Quantifying Routing Asymmetry in the Internet at the AS Level

Yihua He   Michalis Faloutsos   Srikanth Krishnamurthy
yhe@cs.ucr.edu   michalis@cs.ucr.edu   krish@cs.ucr.edu
Department of Computer Science and Engineering
University of California, Riverside

Abstract— **In this paper, our objective is to quantify the extent of the routing asymmetry in the Internet: the measure of the difference between the forward and backward paths between two end points. Routing asymmetry has not been studied extensively. Most of the previous studies only consider asymmetry in terms of length and there is a lack of a systematic approach for quantifying asymmetry. One of the challenges in quantifying asymmetry is the formulation of an appropriate set of metrics that can effectively capture various notions of asymmetry. We point out that asymmetry could be of various types. We propose a framework to quantify the routing asymmetry between end hosts and propose two *new* metrics: *Absolute Asymmetry* and *length-based Normalized Asymmetry*. Our metrics capture the differences in router/AS identities, their appearing sequences and path lengths in a seamless way. We apply our framework to real Internet measurement data and examine routing asymmetry at the Autonomous System (AS) level. We deduce the routing asymmetry distribution based on our framework, and we find that about 14% of pairs of routes considered display AS level routing asymmetry. Furthermore, our studies demonstrate that the routing asymmetry exhibits a skewed distribution, since a few end-points are consistently members of asymmetric pairs.**

I. Introduction

Routing asymmetry is a common phenomenon in the Internet. If we examine the end-to-end paths between a pair of arbitrarily chosen hosts (say A and B), we often see that the path from host A to host B, is different from the path from B to A in the reverse direction. This asymmetry in the Internet can appear at both AS level and router level paths. Clearly, asymmetry at the AS level will also lead to router level asymmetry, while router level asymmetry can exist even if the AS paths are symmetric. Although the routing asymmetry is an important routing phenomenon, it has not been studied extensively.

Routing asymmetry has a significant impact on the network performance and our ability to measure, model, and manage the network. With the presence of asymmetry, identifying paths and estimating delays in the network becomes more challenging. For modeling purposes, identifying paths becomes more difficult, since we need to measure two directions instead of one, between any pair of hosts. Second, asymmetry introduces a significant problems in estimating the one way latency between hosts. Measuring and modeling one-way delay is important for many delay-sensitive applications. Currently, the most common practice is to estimate the one way delay by the half of the easier-to-compute round-trip time delay. Clearly, this estimate becomes worse for increased routing asymmetry. Third, measuring and monitoring the routing asymmetry may potentially be an important indicator of the state of the Internet. For example, a higher than normal asymmetry may suggest changes or even errors in the routing practices.

Permanent asymmetric routes may exist due to the practice of load balancing, policy routing or traffic engineering. At the same time, the Internet may exhibit transient asymmetries as an effect from BGP convergence. An in-depth study of routing asymmetry can undoubtedly aid our understanding of the Internet. However, there have been few studies on routing asymmetry despite its importance. One may attribute this to the challenge in developing a systematic approach for measuring asymmetry. Most previous work focuses on the asymmetry in terms of length between the forward and reverse paths [1][21]. Typically, such efforts simply classify a path as either asymmetric or symmetric. without considering all different types of asymmetry or quantifying the degree of the asymmetry.

The goal of this paper is to capture and quantify the asymmetry in a systematic manner. We want to *measure* the magnitude of routing asymmetry between a pair routes. We start by observing three distinct types of asymmetry. A pair of forward and reverse paths can differ in: a) the identities of their nodes, b) the sequence in which their nodes appear, and c) their length. These three types of asymmetry are interwined, which makes their classification and separation challenging. Here, we develop a framework to quantify routing asymmetry by capturing all of the aforementioned types. We then apply our framework to study real Internet measurements and quantify the AS level routing asymmetry among US higher education institutes.

