Distance Metrics in the Internet

Bradley Huffaker, Marina Fomenkov, Daniel J. Plummer, David Moore and k claffy

Abstract—**We consider and compare four Internet distance metrics and analyze the predictive power of these metrics in selecting, from a given source, the lowest latency destination from among a candidate set. The four metrics are: IP path length; autonomous system (AS) path length; great circle geographic distance; and previously measured round trip time (RTT). We describe general properties of these four metrics and, using an unprecedented volume of real Internet macroscopic topology and RTT data, compare their correlation with actual RTT to the destination. The new methodology we propose for testing different metrics is suitable for evaluating new distance estimation techniques as they become available.**

I. Introduction

A. Background

Packet propagation time between two hosts on the Internet is a simple metric that reflects connection performance as perceived by the user. A packet traverses many links on its way from source to destination, and several parameters of each link, e.g., propagation latency, available bandwidth, queuing delay, and packet loss, contribute to overall end-to-end delay. These parameters are generally unknown and can fluctuate unpredictably over time. Therefore, performing consistently useful quantitative measurements of host-to-host performance, particularly those with predictive power with respect to future performance, is extremely challenging.

In this study we focus on latency as a natural measure of distance in Internet space. By definition, distance is 'the amount of separation between two points' [1]. We seek 'distance metrics' that can accurately and consistently predict latency. In our framework, a distance metric is a quantity, measured or calculated for a given Internet connection between a pair of hosts, that has unambiguous correlation to latency: the lower the value of the metric, the smaller the RTT, and vice-versa.

We consider and compare four possible distance metrics and study their correlation with latency. We emphasize that our approach does not attempt to predict the absolute value of RTT between any two hosts, but only to use available prior information to predict, for a given source, a relatively lower RTT from among a set of candidate destinations. Two of our distance metrics relate to logical topology of the Internet, and two derive from underlying geographic characteristics. Our results suggest that no metric achieves perfect positive correlation to latency, but we have found that metrics derived from physical (geographic) characteristics correlate better with observed RTTs than metrics derived from logical topology information.

B. Related work

Practical distance estimates play an essential role in the nearest server selection problem. Many widely used Internet services are replicated (or mirrored) in different physical locations. The goal of this replication is to provide users with faster access to content by allowing them to select nearby copies and avoid congested paths or servers. In our framework we call a server 'nearest' to a client if it has the lowest RTT from that client.

Heidemann and Visweswaraiah [2] addressed the issue of automatic selection of nearby web servers. They considered different selection algorithms based on domain names, geographic approximation, and ICMP-based (including source-routed) latency probe measurements. Heidemann and Visweswaraiah analyzed the overhead time for each algorithm and compared this cost to the retrieval time for short web documents.

Crovella and Carter [3] studied the effectiveness of hop count and latency metrics, and also explored approaches to bandwidth measurement. They concluded that replica selection based on bandwidth or RTT measurements perform comparably, and both perform substantially better than random selection.

McManus [4] also experimented with a heuristic distance metric derived from BGP AS path length (specifically, the logarithm of the number of ASes in the forward IP path plus one) and found that this metric does not provide a good general purpose mechanism for selecting the nearest server.

Francis *et al.* [5] explored technical issues related to creation of a public infrastructure to provide host-to-host distance information. They proposed to deploy a number of servers that will maintain a virtual topology map of the Internet and distribute it using IP multicast (IDMaps). The authors note limitations of such a service and suggest that building and testing such prototypes is the only way to evaluate its scalability and accuracy.

Ng and Zhang [6] used coordinate-based mechanisms (Global Network Positioning - GNP) in a peer-to-peer architecture to predict Internet network distance, such as RTT. They proposed to model the Internet as a geometric space and characterize the position of a host by a point in this space. They conducted a test study and found that the GNP approach is more accurate and robust than IDMaps.

C. Distance Metrics

In this study we consider four metrics of distance between two Internet hosts: IP path length; autonomous system (AS) path length; actual geographic distance; and previously measured round trip times (RTT).

1. IP path length: the total number of hops traversed by a packet on its forward path from a source to a destination host. If the primary source of network delay is the time it takes routers to process and forward packets (but not including per-hop queuing delays which vary widely), then the number of routers a packet traverses will strongly influence the observed RTT value.

