, then *any* placement of n peering points between the two providers will yield exactly the same cost to the receiver. Nonetheless, because the sender's optimal placement is unique, there can only be one placement which simultaneously satisfies both parties. Notice also that the providers have the option of placing strictly less than n peering points; this is not precluded in the statement of the problem. However, we have been able to show that placing n distinct peering points is an optimal choice for both sender and receiver.

Finally, we note that the sender's optimal placement is independent of the choice of g. This is what we would expect if the sender is using nearest exit routing: as long as the distribution of *origins* of traffic is uniform over S, the sender does not concern itself with the destinations of that traffic. This is reflected in the fact that g does not appear in the sender's optimization problem (5)-(7).

The theorem identifies a kind of equilibrium. Notice that this is not Nash equilibrium of the placement problem in the traditional game-theoretic sense, because the placement of peering points must be agreed upon by both providers. Instead, if we consider a static game where the sender chooses the intervals $I_1, \ldots, I_n$ and the receiver chooses peering points $\mathbf{p}$, then the result above identifies a Nash equilibrium of this game. Given the intervals, a best response for the receiver is to choose peering points equal to the midpoints of the intervals; conversely, given the peering points, the sender chooses the intervals to be Voronoi regions. Both these conditions are met when the peering points are chosen according to the conditions of the theorem.

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.

1.2 More General Linear Networks

In this section, we extend the model of the previous section to include more general traffic distributions on linear networks. While we cannot solve this problem in closed form, we give an efficient dynamic programming algorithm to arrive at the solution.

We consider the situation where both S and R are identified with the interval $[0,1]$. Furthermore, we continue to suppose that a total amount of traffic T is being sent from S to R; without loss of generality, we will assume that $T = 1$. Two assumptions are relaxed from the previous section. First, rather than assuming that traffic originates uniformly across the network S, we assume that there exists a probability density function $\phi(x)$ such that the amount of traffic originating in $[a,b]$ is $\int_a^b \phi(x)\,dx$. Second, we no longer require that each origin send traffic to a destination determined by the same distribution g; that is, we let $f(y|x)$ be an arbitrary probability density function, for each $x \in S$. This formulation allows distinct origins to generate different amounts of outgoing traffic, and also does not make any symmetry assumptions about the destination of the traffic.

Again, we assume the sender and receiver each wish to place n peering points between them, and wish to determine the placements most preferred by each of them. The input to these two problems is the traffic distribution given above.

We first consider the sender S. Suppose we are given a collection of peering point locations $\mathbf{p} = (p_1, \ldots, p_n)$ (where $0 \le p_1 < \cdots < p_n \le 1$), and define the Voronoi region $V_i(\mathbf{p})$ by:

$$\begin{aligned}
V_1(\mathbf{p}) &= \left\{ x \in S : 0 \le x < \frac{p_1 + p_2}{2} \right\} \\
V_i(\mathbf{p}) &= \left\{ x \in S : \frac{p_{i-1} + p_i}{2} \le x < \frac{p_i + p_{i+1}}{2} \right\}, \; i = 2, \ldots, n-1 \\
V_n(\mathbf{p}) &= \left\{ x \in S : \frac{p_{n-1} + p_n}{2} \le x \le 1 \right\}
\end{aligned}$$

All origins in $V_i(\mathbf{p})$ will send their traffic through p_i into R. The sender then solves the following optimization problem:

$$\text{minimize} \quad \sum_{i=1}^{n} \int_{V_i(\mathbf{p})} |x - p_i| \, \phi(x) \, dx \tag{16}$$

$$\text{subject to} \quad 0 \le p_1 \le \cdots \le p_n \le 1. \tag{17}$$

We give an efficient dynamic programming algorithm to solve this problem. First, define the function $\psi(x, p) = |x - p|$. We let the value function $U_k(p_k)$ represent the lowest possible routing cost for traffic originating in $[0, p_k]$ when k peering points are placed in that interval, with the k'th peering point equal to p_k. In other words, $U_k(p_k)$ is the optimal routing cost if the sender's network consisted only of the interval $[0, p_k]$, and k peering points were placed inside $[0, p_k]$—with the last located exactly at p_k. We have the following relations:

$$\begin{aligned}
U_1(p_1) &= \int_0^{p_1} \psi(x, p_1)\phi(x)\, dx \\
U_{k+1}(p_{k+1}) &= \min_{p_k \in [0, p_{k+1}]} \left\{ U_k(p_k) + \int_{p_k}^{(p_k + p_{k+1})/2} \psi(x, p_k)\phi(x)\, dx + \int_{(p_k+p_{k+1})/2}^{p_{k+1}} \psi(x, p_{k+1})\phi(x)\, dx \right\} \\
U^* &= \min_{p_n \in [0,1]} \left\{ U_n(p_n) + \int_{p_n}^{1} \psi(x, p_n)\phi(x)\, dx \right\}
\end{aligned}$$

The last quantity U^* computes the total routing cost to the sender under the optimal placement of all n peering points. It is easily checked that this algorithm does indeed yield the optimal cost, calculated as U^*. Further, by backtracking, one may recover the optimal placements of the peering points $p_1, \ldots, p_n$. For a discrete counterpart of the algorithm (i.e., when the distribution corresponding to ϕ places mass only on finitely many points), this dynamic programming algorithm will run in polynomial time in the number of peering points to be placed and the size of the specification of the distribution given by ϕ.

A similar dynamic programming algorithm exists for the receiver optimal placement of peering points. The receiver solves the following optimization problem:

$$\text{minimize} \quad \sum_{i=1}^{n} \int_{V_i(\mathbf{p})} \int_R |y - p_i| \, f(y|x)\phi(x)\, dy\, dx \tag{18}$$

$$\text{subject to} \quad 0 \le p_1 \le \cdots \le p_n \le 1. \tag{19}$$

Comparing this problem to the sender's problem (16)-(17), notice that the problems are exactly the same, but with the expression $|x - p|$ replaced by the following function of x and p:

$$\psi(x, p) = \int_R |y - p_i| \, f(y|x)\, dy$$

Thus, *exactly* the same dynamic programming algorithm may be used to solve the receiver's optimal placement problem, with the same efficiency properties as well. Notice that just as in the previous section, the sender's optimal placement problem is a special case of the receiver's optimal placement problem, where the distribution corresponding to $f(y|x)$ places all mass at $y = x$.

This is precisely why we may use the same dynamic programming algorithm to solve both problems.

