A Framework to Quantify the Pitfalls of Using Traceroute in AS-Level Topology Measurement

Yu Zhang, Ricardo Oliveira, Yangyang Wang, Shen Su, Baobao Zhang, Jun Bi, Hongli Zhang, and Lixia Zhang

***Abstract*—Although traceroute has the potential to discover AS links that are invisible to existing BGP monitors, it is well known that the common approach for mapping router IP addresses to AS numbers based on BGP routing tables is highly error-prone. We develop a systematic framework to quantify the potential errors of traceroute measurement in AS-level topology inference. In comparing traceroute-derived AS paths with BGP AS paths, we take a novel approach to identifying mismatched path segments and then inferring the causes of these mismatches through a set of tests. Our results show that about 60% of mismatches are due to routers using IP addresses belonging to peering neighbors. This result helps settle a debate in previous works regarding the major cause of errors in traceroute measurement. With the approximate ground truth of the ASes with BGP monitors inside, we identify the inaccuracy of publicly available traceroute-derived topology datasets and find that between 8% and 42% of AS adjacencies on the monitored ASes are false. With a new method to characterize AS links, we show that the derived (false) links between Tier-1/large ISPs and their customers' customers appear more frequently than real links do.**

***Index Terms*—AS topology measurement, traceroute, BGP**

I. Introduction

THE INTERNET is a vast distributed system formed by a myriad of networks called Autonomous Systems (ASes) that exchange routing information using the Border Gateway Protocol (BGP). There have been two basic approaches to measure AS-level connectivity: passive measurement through collecting BGP routing updates, and active measurement using traceroute. In BGP-based measurements, AS adjacencies can be directly extracted from the AS_PATH attribute in BGP updates collected from BGP monitors by Routeviews [1] and RIPE-RIS [2]. But because of policy filters and best path selection, each BGP monitor only provides a limited partial view of the topology [3]–[7]. Most monitors in traceroute measurement projects, such as CAIDA's Ark [8] and DIMES [9], are placed in different ASes than BGP monitors, and thus ideally they can complement the topology inferred from existing BGP monitors [10]. It is also easier to deploy a traceroute monitor on an end-user host than to obtain a new BGP feed from a network operator [9], [11].

However converting router IP addresses in traceroute paths to AS numbers, termed IP2AS mapping, is a difficult problem and remains an open challenge. The most widely adopted method for IP2AS mapping is to look up the origin AS of each IP address from BGP routing tables using the longest prefix matching. Unfortunately this straightforward approach has been known for a long time to be error-prone and generate potentially false AS links [12]–[15]. Although a number of efforts have gone to study the accuracy of traceroute measurement and articulated possible causes for errors, few available traceroute-derived topology datasets have adopted improved methods rather than the naive one. Consequently the following question faces all efforts using the traceroute-derived topology data: *How much impact do the pitfalls of traceroute measurement have on the accuracy of AS-level topology inference?*

Unfortunately, previous works cannot provide an accurate answer to this question on account of the following three reasons. First of all, it remains a debate regarding what is the major cause of errors in converting router paths to AS paths. One opinion states that the major cause is IP addresses assigned from neighbor ASes to enable the point-to-point connection [12], while another is that the most of false links are attributable to the existence of Internet eXchanging Points (IXPs), siblings or Multiple Origin ASes (MOAS) prefixes [14], [15]. Second, previous works did not pin down individual points along AS paths where false AS *links* may be created; most of previous works focused on the accuracy of AS paths, rather than AS links. The typical metric to quantify the impact of a given cause has been the number of incorrect AS paths derived due to this cause. We find that this metric introduces biases into the quantification of the impact of causes on topology inference. Lastly, the accuracy of traceroute-derived topology data has not been fully investigated; previous works have offered few clues on how many AS links in publicly available datasets may be false.

In this paper we aim to quantify the impact of pitfalls of traceroute measurement on topology inference, specifically, *to identify false AS links and find out how they are created*. Our major contribution in methodologies is the development of a systematic framework towards this goal. As illustrated in Figure 1, the framework is composed of four steps: (1) constructing topologies through collecting millions of pairs of traceroute paths and BGP paths from the same AS to the same destination, (2) decomposing mismatched path pairs into *mismatch fragments*, (3) disassembling each mismatch fragment into several one-to-one replacements of ASes, (4) associating

Manuscript received 15 December 2010; revised 2 June 2011. This work was partially supported by the National Basic Research Program of China (973 Program) under grant No. 2005CB321806, the US National Science Foundation under grant No. CNS-0551736, the National Science Foundation of China under grant No. 61073172, the Specialized Research Fund for the Doctoral Program of Higher Education of China under grant No. 200800030034.

Y. Zhang, S. Su, and H. Zhang are with the Harbin Institute of Technology, Harbin, China (e-mail: {yuzhang, zhanghongli}@hit.edu.cn, susheng@pact518.hit.edu.cn).

R. Oliveira and L.Zhang are with the University of California, Los Angeles (e-mail: {rveloso, lixia}@cs.ucla.edu).

Y. Wang, B. Zhang and J. Bi are with Tsinghua University, Beijing, China (e-mail: wangyy-06@mails.tsinghua.edu.cn, zbb@netarchlab.tsinghua.edu.cn, junbi@tsinghua.edu.cn).

Digital Object Identifier 10.1109/JSAC.2011.111007.

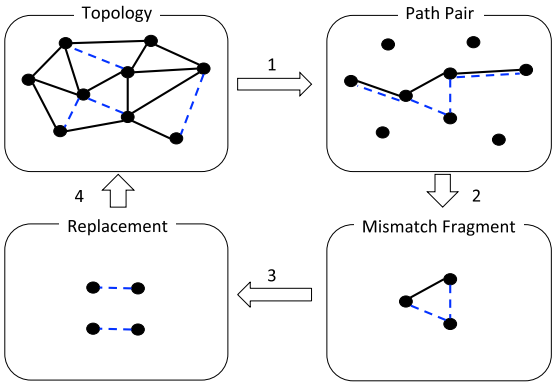


Fig. 1. A methodological framework for the quantification of pitfalls in traceroute measurement. Dash lines represent traceroute-derived AS links.

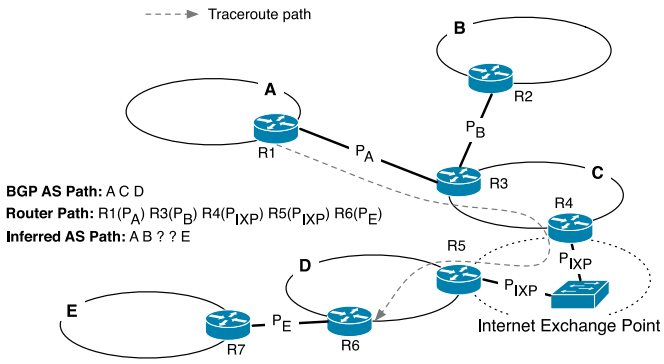


Fig. 2. Illustration of mismatches between BGP path and traceroute path. Performing traceroute from router R_1 to router R_6.

each false link on the topology with the replacement where the false link is created. The second and third steps are the core of our framework:

- We develop a novel method based on the longest common subsequence problem to detect *mismatch fragments* in each path pair systematically. This method allows us to pinpoint multiple mismatches in a single AS path pair, to identify each mismatch point shared by multiple path pairs, and to have local context for inferring the cause of each mismatch. This also gives a possibility of next conversion.
- We convert mismatch cause inferences into AS replacement explanations with a sequence of tests on selected AS pair candidates. This decouples procedures from evidences in cause inference so that we can put the weight-bearing point of cause inference on the datasets which are collected with the state of the art methods more comprehensively than before.

Another goal of this paper is to evaluate the inaccuracy of publicly available traceroute-derived AS topologies. In addition to comparing traceroute-derived AS links with the approximate ground truth of BGP monitored ASes, we also develop a new method for AS link characterization, named *neighbor-to-neighbor pattern*.

With these new methodologies, we harvest new findings:

- Clarifying the argument on the major cause: our results show that about 60% of mismatches occur because of border routers in one AS using IP addresses borrowed from neighbor ASes to enable point-to-point connections, while around 15% are due to the existence of IXPs, siblings or MOAS prefixes.
- In three publicly available, traceroute derived topology datasets, about 8∼42% of AS links on the BGP monitored ASes may be false. Those false links trend to connect Tier-1 and large ISPs with their customers' customers more likely than real links.

This paper is an extend version of our earlier work [16]. In this version, we refine on our framework, provide more details of our methodology and conduct new analyses of mismatch fragments, IP2AS mapping tables and false links.

In the rest of this paper, we first give background information in Section II and briefly review the related works in Section III. We present dataset collection, mismatch detection, cause inference and inaccuracy analysis in turn from Section IV to VII. After discussing limitations of this work in Section VIII we conclude in Section IX.

II. BGP VS. TRACEROUTE

In this section we provide the background information on traceroute- and BGP-based AS-level topology measurements. With an example, we demonstrate a few typical causes of false AS links when using BGP routing tables for IP2AS mapping in traceroute measurement. We then present the taxonomy of possible pitfalls. We also argue that it makes sense to detect errors in a given traceroute path with its BGP counterpart as the reference.

A. BGP VS. Traceroute in An Example

The BGP lists a series of ASes that propagate a given route to a prefix using the AS_PATH attribute, *e.g.* a BGP update shows AS_PATH is $[ABC]$ for a destination prefix P. This *control path* means that AS C announces P, and the traffic directed to P will be forwarded to AS B by AS A and then to C by B.

Different from the control path, the *data path* reflects the actual sequence of networks that traffic follows to reach a given destination. Each traceroute instance returns a series of IP addresses of interfaces on the routers along the data path, and then the AS path can be obtained with IP2AS mapping based on BGP routing tables. Depending on the implementation of ICMP message generation, a router may reply to traceroute probe using any of its multiple interfaces' IP addresses, which may be the incoming interface receiving the probe packet or the outgoing interface sending the response packet [12], [13].

Figure 2 shows an example of the conversion of a list of IP addresses returned by traceroute to an AS path. P_A refers to the address space which belongs to AS A and is used to number the interfaces through which routers R_1 and R_3 are connected. Assume an IXP has reserved a prefix P_{IXP} to number the interfaces of the participating routers. IXPs typically interconnect participants through a layer-2 switching fabric. Suppose a traceroute instance is launched through router R_1 in AS A and the final destination is router R_6 in D.

$R_x(P_y)$ denotes an IP address of an interface on router R_x that is contained in prefix P_y belonging to AS Y.

