Detecting Third-party Addresses in Traceroute Traces with IP Timestamp Option

Pietro Marchetta, Walter de Donato, and Antonio Pescapé

University of Napoli Federico II (Italy)
{pietro.marchetta,walter.dedonato,pescape}@unina.it

Abstract. Traceroute is one of the most famous and widely adopted diagnostic tool for computer networks. Although traceroute is often used to infer links between Autonomous Systems (ASes), the presence of the so-called *third-party* (TP) addresses may induce the inference of false AS-level links. In this paper, we propose a novel active probing technique based on the IP timestamp option able to identify TP addresses. For evaluating both the applicability and the utility of the proposed technique, we perform a large-scale measurement campaign targeting – from multiple vantage points – more than $327K$ destinations belonging to about $14K$ ASes. The results show how TP addresses are very common and affect about 17% of AS-level links extracted from traceroute traces. Compared to a previously proposed heuristic method, our technique allows to identify many more TP addresses and to re-interpret part of its results.

1 Introduction

An accurate knowledge of the Internet topology is essential for a deep understanding of such a complex and ever-evolving system [7, 11, 19, 20]. In the last decade many attempts have been done to overcome the incompleteness of BGP-derived AS-level topologies [12] using traceroute [4, 8, 13]. However, traceroute is known to be inaccurate and to induce errors when its results are used to infer the Internet topology [15, 17, 27].

One source of inaccuracy is represented by the so called *third-party* (TP) addresses [14, 18], i.e. addresses associated to interfaces which are not actually traversed by the IP packets sent toward the traceroute destination. While several other causes may impact the accuracy of AS links derived from traceroute – such as divergence between data and control paths, anonymous hops, unmapped hops, Internet exchange points (IXPs), multi-origin AS prefixes, and siblings – TP addresses (when shared between peering AS neighbors) were recently defined by Zhang et al. [27] as "the last and the most difficult cause to be inferred" and as "a huge obstruction towards the accuracy of traceroute measurements". Several works, by using heuristic methods, tried to deal with such issues with different objectives: to explain the mismatches between BGP- and traceroute-derived AS paths [8, 27], or to complement the AS-level topology inferred from BGP repositories [4, 8, 13]. However, to the best of our knowledge, only two

works tried to isolate and study the phenomenon of TP addresses in order to quantify their impact, achieving different conclusions. By adopting a heuristic method based on IP-to-AS mapped traceroute traces, Hyun et al. [14] conclude that TP addresses mostly appear at the border of multi-homed ASes and cannot be a significant source of AS map distortion. On the other hand, by using pre-computed AS-level graphs and pre-acquired knowledge about routers interfaces, Zhang et al. [27] conclude that TP addresses cause 60% of mismatches between BGP- and traceroute-derived AS paths, where mismatches affect from 12% to 37% of the paths depending on the vantage point.

In this paper, for shedding light on this controversial topic, we propose the first active probing technique able to directly detect the presence of TP addresses in traceroute IP paths. Our technique is based on the IP prespecified timestamp option [5] and requires no previous knowledge about routers interfaces, nor AS paths provided by BGP or IP-to-AS mapping. Performing a large scale measurement campaign, we evaluate the technique showing that: (i) the same IP address may be a TP or not depending on both the source and the destination of the IP path; (ii) TP addresses affect 17% of the AS links extracted from our dataset and (iii) they appear in a significant portion of the detected AS-level loops. We further compare our technique with the method proposed by Hyun et al. [14], which is the only other method not using AS paths extracted from BGP. The comparison reveals that only 1.5% of IP addresses detected as TP by our technique are recognized as such by their heuristic, explaining the underestimation of the phenomenon.

The paper is organized as follows. Sec. 2 introduces TP addresses and explain their effect when traceroute is used to infer topological information; Sec. 3 presents our active probing technique to identify TP addresses in traceroute traces; Sec. 4 describes the methodology adopted to evaluate the proposed technique as well as the main findings; Sec. 5 concludes the paper.

