2009 International Conference on Communications and Mobile Computing

Measuring the impacts of sampling bias on Internet AS-level topology inference

Guoqiang Zhang
Computer Network Information Center, Chinese Academy of Sciences
Beijing, China
zhangguoqiang@cnnic.cn

Abstract

In this paper, we carried out a systematic qualitative study of the impacts of sampling bias over Internet AS-level topology inference. Based on the relatively complete Chinese Internet AS graph derived from our large-scale measurement effort, we further quantified such impacts by showing how much peer-peer and provider-customer links were missed due to sampling bias. Moreover, we brought to light the possible incorrect inferences that can be derived about relationships between ASes if the number of sampling points is limited.

keyword*: Internet topology;sampling bias;autonomous system; BGP;AS relationship*

1 Introduction

Accurate Internet topologies are the basis for a wide array of Internet research activities [10, 17, 20–22, 24]. However, as a constantly evolving and distributedly administered system, obtaining an accurate view of it is not an easy task. Two state-of-art measurement methodologies are the passive measurement that restructures the AS graph from a collection of BGP tables and updates, and the active measurement that infers the AS graph from traceroute mapping. However, these two methods both suffer from sampling bias, i.e, the completeness of the graph and its displayed topological properties are closely related to the number of collecting points and where these collecting points are deployed. For example, the prestigious routeviews [1] project has no data collecting point in China and CAIDA's skitter project [2] has only one data collecting point in China, which will inevitably decrease the data's credibility. Thus, we are curious about the following question "Is the data credible? If not, how it deviates from the real topology?". However, the lack of an accurate benchmark map makes the quantitative evaluation of the data's credibility impossible, and moreover, handicaps the proper analysis of how and why the sampled data deviates from the real Internet map. In addition to this, the handful number of routing tables can also hamper the proper inference of AS relationships, which is an important factor in shaping the Internet structure.

Motivated by the aforementioned problems, we try to gain an in-depth understanding of the impacts of sampling bias on the Internet topology inference. We (a)systematically analyze the potential problems of the passive and active measurement approaches, with focus on the classification of sampling biases; (b)present a systematic study of various scenarios that may cause the failure for discovering certain links and incur inaccurate AS relationship inference; (c)use one subgraph of the Internet AS map–the Chinese Internet AS graph–to quantitatively measure the impacts of sampling bias in reality, unexpectedly finding that sampling bias can cause large number of provider-customer links loss.

The rest of the paper is structured as follows. Section 2 gives a brief introduction to related work. Section 3 studies issues related to the two data collection methods. Section 4 presents various scenarios that may cause the miss of links due to sampling bias. Section 5 gives a quantitative study of the impacts of sampling bias. Finally we conclude our work in Section 6.

2 Related Work

Internet topology is the basis for other network related research. However, its accuracy is often questioned because of the sampling methodologies being used. The most prestigious active measurement project–skitter [2], deploys dozens of dedicated servers worldwide to collect traceroute results, however, compared with the enormous Internet scale, these servers are apparently inadequate. The DIMES [18] project takes advantage of the end users' location diversity to address this limitation, but the end users' behaviors are out of their control. The Routeviews [1] project is a passive measurement project, which sets up several collectors, each peering with dozens of BGP speakers located in different ASes, to collect BGP tables and BGP updates. However, it also faces the problem of inadequate collecting

978-0-7695-3501-2/09 $25.00 © 2009 IEEE
DOI 10.1109/CMC.2009.39

IEEE computer society

Authorized licensed use limited to: SICHUAN UNIVERSITY. Downloaded on March 06,2022 at 02:24:07 UTC from IEEE Xplore. Restrictions apply.

points.

Several works have been carried out regarding the influence of sampling bias over the topological properties of the Internet, most importantly, its degree distribution. In [4], the authors argued that traceroute-like methods for collecting the router-level topology can make the sampled degree distribution differ sharply from the underlying graph, and under some conditions, a tree sampling procedure can produce an artificial power-law. However, R.Cohen [7] showed that when the underlying graph is a power-law with exponent larger than 2, the bias of degree distribution calculated from the BGP sampled data is negligible.

