The Impact of Routing Policy on Internet Paths

Hongsuda Tangmunarunkit[†], Ramesh Govindan[†], Scott Shenker[‡], Deborah Estrin[§]

Abstract—**The impact of routing policy on Internet paths is poorly understood. In theory, policy can *inflate* shortest-router-hop paths. To our knowledge, the extent of this inflation has not been previously examined. Using a simplified model of routing policy in the Internet, we obtain approximate indications of the impact of policy routing on Internet paths. Our findings suggest that routing policy does impact the length of Internet paths significantly. For instance, in our model of routing policy, some 20% of Internet paths are inflated by more than five router-level hops.**

Keywords—**Routing, Routing Policy, Policy Routing, Internet Paths**

I. Introduction

The earliest internet routing protocols attempted to construct lowest delay paths to destinations [1]. Thereafter, based on operational experience with the stability of delay-sensitive routing [2], deployed routing protocols evolved to essentially support global shortest hop-count routing [3].

Today's Internet contains several administrative domains (or Autonomous Systems, ASs). Within a domain, routing uses hop-count as a metric, but because intra-domain routing protocols support hierarchies, the resulting paths are not necessarily shortest in terms of hops. Routing between domains is determined by *policy*. Each autonomous system (AS) can, based on configured policy, independently select routing information from its neighboring ASs, and selectively propagate this information. These policies are not expressed in terms of hop-distance to destinations. Depending on how these policies are constructed, then, the resulting policy-based paths to destinations may incur more router-level hops than shortest-router-hop path routing.

In this paper, we ask the question: By how much does this hierarchical (inter/intra domain) form of routing affect Internet paths? This question was motivated by recent work [4] that observed that, for a significant fraction of Internet paths, there existed an intermediate node such that the composite path through the intermediate exhibited better performance (delay, throughput). In other words, routing in the Internet does not result in delay- or throughput-optimal paths. Perhaps this anomaly can be rectified by changing the Internet's routing infrastructure to be delay or load sensitive. Before we do this, however, it would be appropriate to understand how much of these observations can be explained by the fact that routing hierarchy and policy can result in longer hop paths. Our paper takes the first step towards this goal. Understanding this question can also be important for understanding the overall efficiency of the Internet's routing infrastructure. Finally, an answer to this question can also inform protocol evaluation studies which typically assume shortest-router-hop path (henceforth, shortest path) routing.

To understand how policy routing affects Internet paths, we use a simplified model of inter-AS routing policy that we call *shortest AS path* (Section II). Even though, in theory, routing policy can be completely arbitrary, many—but not all—existing routing policies are based on shortest AS paths. To infer the router-level path corresponding to this policy, we first begin with a router-level map of the Internet. On this map, we assign routers to ASs and obtain an *AS overlay* on top of our router-level map. This construction enables us to compare the router-level *policy path* between any two nodes with the shortest path on the map. Each of the steps in our construction represents a simplification of reality. As such, then, our results are only approximate indications of the impact of policy on Internet paths. However, at each stage, we carefully validate our construction using a collection of actual traceroutes that represent real paths generated by policy-based routing. This gives us some confidence that our conclusions are meaningful.

We find several surprising results (Section III). On average, about 20% of Internet paths are inflated by 50% or more. We also find that about half of the source-destination pairs benefit from a *detour*. For these pairs, there exists an intermediate node—a detour—such that the overall policy path length through this intermediate node is less than the policy path between source and destination.

To our knowledge, no related work has addressed the impact of routing policy on Internet paths. Our work, however, complements many pieces of recent work aimed at understanding the structure of the Internet, and the properties of its paths.

- Several efforts have focused on discovering router level topologies of the Internet [5], [6], [7]. Such mapping efforts are the crucial first step to help us understand the impact of routing policy.
- Other work has empirically studied the availability characteristics of paths [8], the loss and packet delivery performance of paths [9], and the existence of alternate paths with lower delay or higher throughput [4]. By considering hop-

[†] H. Tangmunarunkit and R. Govindan are with USC/Information Sciences Institute, 4676 Admiralty Way, Suite 1001, Marina del Rey, CA 90202. E-mail: {hongsuda,govindan}@isi.edu

[‡] S. Shenker is with ACIRI, 1947 Center St., Suite 600, Berkeley, CA 94704. E-mail: shenker@icsi.berkeley.edu

[§] D. Estrin is with UCLA and Information Sciences Institute. E-mail: estrin@isi.edu

This work was supported by the Defense Advanced Research Projects Agency under grant DABT63-98-1-0007. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the Defense Advanced Research Projects Agency.

distance between network nodes, our work examines the potential inefficiencies resulting from policy routing.

