Internet Path Inflation Due to Policy Routing

Hongsuda Tangmunarunkit
hongsuda@isi.edu

Ramesh Govindan
govindan@isi.edu

Scott Shenker
shenker@icsi.berkeley.edu

ABSTRACT

In our previous work,[1] we used a simplified model of routing policy in the Internet to study the impact of policy routing on Internet path-lengths. This prior work suffered from two shortcomings—it was based on a single snapshot of the Internet topology, and our simplified policy model could generate AS paths that violate peering relationships. In this paper, we address these two shortcomings by re-examining our results with respect to a more recent snapshot of the Internet, and improving the policy model to avoid peering violation. We find that our prior observations regarding the path inflation due to routing policy appear to hold both across time and with respect to a more sophisticated model of routing policy.

Keywords: Internet Paths, Paths, Routing Policy, Policy Routing, Path Inflation

1. INTRODUCTION

Today's Internet contains several administrative domains or Autonomous Systems (ASs). Inside a domain, routing uses router-level hop-counts[†] as a metric to direct packets toward the shortest paths.[2] However, routing between domains is determined by *policy*.[3,4] Based on its configured policy and on received routing information, each AS can independently construct sequences of ASs (AS-paths) to destinations. Routing information is expressed in terms of AS-paths consisting of ASs that agree to carry packets to destinations, as opposed to router-level distance to destinations. Therefore, depending on how the policies are constructed, the resulting policy-based paths (or *policy paths*) may incur more router-level hops than shortest router-hop path routing.

Policy-based routing (or policy routing) is a form of hierarchical routing. Although it has long been accepted that hierarchical routing can *inflate* shortest paths, the extent of this inflation has rarely been examined. Part of the difficulty is the lack of a router-level topological representation of the Internet, which is necessary for studying shortest-router-hop paths. Moreover, obtaining accurate policies from each AS is also difficult since this information is considered proprietary and not normally revealed. Fortunately, due to recent developments in Internet router-level topology discovery[5–9] we now have an access to the router-level representation of the Internet. With this map, we can use an approximate model of routing policy to conduct an initial study of the impact of routing policy on network paths.

Our previous work[1] used a snapshot of a router-level Internet topology and a simple shortest-AS path policy model to study the degree of inflation due to policy routing. We found several surprising results. We found that at least 80% of Internet paths are inflated by the policy and about 20% of paths are inflated by at least 50%. Moreover, about 50% of the source-destination pairs benefit from a superior detour path—a path through an intermediate node such that the combined policy path-length (router-hops) through this intermediate node is less than the direct policy path between source and destination. This finding was in agreement with recent work[10] 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 in terms of delay and throughput.

Although our prior work has quantified the impact of routing policy on Internet paths and hence given researchers some idea of the performance of the current routing scheme, there are two shortcomings in the work itself. First,

This work was supported by the Defense Advanced Research Projects Agency under grant F30602-00-2-055. 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.

Send correspondence to Hongsuda Tangmunarunkit (hongsuda@isi.edu). Hongsuda Tangmunarunkit and Ramesh Govindan are affiliated with ISI/USC 4676 Admiralty Way, Suite 1001, Marina del Rey, CA 90292. Scott Shenker is affiliated with ACIRI 1947 Center St., Suite 600, Berkeley, CA 94704-1198.

[†]More generally, intra-domain routing uses administratively assigned link weights in computing shortest paths within the domain. Lacking detailed knowledge of such weight assignments, we make the simplifying assumption that intra-domain routing is shortest-hop based.

our result was obtained based on a single snapshot of the Internet topology. The Internet is growing and evolving. A snapshot of the Internet topology today might look different from a snapshot taken a year ago. Consequently, our earlier observation might not hold for the current Internet topology. In this paper, we discuss the degree of path inflation due to policy from a more recent snapshot (one taken nearly a year after that described in our previous work) of the Internet. We find that the current Internet path inflation qualitatively agrees with our earlier results.

