Limitations and Possibilities of Path Trading between Autonomous Systems

Yuval Shavitt
School of Electrical Engineering
Tel Aviv University
Email: shavitt@eng.tau.ac.il

Yaron Singer*
Computer Science Division
UC Berkeley
Berkeley, CA 94720
Email: yaron@cs.berkeley.edu

***Abstract*—When forwarding packets in the Internet, Autonomous Systems (ASes) frequently choose the shortest path in their network to the next-hop AS in the BGP path, a strategy known as hot-potato routing. As a result, paths in the Internet are suboptimal from a global perspective. For peering ASes who exchange traffic without payments, *path trading* – complementary deviations from hot-potato routing – appears to be a desirable solution to deal with these inefficiencies. In recent years, path trading approaches have been suggested as means for interdomain traffic engineering between neighboring ASes, as well as between multiple ASes to achieve global efficiency. Surprisingly, little is known on the computational complexity of finding path trading solutions, or the conditions which guarantee the optimality or even approximability of a path trading protocol.**

In this paper we explore the computational feasibility of computing path trading solutions between ASes. We first show that finding a path trading solution between a pair of ASes is NP-complete, and that path-trading solutions are even NP-hard to approximate. We continue to explore the feasibility of implementing policies between multiple ASes and show that, even if the bilateral path trading problem is tractable for every AS pair in the set of trading ASes, path trading between multiple ASes is NP-hard, and NP-hard to approximate as well.

Despite the above negative results, we show a pseudo-polynomial algorithm to compute path trading solutions. Thus, if the range of the instances is bounded, we show one can compute solutions efficiently for peering ASes. We evaluate the path trading algorithm on pairs of ASes using real network topologies. Specifically, we use real PoP-level maps of ASes in the Internet to show that path trading can substantially mitigate the inefficiencies associated with hot-potato routing.

I. Introduction

The Internet is a distributed *multi-tier* routing system, as it forms a network where each node is in itself an autonomous network. While the Border Gateway Protocol (BGP) serves as the binding contract for forwarding traffic between the tens of thousands of autonomously-managed networks, each such Autonomous System (AS) manages its network resources *independently* using interior routing protocols (such as OSPF, IS-IS, or RIP) that determine the path through which the packet travels from its ingress to egress points in the AS. Source to destination routes in the Internet are therefore a concatenation of intradomain paths chosen at the discretion of the ASes committed to forward the packet.

Preliminary results shown in this paper were presented at the NetEcon 2008 workshop [13].

*corresponding author.

Motivated to minimize the bandwidth resources consumed by moving packets to the next AS in the BGP route, ASes frequently engage in *hot-potato* routing - shortest path routing, based on configurable link weights, within the AS to the next-hop AS. As a result of this selfish behavior, routes in the Internet are suboptimal from a global perspective.

As an example, consider a case in which two ISPs AS_i and AS_j connect at two different points x and y, and there is some interdomain path in which traffic is forwarded from AS_i to AS_j. Suppose that for AS_i routing through x is slightly more efficient than routing through y, though for AS_j, it is considerably more efficient to carry traffic from y to the next-hop AS in the BGP path as opposed to x. Currently, under hot-potato routing, AS_i would forward its traffic through x rather than y.

In cases where monetary transfer between ASes occurs, as it does in customer-provider relationships, these inefficiencies can be mitigated through the use of payments. In our example above, AS_j could pay AS_i to forward traffic through y rather than x. However, large ISPs who are responsible for carrying a majority of the Internet traffic, often engage in peering relationships in which there is no monetary transfer [9]. Thus, when two ISPs are peers, in order for one ISP to provide incentives for another ISP to deviate from hot-potato routing, it must deviate from hot-potato in return in a manner which will benefit its peer. A considerable number of coordination protocols aiming to reduce the inefficiencies associated with hot-potato routing have recently been suggested [18], [6], [2], [14], [19], [4]. We refer to such complementary deviations from hot-potato routing as *path trading*, as an AS deviates from a hot-potato path in return for its peer's deviation, from which both ASes aim to benefit.

Since ASes have multiple ingress and egress points, optimal path-trading between peering ASes is a complex optimization task. While large emphasis has been placed on designing protocols that will allow ASes to path-trade without revealing sensitive information like their inner-AS topology, surprisingly little attention has been placed thus far on the *computational feasibility* of obtaining cooperative traffic engineering solutions. While the related problem of optimal establishment of peering points between ASes has been shown to be NP-complete [3], no such results are known for path trading.

This leads to the following questions:

1) Can path trading solutions between peering ASes be computed efficiently?
2) Can path trading solutions be computed on a global scale, i.e. between sets of ASes?
3) Assuming the answer to either of the above is negative, what assumptions on the input guarantee path trading solutions to be efficiently computed?

The goal of this paper is to address the above questions and provide an understanding of the power and limitations of path trading. We show both negative and positive results. We survey our results informally below (precise definitions are provided in the appropriate sections).

We first show that finding a path trading solution among two peering ASes is NP-complete. Furthermore, we show that it is NP-hard to approximate the optimal path trading solution within any factor. Intuitively, this sharp hardness originates from the individual rationality constraint: the requirement that neither ISP is worse off due to the path trading solution. This result corroborates the observations made in [11] and [17] which point to individual rationality as the crux in their discussion on the difficulty in obtaining path trading solutions in practice.

While trading between peering ASes is computationally hard, one may argue that by using heuristics, or under some assumptions, path trading solutions can be computed between pairs of ASes. Under this philosophy, the natural question is then (2) above: given path trading solutions between every pair of ASes, can a path-trading solution be computed among a set of ASes? Path trading between sets of ASes for traffic engineering is the focus of [10], [7], [5], [20]. We show in section IV that even under this assumption, computing path trading solutions between sets of ASes is NP-hard, and that it is NP-hard to approximate the optimal solution within any factor.

The above negative results imply that in order to design path trading protocols among ISPs, certain assumptions must be placed on the input. In section V we present a pseudo-polynomial algorithm for computing optimal path trading solutions, which relies on the assumption that the integer range of the possible path trading alternatives is not large. We then test our algorithm on real ISP topologies obtained from Internet measurements, to witness the benefits of path-trading on real network data.