In more detail, our contributions can be summarized in the following points:

- We develop a framework to quantify asymmetry by proposing two metrics, which capture the magnitude and the relative significance of the routing asymmetry. Consulting the string matching literature, we propose a polynomial-time algorithm based on dynamic programming to estimate the value of the asymmetry.
- We assess the asymmetry of the Internet using our framework. We analyze real traceroute data from NLANR [2] on a specific date (Jan 22, 2004) and deduce the distribution of asymmetry in the Internet on that date.
- We observe that about of 14% the pair of routes exhibit AS level asymmetry, and nearly half of that asymmetry is due to an additional AS hop in one direction.
- We find that the spatial distribution of the asymmetry

is skewed: a few end-points are consistently involved in asymmetric pairs. This implies that the end-to-end routing asymmetry is not uniformly or near-uniformly distributed.

The rest of this paper is organized as follows: In section 2, we present background and related work in brief, on the topic of routing asymmetry. We describe our proposed framework of measuring asymmetry in section 3. We follow with a presentation on an application of our framework on real Internet traceroute data in section 4. Further, in this section we analyze the distribution of Internet asymmetry based on our framework. In section 5, we conclude our work and discuss possible future research directions.

II. Background and Related Work

Routing asymmetry in the Internet has not received much attention despite its importance. To the layperson, routing asymmetry may seem odd: why should not packets follow the same path between end points, preferably the shortest one? In practice however, policy routing and traffic engineering [3] are among the main causes of asymmetry. Internet providers are business oriented and their focus is to implement their policies as per their interests. For example, a packet destined for a different network is moved out of a provider's network as soon as this is possible, even if this means that the packet will experience a longer path or congestion. Traffic engineering may also create asymmetry: a router may attempt to shift traffic from a highly loaded link to a lightly loaded link. In other words, with traffic engineering, the network may potentially alter routing to avoid congested regions. In addition to policy routing and traffic engineering, the absence of a unique shortest path [9] [10] between a pair of hosts could also lead to asymmetric routing. Intra-domain routers use Bellman-Ford or Link State routing [11] algorithms to calculate routing tables. When more than one route between a pair of hosts, reflect the same cost, routers may arbitrarily choose among the equivalent routes.

The seminal work in this area by Paxson [1] studies routing pathologies and defines the problem of routing asymmetry. Subsequently, other studies have addressed this issue partially, looking primarily at the difference in the path lengths (or the difference in delay) between the forward and reverse paths[4]. There has been some work on path inflation due to policies [7][8][13]. Although these topics are related to our work, they are fundamentally different for two reasons: (1) asymmetry does not necessarily result in a difference in path length or delay incurred on the forward and reverse paths ; and (2) path inflation does not necessarily result in routing asymmetry (both the forward and reverse paths may be the same; but they may be inflated in comparison with the shortest paths between the considered end-hosts). We mainly focus on capturing routing asymmetry, and in particular, AS level asymmetry in this work.

Several tools have been built to study the end-to-end routing properties. Mercator [5] is a routing map collection tool run from a single host. Skitter [6] monitors probe the network from about 20 different locations worldwide. Rocketfuel [12] uses a list of public traceroute servers to probe ISP maps. However, those works have not touched routing asymmetry.

III. Framework to quantify asymmetry

In this section, we discuss our approach to quantifying the routing asymmetry of an observed pair of routes between two end hosts. Two metrics are proposed: 1) Absolute Asymmetry (AA), and 2) length-based Normalized Asymmetry(NA). The latter captures the extent of the magnitude of asymmetry relative to the distance between the two end hosts.

In order to quantify the asymmetry between a pair of routes, we basically compares the dissimilarity between the entities on the two routes, considered in sequence. The problem is almost identical to comparing sequences of genes in computational biology[15], where string matching techniques are widely employed. Motivated by this approach, we measure the dissimilarity between a pair of routes by aligning the two routes together and measuring the minimal total cost incurred in aligning them. More formally, we quantify asymmetry as follows:

Consider a pair of paths between end hosts A and B. Suppose we have a route u from A to B and this route traverses a sequence of m entities[1]:

$$u = (u_1, u_2, u_3, ...u_m), \forall u_k \in S, 1 \leq k \leq m, \quad (1)$$

and route v from B to A traverses a sequence of n entities:

$$v = (v_n, v_{n-1}, v_{n-2}, ...v_1), \forall v_k \in S, 1 \leq k \leq n, \quad (2)$$

where, S is the set that includes all possible entities.