- Finally, more recent work has looked at macroscopic properties of the inter-AS topology [10], [11]. By relating the AS structure to the underlying router-level map, our *AS overlay* may be able to explain some of the observed macroscopic properties of the inter-AS topology in terms of the underlying physical structure.

To place our work in context, we point out two important caveats. First, it is well known that hierarchical routing can result in non-optimal paths [12]. Our paper quantifies the extent to which hierarchical routing in the Internet, together with routing policy, affects paths. Second, the correlation, if any, between path length and end-to-end delay is poorly understood. As such, then, our results cannot be directly extrapolated to observed delays on Internet paths. Nevertheless, our results are interesting since path hop count is the yardstick by which today's operational routing protocols are measured. Whether hop-count is the right metric for routing protocols is an orthogonal question that we do not address.

II. Methodology

Our first step in understanding the impact of routing policy on Internet paths was to obtain an Internet map. We used the *Mercator* program [7] for this purpose. Mercator uses hop-limited probes—the same primitive used in traceroute—to infer a router-level Internet map. It employs several heuristics to obtain as complete a representation of Internet topology as possible. One heuristic, informed random address probing, carefully explores the IP address space for addressable routers and hosts. Mercator also exploits source-route capable routers wherever possible to help discovering cross-links and thereby enhances the fidelity of the resulting map. Finally, it implements a technique for resolving interface aliases (interfaces belonging to the same router).

The map we used was collected between March 26, 2000 to April 10, 2000. Our map has 102,639 nodes and 142,303 links. This map is a smaller than the maps reported in [7]. Despite this, we believe we have, as [7] does, captured the transit portion of the Internet core, where policy impacts paths. In addition to the Internet map, we also collected the 61,485 traceroutes used in inferring the map. Because these traceroutes represent actual policy paths, we were able to use them to validate our policy model (see below).

Next, we attempted to compute an *AS overlay* on top of this Internet map. To do so, we assigned an autonomous system (AS) number to each router in our Internet map. For this, we used a BGP routing table from a publicly accessible route server[1]. For a given router interface address, we found the matching route entry in the BGP routing table. We then assigned that router to the origin AS in the AS path associated with the route entry. This technique works for globally routed addresses; [7] describes situations where some ASs number their routers from private address space. These addresses represent potential inaccuracies in our AS overlay computation. There were 3210 such private interface addresses in our map; to these we assigned a designated unused AS number.

Not all IP addresses had matching entries in the BGP routing table. Furthermore, in the routing table, route aggregation can mask the actual origin AS. For these reasons, we also used the RADB[2] to determine the origin AS. Finally, if a router had many interfaces corresponding to different AS numbers, we picked the most frequently assigned AS for that router. In spite of using these two sources of information, we were unable to resolve the AS numbers for 497 non-private IP addresses. To these we assigned another designated unused AS number.

After assigning each node an AS number, we then applied a simple collapsing algorithm to generate the AS overlay. The collapsing algorithm recursively marks neighbors belonging to the same AS with the same "color". Each color represents a node in the AS-level map. However, due to our incomplete information (both in the Internet map [7] and in AS number assignment), there were many disjoint clusters of nodes belonging to the same AS. In most cases, we found that such ASs normally have one large component with many small components each with a small number of nodes. We solved this problem by identifying the biggest cluster of each AS and re-assigning the smaller components (about 20,000 nodes total) to the topologically nearest AS. The resulting AS overlay consists of 2,662 nodes and 4,851 links.