The second shortcoming of our previous study is the simplified policy model. The model does not take the peering relationship among AS domains into consideration; it simply ignores whether two ASs are related by peer-to-peer, provider-to-customer or sibling-to-sibling agreement. Thus it might generate some paths that may violate some peering relationship that would be avoided in reality; an actual AS path should never transit through a customer AS in between two providers, for example, MCI-USC-SPRINT is not a realistic AS path. In this work, we improve our policy model to take into account the peering relationship among ASs and to avoid such peering relationship violation. Surprisingly, we find that the degree of inflation is not very sensitive to the policy model used in the study. Our investigation indicates that when the large-scale statistics of path-lengths due to policy are desired, the shortest AS path is a reasonable model to use. The model is simple and it avoids having to identify the peering relationship among nodes.

Our paper in organized as follows: Section 2 examines the impact of policy routing on Internet paths with respect to a recent snapshot of the Internet topology. Section 3 investigates the sensitivity of the routing policy impact due to a more realistic policy model. Finally, we conclude our work in Section 4.

2. IMPACT OF ROUTING POLICY ON INTERNET PATHS

Recently, we have investigated the impact of policy routing on Internet paths using the Internet router-level map and a simplified routing policy model, the shortest AS path model.[1] We observed that policy routing affects paths significantly. However, our prior study regarding Internet path inflation was based on a single snapshot of the Internet topology. Obviously, the degree of path inflation is influenced by the network map, i.e. different maps will yield different levels of path inflation. The Internet has constantly growing and evolving. Without a better understanding about the Internet topology or how it is evolving, it is difficult to predict whether the level of Internet path inflation remains constant across time. Therefore, in this section, we examine the policy impact on Internet paths on a more recent snapshot of the Internet. In general, we are interested in investigating whether our earlier observations about Internet path inflation were an artifact of the particular snapshot of Internet topology we observed.

To study the degree of Internet path inflation across time, first we needed to collect another snapshot of the Internet topology. We did this by running a slightly more optimized version of Mercator software.[9] The map was collected between May 01, 2001 to May 06, 2001, approximately a year after the previous map was collected. It contains 170,589 and 215,385 links.

The next step was to extract an AS overlay map on top of the Internet router-level map. A node in the AS overlay map represents an AS domain and a link between two ASs represents a connectivity or peering relationship between them. To generate an AS overlay map, we first used the information from the BGP routing table dump[11] and from the RADB database[12] to resolve an AS number for each IP address. Then we applied a simple collapsing algorithm to collapse router nodes with the same AS number to the same domain. Our previous work[1] describes the methodology for generating an AS overlay in more detail [‡]. The resulting AS overlay map for the newer router-level map has 6,645 nodes and 9,706 links (about 2.5 times the size of the previous AS overlay). Figure 1 shows the macroscopic properties—the degree distribution, degree rank distribution, and average neighborhood size[14]—of the AS overlay with respect to the actual AS map obtained from the BGP routing table dump at approximately the same period of time. It is obvious that both maps exhibit qualitatively similar properties.

Once we have the router-level map and the AS overlay map, we then apply a simplified routing policy model—shortest AS path—to examine the path inflation due to policy routing[§]. Essentially, we want to compare the router-level path determined by the routing policy model with the shortest-router-hop path. A policy path between a source-destination pair is the path that traverses through the minimum number of AS hops. A router-level policy path is constructed by first determining the shortest AS path between the source and destination ASs, then a sequence

[‡] Another way of generating an AS overlay was recently proposed by Chang et al..[13] They looked at the router-level path traces more thoroughly and inferred the AS-level topology based on the traces.

[§] Our previous work[1] has more detail about the validation of this simplified routing policy model.