2. Autonomous System (AS) path length. An autonomous system represents either a single network or a group of networks comprising a single administrative entity (e.g., an Internet Service Provider - ISP) that shares a common network adminis-

All authors are with: CAIDA, UC San Diego, San Diego Supercomputer Center MC 0505, 9500 Gilman Drive, La Jolla CA 92093-0505, USA. phone:1 (858) 534 8333 E-mail: {brad,marina,djp,dmoore,kc}@caida.org.

Support for this work is provided by DARPA NGI N66001-98-2-8922, NSF ANIR grant NCR-9711092, and CAIDA members.

trative policy. Using BGP [7][1] routing tables, which articulate inter-domain exchange of traffic among these administrative entities, we can determine the number of unique ASes visited by a packet traversing an Internet path. The main advantage of this metric is that we can derive it at no extra cost to the network, since propagating and updating inter-domain routing (BGP) information is intrinsic to network functioning. If ISP peering points are responsible for a large fraction of overall end-to-end delay (as suggested in [4]), then the AS path length metric would offer good predictive power.

3. Geographical distance. We define the geographical distance between two hosts as the length of the great circle arc connecting their locations on the surface of the Earth. If the topological structure of the Internet and the physical layout of its links were fully consistent with each other and with the straight great circle arc connection, then the routed path would approximately follow the shortest physical path between two hosts. The sum of per-hop distances would not significantly diverge from the actual distance between hosts. However, several studies have shown [9] [10] [11] that the sum of per-hop distances may greatly exceed the great circle distance between two end hosts. Still, geographic distance remains a significant factor in determining RTT between two hosts.

4. RTT. We derive our RTT metrics from previously measured RTTs between the two hosts under consideration; we consider a few different definitions of this metric, e.g., last RTT observed, median derived from a set of previous RTT values. The main advantage of RTT-based metrics is that they directly correlate to our definition of *nearness* and hence can be expected to produce better results. However, measuring RTTs requires active probing and continuous monitoring, which is not always feasible. More importantly, RTTs are influenced by multiple random and variable short-lived factors that are hard to predict, e.g., link congestion, queueing, routing changes.

The contribution of this paper is two-fold. First, we describe properties of these four metrics and, using an unprecedented volume of real Internet macroscopic topology and RTT data, compare their efficacy in solving the nearest server selection problem described above. Second, we propose a new methodology for testing different metrics that can be readily applied to new distance estimation techniques as they become available.

II. METHODOLOGY

A. Data collection

We use two kinds of data for this study: a large volume of forward Internet (IP) path information obtained from CAIDA's macroscopic topology probing project [12]; and inter-domain BGP routing tables from RouteViews project [13].

CAIDA's topology probing tool (`skitter`) is similar to `ping` and `traceroute`, but uses increased timestamp accuracy. This tool iteratively sends 52-byte ICMP echo request packets, incrementally increasing their time-to-live values until a packet reaches the target host. Each trace is a record of the IP addresses of responding intermediate routers on the forward path from the source to the target destination, as well as the RTT to the destination.[2] Such measurements, made from 1 to 15 times daily, characterize connectivity between the topology monitor and the destination hosts on the probe list. For several years CAIDA has collected and analyzed large volumes of Internet topology data from a set of monitor sources to hundreds of thousands of destinations, comprehensively stratifying the Internet address space as well as the Earth [14].

The RouteViews [13] project collects BGP routing perspectives from more than 60 major ISPs worldwide. The combined table typically has nearly 120K globally routable prefixes; we use the combined table to map IP addresses to their origin ASes.

In the classic nearest server selection problem, there are a few servers distributed around the network, and many clients who wish to select the optimal server from their location. However, in practice it is difficult to instrument clients to collect the necessary information to compare different metrics. We use data from CAIDA topology monitors that continuously probe many target destinations and provide information on many thousands of pairwise connections. We treat the destinations as servers and the monitor source as the client in comparing the utility of different distance metrics for the nearest server selection problem.

In this paper we present data from nine topology monitors that poll two different destinations lists - the *IPv4_Space* list (313,471 destinations) and the *DNS Clients* list (58,312 destinations). CAIDA compiled both of these lists with the goal of covering the routable IPv4 address space as representatively as possible but without over-sampling it. Our working definition of 'routable IPv4 space' is the union of IPv4 address prefixes present in the combined RouteViews BGP table [13] at the time we compiled the destination list.

