Quantifying Internet End-to-End Route Similarity

Ningning Hu and Peter Steenkiste

Carnegie Mellon University
Pittsburgh, PA 15213, USA
{hnn, prs}@cs.cmu.edu

Abstract. Route similarity refers to the similarity of two routes between two nodes and an arbitrary third node. This intuitive concept plays an important role in distributed system deployment and path-edge inference. However, route similarity has not been quantitatively studied from an end-node perspective, and its properties are poorly understood. In this paper, we make an initial effort in quantifying route similarity by investigating a simple metric—RSIM. Using two large-scale traceroute data sets, we show that RSIM can be measured using only a small number of random traceroutes, and that it captures the similarity of both upstream and downstream routes. As a case study, we also describe how to use RSIM to infer shared path edges. We show that if a pair of end nodes have an RSIM value larger than 0.8, they have over 80% probability of sharing path edges.

1 Introduction

We define route similarity as the overlap of two end-to-end routes between two nodes and an arbitrary third node. In this definition, we regard routes as a property of end nodes, and route similarity captures the similarity of this property for different end nodes. Although a very simple and intuitive concept, route similarity can be very useful in many applications.

First, route similarity quantifies the difference between end nodes from an Internet routing perspective. Based on that, we can group Internet end nodes into different clusters, and this clustering information can be useful for at least three types of applications: distributed system deployment, vantage point selection, and web server resource management. (1) For distributed systems, being able to tolerate network failures is very important, both in terms of maintaining connectivity and in terms of achieving good performance. Therefore, it is preferable to deploy the system on a set of nodes that are as diverse as possible. Given the clustering information, a simple solution is to avoid using multiple nodes from the same cluster. (2) Similarly, the clustering information can be used to diversify vantage points used for systems like network monitoring. Diverse vantage points can improve both monitoring efficiency and the measurement completeness. (3) Web servers can manage their clients based on client clusters, which can significantly reduce the management overhead, because there generally will be significantly fewer clusters than clients. Web servers can also use techniques

This research was funded in part by KISA (Korea Information Security Agency) and CyLab Korea and in part by NSF under award number CCR-0205266.

like server or content replication to improve the performance of clusters with poor performance.

Second, route similarity can be used for path edge inference. This capability can be useful in systems like BRoute [7]. A recent Internet-scale study on network bottlenecks [6] showed that over 86% of Internet end-to-end paths have bottleneck at the first or last *four* hops of the path. This implies that if two nodes have very similar routes, they will have a high probability of sharing bottlenecks, and we may be able to use the available bandwidth information collected by one node for paths used by the other. This method can significantly reduce available bandwidth measurement overhead for large scale network systems.

We are not aware of any metrics for quantifying route similarity. Related work such as [3] and [4] has studied how to use relay nodes to increase route diversity for link or router fault tolerance, and their route diversity results may be used to measure route similarity. However, our work has a completely different focus—we are interested in edge route similarity, which can not be directly quantified using route diversity. Sometimes IP prefixes are also used to estimate route similarity. This is based on the observation that if two nodes have IP addresses from a common prefix, they often share routes. This approach has three limitations. First, a common IP prefix is not a sufficient condition for route similarity. Nodes from the same prefix, especially those from a large prefix, do not necessary share routes. Second, a common IP prefix is not a necessary condition for similar route, either. For example, from the *Rocketfuel* data set used in [11], we find that node `planet2.cs.ucsb.edu` and node `planetlab1.cs.ucla.edu` have very similar routes although they belong to completely different prefixes–131.179.0.0/16 (AS52) and 128.111.0.0/16 (AS131). Finally, the IP prefix does not *quantify* route similarity, thus it is hard to compare the similarities of different pairs of nodes.

In this paper, we introduce a simple route similarity metric—*RSIM*—that addresses the above problems. RSIM is defined as the ratio between the total number of shared links and the total number of links on two routes. We will show that (1) RSIM is very easy to measure—it only needs a small number of "random" traceroutes; (2) RSIM can capture both upstream and downstream route similarity. Using a case study, we also show that RSIM is very useful for inferring shared path edges. We find that if a pair of end nodes have an RSIM value larger than 0.8, the probability that they share path edges is over 80%.