The inferred AS path in this case differs significantly from the BGP AS path for several reasons. First, routers R_3 and R_6 are replying to the traceroute probe using an interface address that belongs to their other neighbor ASes. Looking up $R_3(P_B)$ in routing tables leads to a false illusion that the path from A to R_6 is going through B, whereas in reality the path is going through C. Second, routers participating in the IXP are replying with an IP address in P_{IXP} reserved for the IXP. Because P_{IXP} is not announced in BGP, there are gaps ('?') in the inferred AS path. Furthermore, R_6 returns an IP address which belongs to E, leading to a false extra AS instead of D.

B. Taxonomy of Mismatch Causes

Besides the above illustration, there are other situations where two paths differ, either the data path is not completely aligned with the control path, or IP2AS mapping pitfalls occur in converting IP paths to AS paths. Here we try to present a taxonomy of reasons why traceroute AS paths may differ from their BGP counterparts, that is mainly credited to the previous efforts [12], [15]. Note that the ICMP outgoing interface issue has not been classified as a separate category, because there will not be a wrong mapping unless it concurs with one of the following causes at least.

1) **Divergence**: There may be divergences between data paths and control paths caused by BGP aggregation, default routing [17], multi-hop sessions, tunneling, layer-2 switching, abnormal routing and even router misconfiguration. The real-world cases of tunneling and abnormal routing will be shown in Section VI.

2) **No-response**: The traceroute path may be incomplete because there is no response on some hops. If such hop is in the middle of path, and its two adjacent hops are mapped to two other different ASes, then an indirect arc between the two ASes may be a false link.

3) **Unmapped-hop**: The IP address on a hop may not be mapped to any AS number, because it's a private IP address, or it isn't announced by its owner. If a prefix isn't announced by its owner, an IP address falling in this prefix may not be unmapped, but be matched wrongly by a less specific (shorter) prefix announced by another AS.

4) **MOAS-prefix**: When the matching prefix is announced by multiple ASes, there is ambiguity in inferring the origin AS. If the occurrence of MOAS prefix is due to multi-homing policies, the prefix may not be actually used by any AS that is announcing it [18]. This is not an issue when using a single routing table from a single router to perform the lookup.

5) **IXP**: These cases correspond to one prefix shared by multiple participants in an IXP, which is a shared infrastructure where multiple networks peer with each other publicly. This prefix may or may not appear in BGP routing tables, and thus they create either an extra AS hop or an unmapped router hop on the inferred path, respectively. It is also possible that the prefix is announced by more than one participants, thus becoming a MOAS prefix.

6) **Sibling**: Siblings are ASes that are owned by the same organization. There may be no clear boundary of IP address spaces or topologies among siblings. One of the ASes may use IP addresses in a prefix that is announced by its sibling AS. It's also possible that siblings share their backbones, so that the routers of one AS alternate with those of its sibling along paths through the backbones.

7) **Neighbor**: A prefix is not always only used by its origin AS exclusively. Between two peering ASes, two neighboring BGP routers share a common prefix (typically, a /30) which is configured as a static route at each end to enable the point-to-point connection. Therefore, a border router may reply to a traceroute probe using one of its interfaces whose IP address is shared and announced by one of neighboring ASes.

C. Argument on Referring to BGP Paths

The interrelations between data plane and control plane is not yet fully understood. We do not claim that the above list has exhausted all possible reasons for mismatches between traceroute paths and BGP paths. Nevertheless, the number of cases not covered could be quite small. The previous work has shown that about 94% of pairs of traceroute and BGP paths are matched with an improved IP2AS mapping method [19]. According to our measurements without the improved IP2AS mapping, 63∼88% of path pairs have a match. In the remaining cases, at most 4% of mismatch path pairs are caused by divergences of BGP paths and traceroute paths, and for the rest we can provide evidences for their occurrences due to errors in converting IP paths to AS paths. Therefore, we believe that it is rational to use BGP paths as the reference for the corresponding traceroute paths by default as [14], [15], [19]. Accordingly, we make an assumption that *traceroute AS path trends to follow BGP AS path as closely as possible*.

III. Related Work

Inferring AS paths with traceroute has attracted many research efforts over recent years. One of the first such studies was done by Chang *et al.* [12], which alerted for possible errors caused by using IP addresses borrowed from neighbor ASes. They presented heuristics based on router alias resolution to identify the ownership of border routers. Later, Amini *et al.* [13] found that a router may reply using any of its multiple interfaces' IP addresses, depending on the implementation of ICMP message generation.

Towards an accurate AS-level traceroute tool, Mao *et al.* [15] compared BGP paths with traceroute paths launched from the same AS. They studied a comprehensive set of possible causes of mismatches. In a following work [19], they presented a dynamic programming algorithm to minimize the number of mismatched path pairs by reassigning /24 prefixes to AS numbers. The main outcome of their works was a method to correct the mismatches due to unmapped hops, MOAS prefixes, IXPs and siblings.

With path pair comparison, Hyun *et al.* [14] quantified the mismatched pairs attributable to IXPs and siblings. They adopted the algorithm for the longest common subsequence problem to describe the pattern of unexplained mismatches. In a later work [20], they presented the concept of *third-party*

addresses, but its definition does not clearly address the issue of IP address sharing between BGP neighbors.

Recently, several works are promising to achieve a more accurate traceroute-derived topology. Pansiot *et al.* [21] did router-to-AS mapping based on the router alias data collected with a multicast tool *mrinfo*. Huffaker *et al.* [22] evaluated the accuracy of some heuristics for mapping routers to AS numbers with router alias data. Tozal and Sarac [23] invented a measurement tool *tracenet* to discover the interfaces on each visited subnetwork on the data path. However, while tackling issues with router aliases, those works did not address other types of pitfalls of using traceroute.

Different with previous works improving IP2AS mapping or searching for real links, the goal of this paper is to dig out false links caused by pitfalls of current practices, that we consider, despite the somewhat negative, an unavoidable step towards the combination of traceroute- and BGP-based measurement. As far as we can tell, our paper is the first to propose a systematic framework to identify mismatches and infer causes, and the first to investigate the inaccuracy of widely used traceroute-based topologies. Our result asserts that the main cause of mismatch is using IP addresses assigned from neighbor ASes in accordance with [12], [13] and departing from [14], [15] that attributed mismatches to the presence of IXPs, siblings and MOAS prefixes.

IV. Data Collection

The data sets used in this work include a mix of traceroute data, BGP data, WHOIS data and Internet Routing Registry (IRR) data. We'll elaborate on the process of collecting each of them in detail in the following paragraphs.

A. AS Path Pair

To detect the possible errors in traceroute-derived data, we collected traceroute raw data and the corresponding BGP routing updates from four ASes. Table I lists the number of destination IP addresses and prefixes, as well as the corresponding BGP information.

UCLA: From a host located at UCLA, we performed probes targeting all /24 blocks in the BGP routing table, using the traceroute tool *scamper* [24] with ICMP-paris [25] from 2009-02-22 to 2009-03-10. We broke each prefix into /24 blocks and selected the IP address $X.X.X.1$ in each block. About 13% of the traceroute probes got a reply from the final destination, while for the remaining cases we only had a partial path.

Ark: There are three CAIDA Ark monitors that happen to be located in ASes which provide a BGP feed to either RouteViews or RIPE-RIS collectors. Each Ark monitor uses a partial destination list (covering a fraction of routable IP address space) that changes over time. To increase the coverage of the traceroute data we cumulated data over a period from 2009-02-01 to 2009-03-12. During this period each monitor covered 66% of all /24 blocks and nearly 80% of all the prefixes in the BGP routing table. For each /24 block, the latest traceroute result was picked. About 7∼8% of the destinations replied to the traceroute probes, which means most paths were partial paths.

TABLE I
INFORMATION OF AS PATH PAIR DATA SOURCES

Monitor	ASN	#pair	#prefix	Collector	Orgnization
ams-nl	1103	5.2M	218K	ris-rrc03	SURFnet
nrl-jp	7660	4.9M	212K	rv2-oix	APAN
she-cn	4538	5.2M	218K	rv-wide	CERNET
ucla	52	7.6M	272K	ucla	UCLA

The IP2AS mapping table is usually constructed by merging a set of routing tables from multiple sources. The result may depend on the number and locations of sources. To provide a baseline for the evaluation of IP2AS mapping data, we generated traceroute AS paths only with the single BGP routing table of the AS from which the traceroute is launched, without the issue of MOAS prefixes. We ignored AS_SETs here on account of a small amount. The sensitivity of IP2AS mapping to different IP2AS data sources (collected in Section IV-B) will be investigated in Section VI-C.

For *ucla*, when mapping router paths to AS paths, we used UCLA BGP tables dumped in the day of the traceroute probing as the reference. UCLA is a small enough network so that all the routers have a consistent view. A traceroute path is paired with its corresponding BGP path to the same prefix, only if there is no change observed in the local BGP routes during the traceroute probe. Otherwise the paths are discarded.

We took a similar approach for Ark's traceroute data with an additional precaution. Since generally there is no guarantee that routers will have consistent tables inside a large AS, the path from the traceroute monitor may or may not follow the path reported by the BGP monitor in the same AS. To reduce this ambiguity, we only paired the paths where both next-hop ASes were the same.

B. IP2AS Mapping Table

To investigate the impact of different IP2AS mapping tables on the topology measurement, we collected the pairs of origin AS numbers and prefixes from BGP routing tables and IRR databases. We also generated a mapping table with the algorithm in [19].

BGP-i2a: From all available routing tables in Routeviews and RIPE-RIS, we extracted 335,953 prefixes announced by 31,593 ASes around 2009-3-1. There were 180 monitors that provide the full tables, *i.e.* the number of prefixes greater than 250,000 at that time.

IRR-i2a: In IRR databases [26] ISPs explicitly insert information such as prefix lists, routing policies, and BGP peers. We extracted 497,422 prefixes originating from 28,376 ASes from IRR databases as of 2009-03-05. This data has a high determinacy with only 89 MOAS prefixes, in contrast to *BGP-i2a* with 4,164 MOAS prefixes.

DP-i2a: With our implementation of Mao's dynamic programming algorithm [19], we obtained another mapping table from *BGP-i2a* and our AS path pair dataset. For the sake of efficiency, we chose one path pair per BGP atom randomly as the input of algorithm. The algorithm reassigned the origin ASes of 8.6% of /24 blocks and corrected nearly 32% of mismatched path pairs.