Clearly, there are several sources of error in both the collected map, and in our AS overlay generation technique. The best we could have hoped for is that our techniques result in a first-order approximation of the actual inter-AS overlay. To verify this, we compared some macroscopic properties [10] (such as the degree distribution, degree rank distribution and hop-pair distribution) of the resulting AS overlay to that of an AS level topology inferred from BGP routing table collected on April 10, 2000[3] (Figure 1). Notice that these macroscopic properties are in qualitative agreement.

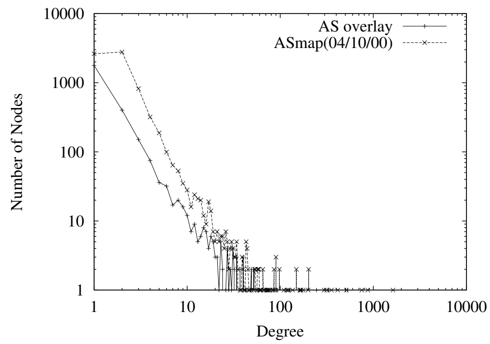
As an additional validation, we compared the collection of AS paths in a BGP routing table dump with the shortest AS path between the two corresponding nodes in our AS overlay. Figure 2 shows the distribution of the path length difference between node pairs in our map and the routing table dump. About 93% of node pairs in our overlay are within one AS hop of the corresponding path in the BGP routing table.

The final step in our methodology is to select a plausible model for Internet routing policy. In general, there exist two kinds of inter-ISP relationships, a provider-customer re-

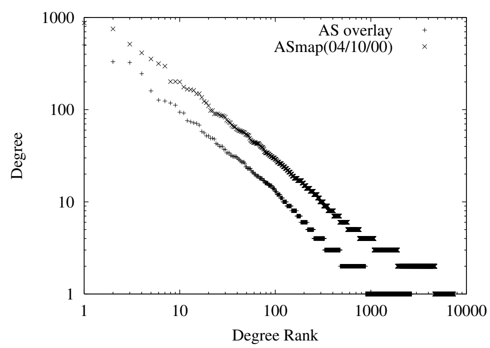
[1] route-views.oregon.ix.net, data used with permission from David Meyer.

[2] whois.radb.net

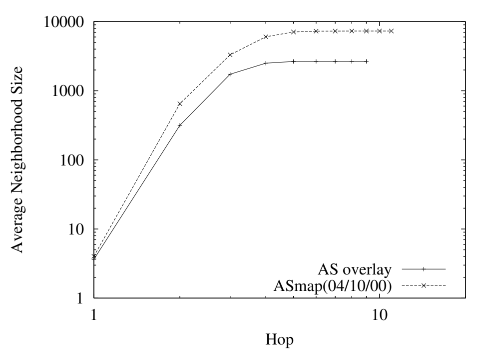
[3] This AS topology has 7,306 nodes and 14,707 links.



(a) Degree Distribution



(b) Degree Rank Distribution



(c) Hop-pair Distribution

Fig. 1. Macroscopic comparison of AS overlay and the AS topology

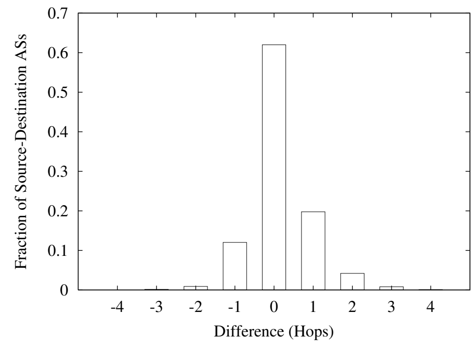


Fig. 2. Distribution of Path length difference between node pairs in AS overlay and AS paths in the routing table dump.

lationship and a peering relationship. To our knowledge, the prevalent and widespread routing policy practice is that an ISP picks the shortest AS path based route for customer and peer supplied routes. Furthermore, an ISP only propagates its customer supplied routes to its peers, never routes supplied by other peers. This latter rule can result in AS-level paths not corresponding to shortest AS path. In the absence of information about the exact nature of inter-ISP relationships, we use a simplified model in which the policy path between two ASs is determined by the shortest AS path between them.