Despite the pure topological properties, connections between ASes often bear some commercial relationships. The seminal work towards inferring AS relationship is made by Gao [11]. In this work, Gao characterized the valid AS path as *valley-free* and *non-step*, and devised a heuristic algorithm to infer AS relationships on the assumption that AS size is typically proportional to its degree. Several follow-up works try to formulate the problem and improve the inference accuracy [9, 12, 14].

In view of the unavoidable incompleteness due to sampling bias, researchers are trying to find those missing links and classify them based on their relationships. In [15], graphs derived from three data sources(whois, BGP, traceroute) were quantitatively compared with each other, and it was shown that BGP and traceroute topologies show similar structural properties but differ substantially from the whois topology. In [5], the authors augmented the Routeviews data set with BGP summary information from a large number of Internet Looking Glass sites and with routing policy information from RIPE database, and obtained a much richer AS graph containing approximately 25% more edges. The patterns or types of missing links were also studied in [8, 9, 13, 23], where all confirmed the fact that peer-peer links form the main part of missing links.

3 Classification of Sampling Bias

This section gives a systematic study of the possible problems arising from the two Internet topology collection approaches.

3.1 BGP-based Measurement

BGP is a path vector protocol. Entries in BGP tables and BGP updates contain the AS path attribute, which is initially intended for routing loop avoidance. When a BGP speaker advertises a route to its neighboring AS peers, it appends its own AS number to the tail of the AS path. Therefore, the most advantageous aspect of BGP-based measurement is its simplicity in extracting the AS connectivity–by simply parsing the BGP AS path attributes. In case of encountering private ASes or AS set, the common way is to exclude these entries from the topology construction.

The most disadvantageous aspect of BGP-based measurement is its sampling bias. We classify the sampling bias into two classes: one is *sampling location bias* and the other is *best path bias*.

As BGP is a path vector protocol, the AS graph viewed from a single AS tends to be a tree rooted from this AS, with most cross links lost. Considering the handful number of BGP tables available, the restructured AS graph will by no means be complete. Furthermore, routing aggregation will amplify the impacts of sampling bias in that it hides the downstream links to the upstream ASes. Finally, private AS links are filtered by export policies from being advertised to other ASes. We call these factors *sampling location bias*.

Since ISPs are often unwilling to reveal their routing tables to the public because routing tables are often considered as business secretes, researchers can only rely on Routeviews-like services to obtain the BGP table dumps. However, the gathered BGP entries by a routeviews collector only comprise a partial view of the BGP routing table maintained by the peering speaker, which corresponds to the best path view(because BGP speaker only advertises its best paths to its peers). Alternative paths maintained by these peers are never advertised to the Routeviews collectors. We call this sampling bias *best path bias*.

3.2 Traceroute-based Measurement

Traceroute based active measurement relies on massive probing of the Internet and infers the AS graph by mapping the IP paths to AS paths [6, 16]. This method also suffers from sampling bias, which can further be classified into *sampling location bias* and *trace-route related bias*.

Similar to the BGP-based measurement, data collected from a single traceroute probe will also produce a tree-like graph with edges emanating from this AS. In addition, the traceroute-base measurement also faces the challenge of choosing an appropriate probing target address list that covers the globally routable address space.

Apart from the sampling location bias, traceroute-based measurement also suffers from problems related to traceroute itself. Traceroute discovers the forwarding IP path to a destination by sending IP packets with successively increasing TTL values and soliciting intermediate routers to reply with TTL expired ICMP messages. However, some networks are not ICMP friendly. For example, ICMP packets may be filtered in some networks, some routers even don't send ICMP replies and ICMP responses may be lost due to network congestion. Usually, peripheral customer networks are motivated to filter ICMP packets or ignore probe packets for security considerations, which makes the sampled traceroute results biased for the core networks.

Authorized licensed use limited to: SICHUAN UNIVERSITY. Downloaded on March 06,2022 at 02:24:07 UTC from IEEE Xplore. Restrictions apply.

Traceroute-based measurement can also subject to other kinds of problems that don't stem from sampling bias.