The remainder of this paper is organized as follows. We define RSIM in Section 2, and we describe the data sets used for our analysis in Section 3. We then discuss RSIM's properties in Section 4, including destination sensitivity, measurability, and symmetry. In Section 5, we describe a case study of using RSIM for path-edge inference. We discuss related work in Section 6 and conclude in Section 7.

2 RSIM Definition

In the remainder of this paper, the term "route similarity" refers specifically to RSIM, which is defined as follows. Let $P(s, d)$ denote the IP level route from node s to node d; $L(s, d)$ denote the number of links on $P(s, d)$; $Total(s_1, s_2, d) =$

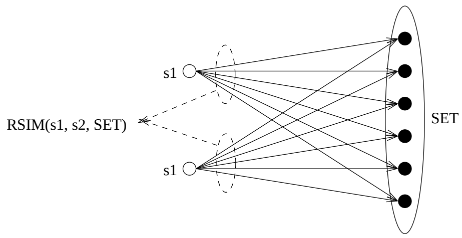

Fig. 1. $RSIM(s_1, s_2, SET)$

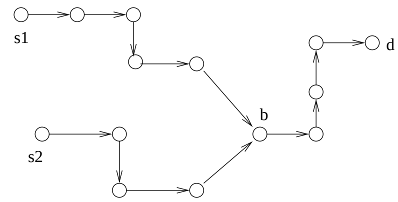

Fig. 2. Example: $RSIM(s_1, s_2, \{d\}) = 8/17$

$L(s_1, d) + L(s_2, d)$; and $Common(s_1, s_2, d)$ denote the total number of links that are shared by $P(s_1, d)$ and $P(s_2, d)$. Let SET denote a set of Internet destinations (see Figure 1), then the route similarity between s_1 and s_2 *relative* to SET is defined as:

$$RSIM(s_1, s_2, SET) = \frac{\sum_{d \in SET} 2 * Common(s_1, s_2, d)}{\sum_{d \in SET} Total(s_1, s_2, d)} \quad (1)$$

Note this definition uses upstream routes from s_1 and s_2, RSIM can be similarly defined using downstream routes. Intuitively, this definition captures the percentage of links shared by the two routes $P(s_1, d)$ and $P(s_2, d)$. In this definition, when SET is obvious, we simplify $RSIM(s_1, s_2, SET)$ as $RSIM(s_1, s_2)$. It is easy to see that $RSIM(s_1, s_2, SET) \in [0, 1]$ for any s_1, s_2, and SET. The larger $RSIM(s_1, s_2, SET)$ is, the more similar the routes of s_1 and s_2 are. Figure 2 shows an example an RSIM computation. Here we have $SET = \{d\}, Common(s_1, s_2, \{d\}) = 4, Total(s_1, s_2, \{d\}) = 9+8 = 17$, so $RSIM(s_1, s_2, \{d\}) = 2 * 4/17 = 8/17$.

3 Data Sets

Our analyses on RSIM are based on two data sets: the *Rocketfuel* data set and the *Matrix* data set.

The Rocketfuel data set refers to the traceroute data collected on 12/20/2002 by the Rocketfuel project [11][2], which used 30 nodes[1] on Planetlab [1] to probe over 120K destinations. This is one of the most complete traceroute data sets publicly available. The destination IP addresses are derived from the prefixes in a BGP table that day. Because the destination IP addresses are randomly selected, they do not necessarily correspond to online hosts, so for many of the destinations the traceroute measurements are not complete. To avoid the impact of incomplete routes on our analysis, we only consider the 5386 destinations which can be reached by at least 28 source nodes using traceroute.