Clearly, this is only an approximate model of routing policy. We validate this model by comparing the length of AS path corresponding to each traceroute in our collection, with length of shortest AS path between the corresponding source and destination. Figure 3 shows the pathlength comparison between the policy paths and the shortest AS paths of the collected traceroutes. We found that the shortest AS path underestimates the traceroute AS path by 1 AS hop or less for 70% of the traces, and less than 2 hops for 95% of the traces. Though the difference seems small, it doesn't represent spectacular agreement with our model since many of the AS paths are relatively small (5 hops or so). Nevertheless, this validation (and another described in Section III-A) is good enough to encourage us to pursue our initial understanding of the impact of routing policy on Internet paths.

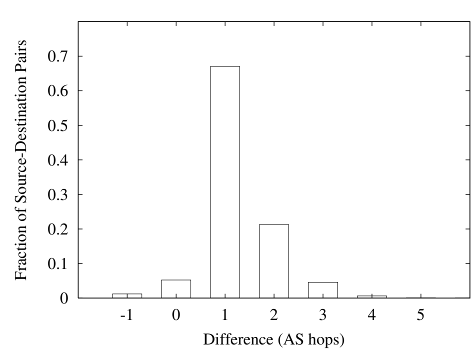


Fig. 3. Validating the shortest AS path model for routing policy.

III. RESULTS

Having a model for AS policy, we can now compute, for any given pair of nodes in the Internet map, the router-level path between them as determined by routing policy. We call such a path the *policy path*. Furthermore, we can also compute the router-level *shortest path* between the same pair of nodes.

Being thus empowered, we analyze the impact of routing policy on Internet paths by asking the following four different questions. These questions are all complementary ways of looking at how routing policy skews paths in the Internet.

- By how much does policy *inflate* unicast paths? That is, how different is a policy path from the corresponding shortest path?
- For nodes A and B in the Internet map, does there exist an intermediate node I such that the sum of the policy paths from A to I and from I to B is less than the policy path from A to B? If such an intermediate exists, A and B can circumvent routing and communicate using fewer hops via I.
- Does policy routing *funnel* Internet paths through larger ASs?
- How does policy routing impact multicast tree sizes?

Before we discuss our results, a couple of caveats are worth mentioning. First, unless otherwise specified, all our results below are derived from our policy model, not from the traceroutes. The model enables us to compute policy paths for arbitrary source destination pairs. Second, our definition of an "Internet path" implicitly considers paths between each pair of nodes in the map equally likely. In practice, for example, there may not exist any end-to-end conversations between two backbone routers. To more accurately model the distribution of Internet paths, it is necessary to have an estimate of how many "hosts" are attached to each router, an estimate we do not have.

A. *Inflating Unicast Paths*

Our first examination of the impact of routing policy considers the difference between policy paths and their corresponding shortest paths. Specifically, we look at two metrics. The first, *inflation ratio* measures the ratio of the length of a policy path to the length of the corresponding shortest path. Figure 4 plots the cumulative distribution of this metric. This figure was obtained by computing the inflation ratio for each random source-destination pair in our Internet map, a total of 100,000 random pairs. It shows a quantatively surprising impact of routing policy. Some 20% of Internet paths are inflated by more than 50%. Some policy paths are inflated by a factor of nearly 4. Finally, for only a fifth of the paths does the policy path length equal the length of the shortest path.

A second metric, the *inflation difference*, provides an alternative view of the impact of routing policy. This metric represents the absolute difference, in terms of the number of router hops, between the policy path and the shortest path.

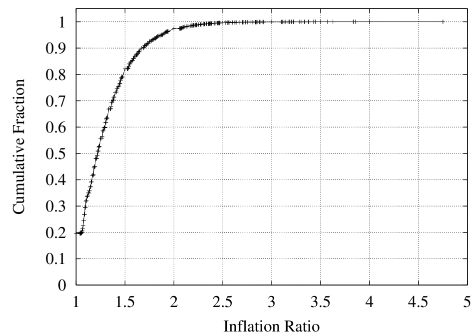


Fig. 4. Cumulative distribution of inflation ratio

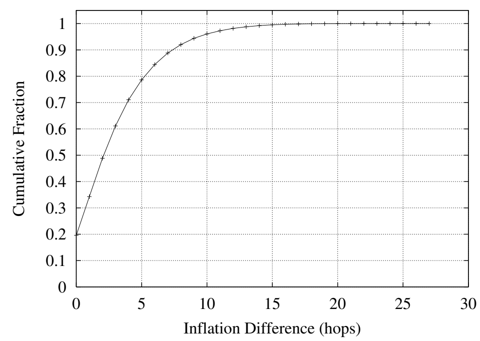


Fig. 5. Cumulative distribution of inflation difference

Figure 5 plots the cumulative distribution of inflation difference. For nearly 20% of the node pairs, the policy path is longer than the shortest path by more than 5 hops. Furthermore, there exist some node pairs for which the policy path is longer by 25 hops than the shortest path.

