Trading Potatoes in Distributed Multi-Tier Routing Systems

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

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

ABSTRACT

The Internet is an example of a distributed system where the task of routing is performed in a multi-tier fashion: interdomain paths between autonomously-managed networks are subject to a global agreement (BGP), and the choice of intradomain paths is left to the discretion of each such network. When forwarding packets, 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. In this paper we explore complementary deviations from hot-potato routing in a manner which benefits both ASes. We show that even for a pair of ASes obtaining such *path trading* solutions is NP-complete, and give pseudo-polynomial algorithms to find them. We use PoP-level maps of ASes obtained from measurements of real AS topologies in the Internet to show that, in comparison to hot-potato routing, path trading can substantially reduce the cost of intradomain routing.

Categories and Subject Descriptors

C.2.2 [**Computer-Communication Networks**]: Network Protocols; C.2.6 [**Computer-Communication Networks**]: Internetworking

General Terms

Economics, Experimentation

Keywords

Bargaining, Hot-potato routing, OSPF

1. INTRODUCTION

In distributed routing systems, packets are being forwarded from one autonomous entity to another in accordance to an agreed contract. The most intuitive example of such systems is perhaps the Internet. In the Internet, tens of thousands of autonomously managed networks form a distributed routing system by forwarding information packets between one another, using the Border Gateway Protocol (BGP) as their binding contract. Each such Autonomous System (AS) manages its network resources independently using interior routing protocols (such as OSPF, IS-IS, or RIP), which 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 the inner-network paths of the ASes committed to forward the packet.

The Internet is an example of a distributed *multi-tier* routing system, as it forms a network where each node is in itself an autonomous network. 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.

Assume that there is a BGP path $\mathcal{P} = \langle AS_1, AS_2, AS_3, AS_4\rangle$ as portrayed in Fig. 1. Under the assumption of hot-potato, a route from AS_1 to AS_4 translates to $\langle s, v_1, u_1, u_2, u_3, u_4, t\rangle$, on the intra-AS level. Note that if AS_2 would deviate from its selfish hot-potato strategy, routing through the alternative path $\langle s, v_1, v_2, u_4, t\rangle$ would increase the overall traffic efficiency in the network by offering a shorter path from s to t, while reducing the number of hops used in AS_3. To execute this alternative, AS_2 must be presented with an appropriate incentive. Suppose another BGP route $\mathcal{P}'$ exists for which AS_3 could deviate in return from hot-potato routing in a manner which could save AS_2 at least one hop in its internal AS-routing on $\mathcal{P}'$. We could then expect this to provide AS_2 with a satisfactory incentive to avoid hot-potato routing, and in this instance, route to AS_3 using v_2 on $\mathcal{P}$.

Such *path trades* between ASes are the focus of our work in this paper. We use a cooperative game theoretic model, and study the dynamics of intradomain and interdomain routing as a two-player game. For two trading ASes, we associate a cost for an AS to deviate from hot-potato for a given route, and a (possibly negative) benefit this AS has for a deviation from hot-potato on some other route by its peering AS. We show that computing an individually rational solution - one in which the costs do not exceed the benefits for the two trading ASes - is NP-complete. We focus on two individu-

*corresponding author

NetEcon'08, August 22, 2008, Seattle, Washington, USA.

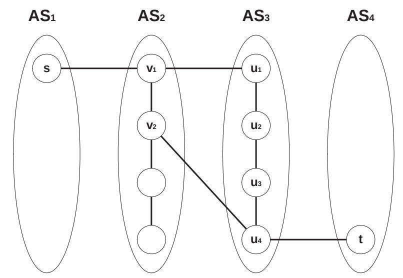

Figure 1: An example for interdomain routing in a 2-tier model.

ally rational objective functions, and a pseudo-polynomial algorithm to obtain them. The first maximizes the social welfare and the second is the Nash Bargaining (NB) solution [5] applied to our problem. Since our set of alternatives is finite, the NB solution does not maintain all of its properties from the Nash bargaining problem where the set of alternatives is convex. To evaluate our path trading method we used recent measurements from the DIMES project [6], and tested path trading between 70 pairs of ASes on real Point of Presence (PoP) level topologies. Our results show that the path trading method indeed provides incentives for ASes to deviate from hot-potato routing, as it enables significant reduction of path costs.