We emphasize that the main goal in this work is to study the computational aspects of path trading. While implementation details play a significant role in path trading protocols and we discuss them briefly in section V, our investigation is focused on exploring the algorithmic barriers.

A. Paper Organization

We begin by presenting our model of a multi-tier routing system in section II, which attempts to capture the dynamics of intradomain and interdomain routing in the Internet. In section III we present impossibility results for path trading between two peering ASes, and in section IV we present analogous results for path trading between multiple ASes. We describe our algorithm for path trading between pairs of ASes in section V which returns the optimal path trading outcome under assumptions discussed, as well as heuristic-based approaches. Section VI contains results from experimentation we conducted on real AS topologies. We conclude with a short discussion in section VII.

B. Related Work

The path-trading approach was first suggested in [18] who identified its potential to benefit neighboring ISPs in optimizing their traffic flow. In [6] Mahajan *et al.* explore the benefits that mutual deviations from hot-potato routing can have for peering ASes, by using a heuristic-based approach which finds solutions which both ASes can benefit from. Their work provides the first evidence that path-trading can be deployed and serve as an attractive alternative to hot-potato routing. Follow up work by [14] improves over the heuristic-based bargaining approach presented by Mahajan *et al.* and suggests implementing the well-known Nash Bargaining solution via methods of convex optimization. To do this, one must make the assumption that the set of possible solutions in convex (while it is in reality not only not convex, but finite). Path trading has also been suggested to be performed on a global scale, i.e. agreements between multiple ASes on their interdomain paths in a manner that benefits all participating ASes [10], [7], [5], [20]. Cooperative approaches have been extended to explore the benefits of coordinated congestion control for multipath routing by Key *et al.* in [4]. Through rigorous analysis, they manage to establish theoretical guarantees on the benefits of coordinated congestion control, thus complementing previous experimental results. More recently, the inefficiency of routing on the intradomain level has been addressed in the context of P2P networks in [19], where the focus is in designing effective architectures for traffic control of Internet applications using cooperation between ASes. The problem of hot-potato routing and inefficiency of P2P systems are closely related, and path trading solutions can be applied in such settings as well. Johari and Tsitsiklis [3] studied a related problem of optimal establishment of peering points between ASes. Motivated by their analytical study of the inefficiency of hot-potato routing in various canonical network topologies, they show that determining the optimal placement of peering points between ASes is NP-complete.

II. The Model

We define a two-tier routing model in the following manner. On the interdomain level is the Internet AS graph denoted $G = (N, L)$ where $N = \{AS_1, \ldots, AS_n\}$ represents the set of ASes and L represents the set of pairs of ASes in N that communicate through an interdomain protocol (announce BGP routes to one another). Each node $AS_i = (V_i, E_i)$ in N is in itself a network, modeled as undirected weighted graph[1].

[1]While any representation of the internal AS structure applies here, we are motivated by the Point of Presence (PoP) level, modeled as an undirected weighted graph, where the weights represent the cost which ASes associate with each link as used by intra-AS routing protocols such as OSPF and IS-IS.

We use $\delta(w, x)$ to denote the shortest path distance between $x, w \in V_i$. [2]

Two ASes $AS_i, AS_j \in N$ are considered to be connected on the interdomain level if and only if there are at least two nodes $x_i \in V_i$ and $x_j \in V_j$ which can forward traffic to one another. In this case $\delta(x_i, x_j) = 0$. We use $\mathcal{I}_i$ to denote the set of ingress nodes in AS_i. We will abuse notation and write $\delta(x_i, AS_j)$ to refer to the shortest distance from a node $x_i \in AS_i$ to AS_j, i.e. $\delta(x_i, AS_j) = \min_{x_j \in I_j} \delta(x_i, x_j)$.

To capture the amount of traffic that goes through an ingress node, we use $f_{ij}(w, AS_k)$ to denote the amount of traffic that enters $w \in \mathcal{I}_i$ going to AS_k through AS_j.[3] For two peering ASes, AS_i and AS_j, the *cost* incurred by AS_i of routing from $x_i \in \mathcal{I}_i$ to AS_j through $x_j \in \mathcal{I}_j$ can be quantified by $f_{ij}(x_i, AS_k) \cdot \delta(x_i, x_j)$.

When AS_i and AS_j are peers, an interdomain routing protocol in AS_i can be written as a function $\varphi_i : \mathcal{I}_i \to \mathcal{I}_j$ which, for each $x_i \in \mathcal{I}_i$, chooses a node in $\mathcal{I}_j$ to forward its traffic to. The *hot-potato* routing function, denoted φ_i^* simply chooses the closest border node in AS_j :

$$\varphi_i^*(x_i) := \text{argmin}_{x_j \in \mathcal{I}_j} \delta(x_i, x_j)$$

Striving to minimize the consumption of their resources, unless given an alternative incentive, ASes practice hot-potato routing. Indeed, there is strong evidence that such a policy is exercised [16].

III. Path Trading Between Two Peering ASes

The main result in this section shows that finding path trading solutions is NP-complete, and that approximating the optimal path trading solution, within any factor, is NP-hard. To prove this, let us first formally define the notion of path trading between two peering ASes.

Two ASes are said to be *path trading* if at least one AS rejects its hot-potato strategy in favor for some alternative routing strategy. For AS_i routing to its peer AS_j, the (non-positive) value of routing traffic from a node $x_i \in \mathcal{I}_i$ to a border node $x_j \in \mathcal{I}_j$ instead of $x^* = \varphi_i^*(x_i)$ is:

$$v_i(x_i, x_j) := \sum_{AS_k} f_{ij}(x_i, AS_k) \cdot \Big(\delta(x_i, x^*) - \delta(x_i, x_j)\Big)$$

This deviation determines a value for AS_j:

$$v_j(x_i, x_j) := \sum_{AS_k} f_{ij}(x_i, AS_k) \cdot \Big(\delta(x^*, AS_k) - \delta(x_j, AS_k)\Big).$$

A path trade between AS_i and AS_j is defined by the routing strategies φ_i and φ_j. Each such path trade defines a utility for AS_i:

$$u_i(\varphi_i, \varphi_j) := \sum_{x \in \mathcal{I}_j} v_i(x, \varphi_j(x)) + \sum_{x \in \mathcal{I}_i} v_i(x, \varphi_i(x))$$

and similarly a utility u_j for AS_j. In the above sum, note that the first term is non-negative and that the second term is non-positive.