The Matrix data set is collected by the authors using 160 Planetlab nodes, each from a different site. In this data set, we let each node run traceroute to

[1] The original data set covers three days and uses 33 nodes. We only use the data from one of the days, and we discarded the data for 3 nodes because they could not get valid traceroute results for a large portion of destinations.

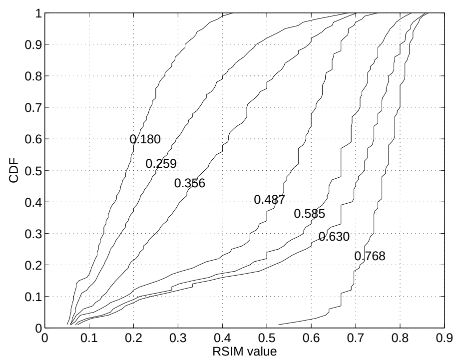

Fig. 3. Destination-sensitivity of RSIM

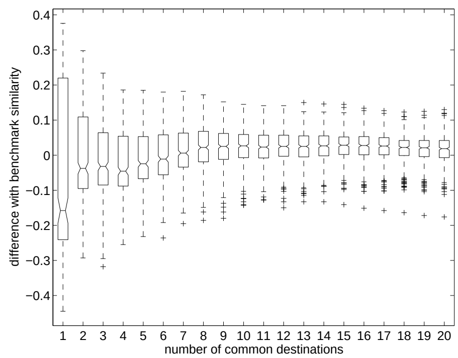

Fig. 4. Measurability of RSIM

all the other nodes, obtaining a route matrix, where we have routes in both directions for all pairs of nodes. This is the largest route matrix we are aware of, and it is very useful for characterizing the symmetry property of RSIM, as discussed in Section 4.3.

4 RSIM Properties

In this section, we analyze the following properties of RSIM to show that it is a useful metric:

- *Destination Sensitivity:* how sensitive is the RSIM value to the choice of destinations;
- *Measurability:* how many measurements are required to compute the value of RSIM;
- *Symmetry:* what is the difference between the upstream route similarity and downstream route similarity?

4.1 Destination Sensitivity

Since RSIM is a function of the destination set SET, the value of RSIM can be different for different destination sets. However, for many applications it is preferable that RSIM is largely independent of the SET parameter, i.e., it is a fundamental property that only needs to be measured once. This property is very important for reducing measurement overhead. In this section, we first focus on the case where SET includes only a single destination. The case where SET includes multiple destinations are discussed in Section 4.2.

We use the *Rocketfuel* data set to study destination sensitivity. We first set SET to include the 5386 reachable destinations in the *Rocketfuel* data set, and compute RSIM values for all the 435 ($= 30 * 29/2$) pairs of source nodes. We will use these RSIM values as benchmark values since they are from the largest destination set possible. The distribution of these RSIM values has a sharp peak around 0.7. Specifically, 85% of the 435 pairs have RSIM values between 0.65 and 0.8. This confirms earlier observations that in 2002, most Planetlab nodes had very little diversity in how they connected to the Internet (most used Abilene). However, there is some diversity: the RSIM values range from 0.1 to 0.8.

To study the destination sensitivity of RSIM for node pairs with different RSIM values, we selected seven source-node pairs with RSIM values roughly evenly distributed in the range [0.1, 0.8]. We calculate their RSIM values for *each* of the 5386 individual destinations for which they have complete route data. These similarity values are plotted in Figure 3. Each curve plots the cumulative distribution for the RSIM values of one source-node pair relative to each individual destination. The numbers marked on the curves are the benchmark RSIM values. The seven curves can be classified into three groups:

1. The first group only includes the rightmost curve. This curve corresponds to a pair of source nodes with the highest benchmark RSIM value (0.768) among the seven pairs, i.e., their routes are very similar. The similarity values of this source-node pair for individual destinations are distributed in a fairly small region—90% of them are in [0.65, 0.85]. That shows that the similarity between this pair of source nodes is not very sensitive to the destination selected.
2. The middle three curves make up the second group. Their benchmark RSIM values are 0.467, 0.585, and 0.630, respectively, and they represent source-node pairs with average similarity values. Clearly, the RSIM values for individual destinations in this group are more diverse than in the first group. The lowest 30% of similarity values are significantly lower than the other 70% of the values. However, the highest 70% of similarity values cluster within a small region with 0.2 width.
3. The three leftmost curves represent the third group, where the node pairs have low similarity—0.180, 0.259, and 0.356. The similarity values of these source-node pairs with respect to individual destinations is almost evenly distribute in a large range with 0.4-0.6 width. That means that their RSIM values are quite sensitive to the SET parameter.