We note one final point: in this case, there is no guarantee that the sender and receiver will agree on a placement of peering points. In fact, it is possible that several optimal peering point placements will emerge for both the sender and receiver; in this case, a difficult problem still remains of determining how the two parties will come to an agreement on placement. As before, we do not address this bargaining game, but stress that it remains an important problem for future research.

1.3 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, $0 \leq p \leq 1$, 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 $1 \leq l < k$, 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.

Intuitively, we may view this process as a random walk on the tree, starting at the origin i_o; refer to Figure 2. Direct all edges upward along the unique path from i_o to the root, and direct all other edges downward. At all nodes with an up transition possible, an upward move is made with probability p. At all nodes with an incoming upward transition, each downward move has probability $(1 - p)/(m - 1)$. At all nodes with no incoming upward transition, each downward move has probability $1/m$. The leaf node first hit by this random walk is the chosen destination node. (Note with probability $1 - p$ the destination chosen is just i_o itself.)

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. We will first show that there exists an optimal level $l^*(p)$, depending on the parameter p, which minimizes the routing cost. The proof is as follows. In an optimal configuration of peering points, let l^* be the "lowest" level (i.e., largest in magnitude) at which a peering point is placed by the receiving provider; and let i be a node with a peering point at level l^*. Now by symmetry of the traffic distribution, if it is optimal to place a peering point at the root of the subtree rooted at i, it must be optimal to place a peering point at each of the siblings of i as well in level l^*. But if every node at level l^* has a peering point, there is no reason to place any peering points "higher" in the tree (i.e., at levels smaller in magnitude) than at level l^* (since, under nearest exit routing, all traffic from the sending provider will enter at level l^*). Consequently, there exists an optimal level l^*, depending on the parameter p, where the receiving provider chooses to place all peering points.

We now proceed to calculate $l^*(p)$; in particular, we would expect that $l^*(0) = k$, and that $l^*(1) = 0$, and further, that l^* is monotonically decreasing in p. The proof requires a careful analysis of the expected routing cost experienced by the receiving provider if all peering points are placed at a fixed level l. For convenience, we begin with a simple example. Suppose we return to the tree of Figure 2, and consider placing peering points at $l = 1$. Each unit of traffic entering at $l = 1$ will be routed within the same subtree with probability $1 - p + p(1-p) = 1 - p^2$, incurring a routing cost to the receiving provider of 1 link. However, it will have to travel to the other subtree with probability p^2, incurring a routing cost to the receiving provider of 3 links. Thus, the expected total cost (per unit traffic) to the receiving provider is: $(1-p^2)(1) + (p^2)(3) = 1 + 2p^2$.

In fact, this analysis is very general. Suppose, in a general tree, we intend to place peering points at level l. Then with probability $1 - p + p(1-p) + \cdots + p^{k-l}(1-p) = 1 - p^{k-l+1}$, incoming traffic is routed within the same subtree, travelling $k - l$ links in the receiver's tree. With probability $p^{k-l+i}(1-p)$, a routing cost $k - l + 2i$ links is incurred to the receiving provider, for $1 \le i \le l - 1$; and with probability p^k, a routing cost of $k - l + 2l = k + l$ links is incurred to the receiving provider. The expected cost per unit traffic is therefore:

$$\begin{aligned} F(l,p) &= (1 - p^{k-l+1})(k-l) + (p^{k-l+1} - p^{k-l+2})(k-l+2) + \cdots + p^k(k+l) \\ &= k - l + 2p^{k-l+1} + \cdots + 2p^k \end{aligned}$$

We note that for each l, the function $F(l,p)$ is strictly increasing in p. We first compute the points at which $F(l,p) = F(l+1,p)$:

$$k - l + 2p^{k-l+1} + \cdots + 2p^k = k - l + 1 + 2p^{k-l+2} + \cdots + 2p^k \implies$$
$$p = \left(\frac{1}{2}\right)^{1/(k-l+1)}$$

Now notice that the right hand side of the last equation is strictly decreasing in l. This fact, together with the fact that $F(l,p)$ is strictly increasing in p for each l, allows us to conclude that:

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

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.

We can extend this analysis to the case where not all links have unit length. Suppose that all links between level i and $i+1$ have length d_i, for $i = 0, \ldots, k-1$ (in both the sending and receiving tree). Then it is not difficult to check that the form of $F(l, p)$ becomes:

$$F(l,p) = \sum_{i=l}^{k-1} d_i + 2d_{l-1}p^{k-l+1} + \cdots + 2d_1 p^{k-1} + 2d_0 p^k$$

We find that $F(l, p) = F(l+1, p)$ implies that $d_{l-1} = 2d_{l-1}p^{k-l+1}$, i.e., $p = (1/2)^{1/(k-+1)}$. In other words, the solution for the optimal level $l^*(p)$ *remains exactly the same as before*—this solution, therefore, is independent of the lengths of the links between the levels.

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. One possible solution is the following: suppose that rather than placing arbitrarily many peering points, the providers agree to place at most $m^{l^*(p)}$ peering points. In this case, the sender would choose to place them all at level $l^*(p)$, and the receiver would, of course, still prefer the level $l^*(p)$. Thus, they would reach agreement on a placement. In practice, therefore, the interesting part of the bargaining procedure for this model is the determination of exactly how many peering points the providers wish to place.

1.4 In General

Under some assumptions on the structure of traffic and topology, the previous three sections have provided insight into the placements most preferred by sender and receiver. In general, computing these optimal placements is computationally intractable; in fact, in this subsection, we will show that appropriately formulated versions of the general peering point placement problems for either the sender or the receiver are NP-complete.