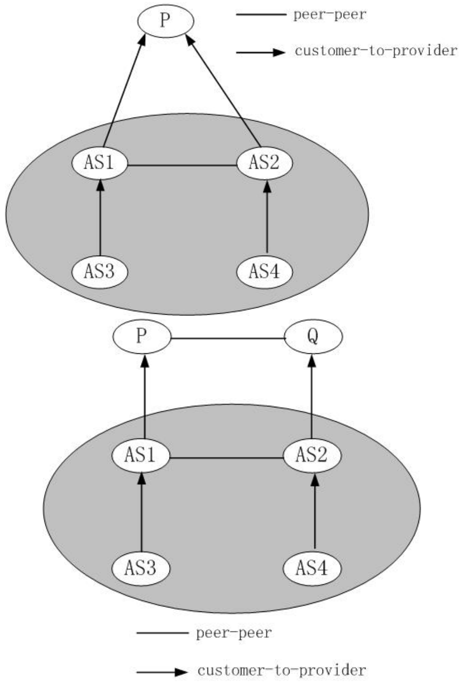


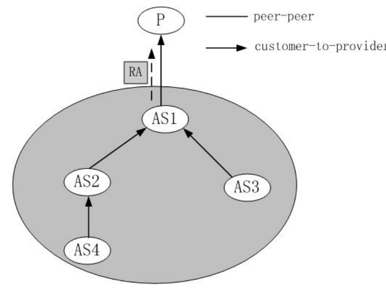
Figure 1: Loss of peer-peer links due to sampling bias

4 Sampling Bias:A Qualitative Study

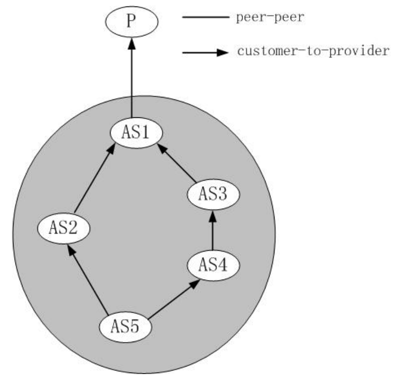
In this section, we will present various scenarios that can cause the loss of links and inaccurate AS relationship inference in BGP-based measurement. To be focused on the sampling bias, we assume that there is no sampling point in the part of network to be discovered, which is common in real Internet AS graph sampling process since most parts of the Internet are measured in a black-box style due to limited sampling points.

4.1 peer-peer links

Due to BGP routing policy constraint, a BGP speaker will never advertise routes learned from one peer AS to other peer ASes or providers, therefore, routing information sampled from an AS's provider or peer will not see the its other peer-peer links. Fig 1 shows the typical scenarios, where the shadowed area is the network to be discovered and *P* is the sampling point. In both cases, *P* will not see the peer-peer link between AS1 and AS2.



(a) routing aggregation



(b) multi-homing

Figure 2: Loss of provider-customer links due to sampling bias

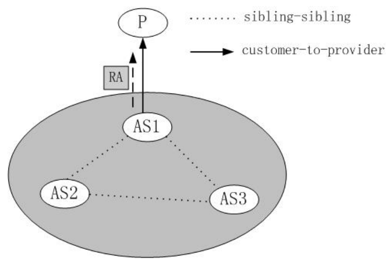
4.2 provider-customer links

The impacts of sampling bias on discovering provider-customer links have not been carefully reviewed yet. One causation is routing aggregation, just as Fig 2(a) illustrates. Routing aggregation will hide the downstream connections to the upstream ASes. In Fig 2(a), AS1 performs routing aggregation when advertising routes to its upstream ASes. Then, routing information sampled from *P* can not determine the provider-customer links AS1-AS2,AS1-AS3 and AS2-AS3. The other cause is best path sampling bias and the existence of multi-homing. As Fig 2(b) shows, AS5 is multi-homed to two providers–AS2 and AS4. Although AS1 can see all the provider-customer links in the shadowed area, path AS1-AS3-AS4-AS5 will be excluded from being advertised to *P* because only the best path AS1-AS2-AS5 will be selected for advertising. Thus, AS4-AS5 will be missed from *P*'s perspective.

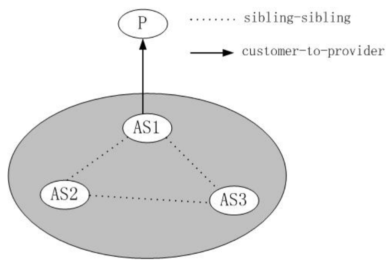
4.3 sibling-sibling links

Similar to provider-customer link loss, routing aggregation can also cause sibling-sibling link loss, just as Fig 3(a) illustrates. Another cause of sibling-sibling link loss is that no path will travel the specific sibling-sibling link. As Fig 3(b) shows, no shortest path from *P* will pass AS2-AS3,

Authorized licensed use limited to: SICHUAN UNIVERSITY. Downloaded on March 06,2022 at 02:24:07 UTC from IEEE Xplore. Restrictions apply.



