# On Third-party Addresses in Traceroute Paths

Young Hyun, Andre Broido, kc claffy
CAIDA, San Diego Supercomputer Center
University of California, San Diego
E-mail: {youngh, broido, kc}@caida.org

***Abstract*— Traceroute IP paths are often used in studies of Internet topology and routing. Though producing one of the best available router-level maps of forward paths, traceroute is susceptible to several types of inaccuracies. However, no one to date has quantified the magnitude of these inaccuracies in real-world traceroute paths. We make an initial attempt by reporting on the observed frequency of one type of inaccuracy—third-party addresses. A third-party address is the address of a router interface that does not lie in the actual path taken by packets. Based on an examination of thousands of traceroute paths from six locations worldwide and the application of several metrics, we find that the situations in which third-party addresses can occur to be relatively uncommon. They mostly occur within a few hops of the destination (that is, at the destination edge of the network), with multihoming being their most likely cause. Our conclusion is that third-party addresses cannot be a significant source of AS map distortions.**

## I. Introduction

A traceroute IP path provides the hop-by-hop router-level forward path to a destination. Traceroute paths are a rich source of data for the study of Internet topology, routing, and performance. Special techiques are employed by tools like traceroute to infer router-level forward paths, since no direct method exists. Traceroute infers an IP path by sending out cleverly crafted probe packets and then analyzing the ICMP error responses. It is commonly believed that, while admittedly producing some of the best available maps, these techniques are susceptible to several types of inaccuracies. However, no one to date has quantified the magnitude of these inaccuracies. We study the frequency of one type of inaccuracy—third-party addresses. A third-party address is the address of a router interface that is not in the true forward path. These arise because RFC1812 [1] dictates that the source address of an ICMP response packet should be set to the address of the *outgoing* interface of the *response* rather than the interface on which the packet triggering the response was received. Nevertheless, traceroute simply considers the source address contained in an ICMP response packet to be the address of the forward path at the probed hop, which is not always correct.

Research on Internet topology occurs mostly at the level of autonomous systems (ASes) rather than at the level of IP addresses. BGP routing tables, such as those available at RouteViews [2] and RIPE [3], are one source of AS-level topology. Another valuable source is AS paths derived from traceroute IP paths. Traceroute IP paths can be easily converted to AS paths by matching addresses to prefixes and prefixes to origin ASes. Because of the importance of obtaining accurate AS paths from traceroute IP paths, we only consider and analyze a restricted type of third-party address—those belonging to ASes that are not a legitimate part of the forward path.

## II. Methodology

All addresses in a path, except the source and the destination, are *intermediate addresses*. An intermediate address that resolves to an AS that differs from the ASes of *both* adjacent addresses[1] in the same path is a *candidate third-party address*, or simply a *candidate address*. Intermediate addresses that do not meet this condition are not candidate addresses. Candidate addresses correspond to situations in which a third-party address can cause an incorrect AS path to be derived from an IP path.[2] This paper covers only candidate third-party addresses.

Our procedure for collecting and identifying candidate addresses is as follows:

1) Perform multiple traceroutes to a large number of addresses.
2) Eliminate all but stable traceroute paths.
3) Convert IP paths to AS paths.
4) Make an initial identification of candidate addresses.
5) Eliminate candidate addresses that have hostnames in the same domain as an adjacent address.
6) Eliminate candidate addresses that resolve to an AS that is under the same ownership as the AS of an adjacent address.

The following sections elaborate on these steps.

### A. Source of Traceroute Paths

CAIDA has deployed around two dozen skitter monitors worldwide [4]. These monitors determine the forward path to thousands of destinations with a technique similar to that used by `traceroute` [5]. Each monitor probes a predetermined set of addresses called the destination list. A monitor makes a complete pass through its destination list before starting again. One pass through a destination list is called a cycle, and the IP path and RTT collected for each destination is called a trace. By overlapping the probes to multiple destinations at a time, skitter can rapidly probe a large number of destinations; for

[1] The adjacent addresses of an intermediate address are those immediately preceding and following that address in the same path. The source and the destination can be adjacent addresses.

[2] It is theoretically possible for two or more third-party addresses to occur successively in a path, with all these addresses mapping to the same AS, but the chances seem so remote that we have ignored this possibility in our analysis.

example, many monitors can determine the full path to several hundred thousand destinations in less than a day.

All traceroute paths of this study were obtained from the following skitter monitors: a-root, k-peer, m-root, champagne, lhr, and sjc. The a-root monitor is located in Herndon, VA, in the network of Verisign. The k-peer monitor is in Amsterdam in the network of RIPE (AS3333) and lies near the Amsterdam Internet Exchange (AMS-IX). The m-root monitor is in Tokyo in the same network as the *m* root DNS server (AS7500) and lies near the NSPIXP of WIDE. The champagne monitor is in Urbana, IL, in the network of the University of Illinois at Urbana-Champaign. The lhr monitor is in London in the network of MFN/AboveNet. The sjc monitor is in San Jose, CA, in the network of MFN/AboveNet.

Two destination lists, covering two different cross sections of the Internet, were active during this study. The DNS list has 200K responding destinations[3] and consists mainly of hosts that have queried one of several instrumented DNS root servers. The IPv4 list has 80K responding destinations[4] and consists of a wide variety of hosts, including web servers, backbone routers, business desktops, and consumer dial-up/broadband desktops. The two lists have around 8K responding destinations in common. The DNS list was active on a-root, k-peer, and m-root, and the IPv4 list was active on champagne, lhr, and sjc.

Skitter trace attempts can fail for a variety of reasons, including packet loss, packet filtering, ICMP rate limiting, and transient routing loops. Because of the importance of intermediate hops in the study of third-party addresses, we only use skitter traces in which all intermediate hops responded to skitter. These are the *complete* traces. Completeness itself, however, does not guarantee that a deduced path represents a snapshot of a single forward path since paths are constructed piecemeal from a series of independent probes made over time (though a short period of time). It is possible for routing to change or fluctuate while a given forward path is being constructed. To avoid errors caused by mid-path routing changes, we distill complete traces to stable paths before performing our analysis. For the purposes of this study, we consider a destination to have a stable IP path if the traces to that destination in several consecutive cycles are complete and yield the same IP path. The persistence of stable IP paths over multiple traces ensures that they have not been distored by mid-path routing changes.[5] Third-party addresses are not routing anomalies or necessarily misconfigurations, and so the use of stable paths does not exclude them, although some instances may be overlooked.

## B. Conversion of IP Paths to AS Paths

The University of Oregon Route Views Project [2] provides ongoing snapshots of default-free routing tables collected from several dozen large ISPs. We use these snapshots, containing announced prefixes and their origin ASes, to convert traceroute IP paths to AS paths. The procedure works by matching addresses to prefixes and then prefixes to ASes. Each address is matched with the most-specific, or longest, announced prefix containing the address. After the AS of every address in a path is determined, the final AS path is obtained by collapsing runs of the same AS to a single AS.

