Technical Report

Department of Computer Science
and Engineering
University of Minnesota
4-192 Keller Hall
200 Union Street SE
Minneapolis, MN 55455-0159 USA

TR 13-031

On Outbound Strategies and the Pareto Efficiency of Multipath Interdomain Routing

Pengkui Luo, Zhi-Li Zhang, and Andrew Odlyzko

September 24, 2013
On Outbound Strategies and the Pareto Efficiency of Multipath Interdomain Routing

Pengkui Luo*, Zhi-Li Zhang*, Andrew Odlyzko†
*{pluo,zhzhang}@cs.umn.edu, †odlyzko@umn.edu

Abstract—Transit providers play a key role in the formation and evolution of the Internet ecosystem. In this paper, we incorporate economical considerations into the study of the dynamics of the transit networks, under the premise of multipath interdomain routing. We model the dynamics as a result of individual profit optimization under capacity constraints, and prove the optimal outbound strategies for an individual transit provider. Furthermore, we formulate the global-scale strategies into a multi-objective optimization problem, from which a Pareto efficient solution that benefits the public good can be obtained.

I. INTRODUCTION

The current Internet ecosystem has been a result of the technological advances of the IT industry and the economic incentives of participating parties over last several decades. As the major constructing part of the Internet, transit providers (TPs) or Internet Serviced Providers (ISPs) play a key role in this process, and help shape the evolution of the Internet. Therefore, studying the dynamics and economic properties of transit providers is of great importance in understanding and predicting the evolution of the Internet, and helping it evolve to a state that better serves the public good.

The current Internet transit providers are organized in a hierarchical structure [13], in which smaller “eye-ball” providers, who provide “last-mile” Internet access for end users but have limited reachability to the rest of the Internet, usually have to pay larger (e.g., Tier-1) providers for carrying their traffic to/from parts of the Internet that are unreachable for them. In the remainder of this paper, when TP-1 provides transit services for TP-2, we call TP-1 is an upstream provider for TP-2, and TP-2 is a downstream customer of TP-1. TP-2 always pays TP-1, no matter the traffic flows from TP-1 to TP-2, or from TP-2 to TP-1. A transit provider usually does not have much control over how traffic should come in, as per the transit agreement it agrees to forward whatever it receives accordingly, except for a few inbound filtering policies, e.g., enforcing the “valley-free” principle [6]. But it does have absolute control over which neighboring ASes it routes traffic to. The outbound strategies it adopts greatly affects the revenues from its customers and the payments to its providers. In this paper, we study the “optimal” outbound strategies, not only from a single transit provider’s perspective, but also from a global scale that benefits the whole Internet ecosystem.

In the standard interdomain routing protocol BGP-4 [11], every flow has only one path. Recent research has proposed to extend it to support multipath routing. To name a few, Xu et al. [15] proposed to extend BGP-4 by enabling ASes to choose multiple routes to one destination. Walton et al. [14] proposed to enable a BGP speaker to advertise multiple routes for the same prefix. We are optimistic about these proposals; in this paper we take multipath interdomain routing as the premise.

The remainder of this paper is organized as follows. We first provide in Section II a generic model for the transit network, and then derive and prove in Section III the optimal local outbound strategy for any individual AS. In Section IV, we extend the findings into the global scale by formulating a multi-objective optimization problem, and solve the Pareto optimum. We discuss our future work in Section V, and an illustrative example in the Appendix.

II. THE NETWORK MODEL

Consider a transit network $G = (V, E)$, consisting of a set of transit providers $V = \{1, \ldots, |V|\}$, interconnected by a set of transit links $E = \{1, \ldots, |E|\}$. We use an incidence matrix $H \in \{0, 1, -1\}^{V \times |E|}$ to record the interconnection, i.e., setting $H_{kl} = +1$ if transit provider $k$ is on the provider end of link $l$, $H_{kl} = -1$ if on the customer end, and $H_{kl} = 0$ if $k$ is not on link $l$. The triple essentially $(V, E; H)$ defines the topology of the whole transit network.