(a) routing aggregation



(b) no shortest path

Figure 3: Loss of sibling-sibling links due to sampling bias

thus making it hidden from the sampled result.

4.4 AS relationship inference errors

Another negative impact of sampling bias is that limited number of sampling points can cause inaccurate AS relationship inference. Particularly, without enough routing information, sibling-sibling links can be mistakenly inferred as provider-customer links. Fig 4 gives an example. In this figure, if routing data is only collected from *P*, AS1-AS2 will be mistakenly inferred as provider-customer link. Only when routing data from AS4 is also collected then can we classify AS1-AS2 as a sibling-sibling link because we now know that AS1 and AS2 provide transit service for each other.

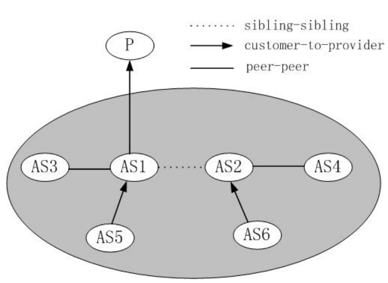


Figure 4: AS relationship inference error due to limited number of sampling points

5 Sampling Bias:A Quantitative Study

In this section, we will take a quantitative study of the impacts of sampling bias. We will use one subgraph of the Internet–the Chinese AS topology–as a case study. Two reasons drive us for the use of Chinese Internet AS graph. First, the routeviews collectors are all out of China, which gives us an opportunity to focus on sampling bias. Actually, most parts of the Internet sampled by routeviews collectors are in black-box style, similar to Chinese Internet. Second, we have the ability to deploy as many probes as possible in China to get a fairly complete Chinese AS graph. This graph can then be considered as a benchmark AS graph and used to quantitatively evaluate the impacts of sampling bias.

ISP or AS name	ASN
CHINA TELECOM	4134(7), 4812
CHINA NETCOM	4808,4814,17620
CERNET	4538(2)
CSTNET	7497
ABITCOOL	9308
USA PENN-STATE	3999
CNNIC-OET-AP	24139
NIBJNET-AP	9298

Table 1: Probing points distribution

5.1 The Chinese AS Subgraph

5.1.1 The Data Sources

Recently, we have undertaken a large-scale active measurement of the Chinese Internet. We have deployed 18 probes, probing to 8757 IP addresses sampled from the routable IP address blocks and web sites of China. Of the 18 probes, one is located in U.S, the other 17 are distributed in geographically different regions of China, covering 10 ASes. Table 1 shows the distribution of the probes(the number within parentheses are the number of probes in the given ASes). The data was collected during the period of April,21 to May,8 in the year of 2007. We applied the technique proposed in [6] to map IP traces to AS-level topology. We also collected the routeviews data from *routeviews2.oregon-ix.net*, which uses 40 peers to collect BGP tables, and the skitter AS link data available from *http://sk-aslinks.caida.org*. We collected ten days of data for both the routeviews and skitter, from 21,April to 30 April. We extracted the Chinese AS subgraph out from the global graph and merged the ten graphs into one. For convenience, we call these three graphs Trace-CN, Routeviews-CN and Skitter-CN respectively throughout the following paper.

Authorized licensed use limited to: SICHUAN UNIVERSITY. Downloaded on March 06,2022 at 02:24:07 UTC from IEEE Xplore. Restrictions apply.

5.1.2 The Representativeness of Trace-CN

We apply the following three steps to verify Trace-CN's relative completeness.

Firstly, we have analyzed the reasons for the missing nodes seen in Routeviews-CN but not discovered by our trace results. It turns out that most of the undiscovered nodes (18 out of 26) are due to the failure of traceroute to reach the target ASes. Table 2 summarizes the detailed failure types.