There are several issues to note about the conversion of IP paths to AS paths. First, a small number of prefixes in a RouteViews snapshot are associated with multiple ASes or with an AS set [6]. Although they are not handled specially by our conversion scripts, these prefixes occur infrequently enough that their impact on our analysis should be small. Second, there are addresses in IP paths that lack a matching prefix [7]. Some belong to address blocks that are never announced globally, and others belong to prefixes that have been withdrawn. Yet others are special addresses, such as multicast and RFC1918 [8] private addresses, that do not belong to any AS.[6] Finally, exchange point addresses are surprisingly common in traceroute paths, though the percentage of paths containing them varies widely among skitter monitors [9]. For k-peer and m-root, located as they are near exchange points, nearly every path contains an exchange point address. A much smaller percentage of the paths of the remaining four monitors contain exchange point addresses, but these still number in the thousands. We use the list of exchange point prefixes compiled by Packet Clearing House, augmented with a few additional prefixes[7], to detect exchange point addresses [10]. In our analysis, we generally ignore any addresses that cannot be matched to a prefix for any of the above reasons, with the sole exception being the refinement of candidate third-party addresses by hostname matching.

## C. Refinement of the Set of Candidate Third-party Addresses

We refined the initial set of candidate addresses with two procedures, both involving the comparison of candidate addresses to adjacent addresses by some attribute. The first is a comparison by hostname. If both hostnames have generic top-level domains (namely, com, net, org, edu, gov, or mil), and the second-level domains are the same, then the two are considered equal, and the candidate address is eliminated from the set.[8] Note that the comparison fails if either hostname has a country-code top-level domain. We also performed further

[3] The actual DNS list has 330K destinations, but no useful paths are obtained for 130K that either never respond or almost never respond to skitter probes.

[4] 133K total destinations, of which 53K do not respond.

[5] Stable IP paths can still be subject to routing anomalies, but the problem of routing anomalies is orthogonal to the problem of third-party addresses, since the former concerns the difference between the actual and the intended while the latter concerns the difference between the actual and the observed.

[6] Nonetheless, there are ASes that announce prefixes for some of these special addresses.

[7] Thanks to Elena Silenok at CAIDA for the additional prefixes.

[8] The assumption that the second-level domain is sufficient to distinguish between two organizations does not always hold for the gov and mil TLDs which may have additional levels of hierarchy. However, for our dataset, this was sufficient for all comparisons involving gov, and the assumption was incorrect for only five comparisons involving mil.

comparisons based on a manually compiled list of equivalent domains.[9]

The second procedure checks for the presence of ASes under the same ownership [9]. Many organizations have more than one AS (and some have dozens) for diverse reasons, such as for convenience in implementing routing policy and for segregating different classes of traffic (e.g., academic vs. commercial). Businesses can also find themselves possessing multiple ASes after an acquisition or merger. We have previously compiled a list of these sets of ASes under the same ownership. If a candidate address and an adjacent address belong to two ASes that are under the same ownership, then the candidate address is eliminated from the set.

## III. Analysis

> *What I tell you three times is true.*
> *—Lewis Carroll, "The Hunting of the Snark"*

### A. Stable Paths and Candidate Addresses

We collected three consecutive cycles of traces on Jan 10-13, 2003 on all six skitter monitors. We determined the destinations that had complete traces in all three cycles and eliminated the traces of all other destinations. We then eliminated the traces of all destinations that had more than one IP path during the three cycles. What remained were the stable paths, one per stable destination. We then converted these stable IP paths to AS paths using RouteViews snapshots taken at about the middle of the time period spanned by the *last* cycle of each skitter monitor. Using a RouteViews snapshot centered on the last cycle, rather than over all three cycles, is justified since only the AS paths of the last (that is, most recent) cycle are needed. The first two cycles are important only in providing a basis for determining stable paths.

Table I provides an overview of the collected traces. We used the three most recent cycles available for each monitor on January 13th to derive the stable paths. The start time is the time of the first trace of the first cycle. The end time is the time of the last trace of the third cycle. Roughly half of the responding destinations of each monitor had stable paths, except those of `champagne`. Considering that traces were collected over such long time periods as 55 hours, in the case of `a-root`, it is somewhat surprising that so many paths have remained exactly the same. We have not examined the causes of path instability in detail, but besides routing anomalies, there is always a certain amount of churn in routing, and technologies like load balancing can cause apparent instability [11] [12] [13] [14]. The percentage of stable paths, if stated in terms of complete traces, is of course higher; for example, 70% of complete `a-root` traces were stable.

Table II provides statistics on the prefix coverage of stable paths. The BGP table time is the time of the RouteViews snapshot used for converting the stable paths of each monitor to AS paths. A BGP prefix is covered by a stable path if the prefix is the most-specific matching prefix for the destination address of the path. The percentage of prefixes covered gives a rough idea of how much of the globally visible networks is reached by the stable paths of each monitor. This is only a coarse measure since prefixes do not equate exactly with networks. A greater prefix coverage typically implies a greater diversity of paths. The union of all stable paths covers 46,467, or 37% of, announced prefixes. The number of prefixes covered by four or more monitors is 14,348, or 11% of prefixes. This suggests that stability is not a function solely of the prefix; one's vantage point is just as important.

Tables III and IV provide statistics on the frequency of candidate third-party addresses in stable paths. These tables present the counts from two different viewpoints. Table III shows the number of *unique* candidate addresses that appear in all stable paths and the percentage of intermediate addresses that are candidate addresses. In contrast, Table IV shows the number of *instances* (appearances) of candidate addresses in stable paths and the number of stable paths that have at least one candidate address. For example, if the address 1.2.3.4 appears in two paths, it will contribute one count to Table III and two counts to Table IV. Both forms of counts are needed: A count of unique addresses provides a sense of the number of distinct locations in the network at which the phenomenon of third-party addresses occurs, and a count of instances conveys a sense of the impact of third-party addresses in real-world paths.

In terms of percentage, few intermediate addresses of stable paths are candidate third-party addresses, but the percentage of stable paths that have candidates is not negligible. However, since third-party addresses always occur in the transit portion of paths, it is not surprising that some candidate addresses appear in more than one path. Those lying closer to the source or at central points in the network will be especially common in paths. In fact, only a small number of candidate addresses are typically responsible for the bulk of the appearances of all candidate addresses. For example, 92 and 146 candidate addresses account for 50% of all appearances in the paths of `a-root` and `lhr`, respectively. This feature is even more pronounced for `m-root`. One address alone accounts for 49% (21,361) of all appearances, and the top three account for 78%. The top three all occur very close to the source at hop two or three. If these top three are excluded, only 9,735 appearances remain, and 8,233 stable paths have candidate addresses. These numbers are more in line with those seen for the other monitors that probe the DNS list. There is less variance across the monitors in Table III, and in particular, `m-root` is no longer an exception.