Perhaps a more reasonable way to consider the impact of routing policy might be to evaluate inflation for node pairs whose shortest paths are of the same length. Thus, Figure 6 plots the cumulative distribution of the inflation ratio for four different shortest path lengths. Not surprisingly, in a distribution sense, longer paths are less inflated in proportion to their lengths—in general, shorter paths have more "room" for inflation than longer paths. Furthermore, the smaller lengths (notably length 5) have a significantly long tail.

Finally, Figure 7 depicts the inflation difference for different shortest path lengths. This shows an interesting trend, namely that longer paths are more absolutely inflated than shorter paths. The exception to this observation is the path length 20. We conjecture that the explanation for this exception is that really long paths have less absolute "room" to "grow". This observation has interesting consequences that we explore in Section III-B.

A subtle point that underlies the results presented in this and the next sections is that, for a given node pair, there may exist many different "shortest" AS paths. The corresponding router-level policy paths for each possible shortest AS path may be widely different. For example, if shortest AS path X

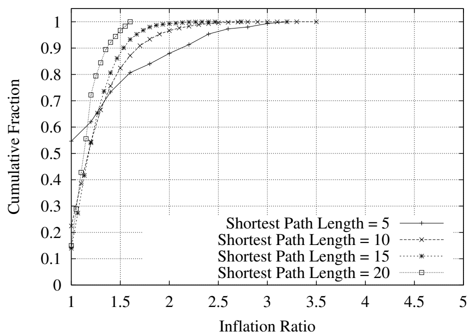


Fig. 6. Cumulative distribution of inflation ratio by path length

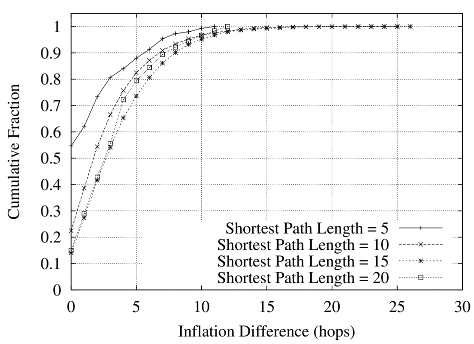


Fig. 7. Cumulative distribution of inflation difference by path length

traverses larger ASs than shortest AS path *Y*, one might expect that the policy path corresponding to *X* might be longer than that for *Y*. Because it is computationally difficult to enumerate all possible shortest AS paths, all results in this paper were derived by essentially randomly selecting shortest AS paths.

Does this discrepancy affect our conclusions? One way to answer this question is to compare the inflation ratio and difference distributions for our collection of traceroutes (Section II) with the distributions presented above. Figure 8 does this. Our shortest AS path computation yields more conservative inflation ratios, and similar inflation differences, to the traceroute data. We also tried other approaches to select shortest AS paths, and found that most of the results presented in this paper are relatively insensitive to the way we pick the shortest AS path. Finally, we also compared our inflation ratios and differences to those obtained from a different mapping effort [13][4]. Figure 9 shows that our policy model results in fairly conservative inflation ratios and differences.