Failure Type	Count
Probing targets not cover this AS	3
Routing loops	2
Trace completed, previous AS not in China	1
Trace not reach the target AS	18
Trace completed, no reply for preceding hops	1
Trace completed, preceding hop private address	1
Total	26

Table 2: Detailed information about the failure types to discover the nodes in Routeviews-CN but not in Trace-CN

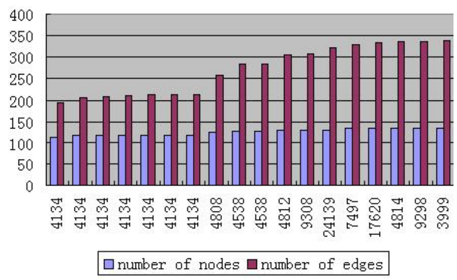


Figure 5: Graph size after merging the frist k trace results

Secondly, we accumulatively merge the AS graphs derived from each probe's trace result, in the sequence of each probe's residing AS's degree, i.e, we merge the AS graphs derived from 7 probes in AS4134 first, one by one, then the AS graph from AS4808, and etc. We measure the number of nodes and number of links after each merging procedure and plot it in Fig 5. The diagram apparently shows a diminishing return after 14 trace results are merged, after which the addition of new probes discover only a few links and contribute little to the final graph.

Finally, we apply the spanning tree covering approach introduced in [8] to simulate the possible coverage of three theorectical networks: the BA model with average degree of 6 and 8, and the PFP model [22]. Each network contains the same number of nodes as Trace-CN, i.e, 135 nodes. In [8], the possible coverage of the network from a single probe is simulated by a shortest path derived spanning tree. We give up the random selection of spanning tree roots, instead, we choose the set of spanning tree roots with their degree ranks in the same order as the AS degree ranks of our probes. In case that several probes being located in the same AS, we choose the same root node several times, but each time generate a spanning tree independently. Fig 6 presents the simulation result, which shows that with the 18 strategically chosen probes, over 95% of the edges can be unveiled.

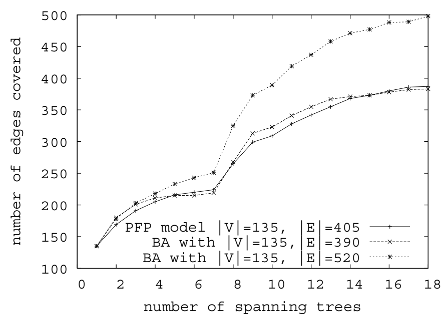


Figure 6: Edge covering with spanning trees of strategically chosen probes where V and E represent set of nodes and edges respectively.

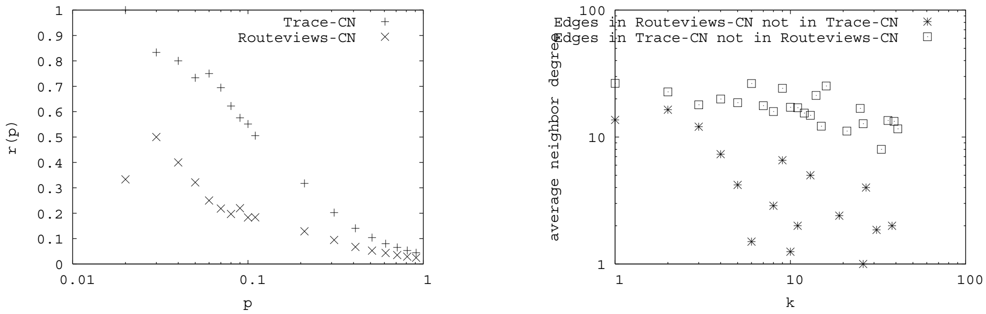
Based on the above analysis, we can make the conclusion that our measurement result can be considered as a relative complete snapshot of the real Chinese AS subgraph.

5.1.3 Elementary Graph Properties

We present the elementary graph properties in Table 3.

We have found Routeviews-CN is indeed unconnected. The relatively lower average degree, clustering coefficient and maximum coreness are all indications of its poorer connectivity. Although the maximum degrees are similar, the maximum coreness of Routeviews-CN is only half of Trace-CN. This underestimation of the Chinese AS graph's core may lead to the wrong interpretation that most of the traffic between Chinese ASes must go through their remote providers(probably in U.S) first. The rich-club phenomenon shown in Fig 7(a), is quite different. The rich-club phenomenon shown by global AS graph is well maintained in the Trace-CN, but is not observed in the Routeviews-CN. We try to explain this by measuring the missing edges' vertex degree correlation. We plot the average neighbor degree of the missing links in Fig 7(b), in which we find that the links seen in Trace-CN but not seen in Routeviews-CN show no obvious mixing pattern, i.e, no strong correlation between two endpoints' degrees of the missing links, whereas links found in Routeviews-CN but missed by Trace-CN tend to show dissortative mixing pattern, i.e, missing links often occur between rich node and poor node.