### B. Distance from the Source

We analyze candidate third-party addresses in terms of their distance from the source. Suppose $S\ A\ B\ C\ D$ is an IP path. Then $S$ is the source address, and $D$ is the destination address. $A$ is an intermediate address at hop 1; $B$ is an intermediate address at hop 2; and so on. The path is 4 hops long. $S\ A$ is

[9] For example: (ameritech.com nvbell.net swbell.net pacbell.net pbi.net sbcglobal.net sbc.com prodigy.net) and (qwest.net uswest.net sni.net superlink.net tamerica.com tamerica.net).

TABLE I

OVERVIEW OF COLLECTED TRACES

| | a-root | k-peer | m-root | champagne | lhr | sjc |
|---|---|---|---|---|---|---|
| **start time (PST)** | 1/11 00:58 | 1/10 16:37 | 1/10 21:43 | 1/11 05:20 | 1/13 00:37 | 1/13 00:42 |
| **end time (PST)** | 1/13 07:43 | 1/12 21:27 | 1/13 00:43 | 1/13 09:02 | 1/13 13:54 | 1/13 12:56 |
| **hours/cycle** | 18 | 17 1/2 | 17 | 17 | 4 1/2 | 4 |
| **destination list** | DNS 200K | | | IPv4 80K | | |
| **complete traces** | 149,761 | 149,279 | 153,591 | 66,148 | 67,098 | 67,262 |
| **% total** | 75% | 75% | 78% | 80% | 82% | 82% |
| **stable traces** | 105,470 | 107,835 | 111,016 | 31,836 | 47,404 | 40,985 |
| **% total** | 53% | 55% | 56% | 39% | 58% | 50% |
| **% comp. traces** | 70% | 72% | 72% | 48% | 71% | 61% |

TABLE II

COVERAGE OF ANNOUNCED PREFIXES (127K) BY STABLE PATHS

| | a-root | k-peer | m-root | champagne | lhr | sjc |
|---|---|---|---|---|---|---|
| **BGP table time** | 1/12 20:00 | 1/13 12:00 | 1/12 20:00 | 1/12 20:00 | 1/13 12:00 | 1/13 12:00 |
| **covered prefixes** | 20,993 | 20,662 | 21,969 | 20,320 | 30,051 | 26,791 |
| **% announced** | 17% | 16% | 17% | 16% | 24% | 21% |

TABLE III

CANDIDATE THIRD-PARTY ADDRESSES IN STABLE PATHS (# UNIQUE ADDRESSES)

| | a-root | k-peer | m-root | champagne | lhr | sjc |
|---|---|---|---|---|---|---|
| **initial candidates** | 2,236 | 2,023 | 2,094 | 1,577 | 1,950 | 1,761 |
| **refinement by name** | −382 | −375 | −388 | −226 | −313 | −358 |
| **refinement by AS** | −237 | −241 | −224 | −206 | −223 | −201 |
| **final candidates** | 1,617 | 1,407 | 1,482 | 1,145 | 1,414 | 1,202 |
| **% initial** | 72% | 70% | 71% | 73% | 73% | 68% |
| **% intermediates** | 1.4% | 1.3% | 1.3% | 2.9% | 2.6% | 2.4% |
| **intermediate addrs** | 113,353 | 111,460 | 117,810 | 39,260 | 54,305 | 49,811 |

TABLE IV

CANDIDATE THIRD-PARTY ADDRESSES IN STABLE PATHS (# APPEARANCES)

| | a-root | k-peer | m-root | champagne | lhr | sjc |
|---|---|---|---|---|---|---|
| **initial candidates** | 14,516 | 31,164 | 49,047 | 6,660 | 6,868 | 6,199 |
| **refinement by name** | −3,319 | −22,132 | −4,365 | −1,802 | −1,795 | −1,744 |
| **refinement by AS** | −1,943 | −1,787 | −1,347 | −836 | −644 | −687 |
| **final candidates** | 9,254 | 7,245 | 43,335 | 4,022 | 4,429 | 3,768 |
| **% initial** | 64% | 23% | 88% | 60% | 64% | 61% |
| **paths with candidates** | 8,266 | 6,253 | 39,479 | 3,337 | 3,800 | 3,222 |
| **% stable paths** | 7.8% | 5.8% | 35.6% | 10.5% | 8.0% | 7.9% |

the path segment of length 1 (“1 hop long”); *S A B* is the path segment of length 2; and so on.

The variation in intermediate addresses is measured by the total number of unique intermediate addresses seen at each hop distance from the source. We compute these per-hop sets of intermediate addresses in a straightforward manner from the stable IP paths. For each hop distance $d$, we first find all addresses, except destination addresses, that appear at hop $d$, and then reduce the set to the unique addresses.[10] Fig. 2 shows the distribution of the unique intermediate addresses that occur at each hop distance in the stable paths. This distribution strongly resembles the path length distribution computed from the same paths, as shown in Fig. 1 [15]. Fig. 3 shows the distribution of the unique candidate third-pary addresses that

[10] We do *not* exclude exchange point addresses, RFC1918 special addresses, or addresses lacking matching prefixes from these sets. Intermediate addresses that appear at different hop distances in different paths are included in the sets of each hop distance in which they appear.

occur at each hop distance.