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:

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

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?

$$\text{minimize} \quad \sum_{(i,j)\in A_S} c_{ij} f_{ij} \tag{20}$$

$$\text{subject to} \quad \sum_{p_k} \sum_{r \in P_S(s,p_k)} y_r = \sum_{d \in R} b_{sd}, \ \ \forall\, s \tag{21}$$

$$\sum_{(s,p_k)} \sum_{r \in P_S(s,p_k):(i,j)\in r} y_r = f_{ij}, \ \ \forall\, (i, j) \in A_S \tag{22}$$

$$y_r \geq 0. \tag{23}$$

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 (20)-(23). 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 (20)-(23). The receiver's placement problem is then:

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

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?

$$\text{minimize} \quad \sum_{(i,j)\in A_R} c_{ij} f_{ij} \tag{24}$$

$$\text{subject to} \quad \sum_{r\in P_R(p_k,d)} y_r = b'_{p_k,d}, \ \ \forall\ (p_k,d) \tag{25}$$

$$\sum_{(p_k,d)} \sum_{r\in P_R(p_k,d):(i,j)\in r} y_r = f_{ij}, \ \ \forall\ (i,j)\in A_R \tag{26}$$

$$y_r \geq 0. \tag{27}$$

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 now show that both *SENDER_PLACEMENT* and *RECEIVER_PLACEMENT* are NP-complete, using a reduction from *VERTEX_COVER*:

VERTEX_COVER(N,A,K):

Given an undirected graph (N,A), does there exist a set of vertices $V\subset N$ of size less than or equal to K such that for all $j\in N$, there exists $i\in V$ with $\{i,j\}\in A$?

Theorem 3 *The problems SENDER_PLACEMENT and RECEIVER_PLACEMENT are NP-complete.*

Proof. Suppose we are given an instance of *VERTEX_COVER*, with input (N,A,K). Create a new directed graph (N',A'), with $N'=N$, and $A'=\{(i,j),(j,i):\{i,j\}\in A\}$. Thus each undirected edge in A becomes two directed edges in A'. Define the source-destination flow vector $\mathbf{b}$ as follows: $b_{sd}=1$ if $s=d$, and zero otherwise. Thus every source in S sends exactly one unit of flow to its mirror image in R. Set $c_{ij}=1$ for all $(i,j)\in A'$, so all links have unit length. Finally, let $n=K$, and let $K'=|N|-K$. Note that the transformation from (N,A,K) to $(N',A',\mathbf{b},\mathbf{c},n,K')$ is obviously a polynomial time operation, in the length of the input (N,A,K).

We claim that the problems *SENDER_PLACEMENT*$(N',A',\mathbf{b},\mathbf{c},n,K')$ and *RECEIVER_PLACEMENT*$(N',A',\mathbf{b},\mathbf{c},n,K')$ are identical to *VERTEX_COVER*(N,A,K); that is, there exists a vertex cover of (N,A) of size $\leq K$ if and only if the sender's (or receiver's) placement problem with input $(N',A',\mathbf{b},\mathbf{c},n)$ yields cost less than or equal to K'.

Given the form of the traffic matrix, we start by noting that the optimal value of the sender and receiver placement problems will be the same. In particular, fix a placement $\mathbf{p}$; and suppose that, using nearest exit routing, the sender experiences a total routing cost L. We claim the receiver experiences a total routing cost equal to L as well. This follows from the diagonal form of the traffic matrix $\mathbf{b}$. Since the sender has already chosen nearest exit routing, the path used by the sender in routing flow from a source s to the nearest peering point p_i must be the shortest available path to p_i; since $s=d$, the receiver can therefore do no better than to use the same path from p_i to d (here we require the assumption that $(i,j)\in A'$ if and only if $(j,i)\in A'$). Since for every placement of peering points $\mathbf{p}$, the receiver and sender experience the same cost, their optimal placements and optimal values will be the same as well.

Now assume that there exists a vertex cover of (N,A) of size less than or equal to K; assume without loss of generality the cover is of size equal to K, and let the cover be $p_1,\ldots,p_n$ (recall $n=K$). Because $p_1,\ldots,p_n$ form a vertex cover, if the sender uses nearest exit routing with the p_i as peering points, it sees a cost of at most $|N|-K$ (all nodes have one unit of traffic to send, and all nodes other than those in the cover are distance one away from a peering point). Thus the answer to *SENDER_PLACEMENT*$(N',A',\mathbf{b},\mathbf{c},n,K')$ is yes; by the preceding paragraph, the answer to *SENDER_PLACEMENT*$(N',A',\mathbf{b},\mathbf{c},n,K')$ is yes as well.

Conversely, suppose that the optimal value of the sender's placement problem with input $(N',A',\mathbf{b},\mathbf{c},n)$ is less than or equal to K'. Let $\mathbf{p}$ be the sender's optimal placement of peering points. Since $K'=|N'|-n$ and $c_{ij}=1$ for all (i,j), all nodes in N' other than $p_1,\ldots,p_n$ must be at most distance one away from a peering point. But this implies that $p_1,\ldots,p_n$ is a vertex cover of size equal to $n=K$, for the graph (N,A); thus the answer to *VERTEX_COVER*(N,A,K) is yes. The same reasoning follows if the optimal value of the receiver's placement problem is less than or equal to K'. Since *VERTEX_COVER* is NP-complete, we have shown that *SENDER_PLACEMENT* and *RECEIVER_PLACEMENT* are NP-complete as well. □

We note here that the computational complexity result of Theorem 3 supports an informal claim made by Awduche et al. in [1]. 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; either closed form solutions or efficient algorithms are presented in the first three parts of this section. In those cases, we might expect the providers to use the outcome of the analysis as a starting point towards a negotiated solution, or, in the case of linear networks in Section 1.1, even as the final implemented solution (since in that special case, the solution given is optimal for both providers).

The fact remains that two forces are opposed: on one side, the derivation of this section shows that finding the optimal placement of peering points is generally intractable; on the other hand, this is a problem that is being addressed on a regular basis in the real Internet, as providers continue to arrange peering agreements with each other. This raises the important question: what types of approximations are being made by providers in determining where they would like to place peering points; and what heuristics would be useful to the providers in this regard? Both these questions—the first empirical, the second theoretical—will prove to be of great practical importance given the results of this section.

2 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 1.4: 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 (20)-(23) for the sender and (24)-(27) 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., Bertsekas and Gallager, [2]):

$$\text{minimize} \quad \sum_{(i,j)\in A_G} c_{ij} f_{ij} \tag{28}$$

$$\text{over} \quad \sum_{r\in P^{\mathbf{p}}(s,d)} y_r = b_{sd}, \ \ \forall\ (s,d) \tag{29}$$

$$\sum_{(s,d)} \sum_{r\in P^{\mathbf{p}}(s,d):(i,j)\in r} y_r = f_{ij}, \ \ \forall\ (i,j) \in A_G \tag{30}$$

$$y_r \geq 0. \tag{31}$$

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 (28)-(31) 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 (20)-(23) and (24)-(27)), 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 4 *Suppose that $S = R$, and both have identical lengths $c_{ij} \geq 0$ for their links. Then given any placement of n peering points, we have the following bound:*

$$\textit{Cost of nearest exit routing} \leq 3 \times \textit{Cost of optimal routing}$$

Further, for every $\varepsilon > 0$ there exists a network such that:

$$\textit{Cost of nearest exit routing} > (3-\varepsilon) \times \textit{Cost of optimal routing}$$

Proof. The proof uses a graphical argument; refer to Figure 3. 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.