Authorized licensed use limited to: SICHUAN UNIVERSITY. Downloaded on March 06,2022 at 02:24:07 UTC from IEEE Xplore. Restrictions apply.



(a) richclub (b) Degree-correlation

Figure 7: rich club and degree correlation between endpoints of lost links for Routeviews-CN and Trace-CN

	Skitter-CN	Routeviews-CN	Trace-CN
Number of Nodes	75	161	135
Number of Edges	159	274	338
Average Degree	4.24	3.40	5.01
Maximum Degree	20	38	41
Characteristic Path Length	2.79	$+\infty$	2.75
Clustering Coefficient	0.179	0.054	0.163
Assortative Mixing	-0.311	-0.385	-0.371
Maximum Coreness	4	3	6

Table 3: Basic graph properties of the three data sources

5.2 Classifying the Missing Links

To give a quantitative impression on what kinds of links will be missed due to sampling bias, we first infer the Chinese AS relationships and then study the discrepancy between Routeviews-CN and Trace-CN in terms of different link types.

5.2.1 Inferring Chinese AS Relationships

Currently, AS relationships are primarily inferred by two category of approaches: the heuristic inference approach based solely on the gathered AS paths [9, 11, 12, 14] and the inference approach relying on the Internet Routing Registries [8, 19]. Unfortunately, neither of these methods fits our needs. To see this, we retrieved the CAIDA's publicly available AS relationship inference result from *http://www.caida.org/data/active/as-relationships/index.xml* and found that it contains only 2 sibling-sibling relationships between Chinese ASes, which is obviously not correct. We also extracted the Chinese AS registration information from the APNIC whois database, and applied the simple method proposed in [5] to check the freshness of these records, i.e, constructing a directed AS reference graph and checking the state of the edges. We found that of the 68bb4 edges, only 15 are bidirectional, which indicates that the APNIC database is quite obsolete.

We then leverage the abundant registration information for Chinese ASes and rely on our knowledge of the Chinese ISPs to help infer the AS relationships. Firstly, the ownership information of ASes is available from IRR or CERNET BGP View project [3]. Just as [9] has mentioned, this information can be used to infer sibling-sibling relationships between Chinese ASes. Secondly, the six major providers in China, namely, China Telecom, China Netcom, China Mobile, Cernet, Unicom and CSTNET are competitors for acquiring customers and their relationships(if two are connected) fall into the peer-peer category. The CRNET, GWBN and DXTNET are customers of some of the six providers and are in turn providers of some other small ASes. Finally, AS4847 is registered as an Internet exchange point, but different from other IXPs, it appears in the collected BGP table. We treat connections with AS4847 separately as IXP type.

Armed with these information, we take the attempt to infer Chinese AS relationships as follows:

1. If one AS is AS4847, infer it as an IXP edge.

2. If two ASes are owned by the same provider, infer it as a sibling-sibling edge.

Authorized licensed use limited to: SICHUAN UNIVERSITY. Downloaded on March 06,2022 at 02:24:07 UTC from IEEE Xplore. Restrictions apply.

3. If two ASes belong to two of the six different major ISPs, infer it as a peer-peer edge.

4. If one AS belongs to the six major ISPs while the other doesn't, infer it as a provider-customer edge.

5. If neither of the ASes belongs to six major ISPs, but one of them belongs to CRNET,GWBN or DXTNET, infer it as a provider-customer edge.

6. All other edges are leaved unresolved.

By applying this simple inference heuristic, of the 338 edges, we resolved 311 edges, leaving only 27 unresolved. And of the 188 inferred provider-customer edges, only one contradicts with the CAIDA's inference result(further analysis shows this contradiction is due to the inference error of the CAIDA), 114 are in agreement with the CAIDA's inference results, and the other 73 edges are the undiscovered edges of Routeviews-CN. We enhance our inference by assigning the unresolved links the relationships inferred by CAIDA's inference result, which produces another 16 provider-customer links, leaving only 11 edges unresolved. This simple validation convinces us of the feasibility and reasonable credibility of our inference result. We apply this approach to the Routeviews-CN graph too and present the detailed inference result in Table 4, with the CAIDA's inference result listed for comparison. From this table, we first see that most of the sibling-sibling AS connections are mistakenly inferred as provider-customer links in CAIDA's published result, indicating that the real Internet may contain much more sibling-sibling AS relationships(14% in Trace-CN) than previously predicated. Another fact is that the Chinese AS graph deduced from Routeviews sampling process does lose many peer-peer links, which conforms to previous studies [8, 9, 13]. The amount of peer-peer links accounts for about 10% of the total links in Trace-CN.