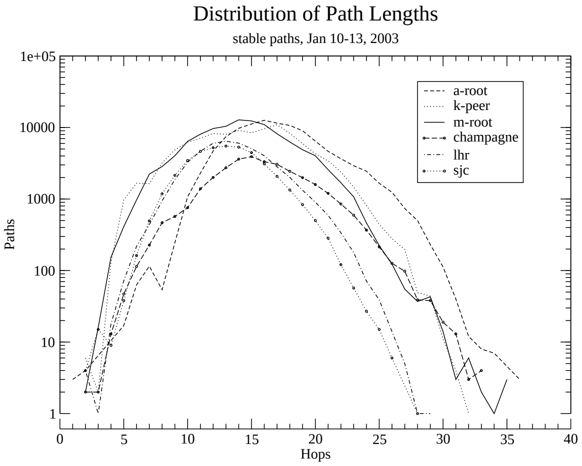


Fig. 1. Distribution of Path Lengths

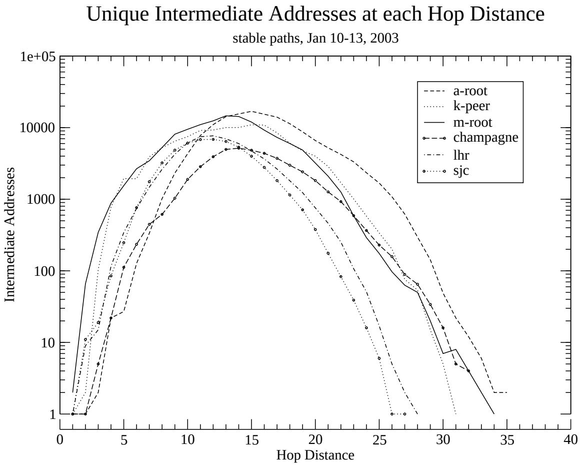


Fig. 2. Distribution of All Intermediate Addresses

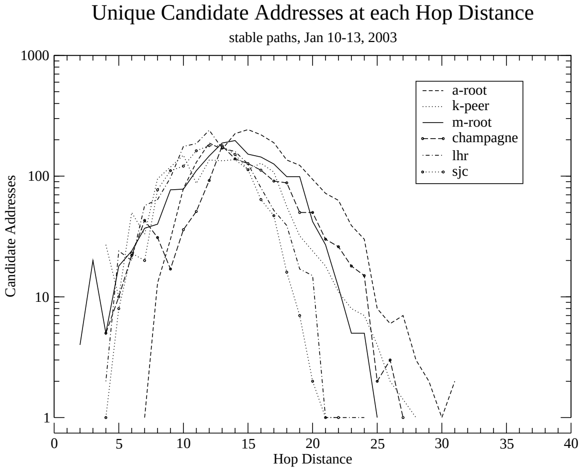


Fig. 3. Distribution of Candidate Third-party Addresses

Fig. 4 shows the ratio, given as a percentage, of the number of unique candidate addresses to the number of unique intermediate addresses at each hop. Apart from a few spikes and dips, the percentages are confined to a narrow band between 1% and 3%. Table V lists the values of these percentages at the 25th, 50th, 75th, and 95th percentiles, as well as the maximum percentage observed. The left group of columns under the heading "Addresses" provides the values for the present context. The 50th percentile of `a-root`, for example, says that at half of the hop distances (at which at least one candidate address appeared) only 1.4% or less of the per-hop unique intermediate addresses are candidate addresses.[11] The 25th and 75th percentiles differ by no more than 1% for nearly all monitors, implying that the percentages of at least half of the hops lie within a narrow band for all monitors except `champagne`. `Champagne` has several hops with above average percentages, and these cause all its percentiles to be higher than for the other monitors. It is interesting to note that the 50th and 75th percentiles naturally suggest two groups and that the monitors falling in each group all run the same destination list. The 50th and 75th percentiles of the monitors running the DNS list all lie near 1.5%, whereas those of the monitors running the IPv4 list are all higher and lie near 2.5%, except for the 75th percentile of `champagne`, which is somewhat higher at 3.8%. The fact that the IPv4 list covers more prefixes than the DNS list, as shown in Table II, explains why the monitors running the IPv4 list see more candidates.

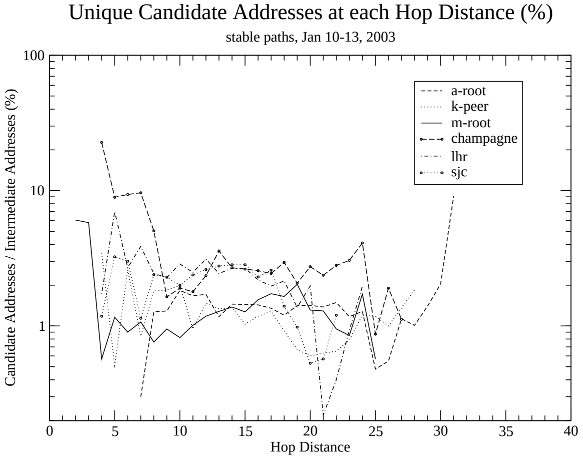


Fig. 4. Distribution of Candidate Third-party Addresses (%)

The wide variation observed in the lowest few and uppermost few hops in Fig. 4 is attributable to the steep decline in the number of intermediate addresses at those hop distances. For example, `a-root` has only 22 unique intermediate addresses at the 31st hop, of which 2 are candidate addresses. Similarly, `champagne` has only 22 addresses at the 4th hop and 5 candidate addresses.

Fig. 4 shows that there is no trend in the distribution of

[11] These percentiles would be somewhat less if the calculations included the hops at which there were intermediate addresses but no candidate addresses.

TABLE V

DISTRIBUTION OF THE RATIO OF CANDIDATE THIRD-PARTY ADDRESSES/PATH SEGMENTS TO PER-HOP INTERMEDIATE ADDRESSES/PATH SEGMENTS (ALL NUMBERS ARE PERCENTAGES)

| | **Addresses** | | | | | **Path Segments** | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | **percentiles** | | | | | **percentiles** | | | | |
| | **25th** | **50th** | **75th** | **95th** | **max** | **25th** | **50th** | **75th** | **95th** | **max** |
| **a-root** | 1.2 | 1.4 | 1.5 | 5.6 | 9.1 | 1.1 | 1.4 | 1.6 | 6.7 | 10.3 |
| **k-peer** | 0.8 | 1.2 | 1.6 | 3.0 | 3.5 | 0.7 | 1.0 | 1.5 | 2.9 | 3.5 |
| **m-root** | 0.9 | 1.2 | 1.6 | 5.9 | 6.1 | 0.8 | 1.3 | 2.0 | 5.9 | 6.1 |
| **champagne** | 2.0 | 2.7 | 3.8 | 16.2 | 22.7 | 1.9 | 2.5 | 4.5 | 16.1 | 22.7 |
| **lhr** | 1.8 | 2.3 | 2.7 | 5.4 | 7.0 | 1.7 | 2.1 | 2.6 | 4.9 | 6.3 |
| **sjc** | 1.2 | 2.3 | 2.8 | 3.2 | 3.2 | 0.8 | 1.9 | 2.6 | 3.2 | 3.2 |

candidate addresses towards any end or center of the plot. Also, as the variation in relative counts is confined within a factor of two in the central region where most addresses are found, it is reasonable to assume that the distribution of candidate addresses is independent from the hop distance from the source.

### C. Source IP Path Segments

We now report on the second type of per-hop variation that we investigated. The variation in path segments is measured by the total number of unique path segments seen at each hop distance from the source. We compute these per-hop sets of path segments in a straightforward manner from the stable IP paths. We first create a set of path segments from each path in the following way. For each path of length $p$, we create a path segment for each hop distance $1 \leq d < p$ by truncating the path at hop $d$.[12] Note that the path segment containing the destination address is excluded. We then reduce the set of all path segments from all paths to just the unique path segments.[13]

Fig. 5 shows the distribution of the unique path segments seen in the stable paths. This distribution and the distribution of the unique intermediate addresses shown in Fig. 2 have nearly the same shape, but the numbers of the path segment distribution are up to 2.6 times higher. Since each unique intermediate address at hop $d$ leads to a new unique path segment of length $d$, there must be at least as many unique path segments at each hop distance as there are unique intermediate addresses at the same hop distances. Load balancing, fluctuating routes, and other causes can increase the number of unique path segments. Fig. 6 shows the distribution of the unique candidate-address path segments—that is, those path segments that end in candidate third-party addresses. This distribution has nearly the same shape as the distribution of unique candidate addresses shown in Fig. 3. Considering these preceding similarities, it is not surprising that the distribution of Fig. 7, showing the ratio of the number of candidate-address path segments to the number of unique path segments at each length, closely resembles the analogous distribution of Fig. 4. Indeed, the percentages listed in Table V for path segments (see the columns under the heading "Path Segments" are nearly the same as those listed for addresses.