The above results show that the larger the benchmark RSIM value is, the less sensitive the RSIM values are to the chosen destination. Specifically, for node pairs with properties similar to the rightmost curve, the RSIM values are not sensitive to the specific single-destination set.

4.2 Measurability

In this subsection, we show that measuring route similarity only needs a small number of traceroute measurements. We again use the *Rocketfuel* data set to demonstrate this property. In this data set, we use different numbers (x) of randomly (uniform) selected destinations to compute route similarity ($RSIM_x$) for each of the 435 source-node pairs. We then compare them with their benchmark values ($RSIM_b$) which is based on all 5386 reachable destinations as discussed in Section 4.1, and compute the relative difference as $(RSIM_x - RSIM_b)/RSIM_b$. Figure 4 plots the distributions of the relative differences from all 435 source-node pairs. The x-axis is the number of routes used in computing $RSIM_x$. The bars are plotted using the `boxplot` function of Matlab, where each bar corresponds to one distribution of the relative difference for all 435 source-node pairs.

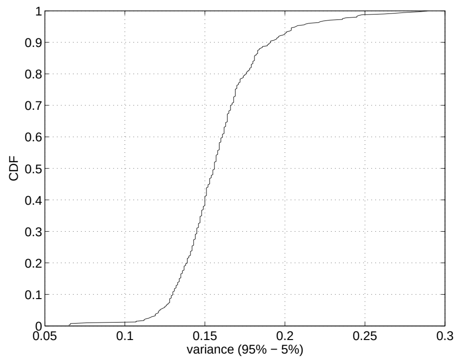

Fig. 5. Impact of randomness

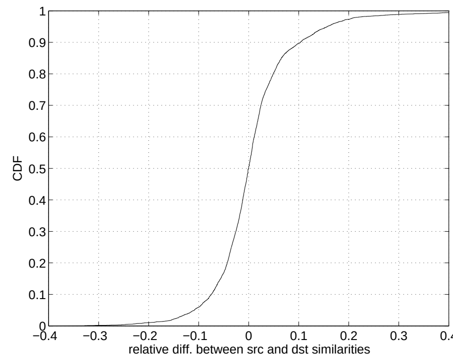

Fig. 6. Symmetry of RSIM

The middle boxes have three lines corresponding to the lower quartile, median, and upper quartile values, and the whispers are lines extending from each end of the box to show the extent of the rest of the data. We can see that the relative difference between $RSIM_x$ and $RSIM_b$ quickly drops as more destinations are used. Once $x \geq 10$, the median difference stays roughly constant at about 5%, although the variance decreases. This result shows that only 10-20 routes are needed to compute the value of RSIM for the *Rocketfuel* data set.

In the above analysis, for each source-node pair and each number of destinations, the similarity value is calculated using one instance of a randomly selected destination set. These values may be different for different randomly selected destination sets. To quantify the impact of this randomness, for each source-node pair, we select 1000 different random 10-destination sets and compare their RSIM values. For each source-node pair, we then record the difference between the 95 percentile similarity value and the 5 percentile similarity value, and use that as a variance measure for different random 10-destination sets. Figure 5 plots the cumulative distribution of this variance for all 435 source-node pairs. We see that the variance for 93% of pairs is less than 0.2, which is small. This tells us that the choice of the randomly selected destination sets does not have a significant impact on the RSIM values.