For any two entities x and y in S, we define a non-negative *base dissimilarity value* $\omega[x, y]$, which represents the magnitude of "how much x is different from y". The value should be set to 0 if $x = y$, and greater than zero if $x \neq y$. The set of such values for all entity pairs in S forms a *base dissimilarity matrix*:

$$\Omega(S) = \{\omega[x, y]\}, x \in S, y \in S \quad (3)$$

A mapping Ψ with function $\psi(i) : (1, 2, ...m) \rightarrow (\psi(1), \psi(2), ..., \psi(m))$, *sequentially* maps all indices in path u to the set of indices of a subset of v:

$$\begin{aligned}\Psi : (1, 2, ...m) \rightarrow (\psi(1), \psi(2), ..., \psi(m)) \text{ such that,}\\ \phi \subset \{v_{\psi(1)}, v_{\psi(2)}, ..., v_{\psi(m)}\} \subseteq v,\\ \forall i \in \{1, 2, ..., m-1\}, \psi(i) \leq \psi(i+1) \quad (4)\end{aligned}$$

.

Here, ϕ represents the empty or null set. Similarly, a mapping Ψ' with function $\psi'(i) : (1, 2, ...n) \rightarrow (\psi'(1), \psi'(2), ..., \psi'(n))$, which *sequentially* maps all indices in path v to the set of indices of a subset of u:

$$\begin{aligned}\Psi' : (1, 2, ...n) \rightarrow (\psi'(1), \psi'(2), ..., \psi'(n)) \text{ such that,}\\ \phi \subset \{u_{\psi'(1)}, u_{\psi'(2)}, ..., u_{\psi'(m)}\} \subseteq u,\\ \forall i \in \{1, 2, ..., n-1\}, \psi'(i) \leq \psi'(i+1) \quad (5)\end{aligned}$$

[1]Here "entities" represent ASes. However this framework could apply to router level analysis. In that case, an entity would represent a router.

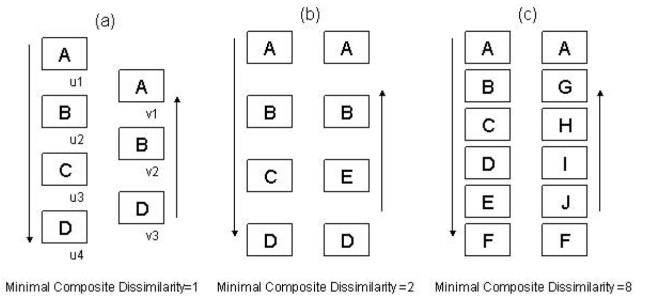

Fig. 1. Some simple asymmetry cases and their optimal mapping cost according to Equation (7)

.

The *composite dissimilarity* of a pair of sequential mappings between two paths u and v is denoted by $C(u, v, \Omega, \Psi, \Psi')$:

$$C(u, v, \Omega, \Psi, \Psi') = \sum_{i=1}^{m} \omega(u_i, v_{\psi(i)}) + \sum_{i=1}^{n} \omega(v_i, u_{\psi'(i)}) \quad (6)$$

Note that when computing the base similarity $\omega(x, y)$ we do a pairwise comparison of the elements x and y.

We define C_{min} as the *minimal composite dissimilarity* for all possible mappings between u and v:

$$C_{min} = min\{C(u, v, \Omega, \Psi, \Psi')\}, \forall\Psi, \Psi' \quad (7)$$

and the mapping pair[2] (Ψ, Ψ') is called the *optimal mapping pair*.

Given a base dissimilarity matrix and a pair of routes u and v, the optimal mapping pair and its composite dissimilarity can be calculated in O(mn) time by means of a dynamic programming strategy; here, m and n are the lengths in terms of hop count, for routes u and v, respectively. In the paper,we use a simple base dissimilarity matrix, in which $\omega(x, y) = 0$ if $x = y$, and $\omega(x, y) = 1$ if $x \neq y$. Note that there are other meaningful ways to define such a matrix. For example, if we want to quantify asymmetry in terms of geographic difference, we can define $\omega[x, y]$ to be the geographic distance between x and y. However, the question of how far (or how 'different') two routers or ASes are, is beyond the scope of this research. For simplicity, we use the aforementioned simple base dissimilarity matrix in this work unless otherwise stated.