1.1 Related Work

A cooperative game theoretic approach has been used in the past by Mahajan *et al.* [4] as a means for AS traffic engineering optimization. Recently, to determine the peering points between ASes, Shrimali *et al.* [7] suggested a Nash Bargaining-based approach which requires introduction of shadow prices. Based on the axioms of Nash bargaining [5], their method is provably efficient and fair. Their model substantially differs from ours as they axiomatically assume a set of convex utilities. In our model, the costs and benefits of peering through different points between the ASes which defines the ASes' utilities, directly follow from the weights ASes associate with their intradomain paths. In [2] Johari and Tsitsiklis study the problem of optimal establishment of peering points between ASes. They analytically study the inefficiency of hot-potato routing in various canonical network topologies and show that determining the optimal placement of peering points for both the sender of the packet and the receiver is NP-complete. In our study the two trading ASes are both the sender and the receiver, and our objectives are concerned with maximizing a function of the utilities of both ASes in an individually rational manner. Cooperative game theoretic approaches have been also extended to explore the benefits of coordinated congestion control for multipath routing by Key *et al.* in [3].

1.2 Paper Organization

We begin by presenting our model of a multi-tier routing system in section 2, which attempts to capture the dynamics of intradomain and interdomain routing in the Internet. In section 3 we give a formal definition of path trading and discuss the objective functions mentioned above. We then show that computing these objective functions is NP-hard, and give pseudo-polynomial algorithms to obtain such solutions. Section 4 contains results from experimentation we conducted on real AS topologies.

2. THE MODEL

We define a 2-tier routing model in the following manner. For a network $G = (V, E)$ in tier 1, V represents the set of nodes and E represents the set of pairs of nodes in V which are connected. Each node $i \in V$ is in itself a network in tier 2, denoted $AS_i = (V_i, E_i)$. Two nodes $i, j \in V$ are considered to be connected in tier 1 if and only if there are at least two nodes $x \in V_i$ and $y \in V_j$ which are connected in tier 2. In tier 2 two nodes are connected if they are a physically connected, and there is some function $w : E^k \longrightarrow \mathbb{R}$ which associates a weight with each link.

Routing in the 2-tier model is performed as follows. The network $G = (V, E)$ chooses some path $\mathcal{P} = \langle 1, \ldots, m \rangle$, $j \in V$, $\forall j \in [m]$, to route data from AS_1 to AS_m. Each node j in $\mathcal{P}$ independently chooses an internal route in its network AS_j to reach the next node $j + 1$ in $\mathcal{P}$. In tier 1, we refer to the routing decision of G as the *global agreement* between the nodes in V, and the path $\mathcal{P}$, as the *interdomain path*. A path within a network in tier 2 is an *intradomain path*.

Our model attempts to capture the two-tier inter-AS and intra-AS routing of information packets in the Internet. Note that routing is more restrictive in our model as we assume interdomain paths are dictated to the participating nodes according to a global agreement. We use this restriction in our model to focus only on optimization of *existing* interdomain paths, rather than optimization of intradomain paths through means of choosing alternative interdomain routes. We discuss this point further in section 5.

Throughout the rest of this paper we use the 2-tier routing model associated with the Internet, yet the discussion which follows can be trivially extended to any distributed multi-tier routing model. We focus our attention on pairs of ASes for which some interdomain path exists, such that the two ASes are adjacent members, or *peers*, on that path. While our discussion can be applied to any representation of the internal AS structure in the Internet, we choose to concentrate on 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 refer to the graph which represents the connections between ASes in the Internet as the *Internet AS graph*. Our tier 1 nodes are therefore ASes, and we often will use the terms inter-AS and intra-AS to refer to interdomain and intradomain, respectively.