Capacity Constraints. Let $x = (x_1, \ldots, x_{|E|})^T$ be the vector of link loads, and $c = (c_1, \ldots, c_{|E|})^T$ be link capacities. Then

$$0 \leq x \leq c.$$  

Pricing Function. A pricing function defines how a provider charges its immediate customers. We adopt a volume-based scheme, and assume a twice-differentiable pricing function $f : \mathbb{R}^+ \rightarrow \mathbb{R}^+.$

$$f'' < 0 < f'.$$  

The derivative $f'$ denotes the marginal cost of Internet transit services. The assumption that $f'$ is monotonically decreasing (i.e., $f$ is strictly concave) captures the economies of scale, i.e., the per-Mbps price decreases as the total charged volume increase. For simplicity, we assume the same pricing function $f$ is adopted by all transit providers. In addition, we also assume that there is only one type of interdomain relationship, customer-provider relationship, between any two interconnected TPs. In other words, we do not consider peering relationship.
Profit Functions. The net profit of the $k$-th transit provider can be formulated as its revenue from charging its customers, minus the payment to its providers, i.e., $P_k(x) = \sum_{i \in H_k} f(x_i) - \sum_{i \in H_k} f(x_i)$). We denote $H_k$ as the $k$-th row of matrix $H$ and $f(x)$ as the column vector $[f(x_1), \ldots, f(x_{|E|})]^T$, and write it in the vector form:

$$P_k(x) = H_k f(x).$$

III. LOCAL OUTBOUND STRATEGY FOR INDIVIDUAL AS

One question that we are particularly interested in is: in general, what are the best outbound strategies for individual ASes to forward traffic? In this section, we study this problem from the perspective of an individual transit provider or an AS. For simplicity we temporarily assume that all traffic is homogeneous, i.e., we don’t care different source-destination pairs and thus all forwarding are link-level instead of path-level, and 2) infinitely divisible, i.e., traffic can be divided into arbitrary parts with arbitrary proportions. We will break these two assumptions when we take a global view in the next two sections.

Downstream traffic forwarding of an transit AS refers to forwarding incoming traffic, which may come from its customers or providers, to its downstream customers through its egress routers; whereas upstream traffic forwarding refers to forwarding incoming traffic to its upstream providers. They are substantially different: in general, the objective of downstream forwarding is to maximize revenues from its customers, whereas upstream forwarding is for minimizing payments to its providers. Therefore, they shall be studies separately. Fig. 1 visualizes these two distinct forwardings.

Optimal Downstream Traffic Forwarding Strategy. Suppose a transit AS needs to forward total demanded volume $z$ to its customers ASes, and would like to maximize its revenue gained from them:

$$\max_x \sum_i f(x_i)$$

$$\text{s.t. } \sum_i x_i = z, 0 \leq x_i \leq c_i^x (i = 1 \ldots m)$$

where $m$ is the number of its customers, $c_i^x$ is the link capacity to its $i$-th customer, and $x_i$ is the traffic forwarded onto that link. Without loss of generality we assume the decreasing ordering of capacities: $c_1^x \geq c_2^x \geq \ldots \geq c_m^x \geq 0$. Its optimal solution $\{x_i^m\}_{i=1}^m$ must satisfy the following condition: there is some $m \in \{0, \ldots, m\}$, denoting the largest index of all unsaturated links, satisfying $^1$

$$\begin{align*}
  c_{m+1}^x &\leq x_1^* = \ldots = x_m^* < c_m^x \\
  x_i^* = c_i^x &\quad \forall i > m
\end{align*}$$

that is, the best downstream strategy is to divide the volume $z$ to its customers as even as possible, so as to gain higher marginal benefit from each.

Proof. Following the standard optimization routine, we construct the corresponding Lagrangian with nonnegative multipliers $\lambda_i$ and $\nu$: $L(x, z, \lambda, \nu) = -\sum_i f(x_i) + \sum_i \lambda_i (x_i - c_i^x) + \nu (\sum_i x_i - z)$. By the KKT conditions we know for all $1 \leq i \leq m$ the optimal $x$ must satisfy the following three equations: $-f'(x_i^*) + \lambda_i + \nu = 0$, $\lambda_i (x_i^* - c_i^x) = 0$, and $\nu (\sum_i x_i^* - z) = 0$. Since $m$ is the largest index of all unsaturated link, we automatically have $x_i^* = c_i^x$ for all $i > m$, and of course $x_m^* < c_m^x$. So we only have to prove two statements: (a) $x_1^* = \ldots = x_m^*$, and (b) $x_m^* \geq c_m^x$.