5.2.2 Clarifying The Discrepancies

To further differentiate between the Trace-CN and Routeview-CN, we obtain the reduced graph Routeviews-CN(V∈ Trace-CN), which is a subgraph of Routeviews-CN containing only those nodes in Trace-CN, and categorize its difference with Trace-CN, i.e, we categorize those edges found in Trace-CN, but not in Routeviews-CN(V∈ Trace-CN) and vice versa. Table 5 shows the detailed result. The peer-peer links account for about 24% of the total missing links of Routeviews-CN(V∈Trace-CN). Similar conclusions have been reported and explained in several previous works [8, 9, 13].

Probably our most unexpected finding is the significant discrepancy in the number of provider-customer links uncovered. In Routeviews-CN(V∈Trace-CN), the lost

	$A\backslash B$	$B\backslash A$
Peer-Peer	31	0
Sibling-Sibling	7	5
Provider-Customer	72	26
IXP	10	2
Unspecified	11	0
Total	131	33

Table 5: The category of differences between Trace-CN(denoted as A) and Routeviews-CN(V∈Trace-CN)(denoted as B)

provider-customer links account for about 35% of the total number of provider-customer links discovered in Trace-CN, and account for 55% of the total missing links of Routeviews-CN(V∈Trace-CN).

As we have said before, loss of provider-customer links can be ascribed to routing aggregation and multi-homing. However, if a provider-customer link loss occurs solely because of routing aggregation, the customer AS will also be unseen from the upstream ASes, which is not true in our Chinese AS subgraph since all the ASes discovered by Trace-CN are also seen in Routeviews-CN. Thus, the primary cause of provider-customer loss is multi-homing. To show the evidence of provider-customer link loss due to multi-homing, we investigate the number of nodes with degree one and two(nodes with low degrees are more likely to be customers) in both graphs. It turns out that Trace-CN contains 24% of nodes with degree one and 30% of nodes with degree two, however, for Routeviews-CN(V∈ Trace-CN), they are 30% and 35% respectively. The higher percentage of nodes with degree one and two in Routeviews-CN(V∈ Trace-CN) means in reality some customers have more providers than what Routeviews has unveiled, and therefore indicates many provider-customer links are lost due to multi-homing.

Unlike the the peer-peer and provider-customer, the uncovered number of sibling-sibling edges shows little difference.

6 Conclusion

In this paper, we undertook an in-depth study of the possible impacts of sampling bias over the Internet AS-level topology inference based on our recent effort of Chinese AS level topology measurement. We applied various techniques to validate its comparative completeness. This Chinese AS subgraph with high credibility will be helpful for Internet researchers who want to get a full understanding of the Internet structure.

We surprising found that sampling bias have significant impact on the discovery of provider-customer links. In the

Authorized licensed use limited to: SICHUAN UNIVERSITY. Downloaded on March 06,2022 at 02:24:07 UTC from IEEE Xplore. Restrictions apply.

	Trace-CN	Routeviews-CN(V∈Trace-CN)	Routeviews-CN	Inference Results from CAIDA
Peer-Peer	34	3	3	0
Sibling-Sibling	48	46	51	2
Provider-customer	204	158	182	272
IXP	41	33	38	0
Unspecified	11	0	0	0
Total Edges	338	240	274	274

Table 4: The AS relationships for Chinese AS graphs

case of Chinese AS subgraph, the lost provider-customer links from Routeviews data account for about 35% of the total provider-customer links discovered by our massive trace result. Sampling bias can also result in AS relationship inference errors for classical AS path based inference algorithms. In particular, without enough information, the sibling-sibling links can be mistakenly inferred to provider-customer links. This, on one hand, cautions the users who want to use this data for other research activities, and on the other hand, calls for reconsideration of the AS relationship inference algorithms when the number of BGP tables is small. The fact that sampling bias may handicap the proper restructuring of the underlying Internet topology calls for heightened awareness of the data set being used.