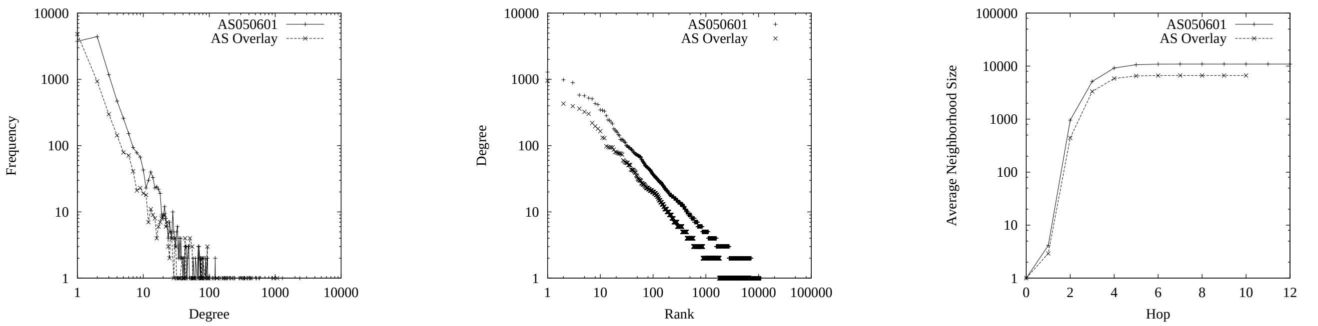

Figure 1. Our AS overlay and the actual AS map have qualitatively similar macroscopic properties.

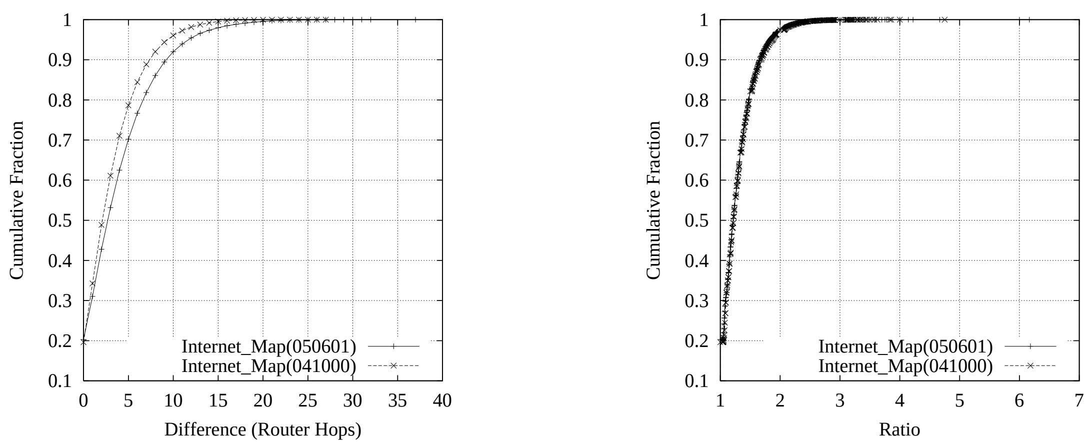

Figure 2. Inflation difference and inflation ratio.

of shortest-router-hop paths connecting a node in the current AS to the closest node in the next-hop AS is combined to form a policy path. A shortest-router-hop path is just a shortest path in terms of router hops that directly connect between a source and destination. We measure the inflation difference and inflation ratio between the policy path and the shortest-router-hop path. The inflation difference is the difference in terms of router-hops between the policy path and the corresponding shortest paths. The inflation ratio is the distance ratio (in terms of router-hop) between the policy path and the corresponding shortest path.

Figure 2 shows the cumulative distribution of the paths with respect to the inflation difference and ratio. We have included the cumulative distribution of inflation with respect to our previous study for comparison. Again, the two data sets were obtained from two snapshots of Internet topology that are approximately one year apart. We observe that the Internet path inflation with respect to the older map, when compared to the newer map, is more conservative according to the Inflation difference and is approximately the same according to the inflation ratio. Moreover, we also find that the average path-length on the newer map is longer than the average path-length on the older map. This is possibly due to the bigger size of the newer map, i.e., the new map is about 50% larger than the older one[¶]. Therefore, it is not so surprising that paths on the newer map reveal a greater inflation difference than the older map. Nevertheless, the cumulative distributions of inflation ratio are approximately the same.

Finally, in the previous study, we also observed that the longer paths, when compared to the shorter ones, are more inflated in terms of the absolute difference but are less inflated in proportion to their lengths. Figure 3 shows the inflation difference and ratio of paths with respect to different shortest-path-length. We notice the similar behavior

[¶]Whether this is the result of the growth in the Internet or a better router discovery technique revealing more information remains an open question.