We divide the proof of statement (a) into two parts: (a1) $x_1^*, \ldots, x_m^*$ are all unsaturated, and (a2) all unsaturated link have the same traffic volume. Proof of part (a1): Since $x_m^* < c_m^x$, we have $\lambda_m = 0$, and therefore $f'(x_m^*) = \nu$. For any saturated link $x_k = c_k^x$, we get $f'(x_k^*) = \lambda_k + \nu = f'(x_m^*)$. Since $f''(x) < 0$, we have $x_k^* \leq x_m^*$. If $k \leq m$, then $x_k^* \leq x_m^* < c_m^x \leq c_k^x$, contradicted with the assumption that $x_k^*$ is saturated. Therefore, $k > m$, i.e., any saturated link must have a larger index than $m$. Proof of part (a2): suppose $x_1^*$ and $x_m^*$ are two unsaturated link, by the KKT conditions we know $\lambda_1 = \lambda_2 = 0$, and therefore $f'(x_1^*) = f'(x_m^*)$. By $f'' < 0$ we know $x_1^* = x_m^*$. We have to show statement (a). To show statement (b), we know $x_m^*+1$ is saturated. By the proof of part (a1), we know $x_m^*+1 \geq c_m^x+1$. $\square$

Optimal Upstream Traffic Redistribution Strategy. Suppose a transit AS needs to forward volume $z$ to its upstream providers, and would like to minimize the total payments incurred by the traffic:

$$\min_x \sum_j f(x_j)$$

$$\text{s.t. } \sum_j x_j = z, 0 \leq x_j \leq c_j^x (j = 1 \ldots n)$$

where $n$ is the number of its upstream providers, $c_j^x$ is the link capacity to its $j$-th provider and $x_j$ is the traffic on that link. Without loss of generality we also assume the decreasing ordering of capacities: $c_1^x \geq c_2^x \geq \ldots \geq c_n^x \geq 0$. Contrast to the formulation in the downstream strategy, here we need to minimize a concave function. Intuitively the optimal point is obtained on the boundary. In other words, the best strategy is to aggregate traffic to as few providers as possible, so as to pay lower marginal cost. Formally, a solution $\{x_j^\star\}_{j=1}^n$ to this minimization problem is such that there is at most one index $^1$

1). $m = 0$ means that every link is saturated. We also introduce two dummy capacity $c_0^x = +\infty$ and $c_m^x+1 = 0$ to make the formulation simpler.
1 ≤ \tilde{n} ≤ n satisfying
\[
\begin{align*}
0 < x^*_{i,j} & \leq c^\alpha_{i,j} \\
x^*_{j,i} & = c^\alpha_{j,i} \forall j < \tilde{n} \\
x^*_{j,i} & = 0 \forall j > \tilde{n}
\end{align*}
\]
that is, the optimal solution consists of at most one unsaturated non-zero link, and all links with higher capacity are saturated, while those with lower capacity do not have any traffic.

**Proof.** First of all, an optimal solution always exists, for the objective function is defined in a closed compact set. Since \(x_{j,i} \leq c^\alpha_{i,j}\) for all \(j, i\)-th largest element among \(\{x^*_{i,j}\}\) is no larger than \(c^\alpha_{i,j}\). Therefore, reordering \(x^*_{i,j}\) decreasingly (i.e., \(x^*_{1,i} \geq \ldots \geq x^*_{n,i} \geq 0\)) does not change the value of the objective function, while keeping all constraints feasible. We thereby add this ordering constraint without loss of generality.