For a network $AS_i = (V_i, E_i)$ and $u, w \in V_i$, we use $\delta(u, w)$ to denote the shortest path distance between u and w in AS_i. The shortest distance, $\delta(u, w)$, can be calculated by 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. For an interdomain path $\mathcal{P}$, and an AS in $\mathcal{P}$, we model the intra-AS bandwidth resources consumed by $\mathcal{P}$ using the aggregate of intra-AS edge weights which are used during routing from an ingress point to the egress point which leads to the next-hop AS in $\mathcal{P}$. The intra-AS routes are chosen from a *finite* set of routes which reach the next-hop AS. intra-AS routing is not required to conform to any global routing policy, and we assume that each AS strives to minimize the consumption of its resources.

To summarize the main points in the 2-tier AS model:

- Inter-AS routing is subject to a global protocol;
- intra-AS paths are chosen from a finite set of paths, according to the preference of each AS;
- Each AS strives to minimize the consumption of its resources.

2.1 Selfish Routing in Multi-Tier Environments

Routing an information packet to the next-hop AS in the interdomain path requires an AS to choose an intra-AS path from the packet's ingress point to its egress point. For the network $AS_i = (V_i, E_i)$, we use B_i to denote the set of border nodes in AS_i, and $B_i(AS_j)$ to denote the subset of vertices in B_i which directly connect to AS_j. A *routing strategy* is a function $X : B_i \to B_i(AS_j)$, which determines an egress node for each border node in AS_i. Since we assume that the shortest path between these two nodes is always used, the routing strategy implicitly determines the intra-AS path of the packet. Note that this is not necessarily the shortest physical, or minimum delay path. For example, when a certain link on the shortest path is to be avoided (e.g. due to congestion), the AS adjusts the link weights inside its network accordingly.

For some interdomain path, the *hot-potato* routing strategy, practiced by AS_i, denoted X_i^*, is the strategy which chooses the shortest path to the next hop AS (for convenience assume that the cost of every two intradomain paths can be distinguished):

$$X_i^*(u) := \text{argmin}_{x \in B_i(AS_j)} \delta(u, x) \tag{1}$$

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 by ASes [9].

Note that under the assumption of hot-potato routing, every interdomain path $\mathcal{P} = \langle AS_1, AS_2, \ldots, AS_n\rangle$, determines a set of n intra-AS vertices $\{v_1, v_2, \ldots, v_n\}$, such that $\forall i \in [n]$ we have $v_i \in B_i$ being the first node of AS_i in the route (w.l.o.g. v_1 is the last node in V_1 which sends a packet to AS_n). We refer to each such node v_i as the *selfish ingress node* in AS_i on the path $\mathcal{P}$. In the example portrayed in Fig. 1 the nodes s, v_1, u_1 and t serve as the selfish ingress nodes of AS_1, AS_2, AS_3 and AS_4, respectively, on the path $\langle AS_1, AS_2, AS_3, AS_4\rangle$.

Lastly, since not all interdomain paths are at equal demand, when we consider routing between two peering ASes AS_i and AS_j we introduce a frequency function $f : B_i \times B_j \to \mathbb{R}$ to quantify the traffic of information packets routed from an ingress node $v \in B_i$ to all ASes through the node $u \in B_j$.

3. PATH TRADING BETWEEN PEERING ASES

We consider two ASes 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, we quantify the cost of routing packets from an ingress node $v \in B_i$ to an egress node $x \in B_i(AS_j)$ by:

$$c_{ij}(v, x) := \sum_{u \in B_j} f(v, u) \cdot (\delta(v, x) - \delta(v, X_i^*(v))) \tag{2}$$

This determines a (possibly negative) benefit for AS_j:[1]

$$b_{ji}(v, x) := \sum_{u \in B_j} f(v, u) \cdot (\delta(X_i^*(v), u) - \delta(x, u)). \tag{3}$$

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

$$u_i(X_i, X_j) := \sum_{v \in B_j} b_{ij}(v, X_j(v)) - \sum_{v \in B_i} c_{ij}(v, X_i(v)) \tag{4}$$

and similarly a utility $u_j(X_j, X_i)$ for AS_j. We will often refer to the utilities simply as u_i and u_j, leaving the routing strategies implicit.