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 4. 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 [2]. 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.

We conclude with some important observations about this model. First, 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.

Such a model makes most sense in analyzing the current Internet backbone in the United States, for example, consisting of a small number of large, national network providers. These providers typically use linear link cost metrics in determining the relative cost of routes through their networks. Further, as discussed in the previous section, while these providers certainly do not share identical topologies, the fact that they are all national networks (with many common points of presence) suggests that our model may be a good first approximation.

One way to refine this approximation is to assume, for example, that the link cost c_{ij}^R of link (i,j) in provider R's network and the link cost c_{ij}^S of link (i,j) in provider S's network satisfy $c_{ij}^R \leq \beta c_{ij}^S$, 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.

3 Pricing and Competitive Routing

In this section, we will investigate the applicability of pricing mechanisms to the peering problem. In particular, we will investigate the consequences of allowing provider R to charge provider S a *price per unit flow* sent through provider R's network. In general situations where R and S do not share the same topology and the same link costs, there will again be a loss of efficiency; our goal in this section is to quantify this loss.

We consider the topology of Figure 5. Both S and R consist of a single link connecting two nodes, s and d. We assume that S has a total amount of flow x_S to send from point $s \in S$ to $d \in R$. Further, we assume peering points have already been placed at both s and d. As a result, two routes exist: S may choose to either send flow out at s to R, then use provider R's link to destination d; or S may use its own link to $d \in S$, then use the peering point at d to send traffic to $d \in R$.

Let f_S and f_R denote the total flow carried by provider S and provider R, respectively. We assume that S has a cost function for the flow on its link, given by $C_S(f_S)$; C_S is assumed strictly convex and strictly increasing with $C_S(0) = 0$. We also assume that C_S has a convex and strictly increasing derivative C'_S, with $C'_S(0) = 0$. We also assume that R has a cost function $C_R(f_R)$, which is assumed convex and nondecreasing with $C_R(0) = 0$; we assume that the derivative C'_R is convex and nondecreasing as well, with $C'_R(0) = 0$.

Without any feedback from provider R, provider S would always exit all traffic at s, thus incurring no routing cost; the entire cost of carrying traffic would be borne by provider R. However, the globally optimal routing solution—the analog of (28)-(31) in this situation—aims to minimize $C_S(f_S) + C_R(f_R)$, subject to $f_S + f_R = x_S$. A simple differentiation establishes that the unique solution to this problem occurs with $C'_S(f_S) = C'_R(f_R)$. Such a point exists since we have assumed that $C'_S(x_S) > 0 = C'_R(0)$. We will denote the globally optimal amounts of flow by f_S^* and f_R^*.

We will consider a two stage scenario where provider R first sets a price p per unit of flow sent on its link, then provider S makes a routing decision about how the x_S units of flow will be split between R and S. (Note that this coincides with the model of *monopoly pricing* in economics [18]: provider R is the monopoly, and provider S determines the demand seen by the monopolist.)

Assume, first, that the price p is fixed. Then provider S will solve the following problem:

$$\begin{array}{ll} \text{minimize} & C_S(f_S) + p f_R \\ \text{subject to} & f_S + f_R = x_S \end{array}$$

If $p \le C'_S(x_S)$, the solution is to set f_R according to $C'_S(x_S - f_R) = p$; if $p \ge C'_S(x_S)$, then $f_R = 0$. For $p \le C'_S(x_S)$, define the function $\phi_S(p)$ by $C'_S(x_S - \phi_S(p)) = p$, so that ϕ_S is the inverse of $C'_S(x_S - \cdot)$. Then in economic terminology, ϕ_S is the *demand function* seen by provider R. In other words, when provider R sets the price to p, it can expect a flow equal to $\phi_S(p)$ to enter via the peering point at s. Alternatively, we may work with the *inverse demand function*:

$$P(f_R) = C'_S(x_S - f_R)$$

The inverse demand function P captures the fact that if an amount of flow $f_R > 0$ is entering provider R through s, then the price must be $P(f_R)$. (If $f_R = 0$, then the price must be at least $P(0)$.) Note from our assumptions on C_S' that P is strictly decreasing and convex, with $P(x_S) = 0$.

We may now use these expressions to write the profit maximization problem facing provider R:

$$\max_{p \in [0, C_S'(x_S)]} p\phi_S(p) - C_R(\phi_S(p))$$

(Notice that the range over which provider R optimizes is $[0, C_S'(x_S)]$, since setting $p > C_S'(x_S)$ will result in zero demand and zero profits.) By the inverse relationship between ϕ and P, we may also write this problem as:

$$\max_{f_R \in [0, x_S]} P(f_R) f_R - C_R(f_R) \tag{32}$$

This optimization problem is the standard profit maximization problem seen by a monopolist [18]. Denote the solution to this problem by f_R^M, and let $f_S^M = x_S - f_R^M$. Under the assumptions made on C_S and C_R, there exists at least one optimal solution where $f_R^M > 0$, identified by:

$$f_R^M P'(f_R^M) + P(f_R^M) = C_R'(f_R^M) \tag{33}$$

In general, the optimal solution is not guaranteed to be unique; for our purposes, it is only essential that at least one optimal solution exist. Now note that because $P' < 0$, at f_R^M we have $C_S'(x_S - f_R^M) = P(f_R^M) > C_R'(f_R^M)$. This implies that $f_R^M < f_R^*$, so that $f_S^M > f_S^*$.

This distortion in the allocation of flow to links indicates an efficiency loss, since the routing cost in the priced peering situation will be higher than that obtained in the globally optimal solution. Our goal will be to attempt to quantify this efficiency loss. We choose a slightly different measure of efficiency than in the previous section: When the flow allocation is (f_S, f_R), we will define the *welfare* to be $C_S(x_S) - C_S(f_S) - C_R(f_R)$. This choice is deliberately made to coincide with economic terminology, where welfare in a monopoly context represents the total systemwide gain in utility from trade; for details, see Mas-Colell et al. [14]. In our situation, note that without the presence of the receiving provider's link, the sender would bear the full cost of routing, $C_S(x_S)$. Under a split of traffic (f_S, f_R), however, the total routing cost becomes $C_S(f_S) + C_R(f_R)$; the difference $C_S(x_S) - C_S(f_S) - C_R(f_R)$, therefore, is the improvement in global cost effected by addition of the link owned by R. Notice that welfare is maximized when global routing cost is minimized, i.e., the point (f_S^*, f_R^*) maximizes welfare; we will refer to $C_S(x_S) - C_S(f_S^*) - C_R(f_R^*)$ as *total welfare*.