For the *IPv4_Space* list, we attempted to find one IP address in each populated /24 prefix. Destinations in this list encompassed 54% (46,726 prefixes) of RouteViews BGP prefixes as of July 2000, which corresponded to 7.5% of the routable /24 prefixes. Three CAIDA topology monitors (in Eugene, OR; London, UK; and Hamilton, NZ) have been probing the *IPv4_Space* list since July 2000. Each monitor probes the entire destination set once per day for dates after March 2001 and approximately one-third of the set per day for dates prior to that.

For the *DNS clients* list, our goal was to have a single destination in each BGP prefix from a list of IP addresses querying DNS root name servers [14]. We collected these addresses from `tcpdumps` taken at the A, F, J, K, and L DNS root servers. This list achieved nearly 50% coverage of RouteViews BGP prefixes as of September 2000 when we deployed it on six topology monitors. These hosts are located in Palo Alto, CA; San Jose, CA; Herndon, VA; London, UK; Amsterdam, NL; and Tokyo, JP. They probe each destination between 8 and 14 times daily.

B. Determining distance metrics from topology probes

CAIDA's topology traces consist of forward IP path and end-to-end RTT between a host monitor and a set of target destinations. We can derive metrics based on IP path length or RTT directly from these data.

We define the geographic distance between a destination and a monitor as the great circle distance computed from the coordi-

[1] An explanation of how BGP calculates and uses these AS paths is beyond the scope of this paper; see [7] [8] for details.

[2] There is no way to capture the reverse path without control of the destination host.

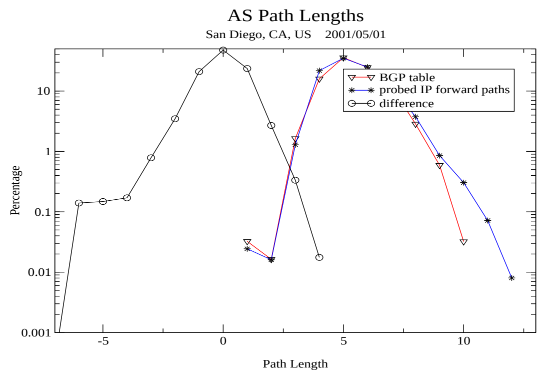

Fig. 1. Histogram of AS path lengths for 1 May 2001. *Difference* is the curve for the difference between AS path seen by the topology probe and that in the BGP table.

nates of the two locations. We know the locations of all topology monitors and use the `IxMapper` [15] service to map IP addresses of destinations to latitude and longitude. If `IxMapper` cannot determine a terrestrial location of a given IP address, we exclude that destination from our geographic distance metric calculations.

While our topology probes provide accurate forward IP path information, there is no record of AS paths traversed. One way to measure an AS forward path for a source-destination host pair is by using the BGP table in the router directly in front of the source, i.e., its gateway to the Internet. However, the IP forward path actually travelled by a packet does not always match the AS path listed in this adjacent BGP table [8] [16], and access to such adjacent routing tables is not always available anyway. Instead we convert each IP address in the forward path to an AS, and count the number of unique ASes in this path. To determine the AS number associated with the IP address at a given hop along the path, we find the longest matching prefix for the IP address in the RouteViews BGP table, and use the origin AS of the path to this matching prefix. This AS is the one who originally announced reachability of the IP address to the global Internet, so it is reasonable to consider that AS the 'home' of the IP address. We next count the number of AS transitions in the path and add one to get the total number of ASes traversed.

We are not able to obtain an AS number for a given hop in three situations: our topology probe did not receive a response from this hop so we do not know its IP address; the IP address does not match any prefixes in the BGP table; or we find more than one 'home' AS for this IP address. In all these cases we omit these IP hops since they likely belong to the same AS as either the preceding or subsequent hop in the path.

How similar are AS path lengths obtained from topology probes and from the BGP table adjacent to the probing source? We have access to the gateway routing table for our San Diego topology monitor, so we compared the distributions of AS path lengths computed by these two methods. Figure 1 presents the results: the 'BGP table' line represents AS path lengths calculated from the routing table for the subset of prefixes probed from San Diego, weighted by number of destinations probed within each prefix. The 'probed IP forward path' line is the distribution of AS path lengths generated from the topology probes. The two curves closely align except in the tail, where the 'forward topology probe' curve is slightly heavier. However, these long paths constitute less than 1% of the paths.