We define a *path trading solution* to be an individually rational path trade between two ASes, i.e., routing strategies for which the utilities of both ASes are non-negative.

3.1 Path Trading Solutions

We focus on two objective functions common in game theoretic literature. The first, which we refer to as the Social Welfare (SW) solution, aims to choose the solution which maximizes the sum of the utilities of the trading ASes. Such a solution is desirable in the interest of optimizing the overall system performance. The second, referred to as the Nash Bargaining (NB) solution, aims to find the solution which maximizes the product of the trading ASes. For 2 players in the Nash bargaining problem [5] we are given a convex set of points in $\mathbb{R}^2$ where a point in the set represent an alternative, and the utilities of player i is the ith coordinate of the point.

The individually rational solution which maximizes the product of the player's utilities on the set of alternatives is proven to be Pareto optimal, symmetric, invariant to affine transformations, and independent of irrelevant alternatives. Furthermore, it is unique. In path trading our set of alternatives is not convex, but finite and discrete. We have that in general, the NB solution for path trading is not unique and symmetric. In our case, it is easy to see that it is still Pareto optimal, invariant under affine transformations, and independent of irrelevant alternatives.

3.2 Hardness of Computing Path Trading Solutions

For two peering ASes AS_i and AS_j in an AS graph, the Path Trading problem (PT) is the task of obtaining an individually rational path trade between the ASes. To do so we must choose routing strategies $X_i : B_i \to B_i(AS_j)$ for each ingress node in AS_i and $X_j : B_j \to B_j(AS_i)$ for each ingress node in AS_j under the constraints:

$$\sum_{v \in B_i} b_{ji}(v, X_i(v)) - \sum_{v \in B_j} c_{ji}(v, X_j(v)) \geq 0. \tag{5}$$

[1]To refrain from unnecessary notation, w.l.o.g. we assume that all egress nodes in AS_i connect with equal distance to AS_j. Note that otherwise it may be possible that $X_i^*(v)$ and x connect at different distances to AS_j and the difference $\delta(X_i^*(v), u) - \delta(x, u)$ may not accurately represent our intuitive concept of AS_j's benefit from v routing via x in AS_i.

and

$$\sum_{u \in B_j} b_{ij}(u, X_j(u)) - \sum_{v \in B_i} c_{ij}(v, X_i(v)) \geq 0. \qquad (6)$$

In Zero Subset Sum (*ZSS*), the well known NP-complete problem, 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. $\sum_{a \in T} a = 0$. By a reduction from *ZSS* we now show that PT is NP-complete as well.

THEOREM 1. *The path trading problem is NP-complete.*

PROOF. First, notice that the problem is clearly in NP. We will show that $ZSS \preceq_p PT$. Given an input $S = \{a_1, \ldots, a_r\}$ to *ZSS*, let $S_- = \{a \in S | a < 0\}$ and $S_+ = \{a \in S | a > 0\}$. We construct the following instance to PT. We construct two ASes, AS_i and AS_j; for every value $a_i \in S_-$ we construct an ingress node v_{a_i} in AS_i, and construct an ingress 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. From each ingress node v_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 ingress node u_j we set $\delta(u_j, x_j) = 0$ and $\delta(u_j, y_j) = 1$.

Lastly, it remains to define the flow of traffic between AS_i and AS_j in our construction. For each $u_{a_j} \in B_j$ the flow is $\sum_{v \in B_i} f(u_{a_j}, v) = a_j$, and in the opposite direction $\sum_{u \in B_j} f(v_{a_i}, u) = -a_i$ for each $v_{a_i} \in B_i$ (the flow is positive since $a_i \in S_-$).