All literature which suggests path trading and its variants, assumes ASes will cooperate in path trading if the implemented routing strategies make neither of them worse off. This individual rationality requirement translates to finding non-trivial routing strategies φ_i and φ_j s.t. both utilities u_i and u_j are non-negative. We therefore say that a pair of strategies φ_i, φ_j is a path trading *solution* if these strategies induce non-negative utilities to both ASes. Note that this is the weakest requirement one can impose. We show that even for this requirement, finding path trading solutions is computationally infeasible.

Theorem 1: Finding a path trading solution is NP-complete.

Proof: The problem is clearly in NP. We reduce from zero subset sum where one is given a set S of positive and negative integers as input and is to determine whether some subset $T \subseteq S$ exists s.t. $\sum_{a \in T} a = 0$.

Given a set of integers $S = \{a_1, \ldots, a_r\}$ denote $S_- = \{a \in S | a < 0\}$ and $S_+ = \{a \in S | a > 0\}$. We construct two ASes, AS_i and AS_j; for every value $a_i \in S_-$ we construct a node v_{a_i} in AS_i, and construct a node u_{a_j} in AS_j for every value $a_j \in S_+$. In each AS we construct exactly two nodes which serve as the connectors between the ASes, $\{x_i, y_i\}$ in AS_i and $\{x_j, y_j\}$ in AS_j, where x_i connects to y_j and y_i connects to x_j. For each node v_{a_i} in AS_i we set $\delta(v_i, x_i) = 0$ and $\delta(v_i, y_i) = 1$; similarly, in AS_j, for each node u_{a_j} we set $\delta(u_{a_j}, x_j) = 0$ and $\delta(u_{a_j}, y_j) = 1$.

There is only a single node u_{d_i} in AS_i that traffic is routed to from AS_j, and similarly, there is such a node u_{d_j} in AS_j. For each u_{a_j} in AS_j the flow is $f_{ji}(u_{a_j}, u_{d_i}) = a_j$, and in the opposite direction $f_{ij}(v_{a_i}, u_{d_j}) = |a_i|$ for each v_{a_i} in AS_i.

First, let $T \subseteq S$ be a non empty subset of integers that sums to 0, and let T_- and T_+ be the negative and positive integers in T. For each u_{a_i} s.t. $a_i \in T_-$ we can reroute the traffic from x_i to y_i, which will have a negative value of $\sum_{a \in T_-} a$ to AS_i and positive value of $|\sum_{a \in T_-} a|$ for AS_j. Similarly, routing all $u_{a_j} \in T_+$ to y_j will have negative value of $-\sum_{a \in T_+} a$ to AS_j and positive value of $\sum_{a \in T_+} a$ to AS_i. Since T is a solution to zero subset sum, $-\sum_{a \in T_-} a = \sum_{a \in T_+} a$, and the solution is thus individually rational.

Conversely, given an individually rational path trading solution, let T_i be the subset of vertices in AS_i that route to y_i rather than x_i, and let T_j be the analogous subset in AS_j. Since the solution is individually rational to AS_j we have that:

$$\sum_{u_{a_i} \in T_i} f_{ij}(u_{a_i}, u_{d_j}) \Big(\delta(y_j, u_{d_j}) - \delta(x_j, u_{d_j})\Big) -$$
$$\sum_{u_{a_j} \in T_j} f_{ji}(u_{a_j}, u_{d_i}) \Big(\delta(u_{a_j}, x_j) - \delta(u_{a_j}, y_j)\Big) \geq 0$$

[2] Note that this shortest distance can be calculated according to any set of rules, e.g., based on additional annotations on the graph edges, and is not limited to minimum hop or minimum weight path.

[3] if no interdomain paths of the form $\langle AS_0 \ldots AS_i, AS_j, AS_k \ldots AS_n \rangle$ exist, then trivially $f_{ij}(x_i, AS_k) = 0$.

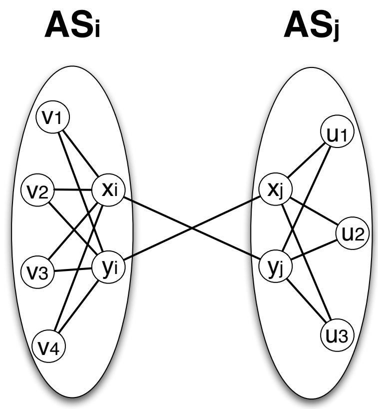

Fig. 1. An illustration of the construction used to show the reduction from zero subset sum to path trading. This illustration is the product of a set S with $|S_-| = 3$ and $|S_+| = 4$.

which implies $|\sum_{u_{a_i} \in T_i} a_i| \geq \sum_{u_{a_j} \in T_j} a_j$. Similarly, the individual rationality to AS_i implies that $\sum_{u_{a_j} \in T_j} a_j \geq |\sum_{u_{a_i} \in T_i} a_i|$. We therefore conclude that $|\sum_{u_{a_i} \in T_i} a_i| = \sum_{u_{a_j} \in T_j} a_j$, and thus that taking the corresponding integers to nodes T_i and T_j give a solution to zero subset sum. □

The optimization version of the above problem can be defined as finding a path trading solution which is individually rational and maximizes the sum of the utilities of the ASes. Since finding individually rational solutions is NP-complete, this optimization problem will necessarily be NP-hard. Intuitively, one may hope that there is room for approximation. That is, while the problem is NP-hard there might still exist a polynomial-time algorithm which guarantees a solution that is some factor of the optimal solution. The theorem below shows that there is no hope in approximation either.

Theorem 2: There is no polynomial time algorithm which approximates the optimal path trading solution within any factor, unless $P = NP$.

Proof: Using a similar construction to the one above, we will show that approximating within any factor implies solving subset-sum where we are given a set of integers $S = \{a_1, \ldots, a_n\}$, a target K, and the objective is to find a subset $S' \subseteq S$ s.t. $\sum_{a_i \in S'} a_i = K$. Given an instance to subset sum we construct two ASes, AS_i, AS_j with connector nodes $x_i, y_i \in AS_i$, and $x_j, y_j \in AS_j$, s.t. (x_i, y_j) and (y_i, x_j) are connected.