TABLE II
CONTRIBUTIONS TO AS DESCRIPTION, IXP AND SIBLING DATASETS

# ASN	Record	Descr.	IXP	Sibling
RIPE	17833	17659	151	3357
ARIN	19983	15201	46	5195
APNIC	4758	4243	78	1317
RADB	3194	2663	31	307
LACNIC	1381	1258	0	292
WCGDB	775	387	4	116
ALTDB	406	349	2	29
NTTCOM	530	344	1	39
AFRINIC	469	334	3	77
JPNIC	581	249	2	71
Others	1135	848	5	383
Total	51045	43535	323	11183

C. IXP and Sibling Lists

To help identify ASes that are IXPs and infer sibling relationships between ASes, we queried registry information for AS numbers from WHOIS databases. The records which include only regional Internet registries' (RIR) name were discarded. Among the records for the same AS number from different WHOIS sources, the last updated one was picked up. From each record, we imported two types of fields: (1) `as-name` (not available in LACNIC) and (2) `descr` (`OrgName` in ARIN, `owner` in LACNIC). The contributions of different WHOIS sources to the AS description, IXP and sibling datasets, as well as the number of records, are listed in Table II. Note that 89% of records are provided by five RIRs.

IXP-list: We collected a list of 404 /24 prefixes belonging to IXPs by crawling three websites: PeeringDB [27], PCH [28] and Euro-IX [29] on 2009-03-09. We also got IXP's ASNs by mapping IXP prefixes to ASNs, and found that about 40% of /24 IXP prefixes are announced by 76 ASes. In addition, we compiled a list of ASNs associated with IXPs based on keywords including IXP names (from the previous websites) and the common words 'internet exchange', 'exchange point', 'access point' and 'gigapop'. We searched the AS name list for these keywords, carefully filtering the false records, *e.g.* a description 'peering at an IXP' is not an IXP. We ended up with a total of 323 ASNs that belong to IXPs.

Sibling-list: We looked for similarities in AS names and descriptions of a given pair of ASes using approximate string matching except the cases that the name is a word appearing in an English dictionary. The acquisition history of all Tier-1 ISPs from wikipedia was also used to group ASes. After computing the transitive closure of sibling relationships, we cleaned up the candidate sibling groups with size greater than 20 manually. A sibling group would be removed if its common string only included the meaningless information, such as that the AS number belongs to some RIR or NIC. Finally, the sibling list contained 3,490 sibling groups with 11,183 ASes, of which 69% announce at least a prefix in BGP.

D. AS Adjacencies

To infer the causes of mismatches and evaluate the accuracy of traceroute-derived AS adjacencies, we collected data from CAIDA Ark, DIMES, Aqualab [11], IRR [26], iplane [30] and UCLA IRL [7], [31].

Ark: CAIDA's Ark project [8] measures the AS-level topology from over 40 monitors with traceroute, using Routeviews data for IP2AS mapping. Two traceroute-derived AS topologies were obtained by merging all snapshots in Feb. 2009. (1) *Ark-direct*: the topology with only direct links, in which every pair of connected ASes have a pair of adjacent hops in the traceroute path at least; and (2) *Ark*: the topology with both direct links and indirect links, in which two connected ASes may be separated by one or more unmapped or non-responsive hops.

DIMES: In DIMES project [9], thousands of volunteers measure the Internet with traceroute. BGP routing tables and WHOIS data are used for IP2AS mapping. We picked DIMES' monthly traceroute-derived AS topology in Feb. 2009, named *DIMES*. This graph includes the AS links which are observed at least once in the given month and at least twice up to the given month. In May 2011, we retrieved DIMES' topology of Feb. 2009 again, named *DIMES-new*, as DIMES had updated their data with new filtering techniques.

Sidewalk: Based on traceroute measurements from clients in a P2P system, Northwestern University Aqualab discovered many AS links invisible to BGP views [11]. In addition to IP2AS mapping with BGP-based data provided by whois.cymru.com, they developed several heuristics to filter out possible false links.

BGP-graph: We take the BGP-derived AS adjacencies available at UCLA IRL [7], [31], which is extracted from RouteViews and RIPE-RIS. For the sake of completeness, BGP updates are cumulated over a period of 9 months at least, from 6 months before to 3 months after the corresponding traceroute measurement.[1]

Ground-truth: According to our previous work [32], the *full* BGP routing table of a monitor should reveal almost all its AS neighbors over time [2]. Such AS is *monitored*. There are 82 monitored ASes in *BGP-graph* for *Sidewalk* and 110 monitored ASes in that for the others. We assembled the adjacencies of those monitored ASes as an extreme approximation of ground truth.

IRR-graph: The IRR databases are constructed and maintained manually by network operators and potentially contain stale or false records. Nevertheless it can provide useful information regarding the completeness of BGP-derived topology and the accuracy of traceroute-derived topology. We extracted 28,700 AS numbers and 156,094 AS adjacencies from all available IRR databases as of 2009-03-05.

E. Miscellaneous

Alias-list: iPlane project [30] provides a list of routers' aliases, *i.e.* a set of interface IP addresses belonging to the same router. The alias data can be used to explain mismatches due to using IP address borrowed from peering neighbors, since we can look up the origin AS number of each interface

[1]There are two samples of *BGP-graph* for the comparison with the traceroute data: for our own traceroute AS path pair data, *Ark* and *DIMES*, the period of cumulating BGP updates is from Sep. 2008 to Jun. 2009; for *Sidewalk*, it is from Jun. 2007 to Dec. 2008.

[2]Here 'full' means that the routing table covers the global routable address space. Some monitors' routing tables may be not full since these monitors tread BGP collectors as "peer" rather than as "customer".

	(a) substitute	(b) end-extra	(c) unmapped	(d) tie-break	(e) loop	(f) missing tail	(g) omission
BGP path	A B C D	A B	A B C	A B C D	A B	A B C	A B
Traceroute path	A E C D	A B C	A ? C	A C B D	A C A B	A D	A * B
LCS solution	=A −B +E =C =D	=A =B +C =$	=A −B +? =C	=A+C=B−C=D,=A−B=C+B=D	=A +C +A =B	=A −B −C +D =$	=A +* =B
Mismatch fragment	=A −B +E =C	=B +C =$	=A −B +? =C	=A +C =B , =B −C =D	=A +C =A	=A −B +D =$	=A +* =B

Fig. 3. Examples of AS path pairs with their LCS solution and mismatch fragments. '∧' and '$' are omitted if not being in mismatch fragments.

alias in BGP routing tables and know which AS lends its own IP address to this router. We extracted a total of 286,043 IP interface addresses on 67,430 routers on 2009-03-05.

AS network types and relationships: In UCLA IRL's AS topology datasets, there are also AS relationships and AS network types [7], which will help us classify traceroute-derived false links. According to the number of downstream customers, the non-tier-1 ASes are classified into three types: *large ISP*, *small ISP* and *stub*.[3] Each AS link is labeled with one of two major types of relationships: *customer-provider*, *peer-peer*. The algorithm for inferring the relationship in [7] follows the *no-valley* and *prefer-customer* policy credited to Gao [33]: an AS path should be *no-valley*, *i.e.* neither a customer-provider link nor a peer-peer link follows a provider-customer link or a peer-peer link in an AS path; when an AS receives path announcements to the same destination from multiple neighbors, the AS prefers the path from a customer over that from a peer, and prefers the path from a peer over that from a provider.

V. Mismatch Analysis: Breaking Paths into Fragments

In this section we detect and describe differences between BGP paths and traceroute paths obtained in Section IV-A. Then we analyze the pattern, the size and the frequency of those differences.

A. Methodology

With the assumption that *the traceroute AS path trends to follow the BGP AS path as closely as possible*, we want to find the *minimum* amount of differences between a pair of AS paths. This is equivalent to finding the *longest common subsequence* (LCS) which is present in both paths. The LCS problem for two sequences can be solved with a dynamic programming, aka the Hunt-McIlroy algorithm, with the complexity of $O(nm)$, where n and m are the lengths of two sequences [34].

Like the classic file comparison utility `diff`, the LCS solution can be described as a minimum array of edit operations needed to transform the BGP AS path into the traceroute AS path: insertion '+', deletion '−', or unmodified '='. The successive '=' operations represent the common segments, while '+' and '−' operations indicate the differences. For example, Figure 3(a) shows a case of one-to-one substitution.

To pinpoint multiple mismatches in the same AS path pair and describe a mismatch in its local context, we define a **mismatch fragment**, for a given AS path pair, as a sequence of successive '−' or '+' operations wrapped around by two '=' operations before and after in the LCS solution. Figure 3(a) shows an example. Five additional steps have been adopted to detect mismatch fragments systematically:

1) **Special tokens**: ' ∧' and '$' represent the beginning and the end of path respectively, which ensure that any difference can be represented by a mismatch fragment, *e.g.* in Figure 3(b) an extra AS at the end of traceroute path. '∗' denotes successive non-responsive hops. '?' denotes successive unmapped hops, *e.g.* one AS replaced with a '?' in Figure 3(c).
2) **Tie-break**: When there are multiple alternative solutions with the same number of modifying operations, the one whose '=' operations appear earlier in the BGP path is picked. The goal of this tie-break is to concentrate the errors in the least number of original hops as possible. In Figure 3(d), the first solution's '$= B$' is earlier than the second solution's '$= C$' in the BGP path, so two mismatch fragments are extracted from the first solution.
3) **Loop**: A mismatch fragment is replaced with its inside loop, if possible, since loops describe differences more properly. In Figure 3(e), the LCS solution is replaced by its inside loop.
4) **Missing tail**: The mismatch fragments with only missing ASes at the end of path is discarded, since our interest is in the extra links brought by traceroute. In the case of the substitution at the end of path, only the first missing AS is kept, and the rest of missing ASes are discarded, because the rest missing ASes in incomplete traceroute path is not due to the errors of IP2AS mapping, but due to the consecutive non-responsive probes. An example in this case is shown in Figure 3(f).
5) **Omission**: The mismatch fragments that do not generate any extra link are discarded. Those fragments include only '+∗' or '+?' between two same AS numbers in traceroute paths. Figure 3(g) shows a mismatch fragment only including an unmapped hop.

TABLE III
Statistics of Mismatch Fragments and Path Pairs

	ams-nl	nrt-jp	she-cn	ucla
# mismatch fragments	11,800	12,543	13,414	16,765
# mismatched pairs	1,940,792	975,757	605,798	1,258,334
% mismatched pairs	37%	20%	12%	17%

We obtained a total of 37,602 unique mismatch fragments. The number of mismatch fragments from different monitors is listed in Table III, as well as the percentage of path pairs that have mismatch fragments. About 32% of unique mismatch fragments are observed by multiple monitors. There are 3.3% of mismatched pairs which include more than one mismatch fragments. The number of mismatched path pairs

[3]Let C be the number of downstream customers. For large ISPs, $C \geqslant 51$; for small ISPs, $5 \leqslant C \leqslant 50$; for stubs, $C \leqslant 4$.