The third curve in figure 1 shows the distribution of a difference between the AS path length derived from the BGP table and the one computed from the topology probes for each individual destination. 90% of paths are in agreement by one AS hop or less. We conclude that our forward topology data yield a good approximation for AS path length.

C. *Scoring distance metrics for nearest server selection*

We now present a methodology for evaluating the ability of various distance metrics to correctly identify the destination with lower latency from a given monitor.

A 'trial' is a pairwise comparison of RTTs from a probe monitor to two destinations. Applying a particular distance metric to the trial yields one of three outcomes: success, failure, or inability to decide. The trial is *successful* if the destination with the lower value of the metric also actually has lower RTT from the monitor. The trial is a *failure* if the destination with a lower value of the metric had higher RTT. If the metric value for both destinations is the same, then the trial had *no predictive value*. The score for each metric is the percentage of *successes*.

Note that RTT values are affected by diurnal network patterns or short-lived events. To avoid unfair comparisons resulting from RTT measurements that are far apart in time, we temporally group destinations in our pairwise trials. We compute the value of a given metric for a given destination and we know the current RTT to this destination from the probe. We then compare the metric value and the RTT for this destination to the metric values and RTTs for destinations in 20 immediately preceding and 20 subsequent probes. All monitors probe 20 destinations in under 3 minutes.

III. Results

A. *General properties of IP and AS path length metrics*

Distributions of IP path length and AS path length have a single mode centered near the arithmetic mean, and are skewed toward higher values. Mean and standard deviation are appropriate statistical measures for such distributions. In particular, during the first five months of 2001 we obtained the following mean length values for measurements of the *IPv4_Space* list and the *DNS Clients* list, correspondingly: $L_{IP} = 15.3 \pm 4.2$; $L_{AS} = 4.1 \pm 1.3$; and $L_{IP} = 14.5 \pm 4.2$; $L_{AS} = 4.4 \pm 1.3$.

Note that IP path length and AS path length may vary over time because of infrastructural or routing changes. We quantified changes in our AS path length measures, to test the articulated hypothesis that Internet AS path lengths have generally increased in the last few years [17].

Our current results are ambiguous. The mean *IP path length* measured over time has increased at 4 monitors, has remained constant at 3 monitors and has decreased at 2 monitors. At the same time, the mean *AS path length* has increased at 2 monitors, remained constant at 5 monitors, and decreased at 2 monitors. The AS path length as seen in RouteViews BGP routing tables has remained relatively flat for the last three years [8]. Over the

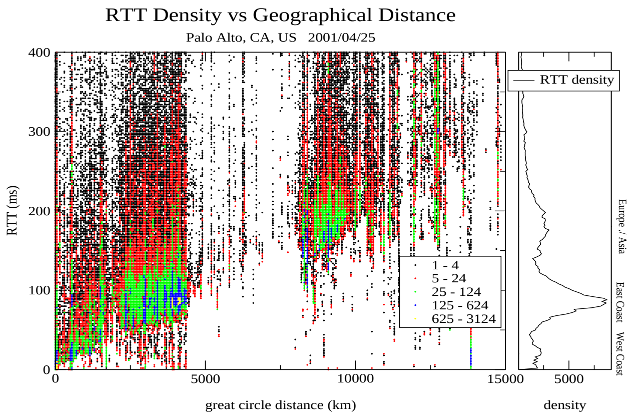

Fig. 2. RTT density vs geographic distance for CAIDA topology monitor in Palo Alto, CA, USA. 25 April 2001.

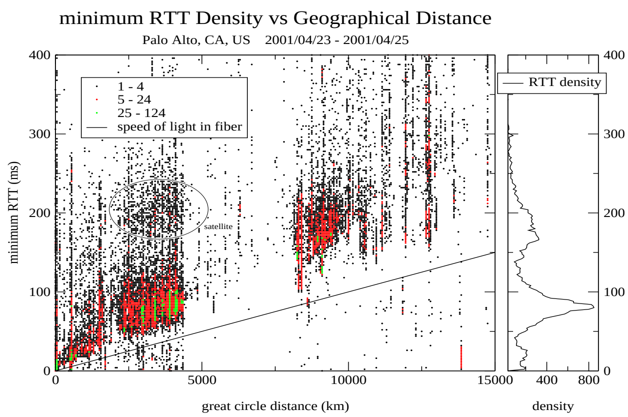

Fig. 3. Minimum RTT density vs geographic distance for CAIDA topology monitor in Palo Alto, CA, USA. 25 April 2001.

same three years, the standard deviation for IP path lengths in our data has remained stable, while the standard deviation for AS path length has increased from 1 to 1.25.