We create a single node in AS_j, denoted z, with $\delta(z, x_j) = 0$ and $\delta(z, y_j) = 1$. For each integer $a \in S$ we create a node u_a in AS_i s.t. $\delta(u_a, x_i) = 0$ and $\delta(u_a, y_i) = 1$ with flow $f_{ij}(u_a, z) = 2a$, and we also create a node t in AS_i s.t. $\delta(t, x_i) = 0$ and $\delta(t, y_i) = \frac{2K+1}{2K}$ and $f_{ij}(t, z) = 0$. The flow from AS_j to AS_i is $f_{ji}(z, t) = 2K$, and $f_{ji}(z, u) = 0$ for all $u \neq t$.

The above construction creates an instance of routing deviations that can be summarized as $(2a_1, -2a_1), \ldots, (2a_n, -2a_n), (-2K, 2K + 1)$, where in each pair the first value is AS_j's value and the second value is AS_i's value. Now observe that any non-trivial solution must incorporate $(-2K, 2K + 1)$ in order to be individually rational to AS_i. In order to be individually rational to AS_j we must choose a subset of routes that sum exactly to $2K$. Thus, the maximal value obtainable is 1. A solution to subset sum implies a value of 1 to the path trading problem, and no solution implies a value of 0. Therefore, the path trading problem cannot be approximated within any factor unless $P = NP$. □

The above results have profound implications on path trading. One may conclude that computing path trading solutions must be conditioned on assumptions on the instances. In section V we present a pseudo-polynomial algorithm for computing optimal path trading solutions, based on the assumption that the integer range of the possible solutions is not large.

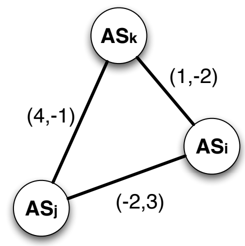

Fig. 2. An illustration of trading with mediators. Here the weighted pair on the edges indicates that values of the pair-wise path trade for the nodes on its ends. The pair $(1, -2)$, for example, indicates that AS_k gains utility of 1 from path trading with AS_i, which in turn receives utility of -2 from this path trade.

IV. Path Trading With Mediators and Sets of ASes

Having shown in Section III that finding path trading solutions between *pairs* of ASes is NP-hard and inapproximable, it is trivially the case for *sets* of ASes as well. However, one may argue that in many cases the number of ingress nodes of an AS is relatively small, and that considering all possible path trading solutions, while exponential in the total number of ingress points, is still computable. The natural question then is whether path-trading policies can be implemented on a global scale, assuming that solutions between pairs of ASes can be efficiently computed.

In the example illustrated in figure 2, we consider three ASes where each pair has a *single* path trading alternative. Each path trade yields a *cost* to one AS, and a *benefit* for the other (denoted as a weighted pair on edges which indicates the value of the path trade to the of the ASes on its ends). In this example, the path trading solutions between the AS pairs are trivial to compute, since there is only one possibility between each pair. While no single pair can trade on its own, the three can mutually trade, and benefit. We can think of each AS as a *mediator* between its neighbors, as it allows them to path trade.

The above example naturally extends to finding sets of ASes that can together trade on the network and form a *routing syndicate*. Assuming that every pair of ASes shares a small number of possible path trades which can be efficiently computed, all possible path trades between ASes can be encoded on a multigraph $G = (N, \bar{L})$ where N represents the set of all ASes in the AS graph, and $\bar{L}$ is the set of possible path trades between every pair of ASes. For each potential path trade between AS_i and AS_j, we have an edge $\ell \in \bar{L}$ and a weighted pair $w_\ell(AS_i, AS_j) = (w_\ell(AS_i), w_\ell(AS_j))$ which represents the value of AS_i and AS_j from path trading. W.l.og. we can assume that in each such pair one of the values is positive and the other is negative.

Similar to path trading, the minimal requirement from a routing syndicate solution is that it is individually rational. Finding an individually rational routing syndicate reduces to finding a non-empty subset of edges L' (which determines a subset of vertices N') s.t. $\sum_{\ell \in L'} w_\ell(AS_i) \geq 0$ for all $AS_i \in N'$. We now show that this task is computationally infeasible.

Theorem 3: Finding an individually rational routing syndicate is NP-complete.

Proof: We reduce from the knapsack problem to show the problem is NP-hard, as it is clearly in NP. In knapsack there are n items, each with a cost and a value, and we are given a budget B and a parameter K. Our objective is to choose a set of items s.t. their sum is at least K and the sum of their cost does not exceed the budget. Given an instance of knapsack, we will construct a graph $G = (N, L)$ in the following manner. Given the budget B and parameter K in knapsack, we construct two nodes AS_B and AS_K with an edge between them with the weight $w(AS_K, AS_B) = (-K, B)$. Each item $i \in \{1, \ldots, n\}$ with cost c_i and value v_i in the knapsack problem will correspond with a node AS_i in G that is connected to AS_K with weight $w(AS_K, AS_i) = (v_i, -\epsilon)$ and connected to AS_B with with weight $w(AS_i, AS_B) = (\epsilon, -c_i)$. We illustrate this construction in Figure 3.

If there is a subset of items S' in knapsack s.t. $\sum_{i \in S'} v_i \geq K$ and $\sum_{i \in S'} c_i \leq B$ there is a routing syndicate which consists of $\{AS_i\}_{i \in S'}$ and $\{AS_B, AS_K\}$ that is individually rational. Conversely, given an individually rational routing syndicate solution, note that it must consist of AS_B, AS_K, and some subset $\{AS_i\}_{i \in S' \subseteq [n]}$, as otherwise the solution cannot be individually rational. Furthermore, observe that an edge $(AS_i, AS_K) \in S'$ implies $(AS_i, AS_B) \in S'$. Thus, $\sum_{i \in S'} v_i = \sum_{\ell \in S'} w_\ell(AS_K) \geq K$, and $\sum_{i \in S'} c_i = \sum_{\ell \in S'} w_\ell(AS_B) \leq B$, and the set of items which corresponds to ASes in S' in the routing syndicate is a solution to knapsack. □

The above theorem shows that even when we consider a single path trade between every two ASes in the network, path trading in sets is hard. Again, the natural question is whether it can be approximated. The most natural measure of the quality of a routing syndicate is the social welfare, i.e. the the total value for its participant. Thus, finding the optimal individually rational routing syndicate can be stated as the following optimization problem:

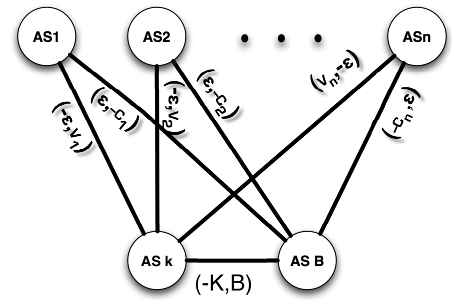

Fig. 3. An illustration of the construction used for the reduction from knapsack to show that finding optimal routing syndicates is NP-hard.