In [14][15], the optimal alignment cost was believed to be a reasonable measure of quantifying the extent to which two strings are different from each other. To demonstrate that the minimal composite dissimilarity is a reasonable measure of asymmetry, we consider examples of routing asymmetry, compute their minimal composite dissimilarity by our framework and discuss the implications. The examples that we consider are shown in Fig 1.

In Fig (1a), one of the optimal mapping pairs (Ψ, Ψ') is found as follows: Ψ is a mapping from $(1, 2, 3, 4) \rightarrow (1, 2, 2, 3)$, which corresponding to comparing $(u_1, u_2, u_3, u_4) and (v_1, v_2, v_2, v_3)$ and Ψ' is a mapping from $(1, 2, 3) \rightarrow (1, 2, 4)$, which corresponding to comparing $(v_1, v_2, v_3) and (u_1, u_2, u_4)$. In this example, the only mismatch is (u_3,v_2)=(C,B). So the minimal composite dissimilarity is 1. In Fig (1b), the optimal mapping pairs (Ψ, Ψ') are such that $\Psi = \Psi' = (1, 2, 3, 4) \rightarrow (1, 2, 3, 4)$. But there are two mismatches: (u_3, v_3) in Ψ and (v_3, u_3) in Ψ'. So its minimal composite dissimilarity is 2. By means of a similar process, we can compute the minimal composite dissimilarity in the case of the example in Fig (1c) to be 8.

We clearly see that our metric computes a larger magnitude for asymmetry in the case of the example in Fig (1c) than for those in Fig (1a) or Fig (1b). This clearly matches with one's intuitions. However, it is debatable as to whether the example in Fig (1b) is more asymmetric than the one in Fig (1a). We argue that our framework and metrics are reasonable because: 1) One would agree that the path length is not the only factor in determining asymmetry; only comparing two paths in terms of the difference in the path lengths will underestimate asymmetry. 2) our framework is flexible enough to address on the debate; a quick fix would be to assign different basic dissimilar matrices depending upon what is important. We consider a case in which we want to measure asymmetry in terms of the geographic difference between nodes on the two paths. We assign the basic dissimilar matrix to be the geographic distance between two nodes on the path. Suppose C and E are very close to each other, but C is far away from B or from D. Then our algorithm would assess that the example case in Fig (1a) is more asymmetric than the one in Fig (1b). Thus, we believe that minimal composite dissimilarity C_{min} is a plausible, reasonable and simple measure of asymmetry. We build our two metrics based on it.

Absolute Asymmetry (AA) is the minimal composite dissimilarity between a pair of forward and backward paths u and v.

$$AA(u, v) = C_{min}$$

length-based Normalized Asymmetry (NA) is the Absolute Asymmetry normalized by the computed round-trip path length.

$$NA(u, v) = \frac{AA(u, v)}{length(u) + length(v)}$$

The absolute asymmetry captures the absolute magnitude of asymmetry inherent in a pair of forward and backward paths, and the length-based normalized asymmetry indicates the extent to which a forward route between a pair of nodes is "off" from its reverse counterpart; in other words, a pair of long routes are considered to be less significantly deviant from symmetry as compared with a pair of shorter routes with the same absolute asymmetry. From the discussion above, we immediately note that for any pair of routes u and v, $0 \leq NA(u, v) \leq 1$ if $\forall\omega(x, y), 0 \leq \omega(x, y) \leq 1$. In the two extreme cases when (a) u and v are exactly the same,

[2]More than one optimal mapping pairs may exist due to a tie in the minimal composite dissimilarity.