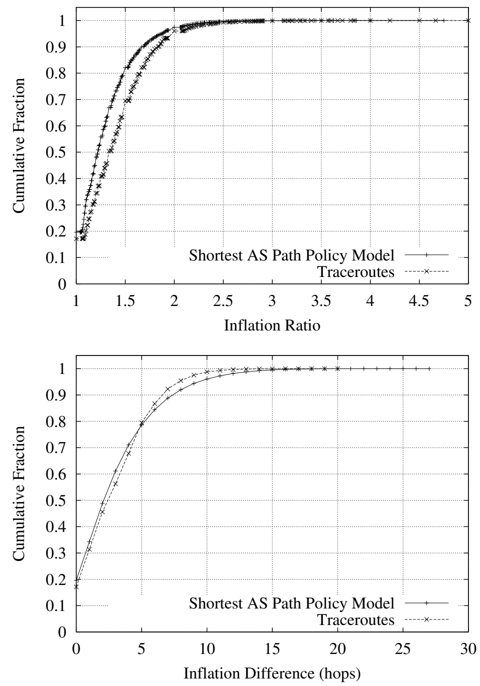


Fig. 8. Comparison of inflation ratio and difference for traceroutes and for shortest AS path

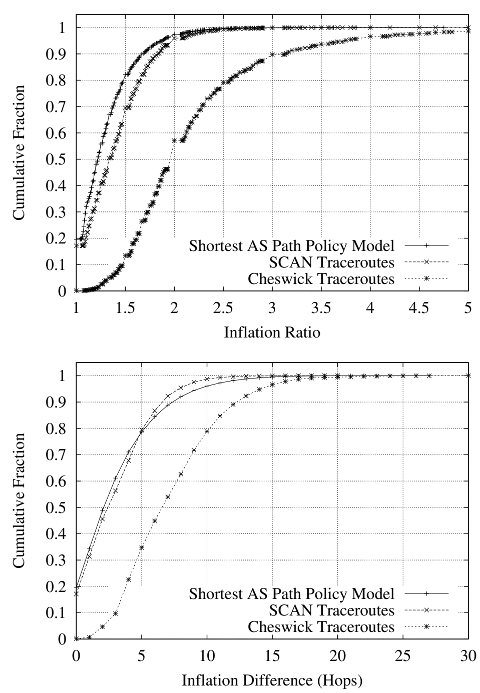


Fig. 9. Comparison of inflation ratio and difference for Lucent traceroutes and for shortest AS path

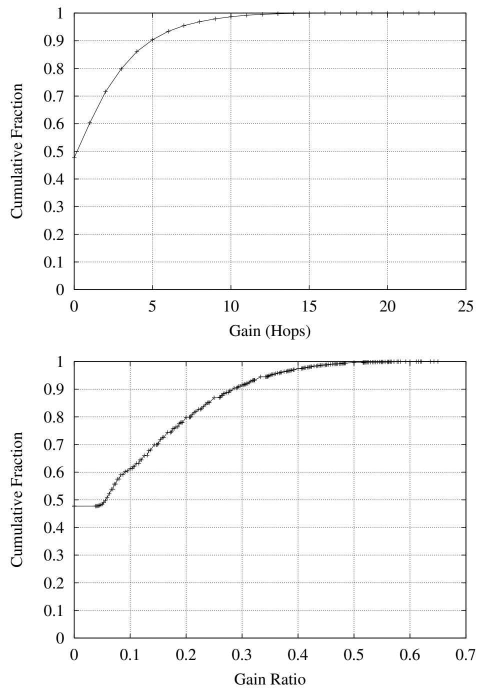


Fig. 10. Detour gain and gain ratio distributions.

B. Finding detours

We have seen that longer policy paths are generally more inflated in the number of hops than shorter ones. This gap motivates the questions we examine in this section: for what fraction of node pairs do there exist better alternate paths (called *detours*) and by how much are these detours better than the policy paths?

Before we answer these questions, we more carefully define a detour. Consider a pair of nodes A and B. We say that there exists a detour between A and B if there exists an intermediate node I such that:

- I lies in a different AS than A or B,
- the AS path from A to I and from I to B is collectively *longer* than the shortest AS path between A and B, and
- the sum of the router-level policy paths between A and I and between I and B is less than the policy path between A and B.

Intuitively, a detour represents a way to circumvent routing by relaying communication between A and B at the application level.

To answer the above questions, we randomly generate a large number of source-destination pairs in which source and destination are selected from different ASs. For each source-destination pair, we find a policy path and the best detour path (if there exists one), then measure the difference between the two paths. We define two metrics for quantifying detours. The detour *gain* is the absolute difference, in router-level hops between the policy path and the best detour path. We define the gain to be zero if there exists no detour for that particular policy path. The detour *gain ratio* is the ratio between the gain and the length of the policy path between source and destination.