An important implication of the results in Figure 4 and 5 is that RSIM is not sensitive to the choice of destination set when SET includes ten or more "random" destinations. The reason is as follows. Although RSIM can be very different for different individual destinations, as illustrated in Figure 3, ten or more different destinations can fairly well "cover" the distribution curve by including most important points. Therefore, even with different destination sets, since they are from the same distribution, which is determined by the node pair, they should converge to the same value.

So far, the results in this section were obtained using routes between 30 source nodes and 5386 destination nodes. We have done a similar analysis using all 120K destinations, i.e., including the incomplete route data. The results we obtained are similar. Even so, the data set used here only covers a limited fraction of nodes on the Internet, and whether or not our conclusion in this section can be extended to the whole Internet should be validated using larger and more diverse data sets.

4.3 Symmetry

It is well known that Internet routes are asymmetric [9], but we find that RSIM values computed using upstream routes and those using downstream routes are very similar. In this sense, RSIM is symmetric, i.e., it captures the similarity of both upstream and downstream routes. In this section, we use the *Matrix* data set to show this property.

For each node pair among the 160 nodes, we calculate their route similarity using both upstream routes ($RSIM_{up}$) and downstream routes ($RSIM_{down}$). We then compute the difference as ($RSIM_{down} - RSIM_{up}$). Figure 6 plots the distribution of this difference for the 14,412 pairs which have at least 10 complete traceroute results to compute both $RSIM_{up}$ and $RSIM_{down}$. We can see that 84% of the pairs have difference within a small range of [-0.1, 0.1], which shows that $RSIM_{up}$ and $RSIM_{down}$ are indeed very similar. That means that, if two nodes have a high probability sharing a large portion of their upstream routes toward a destination, they will also have a high probability sharing a large portion of their downstream routes from that destination.

The implication of this property is two-fold. First, RSIM measurements do not necessarily need upstream routes. If only downstream routes are available, we can still compute RSIM. This property will be used in Section 5 to get a larger sample set for our analysis. Second and more importantly, this property allows us to infer both source-side and destination-side path edges. For example, if s_1 and s_2 have similar routes, they will both have similar destination-side path edges for their routes towards a third node, and have similar source-side path edges for those routes from a third node. The details are discussed in the next section.

5 RSIM Case Study: Path Edge Inference

In this section, we first define what we mean by "path edge", and then describe how to use RSIM to infer path edges. We also present empirical results on how often path edges shared by a pair of nodes can be inferred.

5.1 Path Edge Inference

According to [6], most bottlenecks are located on the last/first four links of a path. The last/first four links are what we call the "path edge". Formally, we call the first four links of an end-to-end route $P(s, d)$ the *source edge (srcEdge*(s, d)*)*, and the last four links the *sink edge (sinkEdge*(s, d)*)*. For example, in Figure 2, path segment (b, d) is the sink edge for both $P(s_1, d)$ and $P(s_2, d)$. Source edges and sink edges together are referred as *path edges*. Ideally, we would like to use RSIM value to predict path edges. This prediction becomes easy if the following two claims are true:

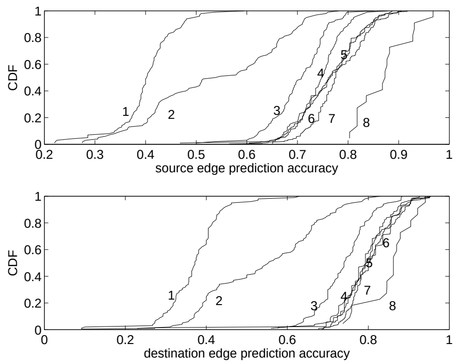

Fig. 7. Edge sharing probability for node pairs with different levels of similarity. The numbers on the curves mark the group ID.