B. Properties of geographical distance and RTT metrics

RTT distributions are typically either bi- or trimodal and heavy-tailed [10]. Peaks in the distribution correspond to clusters of hosts at similar geographic locations. Figure 2 shows all RTT values observed on April 25, 2001 by the topology monitor on the west coast of the United States. On the right side of the figure is a rotated histogram plot of the same data, illustrating a typical multimodal RTT distribution. The three maxima in this curve (from bottom to top) represent destinations on the west and east coasts of the U.S., and in Asia/Europe.

The heavy-tailed RTT distributions for each destination mask the correlation between RTT and geographical distance. By plotting only minimum RTTs, figure 3 omits these long tails from consideration. The filtered data set illustrates that some destinations have minimum RTTs outside the cluster of RTTs for their geographic group. Unusually low RTT values, especially those below the propagation time of the speed of light in fiber, derive from our imperfect IP-to-geographic_location mechanism `IxMapper`, which incorrectly placed these destinations in Europe or Asia when they were likely rather somewhere in the U.S.

There are also destinations with minimum RTTs higher than

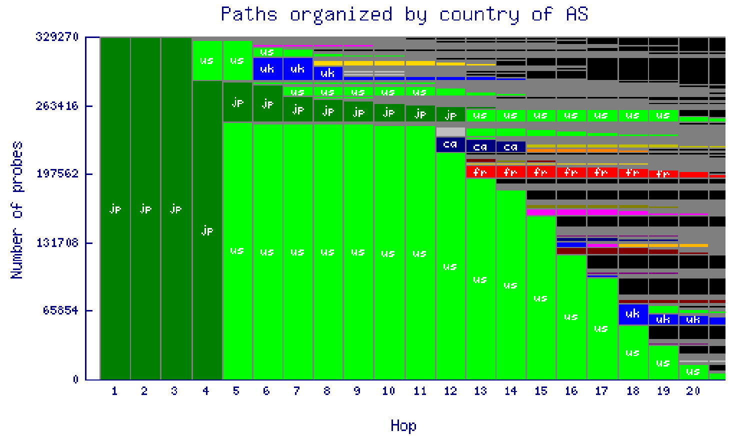

Fig. 4. Distribution of AS paths from CAIDA's topology monitor in Tokyo, Japan by countries. 13 May 2001

the corresponding geographical cluster – the 'satellite' group. We verified that while `IxMapper` accurately located of these nodes, traceroutes to them often show large RTT jumps at a certain hop. We found that interfaces where such jumps occur often belong to companies offering satellite links to customers.

Market realities of international Internet transit also cause violation of correlation between geographic distance and latency. Paths between two hosts that are both outside the U.S. often pass the U.S. as a 'transit country' [9][11]. For example, figure 4 shows the country-based dispersion of paths originating from CAIDA's topology monitor in Japan. Paths from Japan to the United Kingdom, France, and Canada almost always go through the U.S. first. In such cases the sum of the host-to-U.S. plus the U.S.-to-destination distances will have greater predictive power for RTT than the direct source-to-destination distance.

C. Resulting scores of distance metrics

Figure 5 illustrates the ability of various distance metrics to correctly identify the nearest server using the approach described in subsection II-C. Percentage of successful trials is shown for each metric and each monitor. Starting at the top of the figure: 'RTT' refers to the median RTT calculated for a given destination from the previous day's sample; 'Geo' uses the great circle distance between the monitor and a destination; 'IP' is the IP path length; and 'AS' is the AS path length.

Scores vary widely among metrics, but with the exception of the geographic distance metric, variability among monitors is small. The median RTT metric always yields the highest score, although it is slightly lower for the *IPv4_Space* list than for the *DNS Clients* list. We hypothesize that this difference is due to the fact that CAIDA topology monitors poll each destination in the *IPv4_Space* list only once daily. So for destinations in this list we actually use not the median RTT from among a set of samples, but rather a single point measurement of the RTT taken the day before. While less statistically robust than the median, it is remarkable that even a single RTT measurement has greater predictive power for choosing the nearest (shortest RTT) server selection than any other distance metric.