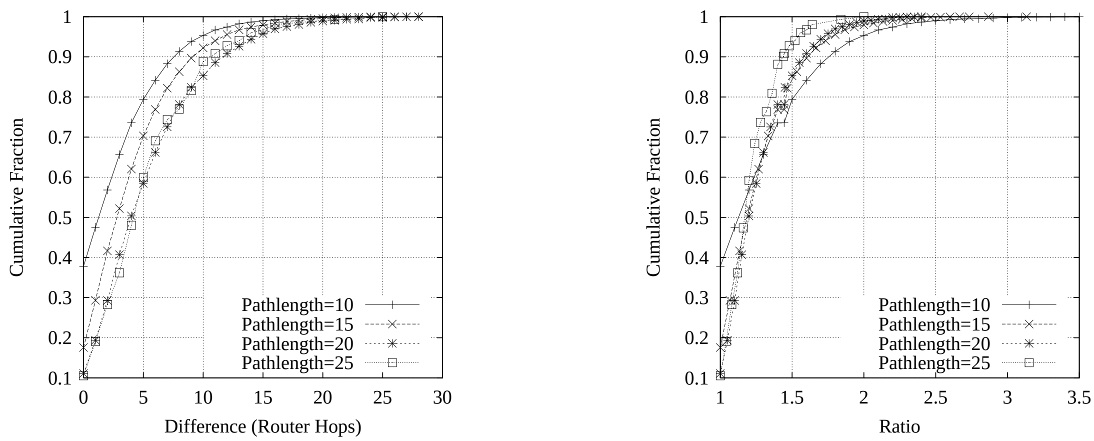

Figure 3. Inflation difference and inflation ratio with respect to different shortest-path-lengths.

in this data set as well. Thus, our recent findings seem to suggest that our observations regarding the Internet path inflation due to policy may hold across different snapshot of the Internet and time.

3. REALISTIC ROUTING POLICY MODEL

In our prior work, we proposed a simplified policy model, namely shortest AS path, for policy-based routing. This policy model can be inaccurate and can violate provider-customer relationships. Therefore, in this section, we improve the accuracy of this simplified model by avoiding such violation in the AS path while maintaining the minimization of the AS hop count. This improved model will allow us to examine the sensitivity of our conclusions in Section 2 to a more accurate routing policy.

Our shortest AS path policy model suffers from one important drawback. Although the model enables us to study router-level paths between any two nodes in the network, it doesn't take into consideration the peering relationship among AS nodes. As a result, some of the generated paths may not be realistic in the Internet context, i.e. they may violate peering relationships by transiting through a stub domain in between two transit-domains and hence are considered invalid. As an example, an AS path traversing through MCI-USC-SPRINT is an invalid path since USC is a customer of MCI and the packets between the national ISPs should never transit through one of their customers.

Fortunately, a recent work by Gao[15] has described a more realistic technique for inferring AS peering relationships, e.g. provider-customer, peer-peer or sibling-sibling relationship. Their work makes two assumptions: 1) that there is a strong correlation between the AS degree and AS size, e.g. an AS with larger degree (i.e. AS with many connectivities to its peers) is a bigger AS domain in size, and 2) that the AS paths are hierarchical. They assume that one signature of a hierarchy is that paths may go up, down, or up and then down the hierarchy. A path connecting two regional ISPs must traverse up the hierarchy to the national ISP, then the two national ISPs exchange packets and packets go down the hierarchy to the destination ISP. They apply the two assumptions to classify the types of AS paths or routes that can appear in BGP routing tables[11,16] and then infer the peering relationship based on this classification. To proceed with our study, we need to annotate our AS overlay map with peering relationships. Once the annotated AS overlay map is obtained, we can improve our policy model to consider only valid shortest AS paths in our study of impact of policy routing on Internet paths.

3.1. Methodology

How do we annotate our AS overlay map with peering relationships? In another words, how do we determine whether a link is a provider-customer, peer-peer link or sibling-sibling link[15] on our AS overlay? Though Gao's algorithm can be used to determine the peering relationship of an actual AS map, one of the major component for the algorithm is a collection of actual AS paths. Since we do not have a collection of paths corresponding to our AS overlay map, we cannot apply this algorithm directly. We solve this problem by first applying Gao's technique to derive peering relationships for the links on an actual AS map. We chose the AS map that was collected on May 06, 2001. This map was generated from a collection of AS paths obtained from a BGP routing table dump.[11] The map consists of 10941 nodes and 22568 links.