$NA(u, v) = 0$; and when (b) u and v contain completely different entities, $NA(u, v) = 1$.[3]

IV. Experimental Evaluations

A. Methodology

We use the data collected by the Active Measurement Project(AMP) from The National Laboratory for Applied Network Research (NLANR)[2] on Jan 22, 2004. The AMP architecture uses a near-full mesh (each monitor sends probes to almost all of the other monitors) that interconnects approximately 135 active monitors deployed at remote sites. These sites are mostly distributed among North American higher education and research institutes. However, sites are also in the Asian Pacific, Latin America, and Europe. A traceroute[17] is performed between each pair of sites once, approximately, every 10 minutes. Some AMP sites can only probe some of the other AMP sites but cannot be probed from those sites. We eliminate such "one-way paths" from our data. The rest of the data volume accumulated over the single day is around 1 GB in size.

There are two reasons, due to which, we choose data from the AMP monitors. First, AMP monitors facilitate traceroutes; the paths obtained represent the actual routes that are traversed by the data packets[4]. The second reason is that AMP monitors send traceroute probes to each other so as to provide data with regards to "almost" the entire mesh; the data, thus, provides a complete set that enables comparisons between pairs of forward and backward paths.

We have extracted traceroute results conducted between the times of 00:05am and 00:15am PST on Jan 22, 2004. The intention is to capture a snapshot of Internet at a particular time. Since traceroute results may differ at different times due to traffic engineering, BGP table convergence and possibly other effects, this intention of constraining ourselves to a time period, minimizes the temporal impact of changes on the traceroute results. The reason for choosing a time that is close to midnight as per the pacific standard time (3AM EST) is that, at this time, the network, for the most part, is considered quiet and this will limit the effects of BGP instability[16].

Note that the traceroute data set, obtained as discussed above, is not perfect. The major problem is that there exist addresses that cannot be resolved, i.e., "*"s and "!"s appear in the traceroute data. In [19], Mao *et al*, describe several heuristics to interpret and thereby reduce the number of "*"s and "!"s in observed traceroute data. These heuristics are based on relatively loose matching criteria; a BGP AS path is classified to match a traceroute AS path if an ordered portion of the latter (created by simply ignoring the "*"s) matches exactly with a similar ordered sequence of ASs in the former path. In our work, however, our goal is to map *every* AS on a forward

[3]In fact, if we assign the simple basic dissimilarity matrix NA can never be equal to 1 since, in all cases, $AA(u, v) < length(u) + length(v)$, due to the fact that the end host pair on the forward and reverse paths are identical.

[4]BGP AS-PATHs are known as policy paths; A BGP AS-PATH between a pair of hosts may not be the identical to the traceroute path between the same pair[18].

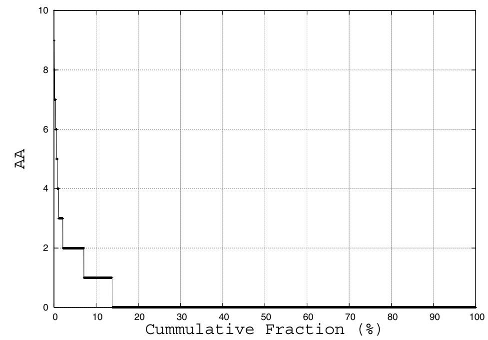

Fig. 2. AA Distribution

path to an AS on its reverse counterpart. Furthermore, since the comparison in [19] simply attempts to declare whether or not a match is seen (a yes or no criterion), a few misinterpreted "*" or "!" may not change the results considerably. On the contrary, our comparisons lead to a *quantitative* result; mis-resolved addresses may *artificially* increase or decrease our estimate of asymmetry. Currently, we are still working on identifying proper heuristics to map unresolved addresses on to specific addresses. In the meantime, our approach is to simply discard data that contains unresolved addresses. As a result, about 27% of the pairs of routes were eliminated; 5616 pairs of routes were left in our data set.

There are two levels of routing asymmetry that can be quantified by our framework: AS level asymmetry and router level asymmetry. In this work, we focus on AS level asymmetry. To identify the ASes via which, paths pass through, we mapped the IP addresses from the traceroute data to AS numbers as per the existing AS-IP mapping policies from [19] and [20]. After this conversion, consecutive entities that reflect the same AS number are collapsed into a single entity. Our methodology is similar to the technique used in [7]. Thus, we obtain AS level data paths, and for these paths we compute the corresponding AA and NA metrics. The results are presented and discussed in the following section.