2 Understanding TP addresses and their impact

The RFC1812 [5] states that the source address of an ICMP error packet should correspond to the outgoing interface of the ICMP reply, rather than the interface on which the packet triggering the error was received [14]. This behavior can cause a traceroute IP path to include addresses associated to interfaces not included in the path actually traversed. For instance, the trace from S to D in Fig. 1 contains the sequence (a, b, c) of IP addresses (hereafter IPs), where a and b are associated to the incoming interfaces of routers A and B respectively, and c is the interface used by router C to send ICMP replies to the traceroute originator. The IP c is a TP address since it is associated – in this specific trace – to an interface not effectively traversed by the packets sent from S to D.

The occurrence of TP addresses can have a significant impact on some traceroute applications. The major impact is related to the inference of AS-level links from traceroute traces: as shown in previous works [14, 27], TP addresses may cause the inference of false AS links. Consider again Fig. 1: if the IP address b

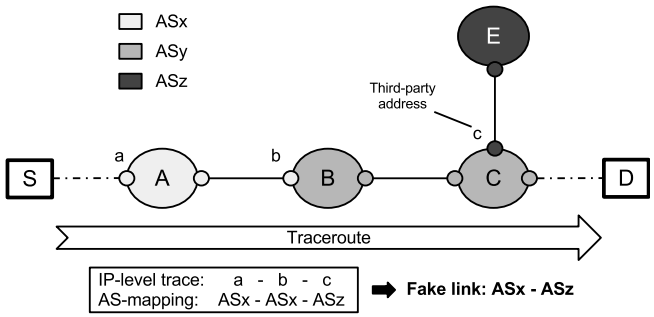

Fig. 1: TP addresses inducing the inference of false AS links

belongs to ASx, and c belongs to the ASz addressing space, then the IP-to-AS mapping of the trace will induce the inference of a false AS link, i.e. $ASx - ASz$. Note also how the TP address hides the ASy which, tough traversed, does not appear in the mapped AS-level trace.

While TP addresses may also impact subnet positioning [26] and alias resolution [25], forcing the adoption of several complex heuristics, in this paper we focus on their impact on the AS-level links inferred from traceroute traces.

3 Detecting TP addresses

Our technique only requires two probes to understand if an IP address discovered by traceroute lies on the path (OP) or not (TP).

Basic principles. Our technique is based on the IP prespecified timestamp (TS) option [23]. It allows to prespecify in a single packet up to four IP addresses from which a timestamp is requested. Hereafter, we adopt the notation PROBE X|ABCD introduced in [24], where PROBE is the probe type, X is the targeted destination and ABCD is the ordered list of prespecified IPs from which a timestamp is requested[1].

Thanks to a large-scale measurement campaign targeting more than 1.7M IP addresses [10], we detected that most routers (including Cisco devices), when processing such option, insert one timestamp every time the probe passes through the interface associated to the prespecified address. Such behavior can be easily detected by targeting Y with an $\text{ICMP}^{echo}_{request}$ Y|YYYY probe. According to [10], if the $\text{ICMP}^{echo}_{reply}$ message contains 1 timestamp, it means that the interface Y was only traversed by the probe when entering the router. If it contains 2 timestamps, Y was traversed by the probe providing either one timestamp when both entering and leaving the router or two timestamps just when entering. Finally, 3 timestamps occur if the probe was stamped twice when entering the router, but only once when leaving it. In such three cases the targeted router exposes a per network interface stamping behavior, which can be exploited to understand if a traceroute hop is part or not of the forward IP path.

[1] The order implies that B cannot insert its own timestamp before A, and so on.