A deviation from hot potato routing (routing to y_i rather than x_i) in AS_i can be modeled as a pair of benefits $\langle a_i, -a_i \rangle$, where a_i is the (negative) benefit of AS_i and $-a_i$ is the (positive) benefit of AS_j. Similarly, this is mirrored by AS_j where each deviation from hot potato is a pair of benefits $\langle a_j, -a_j \rangle$ with a_j as the (positive) benefit of AS_i and $-a_j$ is the (negative) benefit of AS_j. Thus, for $S' = \{\langle a_1, b_1 \rangle \ldots, \langle a_r, b_r \rangle\}$, where $b_i = -a_i$ for every $i \in \{1, \ldots, r\}$, a solution to PT is equivalent to that of choosing a subset $T' \subseteq S'$ s.t. $\sum_{\langle a,b \rangle \in T'} a \geq 0$ under the constraint $\sum_{\langle a,b \rangle} b \geq 0$.

We will show that a subset $T \subseteq S$ such that $T \neq \phi$ is a solution to ZSS if and only if the transformed subset T' is a solution to PT. First, if T is a solution to ZSS, then $\sum_{a \in T} a = 0$, and thus $\sum_{\langle a,b \rangle \in T'} b = 0$, and specifically $\sum_{\langle a,b \rangle} b \geq 0$. Also, $\sum_{\langle a,b \rangle \in T'} a$ is the maximal sum out of all the subsets $U \subseteq S'$ that satisfy $\sum_{\langle a,b \rangle \in U} b \geq 0$, since for any subset $U' \subseteq S'$ for which $\sum_{\langle a,b \rangle \in U'} a > 0$ we have $\sum_{\langle a,b \rangle \in U'} b = -\sum_{\langle a,b \rangle \in u'} a < 0$. Conversely, due to the same reasoning, the maximal subset $T' \subseteq S'$ for which $\sum_{\langle a,b \rangle \in T'} b \geq 0$, necessarily implies that $\sum_{\langle a,b \rangle \in T'} b = \sum_{\langle a,b \rangle \in T'} a = 0$ and therefore the corresponding set $T \subseteq S$ from which T' is constructed is a solution to ZSS. $\square$

The above result immediately gives us the following theorem:

THEOREM 2. *Computing the path trading solutions which maximizes the social welfare as well as the Nash product is NP-hard.* $\square$

3.3 Pseudo-polynomial Algorithms for Path Trading

While the SW and NB solutions are NP-hard to compute, a pseudo-polynomial algorithm can be used to obtain these solutions. We use a dynamic programming procedure which is dependent on the values of the input. To simplify notation, we will only consider costs and benefits in terms of AS_j: for some intra-AS route r, either in AS_i or AS_j, we can consider only benefits $b(r)$ to AS_j with negative benefits to represent costs for AS_j. Similarly, we look at costs $c(r)$ to AS_i, with negative costs representing benefits to AS_i. W.l.o.g we can assume that for every path r_i in AS_i we have $c(r_i) \geq 0$ and $b(r_i) \geq 0$, and similarly, for every path r_j in AS_j we have $b(r_j) \leq 0$ and $c(r_j) \leq 0$. The task of obtaining an individually rational solution thus translates to choosing a set of routes R s.t. $\sum_{r \in R} b(r) \geq 0$ and $\sum_{r \in R} c(r) \leq 0$. We shall now simply refer to benefits and costs of routes.

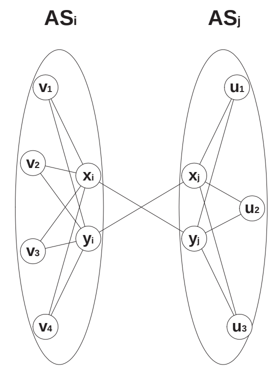

Figure 2: An illustration of the construction used to show $ZSS \preceq_p PT$. This instance is the product of a set S with $|S_-| = 3$ and $|S_+| = 4$.