B. Results

We examine the distribution of asymmetry as it appears in our data. We analyze the nuances of AS level routing asymmetry by asking the following four questions: how many pairs of the considered routes exhibit AS level routing asymmetry? What is the significance of the observed asymmetry? Is a long path more likely to experience asymmetry? As observed from an end-host perspective, are all end hosts experiencing similar levels of asymmetry or do some end hosts experience higher levels of asymmetry as compared with the others?

Towards answering the above questions, we first examine the Absolute Asymmetry. Fig 2 depicts the cumulative distribution of this metric. The figure is obtained by computing the optimal mapping cost over 5000 pairs of routes, and sorting them in the descending order of their Absolute Asymmetry. About 14% of the pairs of routes in our data set display an value of Absolute Asymmetry of one or more at the AS level. In the other

TABLE I

DISTRIBUTION OF ABSOLUTE ASYMMETRY AMONG THE ASYMMETRIC PATHS (14% OF TOTAL PATHS)

Value of AA	Fraction
1	48.7%
2	36.1%
3	7.1%
4	2.2%
5	1.5%
6	1.6%
7	1.7%
8+	1.1%
Total:	100%

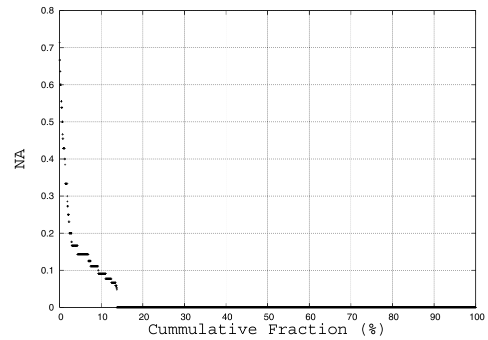

Fig. 3. NA Distribution

TABLE II

AS ROUTE PAIRS GROUPED BY THEIR SHORTER AS PATH LENGTH

Groups	Shorter Length	Fraction
Group 1 (short)	2 – 3	9.4%
Group 2 (median)	4 – 6	78.6%
Group 3 (long)	7 – 10	12.0%

words, most of the end to end paths that were considered were totaly symmetric relative to each other, in our data set. The fraction of the routes that display asymmetry is about half of what was reported in [1]. We attribute this incongruity to two reasons. First, the properties of our data set are different from those of what was used in [1]. In the latter set, a significant number of participating sites were in Europe. This generated a considerable quantum of trans-Atlantic traffic across commercial lines. This in turn, could have potentially resulted in a larger fraction of routes displaying asymmetry. Second, the maximum time interval allowed, between conducting a pair of traceroute instances, in two opposite directions is 10 minutes in our experiments; however this interval varied from 2 hours to 2 days in the experiments in [1]. Our methods *limit* the possibility of *over-estimating* asymmetric routes. This is because, some of the symmetric routes may change over time and may appear to be "asymmetric" if the traceroutes are not all constrained to within a short time span.

Table I reflects the distribution of the Absolute Asymmetry among the asymmetric routes. Almost half of the asymmetric routes show the value of Absolute Asymmetry of 1. Since each element of the base dissimilarity matrix that was used in this research can only take on two possible values, 0 and 1, all of the routes classified as asymmetric with Absolute Asymmetry of 1 are a result of one additional AS in one direction. This implies that about half of asymmetric routes are thus classified, simply because there is a single additional AS in one direction. A few pairs of routes do show higher Absolute Asymmetry; this would imply that, for these routes, the path in one direction differs a lot from the path in the other direction.

Fig 3 shows length-based Normalized Asymmetry (NA) distribution for our data set. The metric provides an intuitive view of the significance of the asymmetry between a pair of routes. We find that most (more than 90%) of the asymmetric routes exhibit low NA (less than 0.1). This means most of the forward routes do not significantly differ from their reverse counterparts. Very few pairs of routes in our data set exhibit high NA (over 0.7). In such cases, one may infer that the forward path is almost totally different from its reverse counterpart.