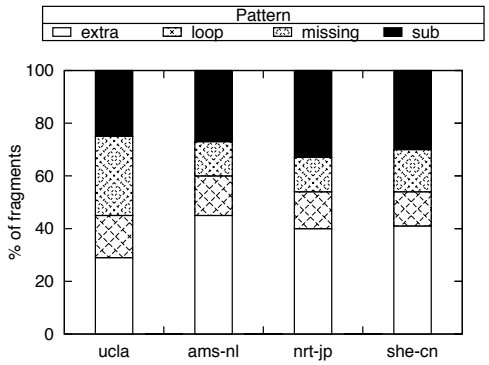


Fig. 4. Pattern of mismatch fragemnts

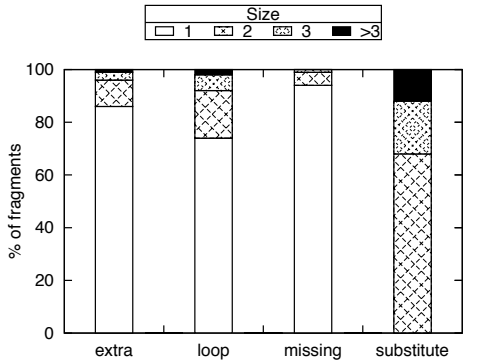


Fig. 5. Size of mismatch fragments

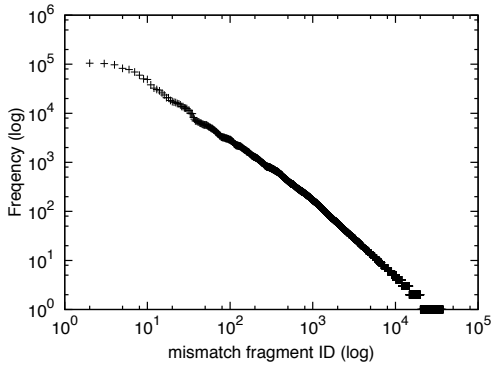


Fig. 6. Frequency of mismatch fragments

is not correlated with the number of mismatch fragments. For instance, *ams-nl* has the fewest mismatch fragments, but the most mismatched pairs. To reason about mismatches, we also extract the corresponding IP-level path segment for each mismatch fragment. Note that one mismatch fragment may have more than one IP-level instance.

B. Analysis of Mismatch Fragments

What do mismatch fragments look like? According to the type of modifying operation, mismatch fragments can be classified into three patterns: *extra* fragments including only '+', *missing* fragments including only '−', and *substitute* fragments including both '−' and '+'. Among the *extra* fragments, those with the same '=' operands are reclassified as the *loop* pattern. The percentages of mismatch patterns in four datasets are shown in Figure 4. We observe that there is no obviously dominant pattern. The breakdown by pattens is 31% extra, 13% loop, 22% missing and 34% substitute. The distribution of patterns is similar for different monitors. There're a larger proportion of missing fragments in *ucla*. Since only the monitor of *ucla* is located in the United States, the neighbor ASes of Tier-1 ISPs in the United States are involved in the missing fragments in *ucla* more frequently than those in other datasets. For instance, the neighbor ASes of AS3356 (Level3) contribute 822 missing fragments to *ucla*, while contributing a total of 643 to the others.

What's size of each mismatch fragment? We define the size of mismatch fragment as the number of modifying operations on AS numbers (not including '∗' or '?'). For example, the size of mismatch fragment in Figure 3(a) is two, and in Figure 3(c) and (e) that is one. Figure 5 shows the distribution of size of mismatch fragments with different patterns. This statistic covers all unique mismatch fragments. Most of extra, loop and missing fragments have a size of one or two, while the substitute fragments have mainly a size of two or three. The small size partially confirms our assumption that the traceroute AS path trends to follow the BGP AS path as closely as possible. The fragments of large size may be caused by the divergences between control paths and data paths or the simultaneous occurrences of multiple mismatch causes.

How frequently does each mismatch fragment occur? We define the frequency of mismatch fragment as the number of path pairs sharing this fragment. Figure 6 shows the frequency of all unique mismatch fragment, where fragments are numbered in descending order of frequency. We note that the curve follows a heavy-tailed distribution, which means that a small number of mismatch fragments are shared by a very large number of path pairs. Since the paths starting from a single monitor form a tree-like structure, the closer a fragment is to the monitor, the higher its frequency is. While 54% of mismatch fragments pop only once, the most frequent mismatch fragments occur up to a hundred thousands of times. If we had been counting mismatches based on paths, we would be counting the same mismatch repeatedly.

In summary, the differences between traceroute paths and BGP paths have simple patterns, small sizes and highly various frequencies. Later we will see that the simple and small differences give a possibility to develop a systematic framework for the cause inference. The heavy-tailed distribution of frequencies indicates that the results based on counting paths would be very sensitive to the relative location of mismatches to the monitor.

VI. Inferring the Causes of Mismatch

In this section we infer causes of mismatches according to the taxonomy in Section II-B. We first articulate our systematic framework to derive the causes of mismatches. Then we describe details on some cause inferences: first, we investigate the divergences between control paths and data paths; then we look into two causes related to the incompleteness and ambiguity of IP2AS mapping tables; third, we study using IP

TABLE IV
MISMATCH PATTERNS CONSTRUCTED BY ONE OR TWO REPLACEMENTS (ONE COMBINATION OF REPLACEMENTS PER ROW)

Name	Pattern	Replacement
1-extra	`=A +B =C`	`(+B, =A)` `(+B, =C)`
1-missing	`=A -B =C`	`(=A, -B)` `(=C, -B)`

1-1-substitute	`=A -B +C =D`	`(+C, -B)`
		`(+C, =A), (=D, -B)`
		`(+C, =D), (=A, -B)`

2-extra	`=A +B+C =D`	`(+B, =A), (+C, =D)`
		`(+B, =A), (+C, =A)`
		`(+B, =D), (+C, =D)`

2-missing	`=A -B-C =D`	`(=A, -B), (=D, -C)`
		`(=A, -B), (=A, -C)`
		`(=D, -B), (=D, -C)`

2-1-substitute	`=A -B +C+D =E`	`(+C, -B), (+D, -B)`
		`(+C, -B), (+D, =E)`
		`(+C, =A), (+D, =B)`

1-2-substitute	`=A -B-C +D =E`	`(+D, -B), (+D, -C)`
		`(+D, -B), (=E, -C)`
		`(+D, -C), (=A, -B)`

2-2-substitute	`=A -B-C +D+E =F`	`(+D, -B), (+E, -C)`

address assigned from peering neighbors. At last, we analyze the results of the impact of each cause.

A. Inference Framework

Before classifying the pitfalls of converting IP paths to AS paths, we have to first discern mismatches due to divergences between control paths and data paths. To bypass this challenge, we draw support again from the main assumption of this paper: *The data path trends to follow the control path as closely as possible*. And hereby we assume that *at most two* errors of IP2AS mapping occur simultaneously in a single mismatch fragment. Our supposition has been partially confirmed by the observation that most of mismatch fragments have a size of one or two (or three for the substitute pattern) in Figure 5.

Since each error of IP2AS mapping involves a pair of ASes (including $*$ or ?) where one AS is misplaced on the position of the other in the path, each mismatch fragment could be created by at most two replacements of ASes. Table IV enumerates all of the mismatch patterns which are constructed by at most two replacements, as well as all possible combinations of replacements, where a replacement `(X, Y)` means that `X` replaces `Y` by mistake. The relationship of replacement between two ASes is not commutative. For the patterns with the combination of two replacements, they can be generated only if both replacements occur at the same time. The loop pattern is omitted here as it is the specific case of 1-extra and 2-extra patterns with the same '=' ASes.

Now we get the first rule of cause inference for *Divergence*:

Rule 1 Divergence: The pattern of mismatch fragment is not in Table IV.

This rule can also be re-described as that the size of mismatch fragment is greater than 3 for the substitute pattern except the 2-2-substitute, or greater than 2 for the loop, missing and extra patterns.

For each of the rest not attributable to *Divergence*, the task of interring its cause is equivalent to finding the reasons for its corresponding replacements. At this point, we have converted the problem of inferring mismatch causes to the problem of explaining replacements of ASes. With the replacement candidates of a given mismatch pattern listed in Table IV, we conduct a sequence of tests on those candidates and then infer mismatch cause from the test results.

TABLE V
THE PRIORITY OF EVIDENCES AND THE CORRESPONDING CAUSES

	Evidence	Cause
1	pattern	1.Divergence
2	pattern	2.No-response
3	IXP-prefix-list	5.IXP
4	pattern	3.Unmapped-hop
5	IXP-ASN-list	5.IXP
6	BGP-i2a	4.MOAS-prefix
7	sibling-list	6.Sibling
8	alias-list	7.Neighbor
9	BGP-graph	7.Neighbor
10	IRR-graph	7.Neighbor

Our test algorithm is comprised of the following six rules, of which each is designed for a single cause of mismatch. Let `(X, Y)` be a replacement candidate, and let `IP(X)` be the corresponding IP addresses mapped to `X`.

Rule 2 No-response: The replacement is `(+*, -Y)`.

Rule 3 Unmapped-hop: The replacement is `(+?, -Y)`.

Rule 4 MOAS-prefix: `IP(X)` belongs to a MOAS prefix announced also by `Y`.

Rule 5 IXP: `IP(X)` or `X` is used by an IXP.

Rule 6 Sibling: `X` and `Y` are siblings.

Rule 7 Neighbor: `IP(X)` is an IP address of an interface on a router owned by `Y`.

The validity of rules are self-evident, because each rule is equivalent straightforwardly to the corresponding definition of the cause in Section II-B. The simplicity of the above rules are benefit from the successful conversion of the scope of problem from the whole topology to AS pairs, while the accuracy of inference fall upon the quantity of datasets supporting those tests. This is exactly our purpose of decoupling procedures from evidences in the inference.

The evidences for each cause are listed in Table V. The evidence *pattern* means that the rule is only based on the pattern of mismatch fragments or the appearance of replacement, without any additional datasets involved. The usages of evidences for *MOAS-prefix*, *IXP* and *Sibling* are self-explaining. For *Neighbor*, we use three evidences as follows: (1) according to the iPlane's router alias list, `IP(X)` belongs to a router that has another interface mapped to `Y`; (2) `X` and `Y` are neighbors in *BGP-graph*; (3) `X` and `Y` are neighbors in *IRR-graph*. The details on *Neighbor* will be present in Section VI-D.

During the process of inference, there are two other issues needed to be addressed: Firstly, one replacement or one mismatch fragment may pass tests for more than one cause. In other words, the replacement or the mismatch fragment can be explained by multiple causes. Secondly, two replacements as a combination in a mismatch fragment may be explained by different types of causes respectively.