Given the close correlation between RTT and geography shown in subsection III-B, it is not surprising that end-to-end ge-

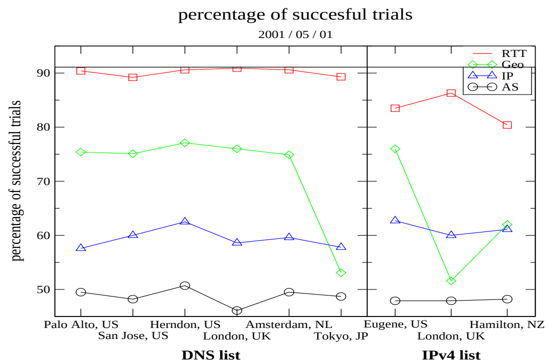

Fig. 5. Success rates for four distance metrics. 1 May 2001.

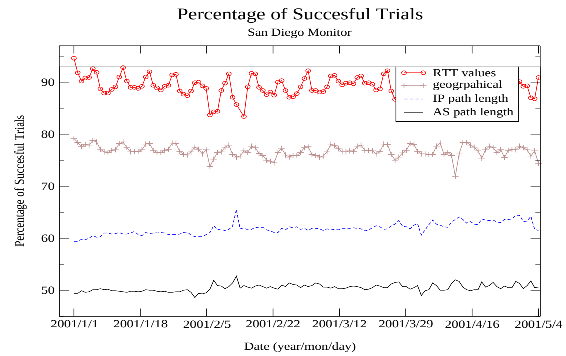

Fig. 7. Temporal variability of success rates for the different metrics during five months of 2001.

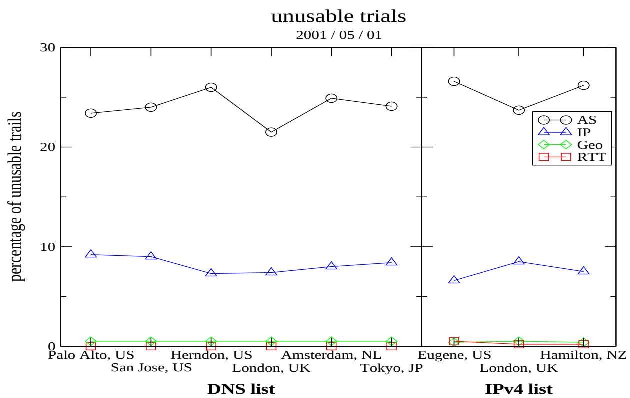

Fig. 6. Non-predictive trial rates for four distance metrics. 1 May 2001.

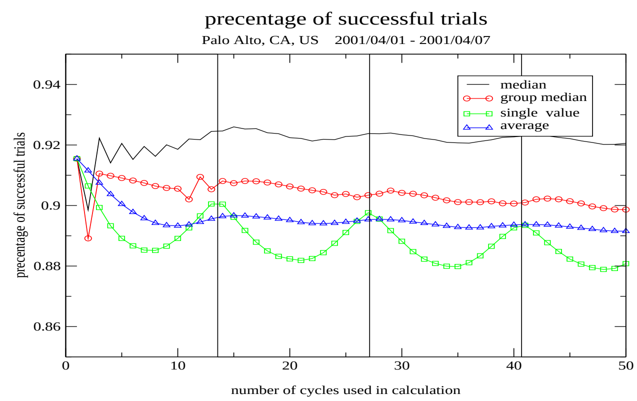

Fig. 8. Success rates for the RTT-based metrics for the monitor in Palo Alto, CA. April 1 to 7, 2001

ographic distance is the second best metric for server selection. However peculiarities in international Internet transit (described above) render this a weaker metric for monitors outside the U.S.

The average score for IP path length is about 60% success. The AS path length metric scores a 50% success rate, which makes it about the same as random selection, i.e. useless.