It is natural for one to raise the question: "Is a longer path more likely to exhibit asymmetry?". To answer this question, we cluster route pairs as follows: those pairs in which, the shorter one way AS paths are of the same length, belong to the same cluster. We then group these clusters to form coarser *groups* as shown in Table II. We then compare the asymmetry among the groups. Fig 4 and Fig 5 depict the AA and the NA distributions for the three groups. Not surprisingly, longer paths, generally, are more likely to experience asymmetry: more than 20% of the route pairs in group 3 experience asymmetry to some degree; this is noticeably higher than the average fraction (14%) of routes (overall) that display asymmmetry. However, the difference between the asymmetry levels displayed by the routes in groups 1 and 2 is insignificant in terms of the AA distribution. On the other hand, we observe from Fig 5 that longer paths (group 3), generally, have lower NA values than the the shorter paths when the range of the NA values is beyond 0.07. These facts suggest that while longer paths are more likely to experience asymmetry, the magnitude of the asymmetries that they experience is generally low in terms of its proportion to their lengths.

Finally, for each AMP site in our data set, we compute the average NA of all routes that start or end from that location.

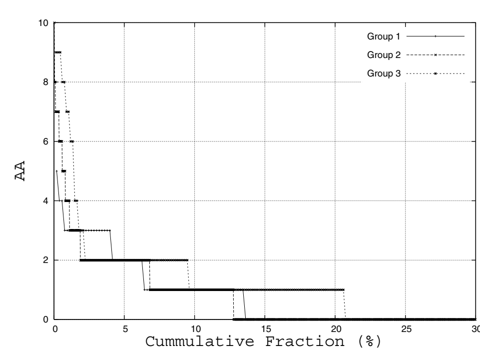

Fig. 4. AA Distribution Grouped by Shorter Path Length

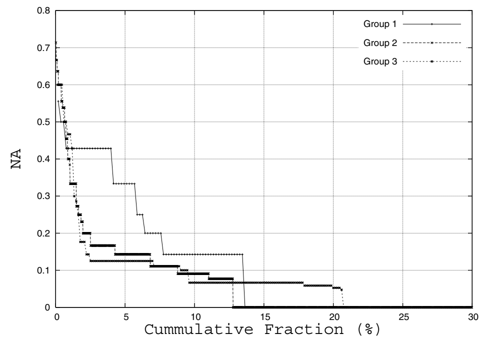

Fig. 5. NA Distribution Grouped by Shorter Path Length

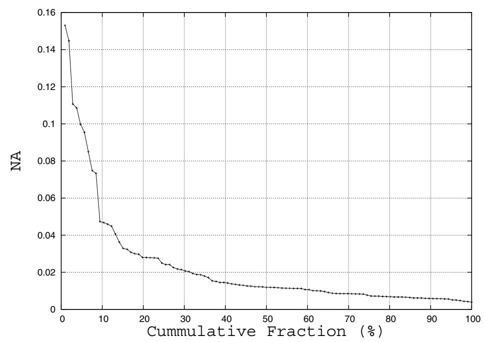

Fig. 6. Spatial NA Distribution

We sort the computed average NA and plot the distribution in Fig 6. We observe a skewed distribution, i.e., a few AMP sites exhibit, on average, a higher NA than the rest. This suggests that end-to-end routing asymmetry is not uniform or near uniform in the Internet.

V. Conclusions

Our key contribution in this paper is a framework to systematically quantify routing asymmetries of the Internet. Our framework uses a string matching technique that is commonly used in pattern matching. Note that our framework can integrate different "cost functions" to capture desired aspects of the asymmetry as we explained in section 3. In addition, we provide two metrics, the absolute asymmetry (AA) and the length-based normalized asymmetry (NA), to reflect the extent and the significance of asymmetry for a given pair of paths. Our quantitative approach would help researchers better study and understand Internet asymmetry.

We apply our framework to real Internet measurement and compute the distribution of the routing asymmetry in the Internet at the AS level. We observe that, about 14% of routes, display AS level routing asymmetry. Furthermore, nearly half of the routes that are classified to be asymmetric, are due to an additional AS in either the forward or the reverse direction. We also find that longer paths are more likely to experience asymmetry. However, the magnitude of asymmetry that they exhibit is low relative to their lengths. We point out that end-to-end routing asymmetry is not uniform since we observe that the asymmetry follows a skewed spatial distribution.