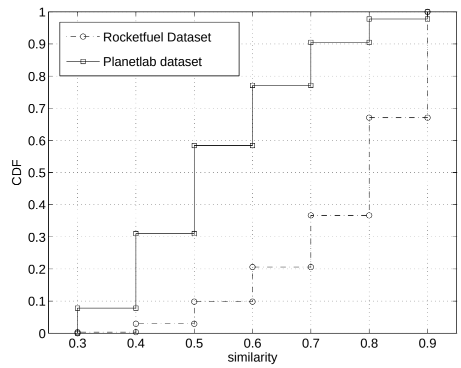

Fig. 8. The probability of a node having a neighbor

> **Claim 1:** $\exists\ TH_{RSIM}$, $\forall\ s_1, s_2$, their upstream routes towards any node d share sink edge iff $RSIM(s_1, s_2) > TH_{RSIM}$.

> **Claim 2:** $\exists\ TH_{RSIM}$, $\forall\ d_1, d_2$, their downstream routes from any node s share source edge iff $RSIM(d_1, d_2) > TH_{RSIM}$.

We call a pair nodes *neighbors* if their RSIM value is larger than TH_{RSIM}. Intuitively, if a pair of nodes s_1 and s_2 are neighbors, and the bottlenecks of $P(s_1, d)$ and $P(s_2, d)$ are on their sink edges, we then can use the bandwidth measured from $P(s_1, d)$ as that of $P(s_2, d)$, because they are very likely to share the bottleneck.

Of course, it is unrealistic to expect that we will be able to find a threshold TH_{RSIM} that gives 100% sharing of the remote path edges. Instead, we now look at whether a threshold TH_{RSIM} exists that indicates a high probability of path-edge sharing. We use the *Matrix* data set for this study. We first take the 14,412 node pairs which have at least 20 complete traceroute results, compute their RSIM values using these routes, then group them into nine groups ($g_i, i = 1..9$) based on their RSIM values: $g_i = \{(s, d)|i * 0.1 \leq RSIM(s, d) < (i + 1) * 0.1\}(1 \leq i \leq 9)$. For each node pair in each group, we calculate the probability of sharing source edges and sink edges. Figure 7 plots the cumulative distribution of source/sink-edge sharing probabilities for each group. In this figure, g_9 is grouped into g_8 because there are only 6 pairs in g_9. The top graph plots the sharing probability for source edges, and the bottom graph plots the results for the destination edge. g_8 stands out distinctively with the best prediction accuracy—almost all node pairs in this group have a sharing probability higher than 0.8. Although not perfect, we think we can use the value 0.8 for TH_{RSIM}. That shows that RSIM can be used for path-edge inference with a high inference accuracy.

5.2 Probability of Sharing Path Edges

The above results suggest that it will be useful for neighbors to share bandwidth information. Whether this is feasible depends on how likely it is that a node can find a neighbor. This is a very difficult question, but we can use the *Rocketfuel* and *Matrix* data set to gain some insights. We use the *Rocketfuel* data set as an example of a system that includes nodes from all over the Internet, while we use the *Matrix* data set to get the view for a real deployed system.

For the *Rocketfuel* data set, we use downstream routes to compute route similarities for the 5386 reachable destinations. The reason that we use destinations instead of sources is to obtain a large scale analysis. The symmetry property of RSIM demonstrated in Section 4.3 allows us to compute RSIM values using downstream routes. For the *Matrix* data set, route similarities are computed for the 160 nodes using upstream routes. Figure 8 plots the distribution of node pairs in each group. In this figure, node pairs are again grouped as we did in Figure 7 (except that g_1 and g_2 are combined with g_3); the x-axis is the smallest RSIM value in each group. The dashed curve shows the result for the *Rocketfuel* data set. We see that 63% of end nodes can find at least one neighbor, i.e., their RSIM value is larger than 0.8. The *Matrix* data set has significantly fewer end nodes, so only 10% of end nodes can find a neighbor.

It is worthwhile to mention that in both of these analyses, destinations are selected from different prefixes, while in reality, many system nodes can come from common prefixes. In this sense, the results presented in Figure 8 provide a pessimistic view.