Once we have the annotated AS map, we then determine a peering relationship for links on the AS overlay map based on the peering relationship information from the actual AS map. The peering relationship is determined as follows. For any link l connecting two ASs in the AS overlay map, if there exists a corresponding link in the actual AS map, the peering relationship of link l is assigned by the relationship of the corresponding link in the actual AS map [‖]. If one of the ASs doesn't exist in the actual AS map, then the existing AS is the provider of the non-existent AS. If both of the ASs on the end nodes do not exist in the actual AS map, the AS with the larger degree is a provider of the AS with the smaller degree. Lastly, if both ASs exist in the actual AS map but there is no corresponding link on the actual AS map, then if both ASs have large degrees (i.e. degree > 60) or both ASs are peers with many other ASs (e.g. the number of peer-peer links that each of them have with other ASs are more than 15 links), then both ASs have a peer-to-peer relationship. However, if the previous condition is not true, then the AS with larger degree is a provider of AS with smaller degree. The pseudo-code for this heuristics is provided at Appendix A.

After each link on the AS overlay map is assigned a peering relationship, we then modify our routing policy model to take the direction of each path into account. A path traverses up the hierarchy through customer-provider links, traverses down the hierarchy through provider-customer links, and traverses across nodes in the same hierarchical level through peer-peer links. A valid path is assumed to be hierarchical: the path should never traverse through a customer-provider link once it traverses through a provider-customer link. Analogously, after the path is traversing down, it will never traverse up again. There may be more than one valid path between any source and destination; our modified policy model will always pick one of the equally shortest ones.

We should point out that the peering relationship assignment of links in both AS maps (i.e. the actual AS map and the AS overlay) is not perfect; there exists a small fraction of node pairs that cannot be reached from one another. However, among all possible node pairs on both maps, these unreachable pairs account for less than 2%. Additionally, the heuristics that we used for assigning peering relationships for links on our AS overlay is conservative. A peer-peer link allows the traffic to flow in both directions without any restriction while a provider-customer link enforces some direction on the path. Since our heuristics only assign peer-peer to links that connect big or important ASs (AS with many peers) and assign provider-customer to links in most cases, the heuristic is conservative.

3.2. Path inflation with respect to realistic policy model

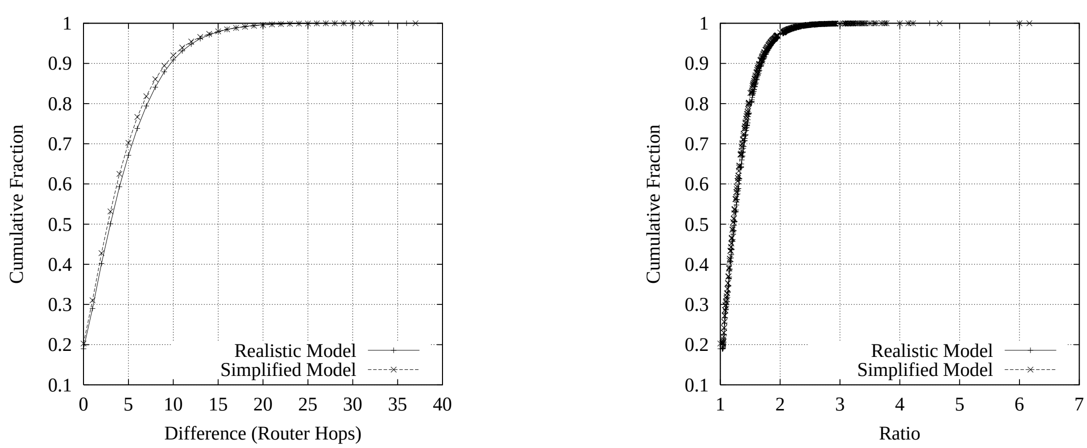

Figure 4. Inflation difference and inflation ratio by the realistic and simplified routing policy model.