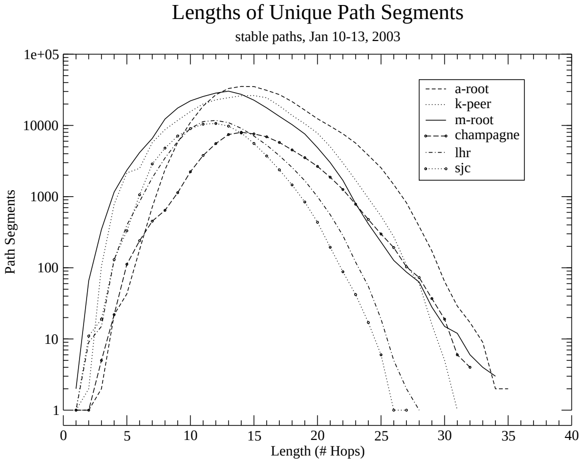


Fig. 5. Distribution of All Path Segments

### D. Distance from the Destination

We analyze candidate third-party addresses in terms of their distance from the destination. Fig. 8 shows the distribution of the unique intermediate addresses that occur at each hop distance *relative to the end* of the stable paths.[14] This distribution looks completely different from the analogous distribution taken relative to the beginning of paths (see Fig. 2). The curves in Fig. 8 are approximately exponential, suggesting that the IP graph implied by our set of stable paths is nearly a tree. Fig. 9 shows the distribution of the unique candidate addresses at hop distances relative to the end of paths. These curves are also approximately exponential for the first dozen hops or so from the end. Consequently, there

[12] We do *not* exclude exchange point addresses, RFC1918 special addresses, or addresses lacking matching prefixes from these path segments.

[13] Two path segments are equal if they have the same length and the same IP addresses at the corresponding positions.

[14] We compute this distribution in the following manner. Working through all stable paths, we first compute the pair (*A*, *D*) for each appearance of a candidate address *A* at a hop distance *D* from the end of the path. Then we reduce the set to the unique pairs. Finally, ignoring the *A* component of each pair, we plot the distribution of the *D* values.

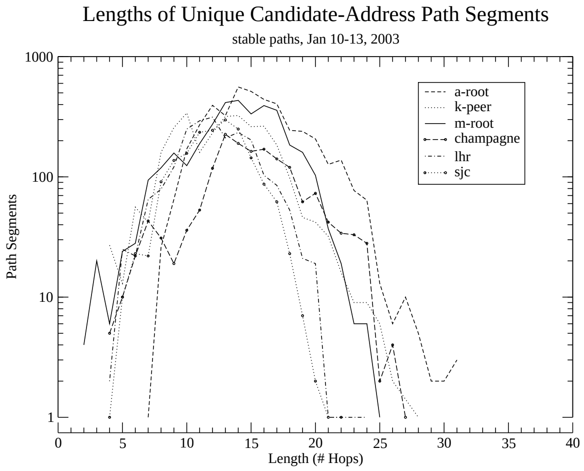


Fig. 6. Distribution of Candidate-Address Path Segments

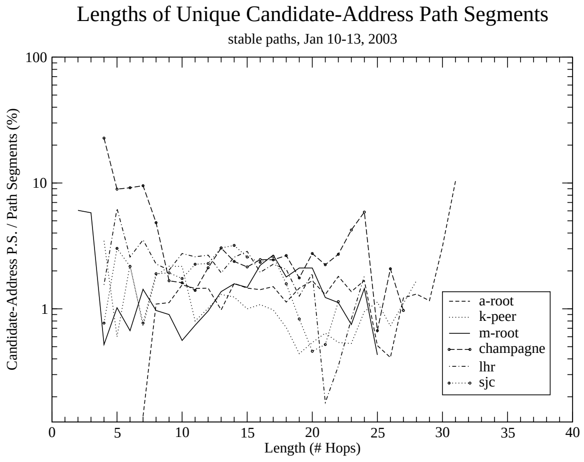


Fig. 7. Distribution of Candidate-Address Path Segments (%)

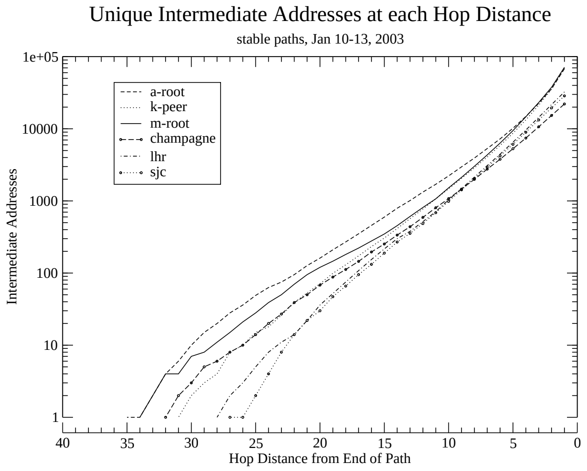


Fig. 8. Distribution of All Intermediate Addresses Relative to Path End

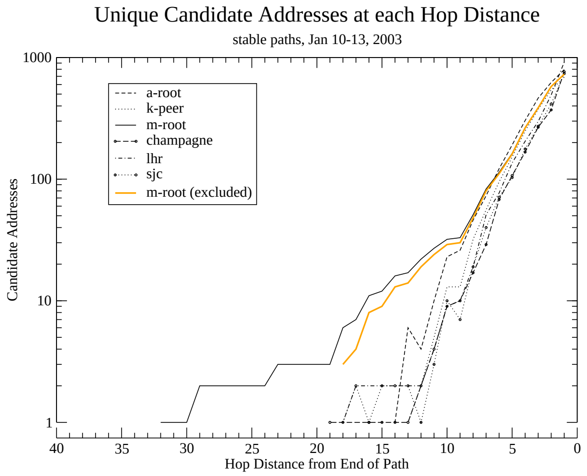


Fig. 9. Distribution of Candidate Third-party Addresses Relative to Path End

are about as many candidate addresses in the first two hops from the end as there are in the remaining hops. This is true across all monitors. Furthermore, three quarters of all candidate addresses appear within four hops of the end for `a-root`, `k-peer`, and `m-root`, and within three hops for the remaining monitors. Hence, the bulk of the candidate addresses appear near the end of paths, or more importantly, at the edge of the network near the destination. The extended tail of `m-root` is attributable to candidate addresses appearing frequently at the beginning of paths. Recall that just three of `m-root`'s candidate addresses lying at the second and third hops are responsible for a large fraction of the total number of appearances of candidate addresses. This fact in combination with the Gaussian-like distribution of path lengths (see Fig. 1) would produce an extended tail like that seen in the distribution. Indeed, the extended tail disappears once the top 3 most frequently occurring candidate addresses are excluded, as shown by the curve labeled 'm-root (excluded)' in Fig. 9.