To be of practical use a metric should be applicable in most cases. Figure 6 shows the percentage of trials with no predictive value for each metric and each monitor in our study. Approximately 23% of the trials were between AS paths of identical length. AS path length is an integer number that is most often of length 3, 4, or 5, which does not present enough variety for differentiation. In contrast, IP path lengths for a pair of destinations were identical in only 7% of trials, and fewer then 0.5% of geographic distances and median RTT trials yielded indistinguishable (non-predictive) results. We conclude that the RTT-based metrics (median or even single previous day measurement) are the most useful for estimating *nearness* in the Internet.

Figure 7 shows the relative stability of different metrics over time. Scores are generally stable but show clear weekly patterns caused by weekday versus weekend traffic differences.

D. Comparison of different RTT-based metrics

We found that RTT-based distance metrics (single value previously observed or median of a set of previously observed values) have higher correlation with latency than any other metric we considered. We next discuss two more questions: 1. What RTT-based distance metric is best? and 2. What is the optimal amount of past data to use in calculating a given RTT-based metric?

We consider the following RTT-based metrics: 'median', 'average', 'single value' previously observed, and 'group median'. For the group median, we used the median of RTTs observed in the 12-hour window centered at exactly 24-hours before the current RTT value. Figure 8 shows results obtained for the topology monitor in Palo Alto, CA, which probes each destination in its list about 13 times daily. The x-axis shows the number of previous probe cycles used to derive a value of a given RTT-based metric. Increasing x-values correspond to using a greater number of previous RTTs in these calculations. The vertical lines designate 24-hour boundaries. The y-axis shows the percentage of successful trials, as defined in section II-C.

Figure 8 shows that all RTT-based metrics have high (around 90%) success rates, and that the difference between maximum and minimum success rates is only about 4%. The 'median' metric performs the best, especially when we calculate it from all data available from the previous 24 hours of measurements. The success rate for the 'average' metric remains nearly constant regardless of how much data is used to calculate it. The 'average' metric yields poorer results than the 'median' and 'single value observed 24 hours ago'. The success rate of the 'single value' metric exhibits a clear diurnal pattern reaching its maximum at the 24-hour boundary. This observation means that a single RTT value measured the day before *at the same time of day* has strong positive correlation with the current latency to the same destination. However when we tried 'group median', hoping to better capture and enhance this 24 hour periodicity, we discovered that the 'median' derived from all available data points from the previous 24 hours works better.

Finally, we see that for all RTT-based metrics, using more of the previous day's data does not necessarily increase its success rate. For some metrics the success rate may even decrease if the sample of past data is not synchronized with current trials at a 24-hour periodicity.

IV. Conclusions and Future Work

We have defined and studied four metrics of distance between Internet hosts: IP path length; AS path length; geographic distance; and RTT-based. We characterized the distributions of these metrics for a comprehensive sample of hundreds of thousands of Internet hosts, and analyzed changes across metrics and over time. In particular, our data do not support the hypothesis of an increase in AS path length over the last few years.

We also presented a novel technique for evaluating the effectiveness of various metrics for solving the *nearest server selection problem*. Metrics derived from previously measured RTTs correctly identified the server with the lower RTT in up to 90% of our trials and yielded the best results for all paths we probed. Great circle geographic distance is the second best metric; for 6 out of 9 monitors it achieves a 75% success rate. Five of these topology monitors are in the U.S.; one is in London. For the remaining 3 monitors (all outside of the US), selection based on end-to-end geographic distance was slightly better than random selection, but not better than selection based on IP path length. The IP path length metric successfully predicts the server with lower latency in 60% of the trials. The AS path length metric is successful in only 50% of cases, no better then random choice. For all four metrics, success rates are stable over time.

When comparing different RTT-based metrics, we found that the median RTT derived from data observed during the previous 24 hours is the best. However, calculating the median metric incurs significant overhead in collection, storage and processing of previous RTT data. Using a single RTT value observed within 24 hours of the current time may provide a much simpler solution that still bears acceptably high positive correlation to current latency (less than a 4% drop in success rate). Furthermore, our data suggests that using previously observed RTT data that is more than 24 hours old to calculate an RTT-based metric does not improve its predictive power.

All four metrics we considered are derived from information local to the topology monitor. This information may potentially be available to any client, allowing its use for nearest server selection. Alternative metrics are possible that could incorporate global information, e.g., average IP path distance from a destination to every other end host in an IP topology graph. Such global metrics are likely more pertinent to other problems, such as server placement, e.g., of DNS root nameservers [18] [19].