In the future, we want to provide an in-depth large-scale analysis of the asymmetry of Internet routing. We want to apply our framework on larger and diverse data sets and study routing properties at different levels from a functional and geographical point of view. Furthermore, we are in the process of using our framework to quantify the asymmetry at the router level.

Acknowledgments

We would like to thank Bradley Huffaker and other CAIDA staff for their inputs during discussions. We also want to thank anonymous reviewers for their help to improve the paper.

References

[1] V. Paxson. End to end behavior in the Internet. In Proceeding of the ACM SIGCOMM, Volume 26, number 4, page 25-38, August 1996.
[2] NLANR active measurement, available at http://watt.nlanr.org
[3] R. Teixeira, K. Marzullo, S. Savage, G. Voelker, Characterizing and Measuring Path Diversity of Internet Topologies, ACM SIGMETRIC, San Diego, CA 2003
[4] S. Savage, A. Collins, E. Hoffman, J. Snell and T. Anderson, The End-to-end Effects of Internet Path Selection, ACM SIGCOMM Conference, Cambridge, MA, September 1999.
[5] R. Govindan and H. Tangmunarunkit, Heuristics for Internet Map Discovery, IEEE Infocom 2000
[6] B. Huffaker, D. Plummer, D. Moore, K. Claffy, Topology Discovery by Active Probing, Symposium on Applications and the Internet (SAINT) 2002
[7] H. Tangmunarunkit, R. Govindan, S. Shenker, D. Estrin, The Impact of Routing Policy on Internet Paths, IEEE Infocom, Anchorage, Alaska. 2001
[8] L. Gao and F. Wang, The Extent of AS Path Inflation by Routing Policies, Global Internet 2002
[9] R. Ghosh and G. Varghese, Symmetrical routes and reverse path congestion control. Technical Report TR-97-37, Department of Computer Science, Washington University, St. Louis, September 1997.
[10] G. Varghese, R. Ghosh, Modifying shortest path routing protocols to create symmetrical routes routing information protocol. UCSD technical report CS2001-0685, September 2001.
[11] R. Perlman. Interconnections: Bridge and routers. Addision-wesley, 1992
[12] N. Spring, R. Mahajan, D. Wetherall, Measuring ISP Topologies with Rocketfuel, ACM SIGCOMM, Pittsburgh, USA 2002
[13] N. Spring, R. Mahajan, T. Anderson, Quantifying the Causes of Path Inflation, ACM SIGCOMM, Karlsruhe, Germany, 2003
[14] A. Aho, Algorithms for finding patterns in strings, Handbook of Theoretical Computer Science, Chapter 5, Elscvier, Amsterdam, 1990
[15] D. Gusfield, Algorithms for Strings, Trees and Sequences: Computer Science and Computational Biology, Cambridge Press, 1997
[16] C. Labovitz, G. Malan and F. Jahanian, Internet Routing Instability, ACM SIGCOMM, Cannes, France, 1997
[17] V. Jacobson, "Traceroute", ftp://ftp.ee.lbl.gov/traceroute.tar.gz
[18] Y. Hyum, A. Broido, K. Claffy, Traceroute and BGP AS Path incongruities, Tech. rep.,Cooperative Association for Internet Data Analysis (CAIDA), 2003
[19] Z. Mao, J. Rexford, J. Wang, R. Katz, Towards an Accurate AS-level Traceroute Tool, ACM SIGCOMM, Karlsruhe, Germany, Aug 2003
[20] Z. Mao, D. Johnson, J. Rexford, J. Wang, R. Katz, Scalable and Accurate identification of AS-level Forwarding Paths, IEEE INFOCOM, Hongkong, China, 2004
[21] L. Amini, H. Schulzrinne, A. Lazar, Observations from Router-level Internet Traces, DIMACS Workshop on Internet and WWW Measurement, Mapping and Modeling, February 13-15, 2002.