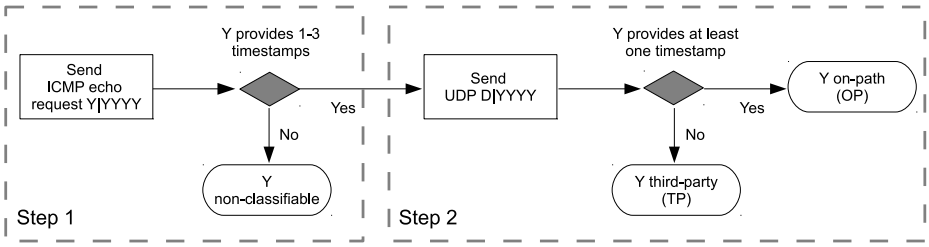

Fig. 2: Classification of the hop Y discovered by traceroute toward D.

TP address detection technique. In order to understand if the hop Y discovered by traceroute toward D is a TP address, the proposed technique works according to the following steps (see Fig. 2): (1.) it targets Y with an $\text{ICMP}^{echo}_{request}$ Y|YYYY probe to verify if it is classifiable or not (see below); (2.) if Y is classifiable, it targets D with UDP D|YYYY[2]: if the TS option brought back into the payload of the $\text{ICMP}^{port}_{unreach}$ message contains at least one timestamp, Y is classified as OP, otherwise it is a TP address.

The first step is necessary because there are other less common router behaviors that may lead the technique to misleading results. Indeed, adopting a conservative approach, a traceroute hop Y is considered *non−classifiable* every time there is no clear evidence that its router has a per network interface stamping behavior, as in the following circumstances:

- **Private address (PVT)**: Y is part of a private addressing block and it may be unreachable by the $\text{ICMP}^{echo}_{request}$ message or it may be employed in different networks along the path toward the destination. In the latter case, a timestamp in the $\text{ICMP}^{port}_{unreach}$ message may be inserted by a different router.
- **Lack of reply (NO−REP)**: no reply is received to $\text{ICMP}^{echo}_{request}$ Y|YYYY, thus either the targeted device dropped the probe or the reply was filtered along the path[3].
- **The TS option is removed (NO−OPT)**: the $\text{ICMP}^{echo}_{reply}$ message received from Y contains no TS option, thus either the targeted hop did not replicate the option in the reply or the option was removed along the path.
- **Zero timestamps (NO−TS)**: the targeted device simply ignores the TS option, without inserting any timestamp in the $\text{ICMP}^{echo}_{reply}$ message.
- **Four timestamps (JUN)**: the targeted device provides 4 timestamps. Such behavior has been already observed in the case of Juniper routers, which insert their timestamp also when the prespecified address is associated to *any* owned interface [10]. Hence, the presence of a timestamp in the $\text{ICMP}^{port}_{unreach}$ message obtained during the second step would not allow to classify Y.

In other words, a traceroute hop Y is considered *classifiable* only if it provides from 1 to 3 timestamps when directly probed with $\text{ICMP}^{echo}_{request}$ Y|YYYY.

[2] UDP probes allow to avoid ambiguities caused by the reverse path [10].

[3] In [10], we observed how equipping classic active probes with the TS option causes a strong reduction of the responsiveness.

We also implemented and made publicly available [4] an enhanced traceroute version, based on paris−traceroute [3], which applies our technique to classify the hops discovered along the path toward the destination.

4 Experimental evaluation

In this section, we describe the large scale measurement campaign conducted to evaluate the proposed technique as well as the main findings.

4.1 Measurement campaign

To evaluate our technique, we selected more than $327K$ destinations in $14K$ ASes among the ones showing stable responsiveness to both ping, according to the PREDICT project [2], and UDP probes carrying the TS option[5]. To perform a large scale measurement campaign, we used 53 PlanetLab nodes [6] located in different ASes as vantage points (hereafter VPs).

In particular, each node was instructed to (1.) send UDP probes toward the destinations and select those which reply and preserve the TS option; (2.) launch UDP paris-traceroute toward the selected destinations; (3.) launch an $\text{ICMP}^{echo}_{request}$ Y$\big|$YYYY toward each intermediate hop Y; (4.) select the classifiable hops as the ones providing 1−3 timestamps; (5.) send an UDP probe toward the traceroute destination prespecifying each time a different classifiable hop collected on the path. In order to avoid ambiguities caused by load balancers, the UDP probes used to classify the hops and the ones generated by traceroute are crafted as part of the same flow according to [3].