Is the inflation degree of Internet paths observed earlier (Section 2) sensitive to a more accurate routing policy model? To answer this question, we plot cumulative distribution of inflation difference and inflation ratio with respect

‖We found that 70% of the links (98.7% of nodes) in the AS overlay map exist a corresponding link (node) in the actual AS map. We have tried to aggregate 6 days of BGP table dumps and generate an annotated aggregate map. The aggregate map improves the link coverage on our AS overlay by less than 1%.

to the modified routing policy model ** in Figure 4. For comparison, we also include the earlier plots of inflation difference and inflation ratio of paths with respect to the simplified policy model in the plot. Our result indicates that the degree of inflation with respect to the two models—the simplified model and the realistic model—are similar. These findings lead us to the next question: how good is the simplified policy model compared to the realistic model?

3.3. Shortest AS path policy model v.s. realistic policy model

Is a simplified routing policy model, namely shortest-AS path routing, a reasonable policy model for simulation studies? To evaluate the performance of our simplified model, we again look at the path-length difference and path-length ratio at both AS level and router-level between the paths according to the realistic routing policy model and the corresponding path according to the simplified model. Note that the path-length difference in the AS level is always greater than zero since by definition the shortest AS path is always the shortest and therefore is equal to or shorter than the realistic path.

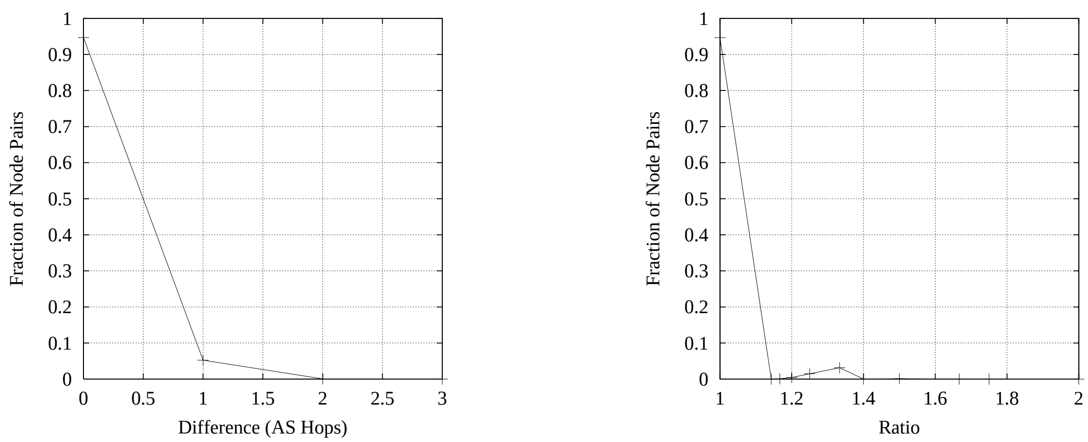

Figure 5. A comparison between the realistic and shortest AS path policy model using AS-level path-length difference and ratio metrics.

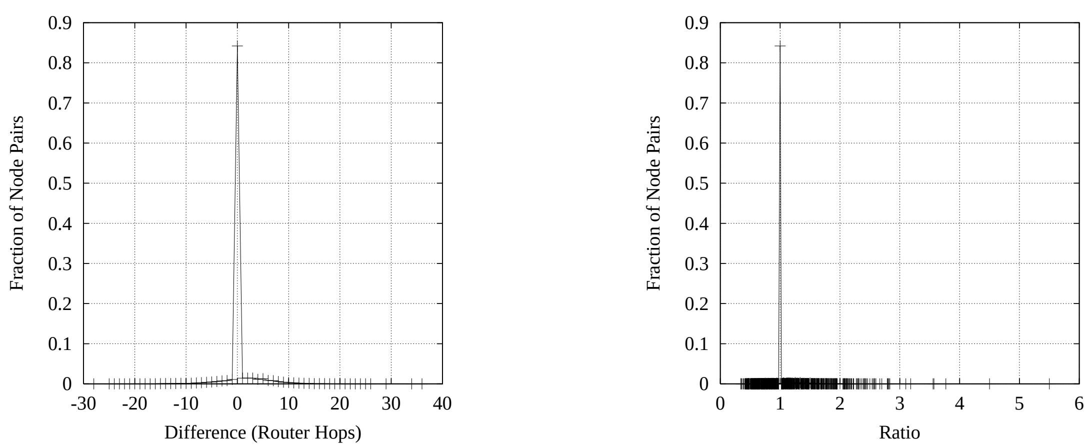

Figure 6. A comparison between the realistic and shortest AS path policy model using router-level path-length difference and ratio metrics.