Welfare may be characterized as the following integral:

$$C_S(x_S) - C_S(x_S - f_R) - C_R(f_R) = \int_0^{f_R} C_S'(x_S - y) - C_R'(y)\, dy$$

Referring to Figure 6, notice if we set $f_R = f_R^*$, the total welfare is simply the area under the inverse demand curve P, and above the *marginal cost* curve C_R', until they intersect at f_R^*; this is the area $W = W_M + L$. In a monopoly, where $f_R^M < f_R^*$, welfare is $C_S(x_S) - C_S(f_S^M) - C_R(f_R^M)$, identified by the area W_M in Figure 6. As a result, there is a *welfare loss* relative to the total welfare given by $C_S(f_S^M) + C_R(f_R^M) - C_S(f_S^*) - C_R(f_R^*)$, and identified by the area of the "triangle" L in Figure 6.

As we will find, in general the *relative welfare loss*, or the ratio L/W of welfare loss to total welfare, can be arbitrarily large. We would like to understand how large, as a function of the operating point f_R^M of the system with pricing, and the optimal value f_R^*. The following theorem gives such a bound:

Theorem 5 *Assume that C_S is strictly convex and strictly increasing, with convex and strictly increasing derivative C_S', and that C_R is convex and nondecreasing, with convex and nondecreasing derivative C_R'; assume also that $C_S(0) = C_R(0) = C_S'(0) = C_R'(0) = 0$. Then the following inequality holds:*

$$\frac{L}{W} \leq \frac{\log\left(\frac{f_R^*}{f_R^M}\right)}{\log\left(\frac{f_R^*}{f_R^M}\right) + 1}$$

Proof. We start by noting that since C_R is convex with $C_R(0) = 0$, we have:

$$C_R(f_R) \leq f_R C_R'(f_R)$$

for all $f_R \geq 0$. Now since f_R^M maximizes profits for the receiver, we have:

$$\pi_R^M = f_R^M P(f_R^M) - C_R(f_R^M) \geq f_R P(f_R) - C_R(f_R)$$

for all f_R, $0 \leq f_R \leq x_S$; we use the notation π_R^M to denote the left hand side, the receiver's optimal profit. Combining this with the previous inequality, we find that:

$$P(f_R) - C_R'(f_R) \leq \frac{\pi_R^M}{f_R}$$

Now we know from the discussion above that the welfare loss is the integral of the left hand side of the previous inequality from f_R^M to f_R^*. Thus, integrating both sides from f_R^M to f_R^*, we find that:

$$L \leq \pi_R^M \log\left(\frac{f_R^*}{f_R^M}\right)$$

Finally, since $W \geq L + \pi_R^M$, we may conclude that:

$$\frac{L}{W} \leq \frac{\log\left(\frac{f_R^*}{f_R^M}\right)}{\log\left(\frac{f_R^*}{f_R^M}\right) + 1}$$

as required. □

The preceding theorem suggests that as f_R^*/f_R^M increases, the relative welfare loss may get arbitrarily large (approaching 1). To demonstrate that this situation is indeed possible, we consider the special case where $C_R'(f_R) = 0$ for all $f_R \geq 0$. In this case, we have the following theorem.

Theorem 6 *Assume that C_S is strictly convex and increasing, with strictly convex and increasing derivative C_S', and that $C_R(f_R) = C_R'(f_R) = 0$ for all $f_R \geq 0$; assume also that $C_S(0) = C_S'(0) = 0$. Then the following inequality holds:*

$$\frac{L}{W} \leq \frac{2\log\left(\frac{x_S}{2f_R^M}\right) + 1}{2\log\left(\frac{x_S}{2f_R^M}\right) + 4}$$

Further, given a fixed ratio x_S/f_R^M, there exists a choice of C_S such that this bound holds with equality.

Proof. We start by noting that $x_S = f_R^* \geq 2f_R^M$. Clearly $x_S = f_R^*$, since f_R^* is identified by $P(f_R^*) = 0$. Note that at the point f_R^M, we have:

$$P(f_R^M) + P'(f_R^M)(2f_R^M - f_R^M) = 0$$

This is just a restatement of the optimality condition. On the other hand, by convexity, we know this tangent to P at f_R^M lies below the curve $P(f_R)$; thus in particular, $P(2f_R^M) \geq P(f_R^M) + P'(f_R^M)(2f_R^M - f_R^M) = 0$, so that $2f_R^M \leq x_S$, or $f_R^M \leq x_S/2$.

Next, define the following function $\hat{P}$:

$$\hat{P}(f_R) = \begin{cases} P(f_R), & 0 \leq f_R \leq f_R^M \\ \frac{f_R^M P(f_R^M)}{f_R}, & f_R^M \leq f_R \leq \frac{x_S}{2} \\ \frac{2f_R^M P(f_R^M)}{x_S} - \frac{4f_R^M P(f_R^M)}{x_S^2}(x_S - f_R), & \frac{x_S}{2} \leq f_R \leq x_S \end{cases}$$

We will show that $\hat{P}(f_R) \geq P(f_R)$ for all $f_R \geq 0$.

This is obvious for $f_R \leq f_R^M$. Noting that by definition of f_R^M, we know $f_R P(f_R) \leq f_R^M P(f_R^M)$ for all $f_R \geq 0$, we also have $\hat{P}(f_R) \geq P(f_R)$ for $f_R^M \leq f_R \leq \frac{x_S}{2}$.

Finally, notice that the slope of $\frac{f_R^M P(f_R^M)}{f_R}$ is given by $-\frac{4f_R^M P(f_R^M)}{x_S^2}$ at $f_R = \frac{x_S}{2}$. Thus $\hat{P}$ is a convex function, which agrees with P at f_R^M and x_S, and is greater than P for $f_R^M \leq f_R \leq \frac{x_S}{2}$. This is sufficient to imply that $P(f_R) \leq \hat{P}(f_R)$ for $\frac{x_S}{2} \leq f_R \leq x_S$, as required.