After removing the traces affected by filtering, the final dataset – publicly available [4] – consisted of ∼12M traces for a total number of $\sim 443K$ addresses.

4.2 Main findings

Since every VP traced IP paths toward the same destinations, a specific IP address may be discovered by multiple VPs: this happens especially for those located close to the destinations. Fig. 3 shows how many distinct VPs discovered the same IP address: more than 96% of IPs were captured by at least two VPs, while about a half were captured by more than 35 VPs.

Hops classifiability. When an IP address is captured by multiple VPs, each node independently states if it is classifiable or not. However, the TS option may trigger the filtering of the $\text{ICMP}^{echo}_{reply}$ message on some paths inducing a subset of VPs to consider the targeted device as non−classifiable (NO−REP). Fig. 4 reports the number of nodes not receiving replies from a device which successfully replied to at least one VP: only 15% of addresses did not experience such in-transit filtering, while on average 4 VPs were forced by filtering to consider a device as non−classifiable. We can conclude that the number of VPs is a key point for applications based on the TS option [16, 24].

[4] `http://traffic.comics.unina.it/tpa/`

[5] According to a campaign conducted from our laboratory at University of Napoli.

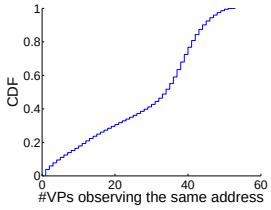

Fig. 3: VPs observing each IP of the dataset.

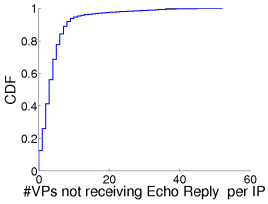

Fig. 4: In−transit filtering.

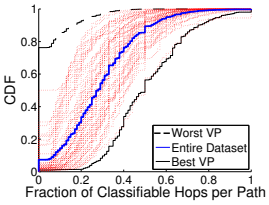

Fig. 5: Classifiable hops per traceroute trace.

When some VPs labeled an IP address as non−classifiable and the other VPs judged the same address as classifiable, we did not consider it as a conflict. Our VPs unanimously agreed about more than 97% of IPs labeling 51% of addresses as classifiable and 47.6% as non−classifiable. Conflicting verdicts regarded a limited number of IPs (1.4%) and were mainly caused by the removal of the TS option on some reverse paths. Tab. 1 reports a breakdown of non−classifiable IPs per category (see Sec. 3): our technique was unable to classify such IPs mostly because of devices not replying (16.4%), ignoring the TS option (14.6%), or belonging to the JUN category (10.4%). We also found 9 IPs exposing multiple behaviors to distinct VPs, mainly caused by non−RFC compliant routers (a phenomenon deeply investigated in [10]).

Besides non−classifiable hops, more than a half of IPs in the dataset were classifiable by our technique. Adopting a per-trace point of view, Fig. 5 shows the fraction of classifiable hops per trace (*i*) for each VP and (*ii*) over the entire dataset: on average 4%, 52% and 30% hops are classifiable in each trace respectively by the most filtered node (Worst VP), the less filtered one (Best VP) and over the entire dataset. As reported in the following, although not all the hops in each trace are classifiable, our technique allows to investigate the TP addresses impact on traceroute applications.

Classification results. Most classifiable hops appeared in several paths from multiple VPs toward multiple destinations. Fig. 6 shows the percentage of classifiable IPs always classified as TP or OP and those classified as both (Mix), on the paths in which they appeared. Such paths are aggregated in three different ways: paths originated (1.) by the same VP toward multiple destinations, (2.) by multiple VPs toward a single destination, (3.) by multiple VPs toward multiple destinations. The obtained results highlight an unexpected general trend: most traceroute traces contain many more TP than OP addresses. Hence, according to the router behavior described in Sec. 3, most of the intermediate routers encountered along the path reply to the traceroute originator using an interface different from the ones traversed by the packets sent to the targeted destination. For both the aggregations 1 and 2, most of addresses were always classified as TP or OP. However, some IPs were also variably classified and this phenomenon is much more important in the aggregation 3. Such an evidence allows to conclude that *the same address discovered with traceroute may lie or not on the IP path depending on the (i) originating node and (ii) the targeted destination, essentially due to both inter- and intra-domain routing.*

Table 1: Root cause analysis of non−classifiable IPs.