Fig. 10 shows the ratio, given as a percentage, of the number of unique candidate addresses to the number of unique intermediate addresses at each hop distance relative to the end of paths. All distributions, except `m-root`'s, are in remarkably good agreement. This was not the case for the analogous distributions made relative to the beginning of paths.

The close agreement of the distributions may suggest that the monitors are sharing a large fraction of the candidate addresses. After all, since the monitors are probing the same (two sets of) destinations, the paths from the monitors must coverge at some point to a small number of alternate paths to the destination. This convergence can happen near the end of paths, exactly the area where most candidate addresses are observed. However, this does not appear to be the case. Table VI shows the distribution of candidate addresses by the number of monitors that see them. Of the total unique candidate addresses seen by all monitors, 59.7% are seen by only one monitor, and 97.0% are seen by three or fewer. There is some overlap but not a great deal. Thus, this observed tendency for candidate addresses to appear near the end of paths is indicative of a genuine phenomenon.

Table VII lists the number of the ASes and prefixes covered by candidate third-party addresses. These were determined by

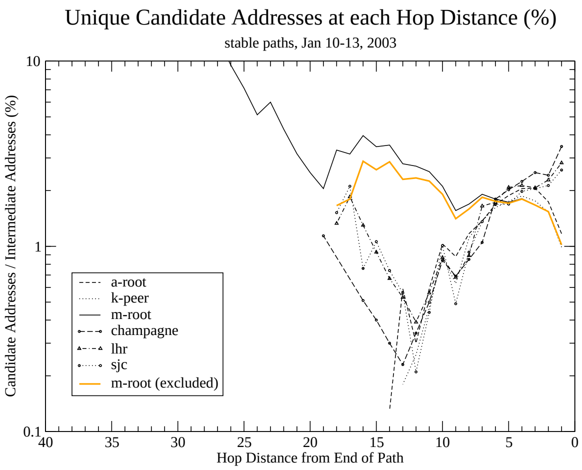


Fig. 10. Distribution of Candidate Third-party Addresses Relative to Path End (%)

TABLE VI
COMMONALITY OF UNIQUE CANDIDATE THIRD-PARTY ADDRESSES ACROSS MONITORS

| # monitors | # addrs | % total | cum.% |
|---|---|---|---|
| **1** | 3,062 | 59.7% | 59.7% |
| **2** | 1,231 | 24.0% | 83.7% |
| **3** | 679 | 13.2% | 97.0% |
| **4** | 82 | 1.6% | 98.6% |
| **5** | 54 | 1.1% | 99.6% |
| **6** | 18 | 0.4% | 100% |

using the RouteViews snapshots described in Table II to match addresses to prefixes. Upon comparing these numbers with the counts of final candidates in Table III, the diversity of ASes and prefixes immediately stands out. There are on average about two candidate addresses per prefix or AS. Of course, these numbers are small compared to the total number of allocated ASes and the total number of announced prefixes, but they do show that candidate third-party addresses are not isolated to just a handful of locations (ASes).

TABLE VII
COVERAGE OF ASES AND PREFIXES BY CANDIDATE THIRD-PARTY ADDRESSES

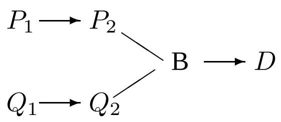
| | prefixes | ASes |
|---|---|---|
| **a-root** | 1,025 | 797 |
| **k-peer** | 942 | 746 |
| **m-root** | 987 | 773 |
| **champagne** | 727 | 574 |
| **lhr** | 866 | 675 |
| **sjc** | 767 | 602 |

### *E. Multihoming*

Fig. 9 and 10 shows that the probability of having a candidate address at a given hop distance increases by an order of magnitude over the last ten hops towards the destination. This strongly suggests that most instances of third-party addresses are found in multihomed stub networks. This may be a simple reflection of the inherent route asymmetry of BGP-based routing.

In August 2002, we performed the analysis described above on the traces of the `sjc` monitor and manually inspected the resulting candidate third-party addresses. We engaged in a variety of (manual) detective work to determine which of these candidate addresses were actual third-party addresses. For example, we obtained traceroute paths to the candidate addresses from a number of traceroute servers (located in different networks) to identify router aliases and to better understand the topology near the candidates. We also made BGP queries to traceroute servers and route servers to obtain the non-best-path routes; this data was also gleaned from RouteViews snapshots. We further made use of `whois` databases and the RADB routing registry to obtain data on prefixes and routes and to better understand the relationships of observed ASes. Some other techniques were also employed.

All in all, we identified about two hundred addresses that seemed likely to be third-party addresses although conclusive evidence was generally lacking. Of these, however, we had strong evidence for about 50 addresses, which appeared to be caused by multihoming.[15] The following diagram illustrates the situation we encountered:



Here, $D$ is the destination, and B its border router. $P$ and $Q$ are the providers of $D$. $P_2$ is the interface of B that faces $P$ and is numbered from an address block owned by $P$; and $Q_2$ is an interface of B that is in an analogous relationship with $Q$. The IP paths one would expect are $P_1$ $P_2$ $D$ and $Q_1$ $Q_2$ $D$, producing the AS paths $P$ $D$ and $Q$ $D$. However, third-party addresses cause us to see $P_1$ $Q_2$ $D$ or $Q_1$ $P_2$ $D$, producing the AS paths P Q D or Q P D, respectively.[16]

In order to confirm that we were observing third-party addresses in the context of multihoming, we made email inquiries to the organizations that owned these addresses. We sent 38 emails, mainly to regional ISPs, and received 11 replies. Ten respondents confirmed that we were seeing third-party addresses and that multihoming of some form was involved. A GRE tunnel [16] was the ultimate cause in two of these cases. In one case, a GRE tunnel was in use to ease the renumbering of hosts while transitioning between providers. In another, a GRE tunnel was in use as a basic component of the organization's unusual network design.

The final respondent described the following situation. Back when they owned a particular address block, they had numbered one of their interfaces from it. They relinquished the address block at some point thereafter, but the interface, being little maintained, retained its address. The address block itself was reallocated to another organization, and they numbered their machines from it. Hence, at the time of our measure-

[15] We do not mean to imply that multihoming is the only significant cause of third-party addresses. Multihoming may simply have been more amendable to investigation with our particular techniques.

[16] Our use of stable paths shields us from a mid-path routing change from one provider to another while performing a traceroute.

ments, an address was allocated to two separate interfaces on two entirely different networks.[17] We saw the address at the old location, but the address mapped to the AS of the new location, causing it to appear to be a third-party address.