We resolve these two issues through comparing the priority of respective evidences supporting the causes. According to our confidence on the accuracy of evidences, the priority

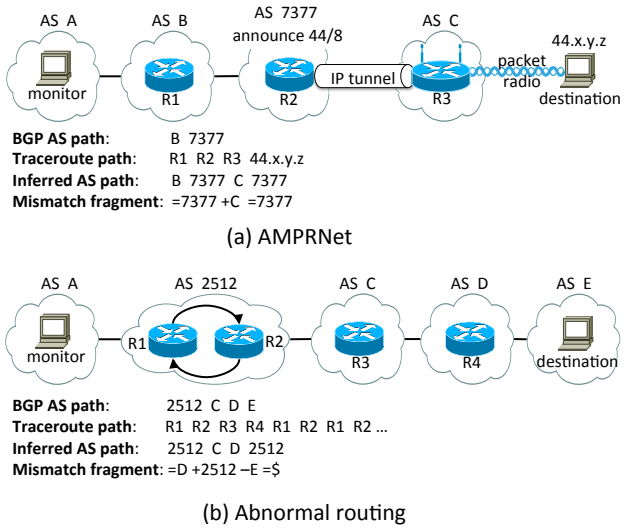


Fig. 7. Illustrations of divergences between control path and data path

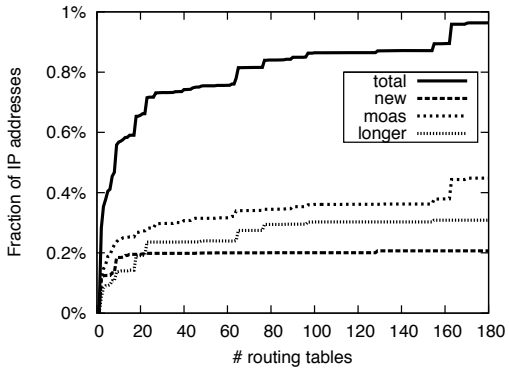


Fig. 8. Sensitivity of IP2AS mapping to multiple BGP routing tables

of evidences is set as their descending order in Table V. The underlying principle of the priority is that the evidence *pattern* is prior to IP-level evidences which are prior to AS-level evidences, except two cases: (1) the mismatch fragment involving an IP address falling in *IXP-prefix-list* is attributed to *IXP* even if it is at an *Unmapped-hop* or belongs to *MOAS-prefix*; (2) two ASes in the same group of *sibling-list* are due to *Siblings* even if the replacement passes the test of Rule 7 according to *alias-list*. Another consideration on the priority is about the inference of *Neighbor* we'll elaborate on in Section VI-D.

So the inference algorithm follows an if-then-else process: It starts by checking whether the mismatch fragment is explained by the evidence 1, if yes then the test stops, otherwise it continues and checks for the evidence 2, and so on until its put in the *Unknown* bin.

Rule 8 Unknown: The mismatch fragments cannot be explained.

B. Divergences Between Control Paths and Data Paths

Besides the supposition of small-size differences, another reason for making our heuristic for inferring *Divergence* so aggressive is that, our focus being on the errors of conventing IP path to AS path, we prefer false positives to false negatives at the current stage in order to reduce the false positives of follow-up inferences. Moreover, this heuristic averts the complication concerning causes for *big-size* fragments. By looking into those mismatches further, we dug out two typical real-world cases, which help us detect a few false negatives.

Amateur Packet Radio Network (AMPRNet): In the AMPRNet, hosts are connected by links over packet radio within subnets and by IP tunnels over the Internet among subnets. The AMPRNet uses the prefix 44/8 announced by AS7377 (UCSD) and connects with the rest of Internet via a router in UCSD. As illustrated in Figure 7(a), the BGP path to 44/8 ends at AS7377, while the traceroute path comes in AS7377, then travels cross the IP tunnel to other ASes, and finally reach the destination whose IP address is mapped to AS7377. By checking all mismatch fragments on the paths to 44/8, we find a total of 42 fragments in this case.

Abnormal routing: In *ucla*, we surprisingly observed that AS 2512 (CalREN), the provider of UCLA network, is appended in some substitute or extra fragments three AS hops away from UCLA. The reason for those mismatches are illustrated in Figure 7(b). During the probe to a destination four AS hops away from UCLA, some of traceroute instances were actually falling in a routing loop within AS 2152 immediately after reaching an AS three AS hops away. By looking into the mismatch fragments where one provider of monitors [4] is involved simultaneously with an IP-level loop, we detect 90 fragments in this case.

Note that some traceroute paths involving the abnormal routing, such as loops, can be easily detected. We still paired those traceroute paths with their BGP counterparts if there were no changes observed in BGP routing tables, because we try to quantify the causes of false AS links which those traceroute instances may generate.

C. Incompleteness and Ambiguity of IP2AS Mapping tables

The incompleteness of IP2AS mapping table may lead to unmapped hops, while the ambiguity comes from MOAS prefixes. We investigate these two issues by comparing different mapping tables. The comparison with *IRR-i2a* and *DP-i2a* will not change the results of cause inference based on *BGP-i2a*. The change of IP2AS mapping stems from three reasons:

1) **New prefix**: A previously unmapped IP address finds a match prefix;
2) **MOAS prefix**: A prefix originally with or without one origin AS now finds more than one origin ASes;
3) **Longer prefix**: An IP address is remapped to another prefix longer (more specific) than the existing one.

In our AS path pair data sets, the initial IP2AS mapping only uses the routing tables of ASes where the monitors are located. The single BGP routing table avoids the ambiguity temporarily, but suffers the incompleteness. Nevertheless, the results on the single table provide us a baseline, *i.e.* the upper bound of incompleteness and the lower bound of ambiguity. Therefore, in our inference framework, both *MOAS-prefix* and

[4]The major providers of 4 monitors: (ams-nl: AS 3257, AS 6453, AS 20965) (nrl-jp: AS 2516, AS 22388) (she-cn: AS 10026, AS 3320) (ucla: AS 2152, AS 2153)

TABLE VI
CHANGES OF IP2AS MAPPING BY ADDING A NEW MAPPING TABLE

% IP addr.	single BGP routing table			BGP-i2a	
	BGP-i2a	IRR-i2a	DP-i2a	IRR-i2a	DP-i2a
New	0.31	1.72	0.59	0.82	0.29
MOAS	0.45	11.20	0.67	11.41	0.55
Longer	0.21	2.71	0.57	2.50	0.62
Total	0.96	15.63	1.83	14.73	1.35

TABLE VII
MISMATCH FRAGMENTS CORRECTED (−) OR GENERATED(+) BY USING DIFFERENT IP2AS MAPPING TABLES

# fragments	BGP-i2a		IRR-i2a		DP-i2a	
	−	+	−	+	−	+
New	508	12	541	60	984	27
MOAS	908	0	1310	0	3601	0
Longer	346	83	1879	8018	2597	53
New+Longer net value	−759 (−2.02%)		+5658 (+15.05%)		−3501 (−9.31%)	
Total net value	−1667 (−4.43%)		+4348 (+11.56%)		−7102 (−18.89%)	

Unmapped-hop should be covered almost completely with few false negatives. By remapping with *BGP-i2a*, we can refine the inference results by detecting the false positives of *Unmapped-hop* and the misclassified cases due to the longer prefix.

We observe how much impact each of three tables has on the results of IP2AS mapping through the following three experiments. The test set in the first two experiments consists of 1,202,475 interface IP addresses from one measurement cycle of CAIDA Ark around 2009-02-26.

On the number of BGP routing tables: We first investigate the sensitivity of IP2AS mapping to the number of BGP routing tables. Starting with a single table from *BGP-i2a*, we cumulate the routing tables one by one. By randomly repeating the entire process 50 times, we compute a final average shown in Figure 8. Only about 1% of IP addresses change the origin ASes, among which the majority (47%) are caused by MOAS prefixes. There are almost no more changes due to new prefixes or longer prefixes after 100 tables are added, while the MOAS prefixes lead to changes continually.

Impact on interface IP addresses: We compare the impact of three tables by adding each of them to a randomly chosen BGP routing table (repeating 50 times). The results are shown in Table VI. The MOAS prefixes contribute the most of changes. Both *BGP-i2a* and *DP-i2a* have a small impact (less than 2% of interface IP addresses), while *IRR-i2a* brings over 10% due to MOAS prefixes. Because *IRR-i2a* has only 89 MOAS prefixes itself, a large amount of new MOAS prefixes means a lot of differences with the BGP routing tables on mono-origin prefixes. That is mainly attributable to the stale information in IRR databases. When we add *IRR-i2a* and *DP-i2a* to *BGP-i2a* separately, the results are similar.

Impact on mismatch fragments: We remap the AS path pair data by adding each IP2AS mapping table to the initial single routing table and see how many mismatch fragments are corrected or generated. Although MOAS prefixes bring the ambiguity in practice, here we temporarily allow a prefix to be mapped to multiple AS numbers simultaneously, and then consider a mismatch fragment to be corrected by a MOAS prefix if one of multiple mapping can correct this mismatch. In the case that one fragment can be corrected by multiple types of causes, we prefer new prefixes first if possible, otherwise we prefer longer prefixes to MOAS prefixes.

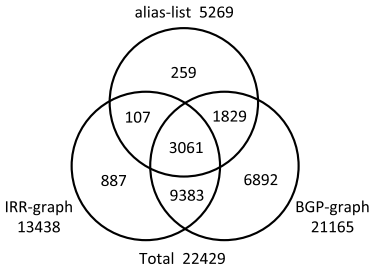


Fig. 9. Venn diagram of contributions to the cause 'Neighbor'

The results are summarized in Table VII. *DP-i2a* can help avoiding nearly 20% of mismatches, while *IRR-i2a* generates over 10% of new mismatches. The results of *BGP-i2a* demonstrate that about 2% of mismatch fragments are the false positive *Unmapped-hop* cases or the misclassified cases due to longer prefixes.

In summary, the number of BGP routing tables has a small impact on IP2AS mapping of interface IP addresses. About 2% of mismatch fragments is corrected by remapping with *BGP-i2a*. *DP-i2a* can improve the accuracy of IP2AS mapping by reducing nearly 20% of mismatches if one-to-multiple IP2AS mapping are allowed, while *IRR-i2a* creates more mismatches due to the stale information.

D. IP Addresses Assigned from Peering Neighbors

Neighbor is the last and the most difficult cause to be inferred, since there are yet no enough complete and accurate dataset for mapping IP addresses to routers or mapping routers to AS numbers, that is also a huge obstruction towards the accuracy traceroute measurement. This topic has been open and hot in Internet measurement research for a long time, and it is beyond the scope of this paper. What we do is to abate the influence of such limitation on our results as much as possible in two ways.