Category (Sec. 3)	IPs	%IPs
PVT	9,428	2.2
NO−REP	72,775	16.4
NO−TS	64,641	14.6
JUN	45,963	10.4
NO−OPT	18,039	4
Multiple Behaviors	9	∼0
Non−classifiable IPs	**210,885**	**47.6**

Impact on derived AS links. While $224K$ IPs were classified at least once as TP address, not all the TP addresses impact the AS-level links derived from traceroute. Mapping each hop to the owner AS [9], we identified in our dataset $14,783$ different ASes. In order to avoid ambiguities caused by the presence of IXPs, we removed from our traces the hops associated to them according to the datasets provided by *peeringDB* [22] and *PCH* [21]. From the resulting $34,414$ AS-level links, we removed 38 links involving sibling ASes according to [1].

Taking into account that the same AS link may appear in several traces toward distinct destinations and depending on the involved IPs, a single AS link may be associated to multiple classifications according to how the two involved IPs were classified each time by our technique. In order to deal with this phenomenon, we applied the following methodology: (1.) if both the involved IPs were classified as OP at least once, we are confident that the corresponding AS link actually exists; else, by adopting a conservative approach, (2.) if both the involved IPs were non−classifiable by our technique at least once, we consider the link as possible; finally, (3.) the AS links which always involved at least one TP address are considered potentially false (see link ASx−ASz in Fig. 1). We counted $1,897$ existing links and $25,990$ possible links. On the other hand, we found $6,299$ potentially false AS links corresponding to about 17% of the links extracted from the dataset.

AS-level loops. False AS links caused by TP addresses may also generate bogus AS-level loops. In our dataset, we registered $587,126$ traces normally reaching

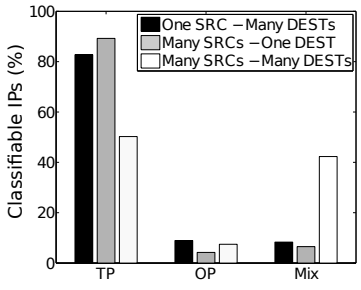

Fig. 6: Addresses classification.

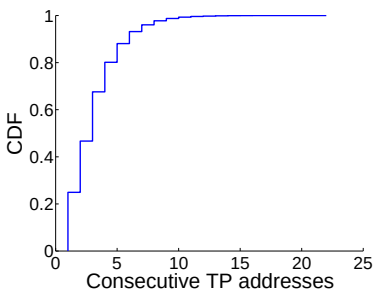

Fig. 7: TP address patterns.

the destination, in which an AS-level loop appeared. Among these traces, about 4,144 loops involved sibling ASes. Thanks to our technique, we discovered that TP addresses are involved in at least 37% of such loops[6]: $105K$ and $149K$ loops respectively started or ended with a TP address, while 6,083 loops involved a sequence of consecutive TP addresses. For instance, considering the AS1 AS2 AS3 AS1 sequence, if AS2 and AS3 are associated to TP addresses, one possibility is that the corresponding path is entirely contained in AS1, thus generating a bogus loop.

4.3 Implications of the results of our technique

The surprising high value of potentially false AS links suggests that TP addresses can be a significant source of AS maps distortion. Such conclusion confirms the one drawn by Zhang et al. [27] and is totally different from the one given by Hiun et al. [14]. Here, we investigated the basic reasons of such contradiction. According to the heuristic method proposed by Hiun et al., a *candidate* TP address is an intermediate hop that resolves to an AS that differs from the ASes of both adjacent IPs in the same path. The method takes into account also path stability, AS ownerships and hostnames.