We now make use of the following bounds. First, again because we know that the tangent to P at f_R^M lies below the curve $P(f_R)$, the following inequality holds:

$$\int_0^{f_R^M} P(f_R) d f_R - f_R^M P(f_R^M) \geq \int_0^{f_R^M} P(f_R^M) - P'(f_R^M)(f_R^M - f_R) d f_R - f_R^M P(f_R^M)$$

Using the condition that $P'(f_R^M) = -P(f_R^M)/f_R^M$, the right hand side is then equal to $\frac{1}{2} f_R^M P(f_R^M)$. We thus have the following inequality:

$$W = L + \int_0^{f_R^M} P(f_R) d f_R \geq L + \frac{3}{2} f_R^M P(f_R^M)$$

We also have the following bound:

$$L \leq \int_{f_R^M}^{x_S} \hat{P}(f_R) d f_R = f_R^M P(f_R^M) \left(\log\left(\frac{x_S}{2f_R^M} \right) + \frac{1}{2} \right)$$

The inequality in the theorem then follows by combining this upper bound on L with the lower bound on W.

Finally, we must show that given x_S/f_R^M, the bound holds with equality for some choice of C_S, or equivalently, some choice of P. We will fix $P(f_R^M) = f_R^M$, which then requires that $P'(f_R^M) = -1$.

We define:

$$P(f_R) = \begin{cases} 2f_R^M - f_R, & 0 \leq f_R \leq f_R^M \\ \frac{(f_R^M)^2}{f_R}, & f_R^M \leq f_R \leq \frac{x_S}{2} \\ \frac{2(f_R^M)^2}{x_S} - 4\left(\frac{f_R^M}{x_S}\right)^2 (x_S - f_R), & \frac{x_S}{2} \leq f_R \leq x_S \end{cases}$$

P clearly satisfies the assumptions of the theorem. We must now check that f_R^M is the optimal choice for the receiver, given this demand function; that the welfare loss then meets the bound follows by the first half of the proof. But to show optimality, it suffices to note that by construction we have $f_R^M P(f_R^M)/f_R = (f_R^M)^2/f_R \geq P(f_R)$ for all $f_R > 0$. As a result, $f_R^M P(f_R^M) \geq f_R P(f_R)$ for all $f_R \geq 0$, so that f_R^M is the optimal choice for the receiver. □

We may draw two important conclusions from the previous theorem and its proof. First, notice that the case where $x_S/2f_R^M = 1$ occurs when the sender has a quadratic cost function; and in this case, the relative welfare loss is equal to $1/4$. Second, the theorem suggests that relative welfare loss increases as C_S' deviates from linearity, since this will lead to a large ratio $x_S/2f_R^M$.

Theorem 5 highlights a fundamental fact that is seen repeatedly in economics, and most simply in the theory of monopoly: selfish agents acting in a market do not necessarily converge to an efficient outcome. We have taken this analysis one step further, and developed novel bounds on the size of the worst case relative welfare loss, when only the operating point of the system is known.

Such a result serves as a warning that simply implementing a priced peering scheme does not necessarily provide the right incentives or feedback necessary to lead the system to an efficient outcome. Again, such a conclusion is seen widely in economics; but in our context, it may perhaps be less obvious. At first glance, the model suggests that the price could be used to relay to the sender the state of the receiver's network, and hence prevent efficiency loss; but at the efficient outcome, the receiver has an incentive to increase his price. Indeed, one can show using analogous techniques that when both providers have traffic to send, and each provider charges a price per unit flow, at least one of the two providers has an incentive to change their price at the system optimal operating point (f_R^*, f_S^*) (except in certain degenerate cases, as when both providers share identical cost functions). Further, our theorems demonstrate that in certain cases, the this incentive leads to a loss of efficiency which can be arbitrarily large.

4 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, depending on the problem we are investigating. When placing peering points between each other, the key problem is that providers must agree simultaneously on the placement. The results of Section 1 show, interestingly, that this precise outcome occurs when both providers share a linear network, and the traffic satisfies a simple symmetry condition. However, in more general cases the results of that section also show that the providers may not agree (as when they are both represented by trees, in Section 1.3, for example). Furthermore, in most cases calculating which peering point placement is optimal for either the sender or the receiver is NP-complete. This poses an important practical challenge: this computational complexity result shows that aligning the interests of the providers with each other will require tractable approximations which still capture their incentives accurately. Developing these heuristics for the peering point placement negotiation process remains an interesting open research direction.

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. Two key conclusions emerge: first, when the sending and receiving network share the same topology and the same linear cost functions, the cost of nearest exit routing is no more than three times the optimal cost. This result extends to a network of multiple providers with the same topology all sending and receiving traffic between each other; and we provide a bound that holds even if the cost functions are not identical.

In general, however, the loss of efficiency may be arbitrarily high. A second model considers pricing traffic routed between domains; the intuition would be that a price feedback signal might mitigate the effects of efficiency loss. Nonetheless, in this case we demonstrate that the efficiency loss may still be arbitrarily high, via a novel bound on the welfare loss due to monopoly pricing. This emphasizes the key point from mechanism design theory in economics that the protocols and system architecture which enable interdomain routing can have a large impact on the efficiency on the resulting network; see, e.g., [3] and [14] for further background on this idea.

More broadly, these analyses highlight a fundamental shift in the nature of network engineering models. Multiple agents in the network—network providers, as well as end users—all act in their own self interest. As a result, protocol designers and network architects can no longer separate the network's lower layer operation from the economics of the higher layers. Results such as those in Theorems 4 and 5 emphasize the fact that we cannot expect distributed selfish agents to replicate the actions of a single network-wide planner. Moving forward, therefore, we expect the line of inquiry considered in this paper to form part of a broader agenda to design networks where consideration of the players' incentives forms an integral part of the design philosophy.

5 Acknowledgements

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] Awduche, D.O., Agogbua, J., and McManus, J. (1998). An approach to optimal peering between autonomous systems in the Internet. In *Proceedings of the 7th International Conference on Computer Communication and Networks*, pp. 346-351.

[2] Bertsekas, D., and Gallager, R. (1992). *Data Networks*. Prentice Hall.

[3] Feigenbaum, J., Papadimitriou, C., Sami, R., and Shenker, S. (2002). A BGP-based mechanism for lowest-cost routing. In *Proceedings of the 21st Symposium on Principles of Distributed Computing*, ACM Press, pp. 173-182.