On the one hand, we reduced the possible false positives by giving *Neighbor* the lowest priority and reducing the false negatives of other causes. Because once a mismatch fragment is needed to be tested for *Neighbor*, *Neighbor* is only one possible cause left. Therefore, the false positives of *Neighbor* must be contained by the union of false negatives of other causes. Keeping this in mind during the inference of previous causes, we have tried to reduce their false negatives by adopting the aggressive heuristics, such as the one for *Divergence*, and choosing the upper bound of impact of a cause, such as *Unmapped-hop*. Moreover, we used the state-of-the-art methods of data collection to guarantee the completeness of inference. For instance, the method of collecting IXP prefix lists from IXP databases was also adopted by the recent IXP-oriented measurement work [35]. In addison, as far as we know, our sibling list collection work was only one until a very recent work on ASN-to-org mapping [36], but it did not provide publicly available data yet as of our writing.

On the other hand, we reduced the possible false negatives by using a comprehensive set of information of AS adjacencies

TABLE VIII
BREAKDOWN OF MISMATCH FRAGMENTS BY CAUSES AND PATTERNS

% fragments	extra	loop	missing	substitute	Total
1 Divergence	1.33	1.06	0.04	4.58	7.03
2 No-response	-	-	6.21	-	6.21
3 Unmapped-hop	-	-	2.25	-	2.25
4 MOAS-prefix	0.86	0.23	0.19	1.05	2.35
5 IXP	1.29	0.34	1.67	2.33	5.66
6 Sibling	2.35	1.22	0.10	2.99	6.56
7 Neighbor	18.71	8.13	11.55	21.57	59.65
8 Unknown	6.87	2.33	-	1.12	10.31
Total	31.41	13.31	22.01	33.64	100

including *BGP-graph*, *IRR-graph* and *alias-list*. When using *alias-list*, we did not infer the owner of router (again, this is beyond the scope of this paper), but checked whether two IP addresses on the same router are mapped to two ASes involved in the replacement, according to *BGP-i2a*.

The contributions of three datasets to the cause *Neighbor* are plotted in Figure 9. *alias-list* supports 24% of inferences. 64% of inferences are drawn from more than one evidence. *BGP-graph* contributes nearly 31% alone while 4% are complemented by the others.

E. Results

After the above progress, we inferred the causes for 89.69% of mismatch fragments in 99.36% of mismatched path pairs. We believe that most of unknown cases may either correspond to BGP sessions not visible in the current BGP view or be due to misclassified *Divergence* cases. The results are presented as the answers to three questions as follows.

How much dose each cause contribute? The breakdown of mismatch fragments by causes and patterns is shown in Table VIII. From the last column, we can see that the majority (60%) of mismatch cases are caused by *Neighbor*, *i.e.* using IP addresses assigned from BGP neighbors, that supports the opinion in [12]. The contribution of IXPs, siblings and MOAS prefixes only sum up to 14.57%, although the previous works [14], [15], [19] considered them as the major causes. In addition, about 7% of mismatch cases are created by gaps in traceroute paths, *i.e.* the unmapped or non-responsive hops.

Is there any correlation between causes and patterns? It is not surprising that every cause can lead to any pattern except that both *No-response* and *Unmapped-hop* only generate the missing pattern. From Table VIII we can observe that every cause generates the substitute pattern the most frequently, while *Neighbor* causes the largest number of mismatch fragments for every pattern. Most of *Unknown* cases have an extra pattern.

What if we quantify the impact by counting paths? In Figure 10, we compare the contribution of causes by counting the number of paths to that by counting the number of fragments side by side. We observe that the diversity among different monitors by counting paths is greater than that by counting fragments. This demonstrates that the results by counting fragments are more robust to the monitor location and the possible flaws in cause inference than those by counting paths. Through further observation, we find two kinds of biases in counting paths: (1) Overestimating the impact of some causes, such as *IXP* (No. 5) in *ucla* and *she-cn*. This is mainly because multiple paths may often share a single point of mismatch close to the monitor. (2) Underestimating the *Unknown* cases. Only about 1% of path pairs contain 9∼14% of mismatch fragments without known causes (No. 8). .

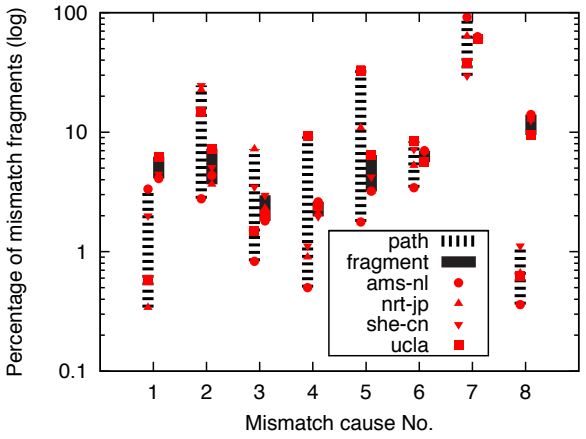


Fig. 10. Contributions of mismatch causes as measured by counting paths and fragments. Note that for two points with Y-values x and y, their vertical distance $x - y$ in log-scale is equal to $log(x/y)$ in normal-scale.

TABLE IX
BREAKDOWN OF FALSE LINKS BY CAUSES AND PATTERNS

# links	extra	loop	missing	substitute	Total
1 Divergence	44	46	5	112	189
2 No-response	-	-	671	-	671
3 Unmapped hop	-	-	155	-	155
4 MOAS prefix	8	0	19	22	40
5 IXP	27	13	49	53	122
6 Sibling	50	17	3	58	106
7 Neighbor	233	117	1271	460	1872
8 Unknown	333	109	-	103	460
Total	654	273	2158	704	3215

In summary, *Neighbor* has the greatest impact on 60% of mismatches, which is quantified by counting the number of mismatch fragments, because counting mismatch fragments is more appropriate and more robust than counting paths. As we will report in Section VII-A that *Neighbor* contributes most of false links, we conclude that the major pitfall of traceroute measurement is *Neighbor*.

VII. ACCURACY OF TRACEROUTE-DERIVED AS-LEVEL TOPOLOGY

In this section we assess the accuracy of traceroute-derived AS adjacencies with the ground-truth data described in Section IV-D. We first investigate the causes of false links to close the circle of our framework in Figure 1, and then evaluate the accuracy of publicly available data. We also attempt to find out some potential patterns among false links with a new method of AS relationship characterization.

A. Causes of False Links

Given the ground truth on a set of monitored ASes with BGP monitors inside, we identify an AS link $X - Y$ as a false link, if either X or Y is in the set of monitored ASes, but $X - Y$ does not exist in the ground truth. We identified 3215 false links out of 49074 traceroute-derived AS links in

TABLE X
INACCURACY OF PUBLICLY TRACEROUTE AS TOPOLOGY DATASETS

# links	All			Monitored ASes			Covered ASes		
	Traceroute	BGP	NOT in BGP	Traceroute	BGP	NOT in BGP	Traceroute	BGP	NOT in BGP
Ark	69962	145440	24846 (35.53%)	24817	60812	9473 (38.12%)	54961	115397	19794 (36.01%)
Ark-direct	56014	145440	14240 (25.42%)	18755	60812	4626 (32.48%)	42919	115397	10430 (24.30%)
DIMES	77358	145440	31773 (41.07%)	27965	60812	11707 (41.86%)	63952	115397	27701 (43.31%)
DIMES-new	74796	145440	27399 (36.63%)	27429	60812	10008 (36.49%)	61883	115397	23776 (38.42%)
Sidewalk	143384	152211	29244 (20.40%)	36007	38945	3024 (8.39%)	85271	93006	11601 (13.60%)

TABLE XI
BREAKDOWN OF FALSE LINKS/MONITORED ASES BY AS TYPES

# links (# monitors)	Tier-1	Large ISP	Small ISP	Stub
Ark	4231 (7)	4939 (34)	389 (34)	100 (35)
DIMES	5050 (7)	6517 (34)	231 (34)	70 (35)
Sidewalk	1484 (6)	1456 (28)	84 (27)	21 (21)

our AS path pair datasets. Since the monitored ASes are only a small part of the Internet, this result is a conservative estimate.

To understand how false links were created, we search for the replacements of those false links in mismatch fragments and group them by the corresponding mismatch causes and patterns. Table IX presents the breakdown of false links. Note that some false links may be counted more than once by different causes and patterns. The results show that *Neighbor* is responsible for 58% of false links, while three causes *IXP*, *Sibling* and *MOAS prefix* only bring 8%. Most of false links (67%) occur in only 22% of mismatch fragments with missing pattern. The gaps in the traceroute paths lead to nearly 26% of false links which could have been avoid easily. So far, we have closed the circle of our framework in Figure 1.

B. Inaccuracy of Publicly Available Data

To evaluate the inaccuracy of publicly available traceroute-derived topology datasets, we compare them with *BGP-graph* on the adjacencies of the following three sets of ASes and quantify the inaccuracy with the fraction of *not-in-bgp* links, *i.e.* the fraction of traceroute-derived AS links which are not in the corresponding BGP topology:

1) **All of ASes**: The not-in-bgp ratio can be considered as an upper bound of inaccuracy.

2) **Monitored ASes**: The corresponding BGP data is the ground-truth data; the not-in-bgp links are false links; the not-in-bgp ratio is a reasonable estimation of actual inaccuracy.

3) **Covered ASes**: If one AS can be reached through the *provider chain*, *i.e.* the successive customer-provider links, from at least one of monitored ASes, the adjacencies of this AS are considered being *covered* by *BGP-graph*. This assumption is deduced from the no-valley and prefer-customer policy [33] and supported by the observation in [7]. Note that the monitored ASes are also covered. Since the relationship inference is not errorless and the no-valley and prefer-customer policy does not always hold up in reality, the not-in-bgp ratio on the covered ASes, also as an estimation of actual inaccuracy, is less reliable, but more representative (because of involving more ASes) than that on the monitored ASes.

The results on three publicly available topology datasets are listed in Table X [5]. The inaccuracy estimated by the not-in-bgp ratio for the monitored ASes is, somewhat disappointing, 8%∼42%. Although *Sidewalk* has a much lower inaccuracy than the others, the number of false links in *Sidewalk* is over 3000 not much less than that in *Ark-direct*.

Nevertheless, *Sidewalk*'s heuristics for filtering out false links have reduced the inaccuracy successfully. Moreover, the effect of removing indirect links is also exhibited by the decrease of inaccuracy from *Ark* to *Ark-direct*, where about 80% ($\frac{9473-4626}{24817-18755}$) of indirect links attached to monitored ASes are false. The new filtering techniques adopted by DIMES reduced about 5% of inaccuracy shown by the comparison between *DIMES-new* and *DIMES*. We also observed that in some cases the degree of AS is inflated by thousands of false links in *Ark* and *DIMES*.