## IV. Previous work

**Data collections.** BGP was defined by Yakov Rekhter and Tony Li in 1995 [19]. Analyzing the BGP routing table as a source of global Internet data was pioneered in mid-90s by Erik-Jan Bos at SURFnet, Tony Bates (then at Cisco) [20], and Geoff Huston [21] [22]. Huston's work analyzes trends in Internet connectivity, including the dynamics of AS degrees, AS path counts, and AS path lengths and presents them as a daily report. David Meyer's RouteViews project [2] has collected BGP tables at the University of Oregon since 1997. RIPE has collected both BGP updates and BGP tables since 1999. The importance of these pioneering work and the public availability of these data collections can hardly be overestimated.

CAIDA's Skitter project has collected traceroute data since 1998 [4]. AS connectivity is derived from traceroutes and published periodically as an AS core poster [23]. Construction of IP-level Internet maps was also done by Cheswick and Burch [24] who made their traceroutes publicly available in 1999-2001.

**Analysis.** Internet path properties were originally studied by Vern Paxson in his Ph.D. thesis [14]. This and later work by Paxson and colleagues focus considerable attention on handling noise in data and upon the sources and consequences of errors in analysis. They also discuss routing pathologies. Their study of stationarity of Internet path properties [12] concluded that many IP paths (70-80% of their data) remained unchanged for longer than one day, suggesting that our analysis of skitter cycles is not significantly affected by routing dynamics. The paper [12] is also one of the first to make extensive use of traceroute servers.

Routing stability is heavily dependent on the richness of AS connectivity. Studies by Labowitz, Ahuja *et al.* [25] and Griffin [26] have confirmed that a rich mesh of connections among ASes can result in an extreme (superexponential in the number of nodes in between) amount of routing traffic caused by a single update. These results underscore the need for complete and accurate sources of AS-level topology so that modeling of routing stability could be performed on realistic data.

Broido and Claffy [27] [28] used systems of AS paths observed from multiple vantage points to classify prefixes into groups called *policy atoms*. Informally, each atom consists of prefixes that are commonly routed in a large portion of the Internet. The authors show that the use of atoms can reduce the size of routing tables by half. The work of Yehuda Afek *et al.* [29] establishes the presence of atoms in BGP updates.

Properties and limitations of AS-level Internet maps have been discussed in several papers. Faloutsos *et al.* [30] presented evidence that the node degree distribution for a BGP AS graph is close to a power function. AS hop distances are studied in Huston's report [21] and in Maennel and Feldmann's work in the modeling of BGP update traffic [31]. Average AS path lengths and peering richness (entropy of outbound link use by routed prefixes for each AS) remained relatively stable in 1999-2001, changes at individual ASes notwithstanding [11].

Lixin Gao [32] classified links in the BGP AS graph as customer, provider, peer and sibling connections. Tangmunarankit *et al.* [33] used this classification to assess hop count differences between router-level and AS-level shortest paths versus shortest paths constrained by policies as inferred in [32]. A related paper [34] evaluates topology generators on the basis of reachability functions, resilience and distortion (tree-likeness) of the resulting graphs.

Research on global IP-level Internet mapping was initiated in [35] and [36]. Router-level maps were collected by the Mercator project [37], which was the first serious attempt at identifying all IP addresses assigned to the interfaces of a single router. To that end, the alias probe heuristic of [38] was used. This router identification technique was also implemented in `iffinder`, a tool developed by Ken Keys at CAIDA [39].

Chang *et al.* tackled the problem of identifying which AS owns a router [40], introducing heuristics that fill in ASes for non-replying hops in the trace. Their paper also discusses third-party addresses. Our analysis suggests that these are a negligible source of AS path incongruity.

Ratul Mahajan *et al.* [41] studied BGP misconfigurations. Third-party addresses are not necessarily the result of misconfigurations. Nevertheless, in our study, we attempted to filter out misconfigurations and other ambiguities so that our set of candidate third-party addresses would not be inflated by that type of noise.

Neil Spring *et al.* [42] discuss how to build a detailed map of an individual ISP's topology using their tool Rocketfuel which employs some 800 traceroute servers to gather as much topological information possible with the minimal amount of measurement. They use a rich set of heuristics for identifying same-router IP addresses, including DNS names, TTLs, IP ID field, and instances of rate limiting triggered by earlier probes. Their detailed analysis results in the discovery of seven times more links than skitter [23] in selected networks of several ISPs. They do not attempt to obtain an AS-level map.

Lisa Amini *et al.* [43] compared properties of traceroute and BGP AS paths. They found that the IP stacks of AIX, FreeBSD, Windows 2000, and the Cisco 7500 set the source address of ICMP reply messages to the interface on which the packet triggering the response arrived. They also found that only Linux sets it to the interface on which the reply is sent, as required by specification [1]. This situation must decrease the number of cases in which traceroutes suffer from ambiguities introduced by third-party addresses. The authors also found a

[17] Similarly, anycast addresses are configured in networks of multiple providers (cf. [17] [18]) even though they are announced in BGP with the same AS.

significant number of cases in which the BGP AS path differed from the traceroute AS path in only one hop, that being a hop of an exchange point. We skip exchange point prefixes in traceroutes when searching for candidate addresses; otherwise the number of candidates would have been much higher.

## V. Conclusions

Traceroute IP paths are a rich source of data for the study of Internet topology, routing, and performance. However, third-party addresses are a potential pitfall in using AS-level topology derived from traceroute IP paths. Based on an examination of thousands of traceroute paths from six locations worldwide and the application of several metrics, we find that the situations in which third-party addresses can occur to be relatively uncommon, with multihoming being an apparent cause for most actual instances of third-party addresses. Our analysis shows that third-party addresses cannot be a significant cause of AS map distortion.

## Acknowledgements

Support for this work is provided by the Defense Advanced Research Project Agency (DARPA) NMS (N66001-01-1-8909) program and DISA's National Communiations System Organization. CAIDA is a collaborative organization supporting cooperative efforts among the commercial, government and research communities aimed at promoting a scalable, robust Internet infrastructure. CAIDA is based at the University of California's San Diego Supercomputer Center (SDSC). www.caida.org.

## References

[1] R.Baker, “Requirements for IP Version 4 Routers, RFC1812,” Jun 1995.

[2] David Meyer, “University of Oregon Route Views Project,” `http://www.routeviews.org`.

[3] RIPE NCC, “Routing Information Service,” `http://www.ripe.net/ris`.

[4] Daniel McRobb and kc claffy, “Skitter,” CAIDA, 1998. `http://www.caida.org/tools/measurement/skitter/`.

[5] Van Jacobson, “traceroute,” `ftp://ftp.ee.lbl.gov/traceroute.tar.Z`.