On the one hand, applying the Hiun's method on our dataset, 7,457 IPs were classified as candidate TP addresses. Such addresses appeared in 56,595 different IP1 IP2 IP3 sequences where all the IPs were mapped to different ASes and IP2 represents the candidate TP address. Each sequence appeared in multiple traces and each time the involved IPs were classified by our technique[7]: (i) 166 sequences resulted as real AS1 AS2 AS3 transitions, since all the three IPs were classified at least once as OP; (ii) although the candidate TP address was non−classifiable by our technique in 39,824 sequences, in 15,850 of them we recognized as TP address the previous or the next hop, which could be the real responsible of a false AS link; (iii) in the remaining 16,605 sequences, our technique always classified the central address as TP in 85% of cases (the two techniques validate each other in such cases) and as OP in 14% of sequences (in contradiction to the response of the Hiun's method). In the last case, we also found 52 sequences classified as both TP and OP depending on the traceroute destination and the VP used.

On the other hand, only 1.5% of the TP addresses identified by our technique is detected by the Hiun's method. The main reason is that a TP address is such independently from the AS point of view. In addition, a traceroute path may contain multiple consecutive TP addresses – a possibility considered *remote* in [14]. Considering the sequences of consecutive TP addresses detected in our traces, Fig. 7 shows the distribution of their lengths. Globally, we registered $680K$ unique sequences: about 25% were isolated TP addresses, but more than a half consisted of more than 3 consecutive TP addresses. As for ASy in Fig. 1,

[6] Since we used a conservative approach, the real impact may be potentially wider.

[7] As described above, the address identified by Hyun as candidate TP address may effectively lie or not on the IP path depending on the source and the destination.

if a traceroute path only crosses border routers exposing TP addresses mapped to other ASes, consecutive TP addresses may entirely hide an AS from the path.

5 Conclusion

In this paper, we presented and evaluated – to the best of our knowledge, for the first time in literature – an active probing technique able to identify TP addresses in traceroute traces. Differently from most previous works, our technique does not rely on information provided by BGP monitors and it allows to conclude that TP addresses can be a significant source of AS map distortion. Thanks to a large scale measurement campaign, we draw the following general conclusions: (i) the same address may be a TP address or not depending on the originating host and the targeted destination; (ii) TP addresses may also be responsible for bogus AS-level loops. We further observed that our technique was able to classify more than half of the total discovered IPs and, surprisingly, about 17% of traceroute-derived AS-level links were affected by TP addresses, being thus potentially false. Finally, our results confirmed the conclusion drawn by Zhang et al. [27] on the severity of this phenomenon and allowed to explain why such conclusion conflicts with the one achieved by Hyun et al [14]: on our dataset, their heuristic method was able to discover only 1.5% of the TP addresses recognized by our technique.

In our ongoing work, we aim at quantifying the magnitude of the map distortion introduced when combining traceroute- and BGP-derived information to infer the AS-level topology of Internet. We also plan to investigate if and how TP addresses can explain known incongruities, such as *extra*, *missing*, and *substitute* hops arising when comparing the AS paths derived from traceroute with the ones extracted from BGP monitors [27].

Acknowledgements. The work of the authors is partially funded by the PLATINO (PON01_01007) and S^2−MOVE (PON04a3_00058) projects financed by MIUR.

References

1. The CAIDA AS Relationships Dataset, June 2012. *http://www.caida.org/data/active/ as-relationships/*.
2. IP Address Hitlist, PREDICT ID USC-LANDER internet- address- hitlist- it47w-20120427, 2010-03-29 to 2012-05-30. *http://www.isi.edu/ant/lander*.
3. B. Augustin, X. Cuvellier, B. Orgogozo, F. Viger, T. Friedman, M. Latapy, C. Magnien, and R. Teixeira. Avoiding traceroute anomalies with paris traceroute. In *Proc. ACM SIGCOMM IMC*, 2006.
4. B. Augustin, B. Krishnamurthy, and W. Willinger. IXPs: mapped? In *ACM SIGCOMM IMC*, 2009.
5. F. Baker. IETF RFC1812: Requirements for IP version 4 routers.
6. A. Bavier, M. Bowman, B. Chun, D. Culler, S. Karlin, S. Muir, L. Peterson, T. Roscoe, T. Spalink, and M. Wawrzoniak. Operating system support for planetary-scale network services. In *NSDI*, 2004.