Maximize

$$\sum_{AS_i \in N'} \sum_{\ell \in L'} w_\ell(AS_i)$$

Subject to:

$$\sum_{\ell \in L'} w_\ell(AS_i) \geq 0 \quad \forall AS_i \in N'$$

We now show that the problem of finding an optimal routing syndicate cannot be approximated within any factor either.

Theorem 4: The optimal individually rational routing syndicate cannot be approximated in polynomial time unless $P = NP$.

Proof: We will use a similar idea from the previous section, and reduce from subset-sum. Given an instance to subset-sum, we construct $n + 3$ ASes. For each integer $i \in S$ we construct an AS, denoted AS_i. We also construct three ASes: AS_x and AS_y, and AS_K. Each AS_i is connected to AS_x with the weighted pair $w(AS_x, AS_i) = (2a_i, -\epsilon)$, and connected to AS_y with the weighted pair $w(AS_i, AS_y) = (\epsilon, -2a_i)$. We connect AS_x to AS_K with the weighted pair $w(AS_x, AS_K) = (-2K, \epsilon)$ and connect AS_K to AS_y with the weighted pair $w(AS_K, AS_y) = (-\epsilon, 2K + 1)$. We illustrate this construction in the figure below. Observe that due to individual rationality, taking an edge (AS_x, AS_i) implies taking the edge (AS_i, AS_y), for all $i \in [n] \cup \{K\}$. Similar arguments as shown in the proof of Theorem (2) show that approximating the optimal routing syndicate solution within any factor implies solving subset-sum. □

V. A Pseudo-Polynomial Algorithm for Path Trading

The results from section III show that computing path trading solutions between a pair of ASes is computationally infeasible and that the optimal path trading solution cannot be approximated either. The next step is therefore to explore what reasonable assumptions one can make that will allow computing path trading solutions. In this section we give a

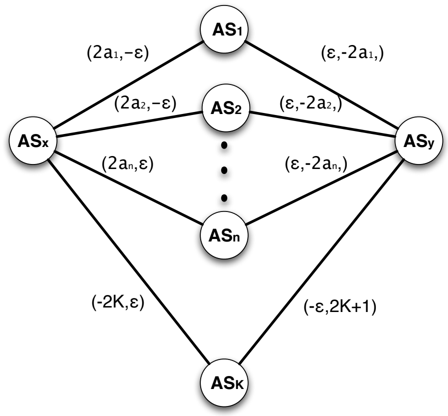

Fig. 4. An illustration of the construction used for showing the inapproximability of optimal individually rational routing syndicates.

pseudo-polynomial algorithm which finds the optimal path trading solution (if one exists). The assumption here is that the range of possible values of the path trading alternatives is bounded. At the end of this section we describe a few approaches that use the algorithm to create heuristic solutions in cases where the possible range of the solutions is large.

A. Computing Path Trading Solutions

Our goal is to construct routing strategies φ_i, φ_j that will be individually rational and maximize the sum of the trading ASes' utilities.[4] To simplify the exposition, we abstract away the computation of shortest path distances in the description of the algorithm. For AS_i, a deviation from hot-potato routing is a route $r = (x_i, x_j)$, for $x_i \in \mathcal{I}_i$ and $x_j \in \mathcal{I}_j$ with the value $v_i(r) = v_i(x_i, x_j)$, as defined in Section III. All possible route deviations from hot-potato in AS_i can therefore be written as the set of routes $\mathcal{R}_i$, and all such possible deviations in AS_j as $\mathcal{R}_j$. We will denote all possible deviations as $\mathcal{R} = \mathcal{R}_i \cup \mathcal{R}_j = \{r_1 \dots r_n\}$. Assuming path trading solutions exist, we wish to find the solution $\mathcal{R}'$ which maximizes:

$$\sum_{r \in \mathcal{R}'} v_i(r) + \sum_{r \in \mathcal{R}'} v_j(r)$$

subject to the constraints:

$$\sum_{r \in \mathcal{R}'} v_i(r) \geq 0, \quad \sum_{r \in \mathcal{R}'} v_j(r) \geq 0.$$

While the inapproximability result in section III was shown for this objective function, we can use a dynamic programming procedure to compute optimal outcomes, assuming that the range of possible values is bounded.

The Algorithm: For $r_M \in \text{argmax}_{r \in \mathcal{R}} v_i(r)$ and $r_m \in \text{argmin}_{r \in \mathcal{R}} v_i(r)$, let $M = n \cdot v_i(r_M)$ and $m = n \cdot v_i(r_m)$. Observe that $v_i(r) \geq 0$ if $r \in \mathcal{R}_j$, and that $v_i(r) \leq 0$ if $r \in \mathcal{R}_i$. For each $\ell \in [n]$ and $v \in \{m, \dots, 0, \dots, M\}$, let $\mathcal{S}_{\ell,v}$ be the subset of $\{r_1, \dots, r_\ell\}$ for which $\sum_{r \in \mathcal{S}_{\ell,v}} v_i(r) = v$, and its value for AS_j is maximal:

[4]Our algorithm also allows maximizing other criteria like the product.

$$\mathcal{S}_{\ell,v} \in \text{argmax}_{\{T \subseteq [\ell] : \sum_{r \in T} v_i(r) = v\}} \sum_{r \in T} v_j(r).$$

If $\mathcal{S}_{\ell,v}$ includes two routes from the same ingress node we say it is illegal (as this would be an illegal output). We define the value $V_j(\mathcal{S}_{\ell,v})$ to be:

$$V_j(\mathcal{S}_{\ell,v}) = \begin{cases} -\infty & \mathcal{S}_{\ell,v} = \emptyset \text{ or } \mathcal{S}_{\ell,v} \text{ is illegal} \\ \sum_{r \in \mathcal{S}_{\ell,v}} v_j(r) & \text{otherwise} \end{cases} \quad (1)$$

Using the above notation, we present the path-trading algorithm below.