[6] Andre Broido, “Multiorigin prefixes in backbone BGP tables,” `http://www.caida.org/~broido/bgp/multiorigin.html`.

[7] E. Nemeth A. Broido and kc claffy, “Fringe Address Spaces,” 7 pp, in preparation.

[8] Y. Rekhter, B. Moskowitz, D. Karrenberg, , G. J. de Groot, and E. Lear, “Address Allocation for Private Internets, RFC1918,” Feb 1996.

[9] Young Hyun, Andre Broido, and k claffy, “Traceroute and BGP AS Path Incongruities,” in preparation.

[10] Sean McCreary and Bill Woodcock, “Exchange Points,” `http://www.pch.net/resources/data/exchange-points`.

[11] A. Broido, E. Nemeth, and kc claffy, “Internet expansion, refinement and churn,” European Transactions on Telecommunications, 13, No.1, Jan-Feb 2002, 33-51.

[12] Y. Zhang, V.Paxson, and S.Shenker, “The Stationarity of Internet Path Properties: Routing, Loss and Throughput,” in *ACIRI Technical Report*, May 2000.

[13] Y. Zhang, N. Duffield, V. Paxson, and S. Shenker, “On the Constancy of Internet Path Properties,” in *Proc. ACM SIGCOMM Internet Measurement Workshop (IMW'2001), San Francisco, California, USA, November 2001*.

[14] Vern Paxson, “End-to-End Routing Behavior in the Internet,” in *IEEE/ACM Transactions on Networking*, Oct 1997, vol. 5, pp. 601–615.

[15] Andre Broido and k claffy, “Internet Topology: connectivity of IP graphs,” in *Scalability and Traffic Control in IP Networks*, Aug 2001, Proceedings of SPIE, vol.4526.

[16] D. Farinacci, T. Li, S. Hanks, D. Meyer, and P. Traina, “Generic Routing Encapsulation (GRE), RFC 2784,” Mar 2000.

[17] “AS112 Project Home Page,” `http://www.as112.net`.

[18] A. Broido, E. Nemeth, and k. claffy, “Spectroscopy of DNS Update Traffic,” in *Proceedings of ACM SIGMETRICS 2003, San Diego, CA*, June 2003.

[19] Y. Rekhter and T. Li., “A Border Gateway Protocol 4 (BGP-4), RFC1771,” Mar 1995.

[20] Tony Bates, “CIDR Report,” `http://www.cidr-report.org`.

[21] Geoff Huston, “BGP Table Data (daily reports),” `http://bgp.potaroo.net`.

[22] Geoff Huston, “Analyzing the Internet's BGP Routing Table,” *The Internet Protocol Journal*, vol. 4, Mar 2001, `http://www.telstra.net/gih/papers/ipj/4-1-bgp.pdf`.

[23] B. Huffaker, A. Broido, k. claffy, M. Fomenkov, K. Keys, Y.Hyun, and D. Moore, “Skitter AS Internet Graph,” Apr 2002, `http://www.caida.org/analysis/topology/as_core_network/`.

[24] H.Burch B.Cheswick, “Internet Mapping Project,” `http://cm.bell-labs.com/who/ches/map`.

[25] Craig Labovitz, Abha Ahuja, Roger Wattenhofer, and Srinivasan Venkatachary, “The Impact of Internet Policy and Topology on Delayed Routing Convergence,” Apr 2001, Proceedings of INFOCOM 2001, Anchorage, AK.

[26] Tim Griffin, “What is the sound of one route flapping?,” `http://www.cs.dartmouth.edu:80/%7Emili/workshop2002/slides/griffin_dart%mouth_20020723.pdf`.

[27] A. Broido and k claffy, “Complexity of global routing policies,” `http://www.caida.org/outreach/papers/2001/CGR/`.

[28] A. Broido and k claffy, “Analysis of Route Views BGP data: policy atoms,” Proceedings of network-related data management (NRDM) workshop, Santa Barbara, May 2001.

[29] Yehuda Afek, Omer Ben-Shalom, and Anat Bremler-Barr, “On the structure and application of BGP policy Atoms,” in *Proceedings of the IMW workshop 2002. Marseille, France*, Nov 2002.

[30] Michalis Faloutsos, Petros Faloutsos, and Christos Faloutsos, “On power-law relationships of the Internet topology,” Aug 1999, Proceedings of ACM SIGCOMM 1999, Cambridge, MA.

[31] Olaf Maennel and Anja Feldmann, “Realistic BGP Traffic for Test Labs,” in *ACM SIGCOMM 2002, Pittsburg, PA*, Aug 2002.

[32] Lixin Gao, “Inferring Autonomous System Relationships in the Internet,” Nov 2000, Proceedings of IEEE Global Internet Symposium, San Francisco, CA.

[33] H.Tangmunarankit, R.Govindan, and S.Shenker, “Internet path inflation due to policy routing,” in *Scalability and Traffic Control in IP Networks*, Aug 2001, Proceedings of SPIE, vol.4526.

[34] H.Tangmunarunkit, R.Govindan, S.Jamin, S.Shenker, and W.Willinger, “Network Topology Generators: Degree based vs. Structural,” in *Proceedings of ACM SIGCOMM 2002*.

[35] k. claffy, T. Monk, and D. McRobb, “Internet Tomography,” Nature Magazine, Web Matters. `http://helix.nature.com/webmatters/tomog/tomog.html`.

[36] Hal Burch and Bill Cheswick, “Mapping the Internet,” in *IEEE Computer*, Apr 1999, vol. 32.

[37] Ramesh Govindan and Hongsuda Tangmunarunkit, “Heuristics for Internet Map Discovery,” in *Proceedings of IEEE Infocom 2000, Tel Aviv, Israel*, 2000.

[38] J.-J. Pansiot and D. Grad, “On Routes and Multicast Trees in the Internet,” in *ACM SIGCOMM Computer Communication Review*, Jan 1998, vol. 28.

[39] Ken Keys, “`iffinder`,” Feb 2001, `http://www.caida.org/tools/measurement/iffinder/`.

[40] H.Chang, S.Jamin, and W.Willinger, “Inferring AS-level Internet topology from router-level path traces,” in *Scalability and traffic control in IP networks*, Aug 2001, Proceeding of SPIE, vol.4526.

[41] R.Mahajan, D.Wetherall, and T.Anderson, “Understanding BGP Misconfiguration,” in *ACM SIGCOMM 2002, Pittsburg, PA*, Aug 2002.

[42] N. Spring, R. Mahajan, and D. Wetherall., “Measuring ISP Topologies with Rocketfuel,” in *ACM SIGCOMM 2002, Pittsburg, PA*, Aug 2002.

[43] Lisa D. Amini, Anees Shaikh, and Henning G. Schulzrinne, “Issues with inferring Internet topological attributes,” in *Internet Performance and Control of Network Systems III. Proceedings of SPIE, vol.4865, Boston, MA*, Aug 2002.