[4] Gao, L., and Rexford, J. (2001). Stable internet routing without global coordination. *IEEE/ACM Transactions on Networking*, 9 (6): 681-692.

[5] Gopalakrishnan, G., and Hajek, B. (2002). Do greedy autonomous systems make for a sensible Internet? Presented at the Conference on Stochastic Networks, June 19-24, 2002.

[6] Griffin, T., and Wilfong, G.T. (1999). An analysis of BGP convergence properties. Proceedings of ACM SIGCOMM 1999, *Computer Communication Review*, 29 (4): 277-288.

[7] Griffin, T., and Wilfong, G.T. (2000). A safe path vector protocol. In: *Proceedings of IEEE Infocom 2000*, pp. 490-499.

[8] He, L., and Walrand, J. (2002). Dynamic provisioning of service level agreements between interconnected networks. Presented at the Conference on Stochastic Networks, June 19-24, 2002.

[9] Hermalin, B.E., and Katz, M.L. (2001). Network interconnection with two-sided user benefits. Working paper, available from:

`http://faculty.haas.berkeley.edu/hermalin/Interconnection_v14.pdf`

[10] Koutsoupias, E., and Papadimitriou, C. (1999). Worst-case equilibria. In: *16th Annual Symposium on Theoretical Aspects of Computer Science*, pp. 404-413.

[11] Labovitz, C., Ahuja, A., Bose, A., and Jahanian, F. (2000). Delayed Internet routing convergence. Proceedings of ACM SIGCOMM 2000, *Computer Communication Review* 30 (4): 175-187.

[12] Laffont, J.-J., Marcus, J.S., Rey, P., and Tirole, J. (2001). Internet interconnection and the off-net cost pricing principle. Presented at the Conference on the Economics of the Internet and Software Industries, January 18-20, 2001.

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

[14] Mas-Colell, A., Whinston, M.D., and Green, J.R. (1995). *Microeconomic Theory.* Oxford University Press.

[15] Mason, R. and Valletti, T. (2001). Competition in communication networks: pricing and regulation. *Oxford Review of Economic Policy*, 17 (3): 389-415.

[16] Norton, W.B. (2000). Internet service providers and peering. Available on request from:

`http://www.equinix.com/press/whtppr.htm`

[17] Roughgarden, T., and Tardos, E. (2002). How bad is selfish routing? *Journal of the ACM*, 49 (2): 236-259.

[18] Tirole, J. (1988). *The Theory of Industrial Organization.* MIT Press.

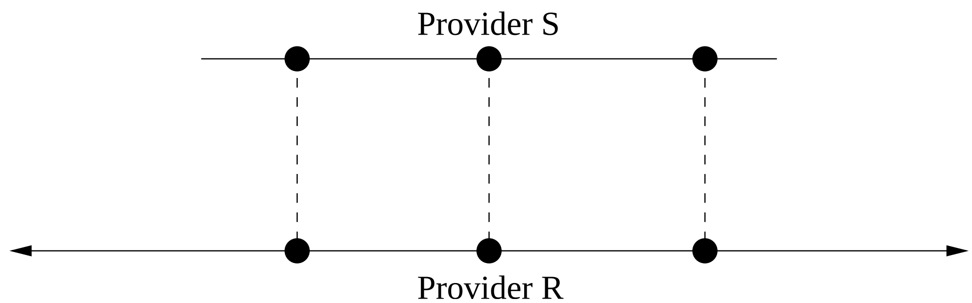

Figure 1: Two overlapping networks. The vertical lines represent links at peering points between S and R; they are drawn as dashed lines because in this model we assume that traffic experiences no cost travelling across a peering point.

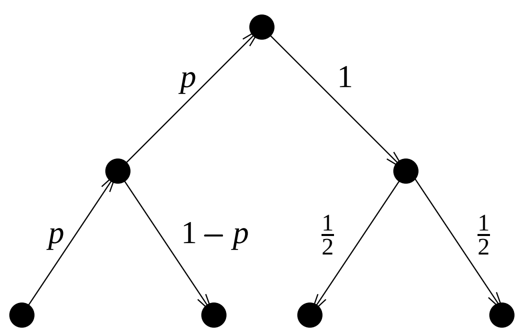

Figure 2: A tree with $k = 2$ levels, and fan-out $m = 2$.To determine the destination given the origin, we direct all edges upward from the origin to the root, and label these transitions with probability p. Transitions downward from a node have probability $(1-p)/(m-1)$ if the node has an incoming upward directed edge, and $(1-p)/m$ otherwise.

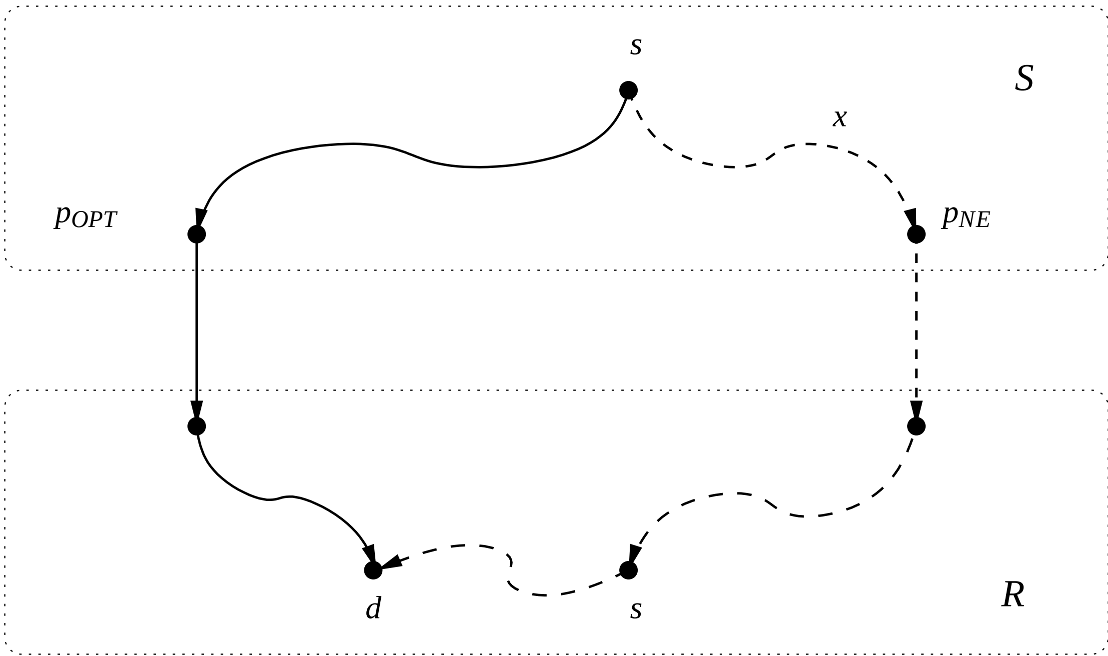

Figure 3: Proof of Theorem 4: Nearest exit routing cost is at most three times optimal routing cost.