References

[1] http://www.routeviews.org.

[2] http://www.caida.org.

[3] Cernet bgp view project. http://bgpview.6test.edu.cn/bgp-view/index.shtml, 2007.

[4] A.Lakhina, J.W.Byers, M.Crovella, and P.Xie. Sampling biases in ip topology measurements. In *IEEE INFOCOM*, 2003.

[5] H. Chang, R. Govindan, and S. Jamin. Towards capturing representative as-level internet topologies. *Computer Networks Journal*, 44(6):737–755, 2004.

[6] H. Chang, S. Jamin, and W. Willinger. Inferring as-level internet topology from router-level path traces. In *Proceedings of the SPIE ITCom 2001*, 2001.

[7] R. Cohen, M. Gonen, and A. Wool. Bounding the bias of tree-like sampling in ip topologies. http://arxiv.org/abs/cs.NI/0611157, 2006.

[8] R. Cohen and D. Raz. The internet dark matter-on the missing links in the as connectivity map. In *IEEE INFOCOM*, pages 1–12, 2006.

[9] X. Dimitropoulos, D. Krioukov, M. Fomenkov, and B. Huffaker. As relationships: Inference and validation. *ACM SIGCOMM Computer Communications Review*, 37(1):29–40, 2007.

[10] M. Gaertler and M. Patrignani. Dynamic analysis of the autonomous system graph. *IPS2004,Inter-Domain Performance and Simulation*, pages 13–24, 2004.

[11] L. X. Gao. On inferring autonomous system relationships in the internet. *IEEE/ACM Transactions on Networking*, 9(6):733–745, Dec 2001.

[12] M. G.D.Battista and M.Pizzonia. Computing the types of the relationships between autonomous systems. In *IEEE INFOCOM*, 2003.

[13] Y. He, G. Siganos, M. Faloutsos, and S. Krishnamurthy. A systematic framework for unearthing the missing links: Measurement and impact. In *4th USENIX Symposium on Networked Systems Design Implementation*, 2007.

[14] L.Subramanian, S.Agarwal, J.Rexford, and R.H.Katz. Characterizing the internet hierarchy from multiple vantage points. In *IEEE INFOCOM*, 2002.

[15] P. Mahadevan, D. Krioukov, M. Fomenkov, and B. Huffaker. The internet as-level topology: Three data sources and one definitive metric. *ACM SIGCOMM Computer Communications Review*, 36:17–26, 2006.

[16] Z. M. Mao, J. Rexford, J. Wang, and R. H.Katz. Towards an accurate as-level traceroute tool. In *Proceedings of the 2003 conference on Applications, technologies, architectures, and protocols for computer communications*, pages 365 – 378. ACM Press, 2003.

[17] R.Albert, H.Jeong, and A.L.Barabasi. Attack and error tolerance of complex networks. *Nature*, 406:378–382, 2000.

[18] Y. Shavitt and E. Shir. Dimes-letting the internet measure itself. http://www.arxiv.org/abs/cs.NI/0506099, 2005.

[19] G. Siganos and M. Faloutsos. Analyzing bgp policies: Methodology and tool. In *IEEE INFOCOM*, 2004.

[20] L. Subramanian, M. Caesar, C. T. Ee, and M. Handley. Hlp: A next generation inter-domain routing protocol. In *ACM SIGCOMM*, 2005.

[21] G. Yan, T. Zhou, B. Hu, Z. Q. Fu, and B. H. Wang. Efficient routing on complex networks. *Physical Review E*, 73(046108), 2006.

[22] S. Zhou. Understanding the evolution dynamics of internet topology. *Physical Review E*, 74(016124), 2006.

[23] S. Zhou and R. J.Mondragon. The missing links in the bgp-based as connectivity maps. http://arxiv.org/abs/cs/0303028v1, 2003.

[24] S. Zhou and R. J.Mondragon. The rich-club phenomenon in the internet topology. *IEEE Communications Letters*, 8(3):180–182, Mar 2004.

Authorized licensed use limited to: SICHUAN UNIVERSITY. Downloaded on March 06,2022 at 02:24:07 UTC from IEEE Xplore. Restrictions apply.