Pseudo-Polynomial Algorithm for Path Trading

Initialize: $V_j(\mathcal{S}_{1,v_i(r_1)}) \longleftarrow v_j(r_1)$
$\mathcal{S} \longleftarrow \emptyset$
$V_j(\mathcal{S}_{1,0}) \longleftarrow 0,$
$V_j(\mathcal{S}_{1,v}) \longleftarrow -\infty \quad \forall v \neq v_i(r_1), 0;$

while $\ell \leq n$
for all $v \in \{m, \dots, M\}$:
if $v \leq v_i(r_\ell)$
$v' \longleftarrow v - v_i(r_\ell)$
$V_j(\mathcal{S}_{\ell,v}) \longleftarrow \max\{V_j(\mathcal{S}_{\ell-1,v}), V_j(\mathcal{S}_{\ell-1,v'}) + v_j(r_\ell)\}$
else $V_j(\mathcal{S}_{\ell,v}) \longleftarrow V_j(\mathcal{S}_{\ell-1,v})$

for all $v \leq \{0, \dots, M\}$
if $V_j(\mathcal{S}_{n,v}) \geq 0$
$\mathcal{S} \longleftarrow S \bigcup \{v\}$

return $\text{argmax}_{v \in \mathcal{S}} V_j(\mathcal{S}_{n,v}) + v$

The algorithm above considers all possible values for AS_i's utility, and at each stage ℓ, finds the optimal subset of $\{r_1, \dots, r_\ell\}$ in terms of AS_j's utility, for each possible value that AS_i may take. The iterative procedure terminates after $O((M - m) \cdot n^2)$ with the optimal values for AS_j, for each possible utility for AS_i. We then choose the outcome that maximizes the sum of the ASes utilities, under the individual rationality constraint.

B. Heuristic-Based Approaches

In cases where the range $\{m \dots, M\}$ is large, there are various heuristics that ASes may apply in order to efficiently compute path trading solutions, using the above algorithm.

1) **Discard high-cost routes:** intuitively routes that have extremely high costs, are likely to be left out of the solution, and removing them from the subset of considered routes can significantly reduce the range.
2) **Partition set of routes:** one can partition the set R into subsets of routes, based on their values, and run the algorithm on each such subset separately. Note that when combining individually rational solutions, the result remains individually rational.
3) **Simplify shortest path metric:** by compromising on the shortest path metric used, the subset of available

alternatives can be significantly reduced. For example, rather than using a complex $\delta(\cdot)$ function, a heuristic based on *unweighted* shortest paths can be applied, and the benefits and costs will be expressed in terms of number of hops saved.

C. A Note on Disclosure of ASes' Sensitive Information and Implementation

ASes are often considered to be private in all that regards their network topology. It suffices for ASes to exchange *opaque preferences* on ingress nodes, as suggested by Mahajan *et al.* in [6], by simply mapping their costs and benefits in our case into some real interval $[-r, r]$. Such a scheme allows path trading between ASes with different objectives (the δ distance metric in our case), as well as prevent disclosure of their topologies or optimization criteria. Note that under the opaque preferences scheme, ASes can obtain individually rational path trading solutions, though we cannot maximize the sum of the ASes utilities. As suggested by Mahajan *et al.* maximizing the social welfare should perhaps be restricted to "friendly ASes" who are willing to agree on an identical mapping. Also noted by Mahajan *et al.* is that an attempt to design a truthful mechanism which elicits the ASes' values fails here due to Myerson-Satterthwaite's impossibility result for bilateral trading [8], which asserts that in general, the trade will not be budget-balanced (i.e. achieving truthfulness requires a third party to sponsor the trade). In conclusion, individually rational solutions can be obtained using the above algorithm and implemented through the discussed work above.

VI. Evaluation

A. Methods

To evaluate path trading on *pairs of peer ASes* we tested our results on the PoP level in the Internet AS graph. We used the DIMES [12] IP level mapping of week 25 of 2007 which provides an IP level graph which includes over 600,000 links. Each such link includes an IP pair and a matching AS pair to which the IPs belong to. On this graph we created a PoP level mapping of each AS. Using the PoP generating algorithm [1] we obtained the mappings of IP addresses to their respective PoPs; we then used the IP level map once again, this time to establish inter-AS connections on the PoP level. We conducted this procedure for 70 ASes which are among the top 100 most connected ASes in the Internet AS graph (there was not enough coverage on the IP level to construct PoP mappings for some of the ASes in the top 100 list). Rocketfuel [15] is yet another available source for the mappings of the Internet on a PoP level, though we found the mappings of the DIMES project to have greater coverage. We show the distribution of the number of PoPs found in each AS in Fig. 5(left), and plot the number of PoPs against the AS degree in Fig. 5(right). In general, the number of PoPs an AS has increases with respect to its degree. This is of course expected since ASes with higher AS connectivity must maintain more PoPs to connect to their customers and peers. Still, one can note that there are ASes with AS degree in the 1000 range, which have more PoPs than ASes with twice their degree. In the range of ASes with degree below 500 there is greater variation in the ratio between AS degree and number of PoPs in the AS. Note that this may be due to inaccuracies in the data which may originate from poor coverage in a certain AS or inaccuracies introduced by the PoP generating algorithm. In table I we show some statistics of the topologies of these PoP level graphs, all with respect to the subgraph of the 70 ASes.

Our maps include the AS topologies alone, without delay measurements nor any weights on the links. We therefore used the measure of minimum hop distance, to estimate the cost of a path. Our maps also do not include detail of the demand matrix between the ASes. For this, we used the assumption that the demand between peering ASes is symmetric. We used the assumption that the probability of entering the AS through a specific ingress border node is uniformly distributed among the ASes' ingress nodes. Again, this assumption is not necessarily true, as we can expect different measures of traffic coming in through PoPs which represent large cities, for example, as oppose to one representing smaller ones. With these assumptions we applied the dynamic programming procedure as specified above.