For $r_M \in argmax_{r \in R} b(r)$ and $r_m \in argmin_{r \in R} b(r)$ let $M = n \cdot b(r_M)$ and $m = n \cdot b(r_m)$. For each $i \in [n]$ and $b \in \{m, \ldots, 0, \ldots, M\}$, let $\sigma_{i,b}$ be the subset of $\{r_1, \ldots, r_i\}$ of minimal cost s.t. $\sum_{r \in \sigma_{i,b}} b(r) = b$. If $\sigma_{i,b}$ includes two routes from the same ingress node we say it is illegal. Thus, we formally define the cost of $\sigma_{i,b}$ to be:

$$C(\sigma_{i,b}) = \begin{cases} \infty & \sigma_{i,b} = \emptyset \text{ or } \sigma_{i,b} \text{ is illegal} \\ \sum_{r \in \sigma_{i,b}} c(r) & \text{otherwise} \end{cases} \qquad (7)$$

We apply the following dynamic programming procedure:

1. Initialize $C(\sigma_{1,b(r_1)}) = c(r_1), C(\sigma_{1,0}) = 0$ and $C(\sigma_{1,b}) = \infty \quad \forall b \notin \{b(r_1), 0\}$;
2. for each $i \in [n]$ compute recursively, for all $b \in \{m, \ldots, M\}$:
   - if $b(r_{k+1}) \leq b$:

     $C(\sigma_{k+1,b}) = \min\{C(\sigma_{k,b}), C(\sigma_{k,b-b(r_{k+1})}) + c(r_{k+1})\}$
   - otherwise $C(\sigma_{k+1,b}) = C(\sigma_{k,b})$
3. Choose the individually rational solution which maximizes the objective (SW or NB) from the sets $\sigma_{n,b}$. If no such solution exists return $\emptyset$.

The iterative procedure terminates after $O((M - m) \cdot n^2)$ steps and outputs the sets $\sigma_{n,b}$ with the minimal costs required for all possible benefits within the integer range of $\{m, \ldots, M\}$. We computed a set of points in $\mathbb{R}^2$ on the Pareto optimum, and thus choosing the (individually rational) solution which maximizes either the sum of the utilities or their product satisfies our objective.

4. EVALUATION

4.1 Methods

To evaluate path trading we tested our results on the PoP level in the Internet AS graph. We used the DIMES [6] 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 [8] is yet another available source for the mappings of the Internet on a PoP level, though we found mappings for only ten ASes and therefore chose to use a map generate by the DIMES project. We show the distribution of the number of PoPs found in each AS in Fig. 3(left), and plot the number of PoPs against the AS degree in Fig. 3(right). In table 1 we show some statistics of the topologies of these PoP level graphs, all in 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 is 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 was also used in [7]. Also, the intention of our experiments is to show that ASes can find incentive to conduct path trade, as we consider real intra-AS topology and inter-AS links.

4.2 Results

The results presented here are of application of the SW solution only. The results of applying NB solution were similar, and are not presented in this paper. We first investigate the number of hops which can be saved as a function of the path trades conducted, as presented in Fig. 4(right), and in Fig. 4(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.

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. 5(right) we plot the number of hops which can be saved as a function of the ASes' PoP degree, and in Fig. 5(left) we plot a histogram of the number of hops saved normalized by the PoP degree. 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 mostly connected to other ASes through PoPs (one is not even among the top 10 most connected ASes). This is a testament that the inner AS topologies affect the number of hops which can be saved through path trade. For example, ASes with star topologies, or with a small number of border nodes, are examples where benefits from path trading solutions are to less likely to be found.

5. DISCUSSION

As our experimental evaluations suggest, path trading can considerably reduce the number of intradomain hops between an ASes' PoP nodes. Of course, 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. On the algorithmic front, although we found the pseudo-polynomial quite efficient, finding approximation algorithms which do not depend on the size of the input values would be a valuable addition to our existing work, as well as for further research on similar bargaining problems.

6. ACKNOWLEDGMENTS

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 Christos Papadimitriou and Jennifer Rexford for referring us to previous related work.

7. REFERENCES

[1] Dima Feldman. Automatic large scale generation of internet PoP level maps. *Masters Thesis, Department of Electrical Engineering - Systems, Tel Aviv University, Israel*, 2007.