Figure 10 shows the cumulative distribution of the two metrics. Surprisingly, for 50% of the sampled paths, there exist superior detour paths. Furthermore, 20% of the sampled paths have a detour that is 3 hops or more less than the corresponding policy path. An alternative result is that 20% of the paths have a detour that is more than 20% shorter. Our finding is in quantitative agreement with that of [4], although their findings were a more direct measure of path performance whilst ours is based only on path length. To what extent our finding is an explanation for theirs is a little unclear.

C. Path Concentration Around Large ASs

The previous two sections have studied the impact of routing policy on Internet paths. In this section, we look at a slightly different question: does routing policy force Internet paths through larger ASs? This question is one aspect of a larger, more general, question: Is topological connectivity rich enough that the logical connectivity imposed by policy routing significantly skews the paths Internet traffic takes?

To study this question, we define for each node pair the *dominant AS* to be the largest (by size[5]) transit AS encountered in the path between the two nodes. For each AS, we then define a *dominance fraction*; the fraction of node pairs for which that AS is the dominant one. We are interested in the correlation between size and the dominance fraction. In computing our dominance fraction, we do not consider node pairs which lie within a single AS, or in adjoining ASs. These, by definition, do not have any transit ASs between them.

We measure this metric both for our policy approximation, as well as for shortest router-level paths. Figure 11 shows that, regardless of whether policy routing is used, or shortest paths are used, the top 15 ASs by size dominate about 90% of the paths. This number is surprising not only in its magnitude, but also in the absence of any qualitative difference in dominance correlation between policy routing and shortest path routing.

In Section III-A, we discussed the existence of multiple shortest AS paths. Unlike other results, it turns out that the dominance correlation is at least quantitatively affected by our method of selecting the shortest AS path. In particular, if instead of picking larger ASs to explore in our breadth

[4]To measure the inflation metrics on this collection of traceroutes, we need to create an instance of Internet map corresponding to the traceroute collection. For this, we applied the alias resolution technique described in [7], then synthesized the Internet map.

[5]An alternative way to define a dominant AS is by its degree in the AS topology. We find that the largest ASs by size are well-correlated with the largest ASs by degree. It is therefore not surprising that when we use this alternative definition for dominant AS, our results do not change qualitatively.

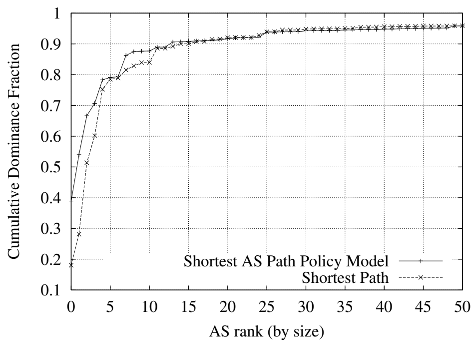


Fig. 11. Cumulative dominance fractions for the top 50 ASs.

first search for shortest AS paths, we choose smaller ASs, then, we find the dominance correlation curve to be different. Even by pessimally picking the shortest AS path, we find that almost 70% of the paths are dominated by one of the top 25 ASs. We also checked if our results were sensitive to the method of selecting router-level shortest paths. They were not.

D. Multicast

A final question we look at is the effect of policy on multicast tree sizes. For this, we assume a simple random source and receiver placement, and compute source-rooted multicast trees using shortest paths and policy paths. We then compare the relative sizes of the shortest-path tree and the policy tree.

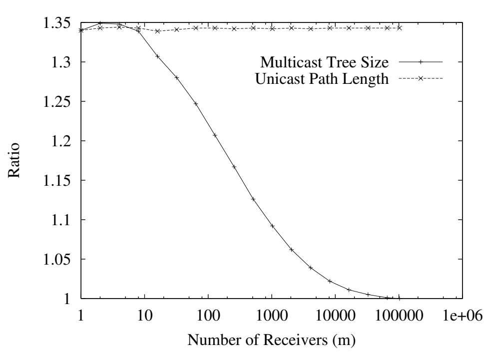


Fig. 12. The effect of policy on multicast tree sizes

Figure 12 shows the ratio between the size of a policy tree and the corresponding shortest-path tree as a function of the number of receivers. In the same figure, we also show the average ratio between unicast path length of a policy path versus a shortest path from the source to all receivers. As expected, the average unicast ratio is independent of the number of receivers in the group. This number is consistent with the average inflation ratio shown in Figure 4.