**There are about 2% of sampled node pairs that are not reachable. We ignore these pairs in our inflation distribution.

Figure 5 shows the path-length difference and ratio in terms of AS hops. We found that 95% of the paths have the same length for both policy models. However, there exist some realistic paths that are 3 AS hops longer. And there exist some realistic paths that are 2 times longer. Figure 6 shows the path-length difference and ratio in terms of router hops. Again, about 84% of the paths show the same length. There exist some realistic paths that are shorter, and some longer, than their corresponding shortest AS paths. This is not surprising given that we do not find any correlation between the AS-level path-length and router-level path-length. Moreover, our previous result regarding the detour paths suggests that there are many paths (about 50%) that are longer in AS hops but shorter in router-level hops.

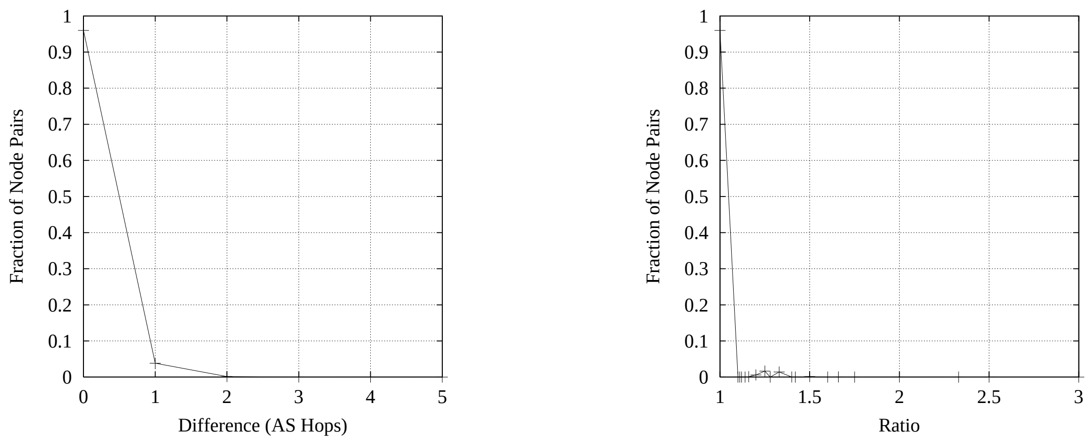

Figure 7. A comparison between the realistic and shortest AS path policy model using AS-level path-length difference and ratio metrics on the actual AS map.

To show that our earlier findings regarding the two policy models are not due to the heuristics that we use in determining the peering relationship on our AS overlay, we plot the path-length difference and ratio (in AS hops) of paths on the actual AS map generating according to the two models. Figure 7 shows the path-length difference and ratio. We find that 96% of paths have the same length for both policy models. The plots are quantitatively similar to those in Figure 5. Our findings suggest that the shortest AS path is a good routing policy model when we are interested in large-scale statistics of paths due to policy.

4. CONCLUSION

In this paper we have addressed the two shortcomings of our previous work in studying the impact of policy routing on Internet paths. First, we have re-examined the Internet path inflation due to the simplified policy routing model using a more recent snapshot of the Internet topology. We found that, our observations regarding the Internet path inflation with respect to both snapshots of the Internet topology are in qualitatively agreement.