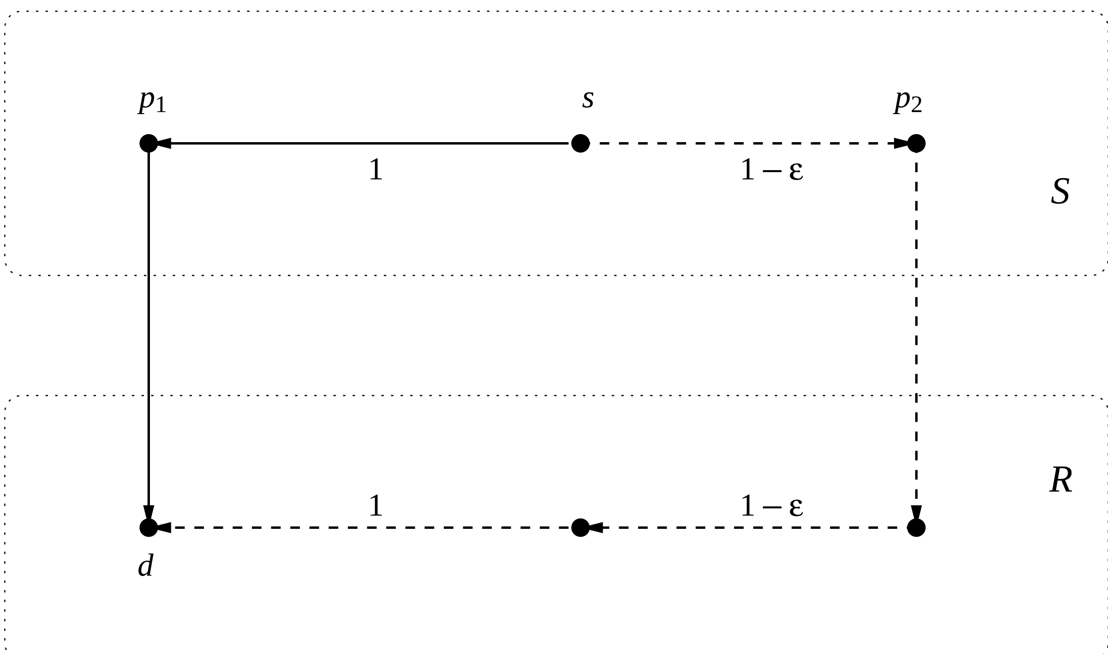

Figure 4: Proof of Theorem 4: Example where the upper bound is tight.

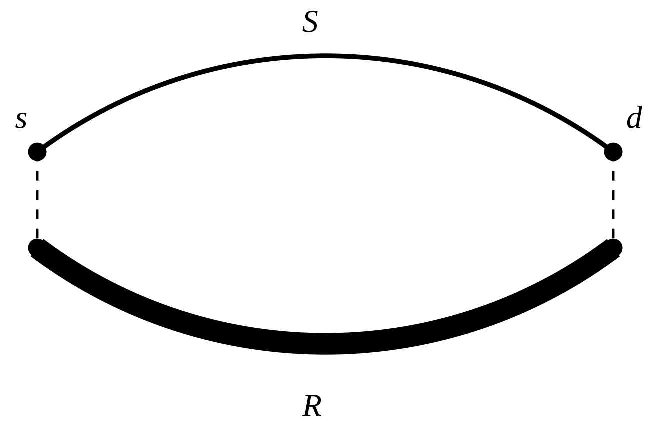

Figure 5: Provider S pays a price per unit flow sent across the link owned by provider R.

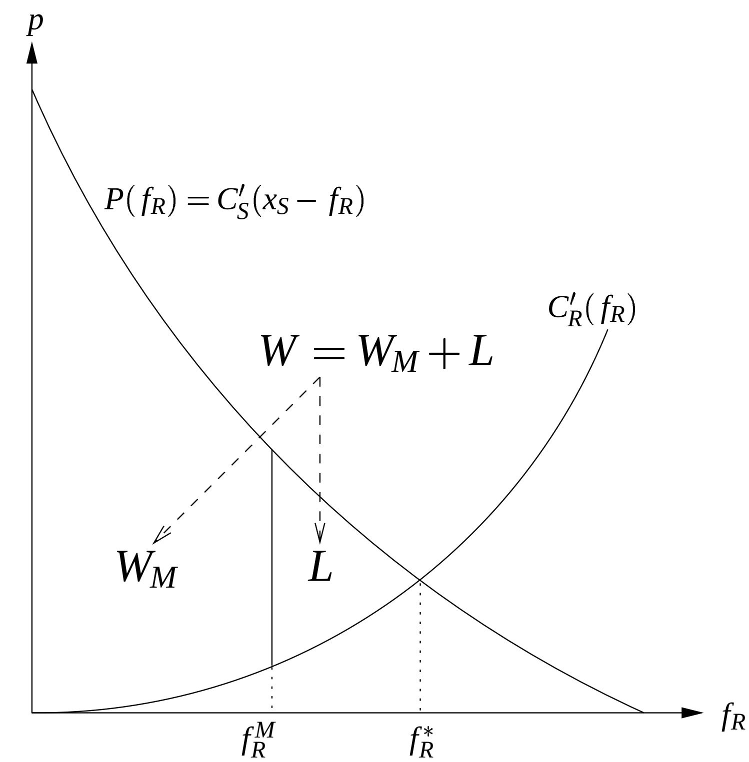

Figure 6: The total welfare $C_S(x_S) - C_S(f_S^*) - C_R(f_R^*)$ is given by W. When provider R charges provider S, the resulting welfare is given by $C_S(x_S) - C_S(f_S^M) - C_R(f_R^M)$. This results in a welfare loss "triangle" with area L.