Comparing the not-in-bgp ratios in different AS sets for the same topology, we find that, for *DIMES* and *Ark*, their respective not-in-bgp ratios are similar. In other words, the not-in-bgp ratio seems to be independent of the set of ASes in these two datasets. If so, our estimation of upper bound of inaccuracy will be very close to the actual one.

For *Sidewalk*, the not-in-bgp ratio in the monitored ASes is much less than that in all of ASes. One reason may be that Sidewalk's heuristics for filtering out false links is biased in favor of false links attached to large ASes. Those heuristics utilize BGP AS paths to filter out false links, and it is well known that the BGP data have a better coverage on large ASes than small or stub ASes. That is also why it is difficult to identify the false links in the lower level of Internet hierarchy.

To demonstrate the above opinion further, we break down the false links by the network types of monitored ASes which false links are attached to. Table XI shows the results, as well as the number of monitored ASes of each type. We can see that the most (over 95%) of false links are associated with Tier-1 or large ISPs. Although the average number of false links attached to small ISPs and stubs are quite small, the total number of small ISPs and stubs are very large (about 98% of ASes). Therefore, there may be a large amount of false links hidden in the lower level of topology which we cannot identify.

In summary, the inaccuracy of publicly available traceroute-derived topology datasets, as measured by the not-in-bgp ratio on the ground truth, ranges from 8% to 42%. While the heuristics applied to *Sidewalk* and removing the indirect links can reduce the inaccuracy, there may be a large amount of false links which cannot be identified yet.

C. Characterizing False Links

One may also wonder the relationships between two ends of a false link. It is too obvious to mention that the AS relationship is not available for an AS link even not seen

[5]Note that the size of ground truth for *Sidewalk* is about 64% of that for the others, because the number of monitored ASes for *Sidewalk* is 80, much smaller than 110, that for the others.

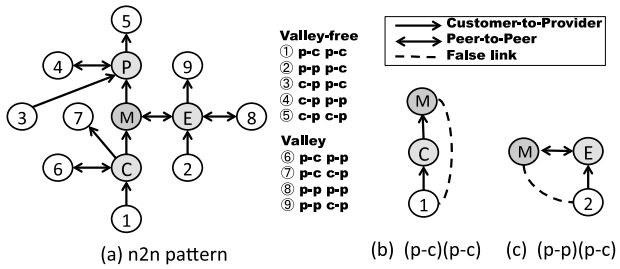


Fig. 11. Illustration of n2n patterns from an AS (M) trough its customer (C), provider (P) and peer (E) to other ASes. The number of each AS represents its priority under prefer-customer and no-valley/valley policy.

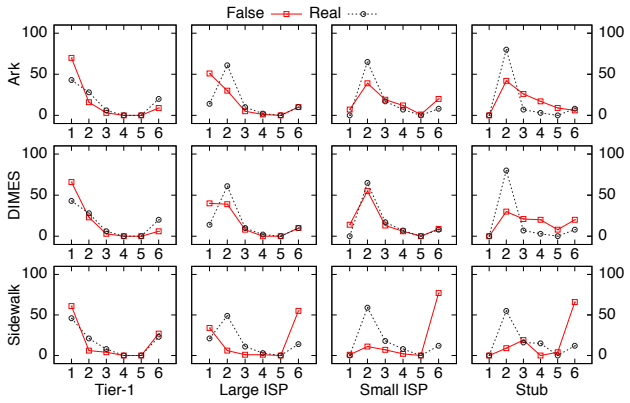


Fig. 12. Distribution (in percentage %) of n2n patterns of false/real links attached to different network types of monitored ASes. The number 1∼5 on X-axis are the ID of no-valley n2n patterns as shown in Figure 11, where AS (M) represents the monitored AS, while the number 6 is for the rest.

in BGP. To characterize the relationship between a given AS and its real or false neighbors, we present a new classification method for characterizing relationships between two ASes: **neighbor-to-neighbor (n2n) pattern**. Given two ASes X and Y, if there is a third AS Z which is the neighbor of both X and Y, the *n2n pattern from X to Y* is defined as a sequence of two relationships $R(X,Z)$ and $R(Z,Y)$, where $R(\cdot)$ refers to the relationship between two ASes. When there are multiple common neighbors of X and Y, we choose the one according to the prefer-customer and no-valley policy. If there is no no-valley path, we then only follows the prefer-customer policy.

The idea of n2n pattern comes from the observation that many mismatch fragments, which has the missing pattern or is due to *Neighbor*, are usually accompanied with a replacement of one AS Z by its neighbor X. And thus X may adjoin Z's another neighbor Y on a traceroute path so that a false link $X-Y$ comes into being. From this point, the n2n pattern of a false link starting from a monitored AS can be looked as the expected BGP counterpart of this false link traveling through the two ends with an interval of one AS hop.

In Figure 11(a), we enumerate all of the n2n patterns from one AS to its neighbor's neighbor, where each AS is numbered according to its corresponding priority. The accuracy of n2n patterns on no-valley paths and some valley paths from monitored ASes do not suffer from the incompleteness of BGP view, but the pattern 6 and pattern 8 in valley paths do, since the peer-peer links in those patterns are prone to be invisible in the BGP view.

The distribution of n2n patterns of false links grouped by the network type of the monitored ASes is plotted in Figure 12. The false links without n2n pattern, *i.e.* without any common neighbor between two ends, is also counted in (6 on X-axis). To look for the significance in those distributions, we also draw the distribution of n2n patterns of real links attached to monitored ASes, *i.e.* links in the ground truth. The observations we obtain from Figure 12 are as follows:

Existence of pattern: Most of false links have a no-valley n2n pattern (small value when X=6), that confirms our observation that the false links mostly occur among neighbors. It also confirms that in most of cases there does exist a no-valley path traveling through the two ends of a false link with an interval of one AS hop. The exception is in false links in *Sidewalk*, out of which 1135 (38%) have no n2n pattern, *i.e.* the distance between two ends of each false link is more than two hops away. This indicates that the heuristics of *Sidewalk* trend to delete 'close' false links instead of 'long-distance' ones.

Major pattern: The majority (40%∼70%) of false links on Tier-1/large ISPs have a n2n pattern of (p-c)(p-c). In other words, as illustrated in Figure 11(b), false links trend to occur when the immediate customers of Tier-1/large ISPs are replaced by their own customers or those Tier-1/large ISPs. A typical scenario is that customers use the IP addresses assigned from their providers to number their own routers' interfaces which are connected to their provider's routers. For false links on small ISPs and stubs, the major pattern (30%∼55%) is (p-p)(p-c). Similarly, the scenario is that between two ASes with peer-peer relationship, one numbers its interfaces with IP addresses shared by the other, leading to a false link as illustrated in Figure 11(c).

The difference in major patterns of false links between Tier-1/large ISPs and small ISPs/stubs can be explained by the difference in the number of downstream customers, which is coincident with the metric for our classification of network types. More downstream customers one has, more opportunities one has to replace its own customers. Alternatively, one may replace its peers more probably.

Comparison with real links: Focusing on real links, we see that most (61%∼80%) of links on non-tier-1 ASes have a (p-p)(p-c) pattern, that reveals an interesting phenomenon that, roughly speaking, the most neighbors of a non-tier-1 AS are its peers' customers. At the same time, there are much fewer (0%∼14%) real links with (p-c)(p-c) patten, that indicates that non-tier-1 ASes are unlikely connected with its customers' customers.

These observations are different with those on false links attached to large ISPs, where the major pattern is (p-c)(p-c). Specifically, on Tier-1s in *Ark* and *DIMES*, the percentage of false links with a (p-c)(p-c) patten is 66%∼70%, while that of real links is 43%; on large ISPs, this percentage of false links is 40%∼51%, while that of real links is 14%. In addition, the false links left in *Sidewalk* have a 'long-distance', contrast to real links with a 'short-distance'.

In summary, the false links in traceroute measurements usually involve neighboring ASes. False links between Tier-1/large ISPs and their customers' customers appear more

frequently than real links do. The heuristics of *Sidewalk* trend to delete 'close' false links rather than 'long-distance' ones.

VIII. Limitations

The limitations of this work mainly come from two aspects: (1) *flaws* in the datasets that are used in cause inference, and (2) *violations* of our assumption that traceroute AS path trends to follow BGP AS path as closely as possible. First, the accuracy of cause inference heavily depends on the quality of datasets, even though our results have shown that counting mismatch fragments is more robust to *flaws* than counting AS paths. Even if we assume new datasets become available which prove that some *flaws* are not negligible, it will be easy to feed our framework with those new datasets to correct the results. Second, *violations* which were not inferred as *Divergence* cases have brought the false positives to the results of cause inference, even though our heuristic for *Divergence* is aggressive. This limitation remains a challenge to the research community, as it seems out of reach at this time to either obtain access to the routing tables of all ASes to get the complete routing paths, or infer data paths with accurate IP2AS mapping.

In spite of the above limitations, we believe the two main conclusions in this paper are sound. First, the major pitfall is *Neighbor* whose contribution is nearly five times greater than the runner-up. It seems unlikely that the impact of *flaws* and *violations* is great enough to change the major cause. Second, our estimation of the inaccuracy of publicly available topology datasets does not relate to *violations*, but may suffer from *flaws* in the ground-truth on BGP monitored ASes. According to our previous work [7], the *flaws* are small, *e.g.* 1.5% of links of a Tier-1 missing from the ground truth due to the aggregation of small prefixes.

IX. Conclusion

In this paper we have developed a systematic framework to identify and quantify pitfalls in traceroute-based AS-level topology measurement. Decoupling procedures from evidences, this framework enabled us to obtain more accurate evaluation on errors in topology inference than previous works. Our results shed light into the major pitfalls of current practices in AS topology inferences and the limitations of publicly available AS topology sets. Since most of the inconsistencies originate from IP address sharing between neighbor ASes, we believe that building an accurate database of router interface aliases can bring significant improvement to the accuracy of the router path to AS path conversion process. We hope that, while revealing current imperfections, this work can prompt future works towards the goal of combining BGP and traceroute data to obtain a more accurate AS-level topology.

An enlightenment we acquire from this work is that the structure of Internet is so huge and complex that '*there are all kinds of fish in the sea of Internet*'. This inherent nature of Internet has made many measurement researches develop complicated methodologies to take care of various corner cases, even including fictive ones, to guarantee the soundness of results. We argue that developing a systematic framework which can decouple approaches from datasets can be an effective means in handling the complicated reality of Internet.