Now the remaining statement we need to show is that “the optimum consists of at most one unsaturated non-zero link.” Suppose there are more than one \(j\) satisfying \(0 < x^*_{j,i} < c^\alpha_{i,j}\). We denote two of them as \(j_1, j_2\) where \(1 \leq j_1 < j_2 \leq n\). By the aforementioned reordering of \(x^*_{i,j}\) we know that \(x^*_{j_1,i} \geq x^*_{j_2,i}\). Since \(f'' < 0\), we have \(f'(x^*_{j_1,i}) \leq f'(x^*_{j_2,i})\), interpreted as \(|f(x^*_{j_1,i} + \Delta x) - f(x^*_{j_1,i})|/\Delta x \leq |f(x^*_{j_2,i} - \Delta x) - f(x^*_{j_2,i})|/\Delta x, for some \(\Delta x \to 0\). Hence we have \(f(x^*_{j_1,i} + \Delta x) + f(x^*_{j_2,i} - \Delta x) \leq f(x^*_{j_1,i}) + f(x^*_{j_2,i})\). Therefore, we can always choose a small amount \(\Delta x\) from \(x^*_{j_1,i}\) and add to \(x^*_{j_2,i}\), so as to further decrease the value of objective function, until \(x^*_{j_1,i}\) becomes zero or \(x^*_{j_2,i}\) saturates \(c^\alpha_{j_2,i}\), while keeping other links unchanged. □

**IV. GLOBAL MULTIPATH ROUTING**

In this section, we will extend our local insights gained in the previous section into a global perspective, namely, the whole transit network \(G = (V, E)\). We preserve all the general settings and assumptions made in section II, but throw away all additional assumptions in section III.

**Global Traffic Model.** Now let us model flows in the network. Denote \(W = \{1, \ldots, |W|\}\) as the set of flows in the network. Since each flow may take different paths simultaneously, we denote all valid or valley-free \(|W|\) paths in the network as \(P = \{1, \ldots, |P|\}\), and introduce a matrix \(R \in \{0, 1\}^{|W| \times |P|}\) to map a path onto its belonging flow: setting \(R_{r,j} = 1\) if path \(j\) belongs to flow \(r\), and setting to 0 otherwise. Note that each column of \(R\) only have one 1 entry.

We associate paths with links by introducing matrix \(Q \in \{0, 1\}^{|E| \times |P|}\), where \(Q_{ij} = 1\) if link \(l\) is on path \(j\), and 0 if otherwise. Tuple \((W, P; R, Q)\) defines the flow/path information. Let \(y = (y_1, \ldots, y_P)^T\) be the volumes on all paths, and \(d = (d_1, \ldots, d_W)^T\) be the demanded volumes of all flows.

We aggregate path volumes into end-to-end flow volumes
\[
d = R y
\]

By aggregating path volumes \(y\) into link loads \(x\), we have \(x_l = \sum_{j=1}^{|P|} Q_{ij} y_j\) for all \(l \in E\), or equivalently
\[x = Q y\]  \hspace{1cm} (9)

Fig. 2 shows an example for this setting. It consists of \(|V| = 5\) vertices or transit providers, connected by \(|E| = 6\) transit links. Fig. 2(a) annotates the indexes of vertices and edges. A directed edge always points from a customer to a provider. Fig. 2(b) annotates the undirected capacity of each link, represented in vector \(c\). Suppose there are a total of two end-to-end flows in the network: flow 4 → 5 with demanded traffic volume 6, and flow 1 → 4 with demanded volume 2. Flow 4 → 5 can be decomposed into four possible valley-free paths, whose volumes are denoted from \(y_1\) to \(y_4\); whereas flow 1 → 4 can be decomposed into two, denoted as \(y_5\) and \(y_6\). Matrices or linear operators \(H, Q\) and \(R\) can be obtained by standard graph traversal algorithms, shown in Fig. 2(c).

**Multi-objective Optimization.** Substituting \(x = Q y\) in Eq. (3) with \(x = Q y\), we express the profit of transit AS \(k\) in terms of path volume vector \(y\): \(\Pi_k(y) = H_k f(Q y)\). For all \(|V|\) transit providers, we write it in the matrix/vector form as a \(|V|\)-dimensional vector \(\Pi(y) = H f(Q y)\). As a reasonable profit-maximizer, each transit provider is supposed to maximize its own profit. Here we have \(|V|\) players, and formulate the problem as multi-objective optimization, a set of non-linear
objective functions with linear constraints.

\[
\begin{align*}
\text{"max"}_{y \geq 0} & \quad f(Qy) \\
\text{s.t.} & \quad Qy \leq c, \text{~} Ry = d.
\end{align*}
\]

Refer to the Appendix for the expanded formulation of the particular example we mentioned earlier.