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

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

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

[5] John F. Nash. The bargaining problem. *Econometrica*, 18:155–162, 1950.

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

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

[8] 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.

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

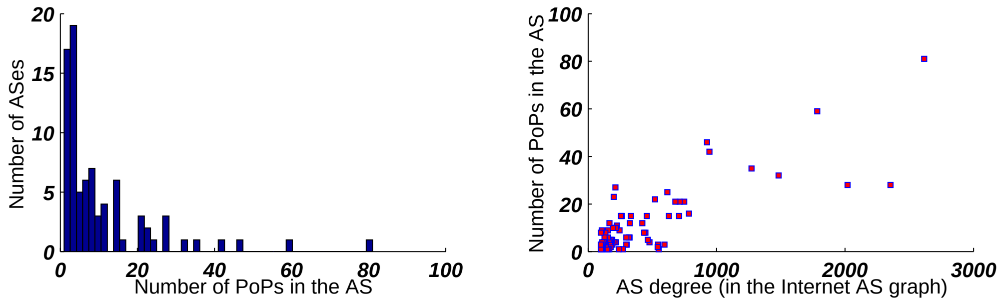

Figure 3: Statistics of the PoP level graphs. On the left, the distribution of the number of PoPs found in the 70 ASes for which we conducted are experiment is shown. On the right, the number of PoPs is plotted as a function of the AS degree in the Internet AS graph.

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

Table 1: statistics of the intra-AS PoP graph topologies.

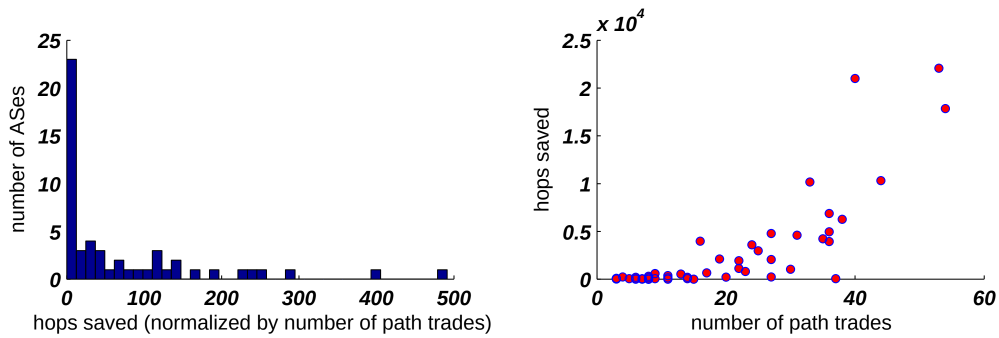

Figure 4: The total number of hops saved as a function of the number of path trades.

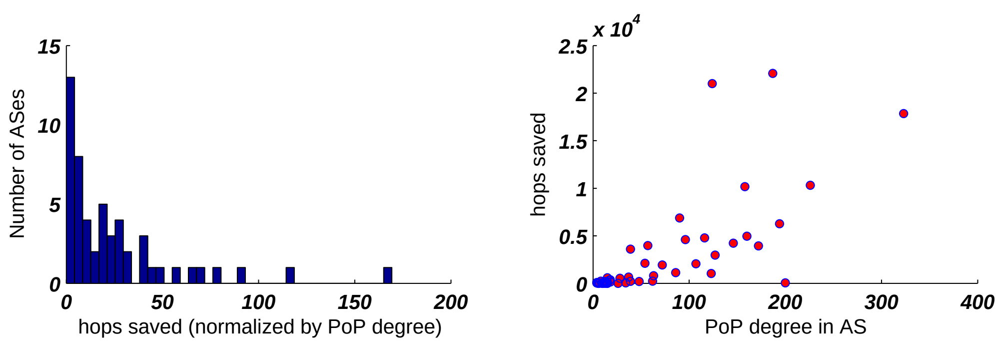

Figure 5: The total number of hops which can be saved as a function of the PoP degree.