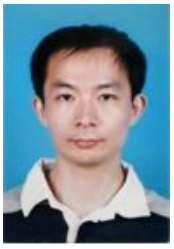
References

[1] "RouteViews routing table archive," http://www.routeviews.org/.
[2] "RIPE routing information service project," http://www.ripe.net/.
[3] H. Chang, R. Govindan, S. Jamin, S. J. Shenker, and W. Willinger, "Towards capturing representative AS-level Internet topologies," *Elsevier Computer Networks Journal*, vol. 44, no. 6, pp. 737–755, 2004.
[4] H. Chang and W. Willinger, "Difficulties measuring the Internet's AS-level ecosystem," in *Annual Conference on Information Sciences and Systems (CISS'06)*, 2006, pp. 1479–1483.
[5] D. Raz and R. Cohen, "The Internet dark matter: on the missing links in the AS connectivity map," in *Proc. IEEE INFOCOM*, 2006.
[6] M. Roughan, S. J. Tuke, and O. Maennel, "Bigfoot, sasquatch, the yeti and other missing links: what we don't know about the AS graph," in *Proc. 8th ACM SIGCOMM conference on Internet measurement*, ser. IMC '08. ACM, 2008, pp. 325–330.
[7] R. Oliveira, D. Pei, W. Willinger, B. Zhang, and L. Zhang, "The (in)completeness of the observed Internet AS-level structure," *IEEE/ACM Trans. Netw.*, vol. 18, no. 1, pp. 109–122, 2010.
[8] "Archipelago Measurement Infrastructure," http://www.caida.org/.
[9] Y. Shavitt and E. Shir, "DIMES: Let the Internet measure itself," *ACM SIGCOMM Computer Comm. Review (CCR)*, 2005.
[10] Y. He, G. Siganos, M. Faloutsos, and S. Krishnamurthy, "Lord of the links: a framework for discovering missing links in the Internet topology," *IEEE/ACM Trans. Netw.*, vol. 17, pp. 391–404, April 2009.
[11] K. Chen, D. R. Choffnes, R. Potharaju, Y. Chen, F. E. Bustamante, D. Pei, and Y. Zhao, "Where the sidewalk ends: extending the Internet AS graph using traceroutes from P2P users," in *Proc. 5th international conference on Emerging networking experiments and technologies*, ser. CoNEXT '09. ACM, 2009, pp. 217–228.
[12] H. Chang, S. Jamin, and W. Willinger, "Inferring AS-level Internet topology from router-level path traces," in *SPIE ITCom*, 2001.
[13] L. Amini, A. Shaikh, and H. Schulzrinne, "Issues with Inferring Internet Topological Attributes," in *Proc. SPIE*, 2002.
[14] Y. Hyun, A. Broido, and kc claffy, "Traceroute and BGP AS path incongruities," CAIDA, Tech. Rep., 2003.
[15] Z. M. Mao, J. Rexford, J. Wang, and R. H. Katz, "Towards an accurate AS-level traceroute tool," in *Proc. ACM SIGCOMM*, 2003.
[16] Y. Zhang, R. Oliveira, H. Zhang, and L. Zhang, "Quantifying the pitfalls of traceroute in AS connectivity inference," in *Proc. Passive and Active Measurement Conference (PAM) 2010*. Springer, 2010.
[17] R. Bush, O. Maennel, M. Roughan, and S. Uhlig, "Internet optometry: assessing the broken glasses in Internet reachability," in *Proc. 9th ACM SIGCOMM conference on Internet measurement conference*, ser. IMC '09. ACM, 2009, pp. 242–253.
[18] X. Zhao, D. Pei, L. Wang, D. Massey, A. Mankin, S. F. Wu, and L. Zhang, "An analysis of BGP multiple origin AS (MOAS) conflicts," in *Proc. 1st ACM SIGCOMM Workshop on Internet Measurement*, ser. IMW '01. ACM, 2001, pp. 31–35.
[19] Z. M. Mao, D. Johnson, J. Rexford, J. Wang, and R. H. Katz, "Scalable and accurate identification of AS-level forwarding paths," in *INFOCOM 2004*, 2004.
[20] Y. Hyun, A. Broido, and kc claffy, "On third-party addresses in traceroute paths," in *Proc. Passive and Active Measurement Workshop (PAM)*, 2003.
[21] J.-J. Pansiot, P. Mérindol, B. Donnet, and O. Bonaventure, "Extracting intra-domain topology from mrinfo probing," in *Proc. Passive and Active Measurement Conference (PAM) 2010*. Springer, 2010.
[22] B. Huffaker, A. Dhamdhere, M. Fomenkov, and kc claffy, "Toward topology dualism: Improving the accuracy of AS annotations for routers," in *Proc. Passive and Active Measurement Conference (PAM) 2010*. Springer, 2010.
[23] M. E. Tozal and K. Sarac, "Tracenet: an internet topology data collector," in *Proc. 10th annual conference on Internet measurement*, ser. IMC '10. New York, NY, USA: ACM, 2010, pp. 356–368.
[24] M. Luckie, "Scamper: a scalable and extensible packet prober for active measurement of the Internet," in *Proc. 9th ACM SIGCOMM conference on Internet measurement conference*, ser. IMC '10. ACM, 2010.
[25] B. Augustin, X. Cuvellier, B. Orgogozo, F. Viger, T. Friedman, M. Latapy, C. Magnien, and R. Teixeira, "Avoiding traceroute anomalies with paris traceroute," in *IMC '06*, 2006.
[26] "Internet Routing Registry," http://www.irr.net/.
[27] "PeeringDB website," http://www.peeringdb.com/.
[28] "Packet clearing house IXP directory," http://www.pch.net/ixpdir/.

[29] "European Internet exchange association," http://www.euro-ix.net.
[30] H. Madhyastha, T. Isdal, M. Piatek, C. Dixon, T. Anderson, A. Krishnamurthy, and A. Venkataramani, "iPlane: an information plane for distributed services," in *Proc. of OSDI*, 2006.
[31] B. Zhang, R. Liu, D. Massey, and L. Zhang, "Collecting the Internet AS-level topology," *ACM SIGCOMM Computer Comm. Review (CCR)*, vol. 35, no. 1, pp. 53–61, 2005.
[32] R. Oliveira, D. Pei, W. Willinger, B. Zhang, and L. Zhang, "In search of the elusive ground truth: The Internet's AS-level connectivity structure," in *Proc. ACM SIGMETRICS*, 2008.
[33] L. Gao, "On inferring autonomous system relationships in the Internet," *IEEE/ACM Trans. Netw.*, vol. 9, no. 6, pp. 733–745, 2001.
[34] J. W. Hunt and M. D. McIlroy, "An algorithm for differential file comparison," Bell Laboratories, Tech. Rep., 1976.
[35] B. Augustin, B. Krishnamurthy, and W. Willinger, "IXPs: mapped?" in *Proc. 9th ACM SIGCOMM conference on Internet measurement conference*, ser. IMC '09. ACM, 2009, pp. 336–349.
[36] X. Cai, J. Heidemann, B. Krishnamurthy, and W. Willinger, "Towards an AS-to-organization map," in *Proc. 9th ACM SIGCOMM conference on Internet measurement conference*, ser. IMC '10. ACM, 2010.

Yu Zhang received the B.S., M.S., and Ph.D degree in Computer Science from Harbin Institute of Technology (HIT), Harbin, China, from 1998 to 2009. From 2008 to 2009, he was a visiting scholar at University of Massachusetts, Amherst, and University of California, Los Angeles. He currently is a lectuer in Computer Science and Technology Department of HIT. His research interests include Internet topology measurement, inter-domain routing and complex network.

Ricardo Oliveira received the B.S. in Electrical Engineering from the Engineering Faculty of Porto University (FEUP), Portugal in 2001, the M.S. degree in Computer Science from University of California, Los Angeles in 2005, and the Ph.D degree in Computer Science from UCLA in 2009. During his PhD he did several internships, including AT&T Labs, Thomson and Juniper Networks in Silicon Valley. During his PhD he developed research in inter-domain routing, network security and Internet topology. He is the main author of 4 patents and dozens of papers in international conferences and journals, including ACM SIGCOMM, ACM SIGMETRICS and IEEE Transactions on Networking. He is a member of the ACM and the IEEE. In 2009 he was the recipient of the Cisco Outstanding Graduate Student Research Award and the Edward K. Rice Outstanding Doctoral Student Award. In 2009 he co-founded ThousandEyes, a startup in the application performance space.



Yangyang Wang received his B.S. degree in Computer Science and Technology from Shandong University, Jinan, China, in 2002 and his M.S. degree from Capital Normal University, Beijing, China, in 2005. He is currently a Ph.D. candidate in Department of Computer Science and Technology at Tsinghua University, Beijing, China. His research interests include Internet routing architecture and Future Internet design.

Shen Su received his B.S. degree and M.S degree in Computer Science and Technology from Harbin Institute of Technology (HIT), Harbin, China, from 2004 to 2010. He is currently a Ph.D. candidate in Department of Computer Science and Technology at HIT. His research interests include network measurement and wireless network.

Baobao Zhang received his B.S. degree in Computer Science and Technology from Beijing Jiaotong University, Beijing, China, in 2010. He is currently a Ph.D. candidate in Department of Computer Science and Technology at Tsinghua University, Beijing, China. His research interests include network architecture and IPv6.

Jun Bi received the B.S., M.S., and Ph.D. degree in Computer Science from Tsinghua University, Beijing, China, from 1990 to 1999. From 2000 to 2003, he was a research scientist of Bell Labs Research Communication Science Division and Bell Labs Advanced Communication Technologies Center. Currently he is a professor and the director of Network Architecture & IPv6 Research Division, Network Research Center of Tsinghua University. He served as chair and member of technical program committee of many conferences, including ICNP, CoNext and NGI. He is a senior member of ACM, a fellow of IARIA, and the secretariat director of Asia Future Internet Forum (AsiaFI).

Hongli Zhang received her B.S. degree in Computer Science from Sichuan University, Chengdu, China in 1994, and her Ph.D. degree in Computer Science from Harbin Institute of Technology (HIT), Harbin, China in 1999. She is a professor in Computer Science and Technology Department of HIT and the vice director of National Computer Information Content Security Key Laboratory. Her research interests include network and information security, network measurement and modeling, and parallel processing.

Lixia Zhang received her Ph.D in computer science from the Massachusetts Institute of Technology and joined the Xerox Palo Alto Research Center as a member of the research staff. She joined the faculty of UCLA's Computer Science Department in 1996. In the past she has served as the vice chair of ACM SIGCOMM, member of the editorial board for the IEEE/ACM Transactions on Networking, and the Internet Architecture Board. Zhang is a fellow of ACM and a fellow of IEEE, and recipient of 2009 IEEE Internet Award. She is currently leading a research team working on a new Internet architecture design, Named Data Networking (http://named-data.net).