**Pareto Optimum.** The solution to the above problem is a set of Pareto points. Pareto solutions are those for which improvement in one objective can only occur with the simultaneous worsening of at least one other objective. Formally, \( y^* \) is Pareto optimal if there does not exist another feasible vector \( y \) such that \( \Pi_k(y) \geq \Pi_k(y^*) \) for all \( k \in V \), and \( \Pi_j(y) > \Pi_j(y^*) \) for some \( j \in V \); or equivalently, for every \( y \) in the feasible region, either \( \Pi_k(y) = \Pi_k(y^*) \) for all \( k \in V \), or \( \Pi_j(y) < \Pi_j(y^*) \) for some \( j \in V \).

The Pareto optimum almost always gives not a single solution, but rather a set of solutions called Pareto optimal set, or non-dominated solutions. The maxima lie in the boundary of the feasible region, or in the locus of the tangent points of the objective functions, called the Pareto Front. It is usually not easy to find an analytical expression of the surface that contains these points. [4] provides a comprehensive survey on evolutionary-based multiobjective optimization techniques.

For the aforementioned example, we use the genetic-algorithm-based gamultiobj solver in Matlab Global Optimization Toolbox to calculate the Pareto optimal set numerically. For the particular example we mentioned earlier, we set the pricing function as \( f(x) = \ln(x + 1) \) and obtain a Pareto optimal set consisting of 90 points (they are 6D in the parameter space, and 5D in the objective space). They are "equally optimal" if we don’t prioritize them based on our domain knowledge. Here we would like to emphasize representing Pareto points.

The Pareto point that maximizes transit provider 1’s profit \( \Pi_1 \) is denoted as (path-load, profit) pair \( (y^{*(1)}, \Pi^{*(1)}) \). We also derive the corresponding link load \( x^{*(1)} \).

\[
\begin{align*}
\Pi^{*(1)} &= [3.93, 1.19, 0.85, -3.20, -2.77]^T \\
y^{*(1)} &= [0.9, 3.0, 0.0, 2.1, 0.9, 1.1]^T \\
x^{*(1)} &= [6.0, 6.2, 4.8, 3.0, 3.2, 3.0]^T
\end{align*}
\]

Similarly, we have the following Pareto points that maximize \( \Pi_2 \) to \( \Pi_5 \) respectively. Take vertex 2 for example, the "best" Pareto point for it is \( (y^{*(2)}, \Pi^{*(2)}) \), in which it achieves its highest possible profit \( \Pi^{*(2)} = 2.29 \) in the Pareto optimal set. If we look at it from the perspective of the link load \( x^{*(2)} = [1.0, 2.5, 4.0, 3.0, 4.0, 3.0]^T \), we can see that vertex 2’s upstream link (i.e., link 1) has volume as low as 1.0, but its downstream links (i.e., links 3 and 4) have volumes as high as 4.0 (almost saturated) and 3.0 (saturated).

\[
\begin{align*}
\Pi^{*(2)} &= [1.95, 2.29, 1.74, -3.22, -2.77]^T \\
y^{*(2)} &= [3.0, 0.8, 2.2, 0.0, 0.3, 1.7]^T \\
x^{*(2)} &= [1.0, 2.5, 4.0, 3.0, 4.0, 3.0]^T
\end{align*}
\]

\( \Pi^{*(3)} = [2.03, 2.13, 1.81, -3.20, -2.77]^T \)

\( y^{*(3)} = [3.0, 0.8, 2.2, 0.0, 1.0, 1.0]^T \)

\( x^{*(3)} = [1.7, 1.8, 4.7, 3.0, 3.3, 3.0]^T \)

\( \Pi^{*(4)} = [2.83, 1.97, 1.15, -3.18, -2.77]^T \)

\( y^{*(4)} = [2.8, 2.0, 1.0, 0.2, 0.2, 1.8]^T \)

\( x^{*(4)} = [2.4, 4.0, 5.0, 3.0, 3.0, 3.0]^T \)

\( \Pi^{*(5)} = [2.36, 2.12, 1.50, -3.21, -2.77]^T \)

\( y^{*(5)} = [2.9, 1.2, 1.8, 0.1, 0.3, 1.7]^T \)

\( x^{*(5)} = [1.6, 3.0, 4.5, 3.0, 3.5, 3.0]^T \)

Note that provider 5’s upstream links (i.e., links 4 and 6) are all saturated in any case, so \( \Pi_5 \) remains constant.