Our assumptions indeed introduce inaccuracies, however, using the highest degree ASes in the Internet map, we can expect that the amount of traffic which flows between two peering ASes in one direction, is balanced by the amount of traffic which travel in the opposite direction. This assumption of symmetry is common in such experiments also used in [6], [14]. Also, the intention of our experiments is to show that ASes can find incentive to conduct path trading, as we consider real intra-AS topology and inter-AS links.

B. Results

The results presented here are of application of the solution which maximizes the sum of the ASes' utilities, as discussed throughout the paper. We first investigate the number of hops which can be saved as a function of the total number of path trades conducted, as presented in Fig. 6(right), and in Fig. 6(left) we plot a histogram of the number of hops saved normalized by the number of path trades. As one may expect the number of hops potentially saved with path trading increases with respect to the number of possible path trades. Note that there is a small fraction of ASes which can conduct many path trades, though the total hops they save is small. This typically occurs in ASes that connect to their peers through a single PoP, and more specifically a one-to-many connection. These ASes can then save a significant number of hops to their peers without loss, though as they connect to many peers through a single PoP, hot-potato routing does not greatly affect such ASes, and their benefit from path trading is relatively small. This observation is fortified in our following discussion of the relationship between hops saved and PoP degree. In the economic market of the Internet, one may expect such ASes to leverage this fact and charge monetary payments for their services.

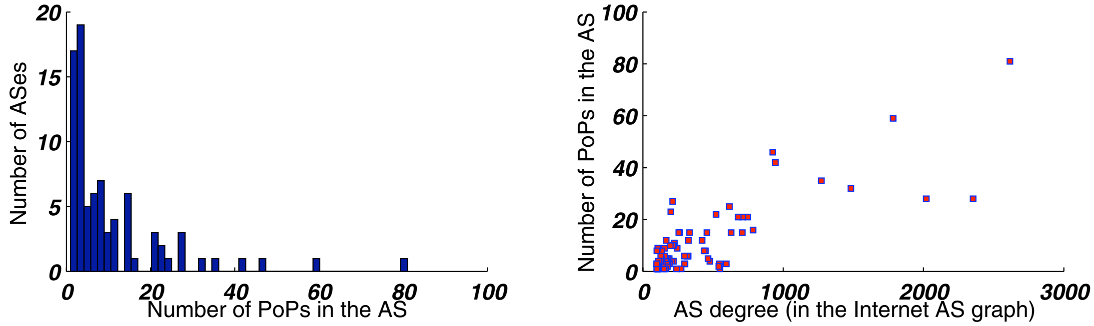

Fig. 5. Statistics of the PoP level graphs. On the left, the distribution of the number of PoPs found in the 70 ASes for which we conducted are experiment is shown. On the right, the number of PoPs is plotted as a function of the AS degree in the Internet AS graph.

Statistic	AS degree (sub graph)	number of PoPs	PoP degree
median	9	8	15
mean	15.57	12.67	51.01

TABLE I
STATISTICS OF THE INTRA-AS POP GRAPH TOPOLOGIES.

We also studied the relationship between the benefit from path trading and the ASes' PoP degree - the number of PoPs in *different* ASes an AS is connected to. In Fig. 7(right) we plot the number of hops which can be saved as a function of the ASes' PoP degree, and in Fig. 7(left) we plot a histogram of the number of hops saved normalized by the PoP degree. We have seen above that there were ASes which participated in relatively many path trades, though saved little or no hops in return. By plotting the number of hops saved as a function of the PoP degree we can see that there are very few high-PoP-degree ASes who benefit little from path trading. This accents the importance of rich inter-AS PoP level connectivity for benefiting from path trading.

Although there is evidentially a general increase in the number of hops saved, it is interesting to see that the two ASes which benefit the most from path trading are not necessarily the ones with the highest inter-AS PoP connectivity (one is not even among the top 10 most connected ASes). This is further testament to the fact that intra-AS topologies affect the number of hops which can be saved through path trading. For example, ASes with star topologies, or with a small number of border nodes, are examples where benefits from path trading solutions are less likely to be found.

VII. DISCUSSION

In this paper we studied the computational feasibility of finding path trading solutions between ASes, and examined their potential contribution to reducing the inefficiency caused due to hot-potato routing. We have shown that path trading between pairs of ASes is NP-hard and NP-hard to approximate within any factor. Analogous results were shown for path trading between sets of ASes, even in the case where each pair of ASes in the set shares a single path trading solution. Our impossibility results show that finding path trading solutions is hard due to the individual rationality constraint, which is the basic requirement from any reasonable path trading protocol.

On the constructive end, we presented a pseudo-polynomial algorithm for obtaining path trading solutions, which can be computed in time proportional to the range of the ASes' values from the possible path trading solutions. To test whether path trading can provide appropriate incentives for networks with real-world topologies, we applied our algorithm for pairwise path trading on the PoP level intra-AS graphs in the Internet and have shown it can substantially reduce the number of intra-AS hops used for routing. Our results suggest that indeed there is correlation between intra-AS topologies and potential benefits from path trading.

This work leaves open questions that can be addressed both experimentally and theoretically. On the experimental front including link delays to the PoP level maps would allow better prediction to the benefits which ASes can expect as a result of path trading. Also, additional PoP mappings of ASes will enable analyzing path trading on a larger scale.

On the theoretical end, the impossibility results imply there is great value in understanding the assumptions which allow for efficient computation of path trading solutions. Since our impossibility results seem independent of the intra-AS topology, assumptions on the shortest-path functions used for intradomain paths can be explored.

VIII. ACKNOWLEDGMENTS

We wish to thank Christos Papadimitriou for his contribution to the proof of theorem 2. We wish to thank Brighten Godfrey, Michael Schapira, Guy Rothblum and Gregory Valiant for their many helpful comments and suggestions. We also wish to thank Jennifer Rexford for referring us to previous related work.