5.3 Discussion

RSIM can be used for path-edge inference which is closely related to the goal of the BRoute system proposed in [7]. In BRoute, each node collects AS-level source tree and sink tree information to infer path edges, which are further used to infer path bottlenecks and available bandwidths. The difference between RSIM and BRoute is that RSIM is a general metric that can be used by different applications, while BRoute focuses on path bandwidth inference by only characterizing routes of each individual node, i.e., BRoute does not directly quantify the similarity of routes from two different nodes. However, BRoute is not sensitive to system node distribution, while using RSIM for path-edge inference depends on the chances that system nodes can find neighbors.

6 Related Work

We are not aware of related work that directly quantifies route similarity from an end-node perspective, but there are quite some that is related with different aspects of the RSIM metric. Besides the [3] and [4] on route diversity and redundancy, which are mentioned earlier in this paper, there are also a lot of research focusing on identifying congestion sharing [10], bottleneck sharing[8] and packet loss sharing [5]. [10] used loss or delay observations at end hosts to infer whether or not two flows experiencing congestion are congested at the same network resources; [8] proposed a passive measurement approach for learning Internet path characteristics, which can be used for bottleneck sharing inference. Finally, [5] proposed a Bayesian probing technique to determine whether a pair of connections from a same node experience shared loss.

In terms of path-edge inference, we have discussed the possibility of using RSIM for BRoute bandwidth estimation. Another closely related result is the inference algorithm proposed by [9], which can infer the AS path for a pair of nodes when only having access to the destination node. The idea is to first infer the first AS hop with the help of a public traceroute server, and use that information in an more general AS-path inference algorithm to obtain an AS-path inference with higher accuracy.

7 Conclusion and Future Work

In this paper, we present a study of Internet end-to-end route similarity by studying a specific route similarity metric—RSIM. It quantifies the similarity of two routes between two nodes and an arbitrary third node. Based on two large traceroute data sets, we show that for our data sets RSIM can be measured using a small number of random traceroutes, and it captures the similarity of both upstream routes and downstream routes. When using RSIM for path-edge inference, we show that if a pair of nodes have an RSIM value larger than 0.8, they have a high probability of sharing path edges of routes with an arbitrary peer node.

The work presented here is only a first step in understanding Internet route similarity. There could be better metrics in terms of destination sensitivity, measurability, and symmetry. Moreover, more detailed studies are needed to quantify the properties of route similarity.

References

1. Planetlab. http://www.planet-lab.org.
2. Rocketfuel Data Sets. http://www.cs.washington.edu/research/networking/rocketfuel.
3. M. Cha, S. Moon, C.-D. Park, and A. Shaikh. Positioning Relay Nodes in ISP Networks. In *Proc. IEEE INFOCOM*, March 2005.
4. J. Han, D. Watson, and F. Jahanian. Topology aware overlay networks. March 2005.
5. K. Harfoush, A. Bestavros, and J. Byers. Robust identification of shared losses using end-to-end unicast probe. Technical Report BUCS-2000-013, Boston University.
6. N. Hu, O. Spatscheck, J. Wang, and P. Steenkiste. Optimizing network performance in replicated hosting. In *the Tenth International Workshop on Web Caching and Content Distribution (WCW 2005)*, September 2005.
7. N. Hu and P. Steenkiste. Exploiting Internet route sharing for large scale available bandwidth estimation. In *Proc. of ACM/USENIX IMC*, October 2005.
8. D. Katabi, I. Bazzi, and X. Yang. A passive approach for detecting shared bottlenecks. In *IEEE International Conference on Computer Communications and Networks*, 2001.
9. Z. M. Mao, L. Qiu, J. Wang, and Y. Zhang. On AS-level path inference. In *to appear in SIGMETRICS'05*, June 2005.
10. D. Rubenstein, J. Kurose, and D. Towsley. Detecting shared congestion of flows via end-to-end measurement. In *IEEE/ACM Transactions on Networking*, volume 10, June 2003.
11. N. Spring, R. Mahajan, and D. Wetherall. Measuring ISP topologies with Rocketfuel. In *Proc. ACM SIGCOMM*, August 2002.