7. A. Botta, W. de Donato, A. Pescapè, and G. Ventre. Discovering topologies at router level: Part ii. In *GLOBECOM*, pages 2696–2701, 2007.
8. K. Chen, D. Choffnes, R. Potharaju, Y. Chen, F. Bustamante, D. Pei, and Y. Zhao. Where the sidewalk ends. *Proc. ACM CoNEXT*, 2009.
9. T. Cymru. IP to ASN mapping. *http://www.team-cymru.org/Services/ip-to-asn.html*, 2012.
10. W. de Donato, P. Marchetta, and A. Pescapé. A hands-on look at active probing using the ip prespecified timestamp option. In *Proc. PAM*, 2012.
11. B. Donnet and T. Friedman. Internet topology discovery: a survey. *IEEE Communications Surveys and Tutorials*, 2007.
12. E. Gregori, A. Improta, L. Lenzini, L. Rossi, and L. Sani. On the incompleteness of the AS-level graph: a novel methodology for BGP route collector placement. In *Proc. ACM SIGCOMM IMC*, 2012.
13. 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 Transactions on Networking*, 2009.
14. Y. Hyun, A. Broido, and K.C. Claffy. On third-party addresses in traceroute paths. In *Proc. PAM*, 2003.
15. Y. Hyun, A. Broido, and K.C. Claffy. Traceroute and BGP AS path incongruities. Technical report, CAIDA, 2003.
16. E. Katz-Bassett, H. V. Madhyastha, V. K. Adhikari, C. Scott, J. Sherry, P. van Wesep, T. E. Anderson, and A. Krishnamurthy. Reverse traceroute. In *Proc. NSDI*, 2010.
17. M. Luckie, A. Dhamdhere, D. Murrell, et al. Measured impact of crooked traceroute. *ACM SIGCOMM Computer Communication Review*, 2011.
18. P. Marchetta, W. de Donato, and A. Pescapé. Detecting third-party addresses in traceroute ip paths. In *Proc. ACM SIGCOMM*, 2012.
19. P. Marchetta, P. Mérindol, B. Donnet, A. Pescapè, and J-J. Pansiot. Topology discovery at the router level: A new hybrid tool targeting isp networks. *IEEE JSAC*, 2011.
20. P. Marchetta, P. Mérindol, B. Donnet, A. Pescapé, and J.J. Pansiot. Quantifying and Mitigating IGMP Filtering in Topology Discovery. In *Proc. IEEE GLOBECOM*, 2012.
21. Packet Clearing House. IXP directory. *https://prefix.pch.net/*.
22. PeeringDB. Exchange points list. *https://www.peeringdb.com/*.
23. J. Postel. Internet Protocol. RFC 791 (Standard), September 1981.
24. J. Sherry, E. Katz-Bassett, M. Pimenova, H. V. Madhyastha, T. Anderson, and A. Krishnamurthy. Resolving ip aliases with prespecified timestamps. IMC '10, pages 172–178, New York, NY, USA, 2010. ACM.
25. M. Tozal and K. Sarac. Palmtree: An ip alias resolution algorithm with linear probing complexity. *Computer Communications*, 2010.
26. M. Tozal and K. Sarac. Tracenet: an internet topology data collector. In *Proc. ACM SIGCOMM IMC*, 2010.
27. Y. Zhang, R. Oliveira, Y. Wang, S. Su, B. Zhang, J. Bi, H. Zhang, and L. Zhang. A framework to quantify the pitfalls of using traceroute in as-level topology measurement. *IEEE JSAC*, 2011.