**V. Discussions and Future Work**

There are several future directions we would like to continue exploring. The first direction is designing a unified downstream-upstream strategy for an individual AS. Although the downstream and upstream traffic forwarding strategies derived in Section III give us enough insights on the best strategies towards profit maximization for an individual AS, the two strategies are studied separately. We would like to explore a unified strategy for any ASes to do local decision makings. To achieve this, we plan to deduct all inbound traffic volumes from corresponding link capacities, and apply the two strategies simultaneously based on the remaining capacities. We expect that the two strategies do not interfere with each other, for they deal with two different sets of links. In addition, we would like to study the convergence and stability of the equilibrium state.

The second direction is multipath source routing. This globally planned path-volume assignment used in this paper is perhaps more suitable to be implemented as source routing, rather than the current path vector routing like BGP [11]. In addition, the commercial relationships and AS-level capacity information may need to be revealed to end-hosts for decision making. Alternative, we can explore under the assumption of Pathlet routing [7], in which “networks advertise fragments of paths, called pathlets, that sources concatenate into end-to-end source routes.”

The third possible direction is to add content providers into the context, so as to form a two-sided market in which the transit providers charge both the content providers and the consumers [1, 5, 9, 12]. By doing this, we can analyze various attaching and peering strategies from the perspective of content providers, and enable the study of end-to-end demand distribution on this transit ecosystem, and explore transport providers’ incentives of investing in network capacity and infrastructure [2, 8].

The fourth possible direction is to propose a framework to evaluate a given topology of the transit network, by incorporating the traffic model, the demand model, the charging
model and the investment model of ASes, as sketched in Fig. 3. In general, we are interested in exploring how changing the transit structure would impact the distribution of traffic and profits/revenues. Furthermore, by comparing the results when individual ASes try to maximize their own profits/utilities with the results when the overall “social welfare” is achieved, under the same transit network structure, we can see what kinds of structures would encourage the maximization of individual profits to approach to the social optimality. This is useful in further studying/modeling the evolution of AS-level topology [3], [10].

REFERENCES


APPENDIX

The expanded formulation of Eq.(10) for the example in Fig. 2.

\[
\begin{align*}
\max & \ [f(y_2 + y_4 + y_6) + f(y_2 + y_4 + y_6)] \\
\max & \ [f(y_1 + y_2 + y_5) + f(y_1 + y_4) - f(y_3 + y_4 + y_6)] \\
\max & \ [f(y_3 + y_4 + y_6) + f(y_2 + y_3) - f(y_2 + y_4 + y_6)] \\
\max & \ [-f(y_1 + y_2 + y_5) - f(y_3 + y_4 + y_6)] \\
\text{s.t.} & \ y_1, y_2, y_3, y_4, y_5, y_6 \geq 0 \\
& \ y_2 + y_4 + y_5 \leq 9 \\
& \ y_2 + y_4 + y_6 \leq 8 \\
& \ y_1 + y_2 + y_5 \leq 5 \\
& \ y_1 + y_4 \leq 3 \\
& \ y_3 + y_4 + y_6 \leq 4 \\
& \ y_2 + y_5 \leq 3 \\
& \ y_1 + y_2 + y_3 + y_4 = 6 \\
& \ y_5 + y_6 = 2
\end{align*}
\]

We can see that the profit sum of all transit providers in this closed system equals zero, and the profits of the 4th and 5th transit providers are negative. This is reasonable because they are the “bottommost” providers (i.e., access providers) in this transit hierarchy, thus cannot charge other transit providers in this closed system. We remark that the bottommost providers do earn revenues from end users, who usually pay monthly flat-rate subscription fees. Modeling this part of revenue would involve another pricing model or even a demand model for end users, and certainly go out of the scope of our paper. Our paper focuses on the traffic distribution and the payment flows within the transit network, with a given end-to-end demand vector \( d \) as input, and therefore unnessitates the demand model for end users. If we were to bother this part of revenue, it would add nothing but a positive offset for the profits of bottommost providers and distract readers from our main discussion about the insight we would like to get from the transit network itself.