We are also exploring composite metrics that will combine multiple metrics and can potentially be incorporated into a framework for global Internet distance estimation. More broadly, the concept of distance between Internet hosts is relevant to network optimization and traffic engineering problems. The number of possible paths between two hosts is practically unlimited but some paths will perform better than others at different times. A metric that reflects performance 'distance' between two arbitrary hosts within the network, and is readily available at little cost in measurement, storage, and processing, would be of great benefit to traffic engineering. In a variety of ways, a global architecture for host distance estimation based on empirical data and proven methodology could help mitigate scalability problems as the Internet continues its stunning growth.

References

[1] H. Woolf, Ed., *Webster's New Collegiate Dictinoary*, G. & C. Merriam Co., 1981.

[2] J. Heidemann and Visweswaraiah V., "Automatic Selection of Nearby Web Servers," in *Technical Report 98-688, University of Southern California*, Dec. 1998.

[3] E. Corvella and R. Carter, "Dynamic server selection in the Internet," in *IEEE Infocom*, Aug. 1995.

[4] P. McManus, "A passive System for Server Selection within Mirrored Resource Environments Using AS Path Length Heuristics," in *AppliedTheory Communications, New York*, Apr. 1999, http://www.gweep.net/ mcmanus/proximate.pdf.

[5] P. Francis, S. Jamin, V. Paxson, L. Zhang, and D. Gryniewicz, " An Architecture for a Global Internet Host Distance Estimation Service ," in *IEEE INFOCOM*, June 2002, http://citeseer.nj.nec.com/francis99architecture.html.

[6] T. S. Eugene Ng and H. Zhang, "Predicting Internet Network Distance with Coordinates-Based," in *IEEE INFOCOM*, March 1999, http://citeseer.nj.nec.com/502616.html.

[7] K. Lougheed and Y. Rekhter, "RFC 1106: Border Gateway Protocol (BGP)," June 1990.

[8] A. Broido, E. Nemeth, and k. claffy, "Internet expansion, refinement, and churn," in *European Transactions on Telecommunications*, Jan. 2002, http://www.caida.org/outreach/papers/2002/EGR/.

[9] B. Huffaker, M. Fomenkov, D. Moore, E. Nemeth, and k. claffy, "Measurements of the Internet topology in the Asia-Pacific Region," in *INET*, July 2000, http://www.caida.org/outreach/papers/asia_paper/.

[10] B. Huffaker, M. Fomenkov, D. Moore, and k. claffy, "Macroscopic analyses of the infrastructure: measurement and visualization of Internet connectivity and performance," in *PAM*, October 2001, http://www.caida.org/outreach/papers/2001/Rssac2001a/.

[11] F. Junqueira and R. Teixeira, "Connectivity in the South American Internet," 2000, http://pontal.ucsd.edu/ flavio/Docs/proj.ps.

[12] B. Huffaker, D. Plummer, D. Moore, , and k. claffy, "Topology discovery by active probing," in *SAINT*, Jan. 2002, http://www.caida.org/outreach/papers/2002/SkitterOverview/.

[13] D. Meyer, "University of oregon route views project," http://www.antc.uoregon.edu/route-views/.

[14] "topology monitors status," topology monitors status, http://www.caida.org/cgi-bin/skitter_monitors/main.pl.

[15] "Ixia's geographic mapping service: Ixmapping," http://www.ixiacom.com/products/paa/netops/IxMapping.php.

[16] A. Broido and k. claffy, "Internet topology: connectivity of IP graphs," in *SPIE International symposium on Convergence of IT and Communication*, Sept. 2001, http://www.caida.org/outreach/papers/2001/OSD/.

[17] G. Huston, "The Changing Structure of the Internet," http://www.telstra.net/gih/papers/apectel23.pdf.

[18] M. Fomenkov, k. claffy, B. Huffaker, and E. Moore, D. Nemeth, "Macroscopic Internet Topology and Performance Measurements from the DNS Root Name Servers," in *Usenix LISA*, Dec. 2001, http://www.caida.org/outreach/papers/2001/Rssac2001a/.

[19] N. Brownlee, E. Nemeth, and k. claffy, "DNS Root/gTLD Performance Measurements," in *Usenix LISA*, Dec. 2001, http://www.caida.org/outreach/papers/2001/DNSPerfMeas/.