However, for a small number of receivers, the policy trees are actually larger by about 30% than shortest-path trees. As the number of receivers increases, policy trees continue to grow larger than shortest-path trees. This result is somewhat counter-intuitive since one might expect more path sharing with policy routing between receivers in the same AS, and hence smaller policy trees. However, the reduction of path sharing is probably offset by the overall increase in path length due to inflation.

IV. Conclusions

Does policy have an impact on Internet path length? Our paper clearly makes the case that it does, albeit for a simple model of prevailing routing policies. In our model, nearly 50% of paths benefit from a detour. Some small multicast trees are enlarged almost 30% by policy.

While our shortest AS path approximation may be rendered obsolete by more complicated routing policy, there exists a more enduring representation of our work. Shortest AS path represents the routing that would have resulting from a pure (policy-free) hierarchical routing in the Internet. In this sense, our paper quantifies the impact on Internet paths of the particular instance of hierarchy that the Internet has evolved to today.

Acknowledgements

We would like to thank Pavlin Radoslavov for his help in obtaining BGP and RADB data, and Cengiz Alaettinoglu for discussions regarding prevalent Internet routing policies.

References

[1] J. M. McQuillan, I. Richer, and E. C. Rosen, "The New Routing Algorithm for the ARPANET," *IEEE Transactions on Communications*, vol. 7, no. 1, pp. 1–7, 5 1980.

[2] A. Khanna and J. Zinky, "A Revised ARPANET Routing Metric," in *Proceedings of ACM SIGCOMM*, 1989.

[3] C. C. Cheng et. al., "A Loop-Free Extended Bellman-Ford Routing Protocol Without Bouncing Effect," in *Proceedings of ACM SIGCOMM*, 1989, pp. 224–236.

[4] S. Savage, A. Collins, E. Hoffman, J. Snell, and T. Anderson, "The End-to-End Effects of Internet Path Selection," in *Proceedings of ACM SIGCOMM*, Boston, MA, September 1999.

[5] Hal Burch and Bill Cheswick, "Mapping the Internet," *IEEE Computer*, vol. 32, no. 4, pp. 97–98, April 1999.

[6] K. C. Claffy and D. McRobb, "Measurement and Visualization of Internet Connectivity and Performance," http://www.caida.org/Tools/Skitter/.

[7] Ramesh Govindan and Hongsuda Tangmunarunkit, "Heuristics for Internet Map Discovery," in *Proceedings of IEEE Infocom*, Tel-Aviv, Israel, April 2000.

[8] V. Paxson, "End-to-end Routing Behavior in the Internet," in *Proceedings of the ACM SIGCOMM Symposium on Communication Architectures and Protocols*, San Francisco, CA, September 1996.

[9] V. Paxson, "End-to-end Internet Packet Dynamics," in *Proceedings of the 1997 ACM SIGCOMM Conference on Communication Architectures and Protocols*, September 1997.

[10] C. Faloutsos, M. Faloutsos, and P. Faloutsos, "What does Internet look like? Empirical Laws of the Internet Topology," In Proceedings of ACM SIGCOMM 1999, Boston, MA, September 1999.

[11] G. Phillips, H. Tangmunarunkit, and S. Shenker, "Scaling of Multicast Trees: Comments on the Chuang-Sirbu Scaling Law," In Proceedings of ACM SIGCOMM 1999, Boston, MA, September 1999.

[12] L. Kleinrock and F. Kamoun, "Hierarchical Routing for Large Networks: Performance Evaluation and Optimization," *Computer Networks*, vol. 1, pp. 155–174, 1977.

[13] Bill Cheswick, Hal Burch, and Steve Branigan, "Mapping and Visualizing the Internet," in *USENIX*, San Diego, CA, June 2000.