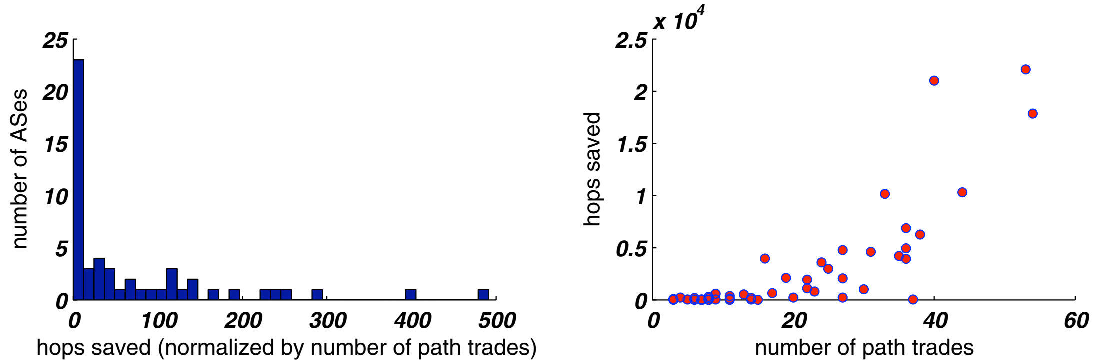

Fig. 6. The total number of hops saved as a function of the number of path trades.

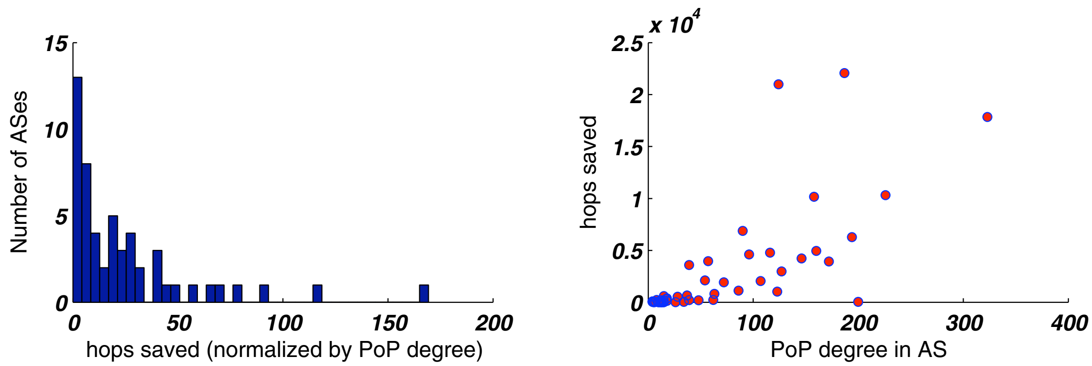

Fig. 7. The total number of hops saved as a function of the PoP degree.

References

[1] Dima Feldman and Yuval Shavitt. Automatic large scale generation of internet PoP level maps. In *GLOBECOM*, december 2008.

[2] Wenjie Jiang, Rui Zhang-Shen, Jennifer Rexford, and Mung Chiang. Cooperative content distribution and traffic engineering in an isp network. In *SIGMETRICS/Performance*, pages 239–250, 2009.

[3] R. Johari and J. Tsitsiklis. Routing and peering in a competitive internet. In *Decision and Control*, volume 2, pages 1556–1561, 2004.

[4] Peter B. Key, Laurent Massoulié, and Donald F. Towsley. Path selection and multipath congestion control. In *INFOCOM*, pages 143–151, 2007.

[5] Yong Liu and A.L. Narasimha Reddy. Multihoming route control among a group of multihomed stub networks. *Comput. Commun.*, 30(17):3335–3345, 2007.

[6] R. Mahajan, D. Wetherall, and T. Anderson. Negotiation-based routing between neighboring ISPs. In *Second Networked System Design and Implementation*, may 2005.

[7] Ratul Mahajan, David Wetherall, and Thomas E. Anderson. Mutually controlled routing with independent isps. In *NSDI*, 2007.

[8] Roger B. Myerson and Mark A. Satterthwaite. Efficient mechanisms for bilateral trading. *Journal of Economic Theory*, 29(2):265–281, April 1983.

[9] William B. Norton. Internet service providers and peering. In *Proceedings of NANOG 19*, 2000.

[10] Bruno Quoitin and Olivier Bonaventure. A cooperative approach to interdomain traffic engineering. In *In Proceedings of EuroNGI*, 2005.

[11] Jennifer Rexford. Route optimization in IP networks. *Handbook of Optimization in Telecommunications*, 2006.

[12] Yuval Shavitt and Eran Shir. DIMES: Let the internet measure itself. *ACM SIGCOMM Computer Communication Review*, 35(5), October 2005.

[13] Yuval Shavitt and Yaron Singer. Trading potatoes in distributed multi-tier routing systems. In *NetEcon*, August 2008.

[14] Gireesh Shrimali, Aditya Akella, and Almir Mutapcic. Cooperative inter-domain traffic engineering using nash bargaining and decomposition. In *INFOCOM*, pages 330–338, 2007.

[15] Neil T. Spring, Ratul Mahajan, David Wetherall, and Thomas E. Anderson. Measuring ISP topologies with rocketfuel. *IEEE/ACM Trans. Netw.*, 12(1):2–16, 2004.

[16] Renata Teixeira, Aman Shaikh, Timothy Griffin, and Jennifer Rexford. Dynamics of hot-potato routing in IP networks. In *SIGMETRICS*, pages 307–319, 2004.

[17] Ning Wang, Kin Hon Ho, George Pavlou, and Michael P. Howarth. An overview of routing optimization for internet traffic engineering. *IEEE Communications Surveys and Tutorials*, 10(1-4):36–56, 2008.

[18] Jared Winick, Sugih Jamin, and Jennifer Rexford. Traffic engineering between neighboring domains. Technical report, 2002.

[19] Haiyong Xie, Yang Richard Yang, Arvind Krishnamurthy, Yanbin Liu, and Avi Silberschatz. P4P: Provider portal for applications. In *Proceedings of ACM SIGCOMM*, Seattle, WA, August 2008.

[20] Yang Richard Yang, Haiyong Xie 0002, Hao Wang, Abraham Silberschatz, Arvind Krishnamurthy, Yanbin Liu, and Erran L. Li. On route selection for interdomain traffic engineering. *IEEE Network*, 19(6):20–27, 2005.