Additionally, we have addressed the second shortcoming due to the simplified policy model by examining a more realistic model. The improved model considers the peering relationships among ASs and avoids paths that violate the peering relationship. We find that the degree of inflation with respect to the two models—a simplified model and a realistic model—are similar. We also found that a significant fractions of paths yield the same path-length in both AS-hops and router-hops with respect to the two models. Finally our work indicates that in a simulation study, if the goal is to study the large-scale statistics of paths due to policy (e.g. a study that look at a collection of paths rather than an individual path), the shortest AS path policy model is a reasonable model. The model is simple. It allows us to compute a path from any source-destination pair and it doesn't require the knowledge of peering relationship which might be difficult to obtain.

APPENDIX A. PEERING RELATIONSHIP IDENTIFICATION ON THE AS OVERLAY

```
/* Given a link(A,B), Get_Peering_Relationship returns the peering relationship between A and B.
   Assumption: the peering relationship of the actual AS map (obtained from the BGP table)
   is already determined.
   degree(X) : returns the degree of node X.
*/
Get_Peering_Relationship(link(A,B)) {

  if link(A,B) exists in the actual AS map
    then link_type = the relationship of (A,B) in the actual AS map
  elseif node A and node B are not in actual AS map
        then if degree(A) > degree(B)
               then link_type = PROVIDER_CUSTOMER
             else link_type = CUSTOMER_PROVIDER
  elseif node A is not in the actual AS map
        then link_type = CUSTOMER_PROVIDER
  elseif node B is not in the actual AS map
        then link_type = PROVIDER_CUSTOMER
  else /* node A and B are in the actual AS map but there is no link
          connecting A and B in the actual AS map */
        if ( ((degree(A) > 60) && (degree(B) > 60)) &&
             ((A has more than 15 peers) && (B has more than 15 peers)) )
          then link_type = PROVIDER_PROVIDER
        else if (degree(a) > degree(b))
               then link_type = PROVIDER_CUSTOMER
             else link_type = CUSTOMER_PROVIDER

  return(link_type);
}
```

REFERENCES

1. H. Tangmunarunkit, R. Govindan, S. Shenker, and D. Estrin, "The Impact of Policy on Internet Paths," in *To appear, Proceeding of IEEE INFOCOM*, (Anchorage, AK), 2001.
2. C. Cheng, R. Riley, S. P. Kumar, and J. G. Aceves, "A Loop-Free Extended Bellman-Ford Routing Protocol Without Bouncing Effect," in *Proceedings of ACM SIGCOMM*, pp. 224–236, 1989.
3. Y. Rekhter and T. Li, *A Border Gateway Protocol 4 (BGP-4)*, March 1995. RFC 1771.
4. B. Halabi, *Internet Routing Architectures*, CISCO Press.
5. J.-J. Pansiot and D. Grad, "On routes and multicast trees in the Internet," *ACM SIGCOMM Computer Communication Review* **28**, pp. 41–50, January 1998.
6. K. C. Claffy and D. McRobb, "Measurement and Visualization of Internet Connectivity and Performance." http://www.caida.org/Tools/Skitter/.
7. R. Siamwalla, R. Sharma, and S. Keshav, "Discovering Internet Topology." Unpublished manuscript.
8. H. Burch and B. Cheswick, "Mapping the Internet," *IEEE Computer* **32**, pp. 97–98, April 1999.
9. R. Govindan and H. Tangmunarunkit, "Heuristics for Internet Map Discovery," in *Proceeding of IEEE INFOCOM*, (Tel-Aviv, Israel), March 2000.
10. 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.
11. "Bgp routing table dump archive." University of Oregon Route Views Project. http://rv-archive.uoregon.edu/oix-route-views/.
12. "Radb database services." Merit Network, Inc. http://www.radb.net.
13. H. Chang, S. Jamin, and W. Willinger, "Inferring AS-level Internet Topology from Router-Level Path Traces," in *Proceedings of SPIE ITCom*, (Denver, CO), August 2001.
14. C. Faloutsos, P. Faloutsos, and M. Faloutsos, "On Power-Law Relationships of the Internet Topology," in *Proceeding of ACM SIGCOMM*, Sept. 1999.
15. L. Gao, "Inferring autonomous system relationships in the internet," in *Proceeding of IEEE Globecom*, (San Francisco, CA), 2000.
16. "Raw routing table information." National Laboratory for Applied network Research. http://moat.nlanr.net/Routing/rawdata/.