# China's Internet: Topology Mapping and Geolocating

Ye Tian[‡], Ratan Dey[§], Yong Liu[§], Keith W. Ross[§]
[‡]Anhui Key Lab of High Performance Computing, Univ. of Science and Technology of China, Hefei, Anhui 230026, China
[§]Polytechnic Institute of New York University, Brooklyn, NY 11201, USA
Email: yetian@ustc.edu.cn, ratan@cis.poly.edu, yongliu@poly.edu, ross@poly.edu

***Abstract*—We perform a large-scale topology mapping and geolocation study for China's Internet. To overcome the limited number of Chinese PlanetLab nodes and looking glass servers, we leverage several unique features in China's Internet, including the hierarchical structure of the major ISPs and the abundance of IDCs. Using only 15 vantage points, we design a traceroute scheme that finds significantly more interfaces and links than iPlane with significantly fewer traceroute probes.**

**We then consider the problem of geolocating router interfaces and end hosts in China. We develop a heuristic for clustering the interface topology of a hierarchical ISP, and then apply the heuristic to the major Chinese ISPs. We show that the clustering heuristic can geolocate router interfaces with significantly more detail and accuracy than can the existing geoIP databases in isolation, and the resulting clusters expose the major ISPs' provincial structure. Finally, using the clustering heuristic, we propose a methodology for improving commercial geoIP databases.**

## I. INTRODUCTION

China[1] is the country with the largest number of Internet users and the second largest IP address space [1]. With its complex and unique structural features, China's Internet is very different from the Internet in US and Europe. Nevertheless, China's Internet has received relatively little attention in the measurement community to date. This is perhaps because it lacks the infrastructure and resources that are essential for large-scale Internet measurement studies, such as Rocketfuel [2] and iPlane [3]. For example, China has few PlanetLab nodes and looking glass servers. Moreover, whereas many routers outside of China have names from which geolocation can be inferred, few router interfaces have names in China.

Of particular interest is geolocation services for China's Internet. Many automatic IP address geolocation techniques based on landmarks and active delay measurement have been proposed in recent years [4]. However, Li *et al.* [5] show that the delay-distance correlation, which is a foundation for many delay based geolocation techniques, is weak in China's Internet. In addition, as we will show in this paper, existing commercial geoIP databases for Chinese IP addresses have many incomplete and erroneous entries.

In this paper, we carry out a large-scale topology mapping and geolocation study for China's Internet. To overcome the small number of Chinese PlanetLab nodes, looking glass servers, and router interfaces with geographical names, we leverage several unique features in China's Internet, including the hierarchical structure of the major ISPs and the abundance of IDCs. The contributions of this paper are as follows:

- We develop two techniques, namely *nested IP block partitioning* and *collaborative tracerouting*, which allow us to perform a comprehensive and efficient traceroute measurement study of China's Internet using only 15 internal vantage points. In particular, our approach discovers significantly more interfaces and links than iPlane with significantly fewer traceroute probes.
- We develop a heuristic for clustering the interface topology of a hierarchical ISP, so that each cluster is a connected component within a city. We show that this clustering heuristic can geolocate interfaces with significantly more detailed location information than the existing geoIP databases in isolation. We also analyze the clusters generated by our clustering heuristic, and show that they expose several characteristics of the Chinese Internet, including recent mergers of ISPs, and the ISPs' networks centered around provincial capital cities.
- Using the geo-clustering heuristic, we propose a methodology for improving commercial geoIP databases. By evaluating with datacenter landmarks, we show that our approach is able to provide more detailed and accurate location information as compared with the original geoIP database.

## II. TRACEROUTE MEASUREMENT

When attempting to map China's Internet with traceroute, we face two challenges. The first is to identify a set of target IP addresses that is sufficiently, but not overly, dense within the Chinese Internet from public BGP snapshots (e.g., from Oregon Routeviews [6] and RIPE RIS [7]). The other challenge is efficiency. In our measurement we only use 15 stable vantage points located in China (7 PlanetLab nodes and 8 web-based traceroute servers). Our objective is to devise a traceroute strategy that sufficiently covers the Chinese Internet without overly burdening these vantage points. To address these two challenges, we devise two techniques, namely, *nested IP block partitioning* and *collaborative tracerouting*.

### A. Nested IP Block Partitioning

When partitioning the IP space and derive traceroute targets from BGP snapshots, we find that *block nesting* [8], where a block from one BGP routing table entry resides in another block from a different entry, is very common; moreover, there are often several levels of nesting. An example of nested IP

---

This work was supported by the Fundamental Research Funds for the Central Universities (No. WK0110000007 and WK0110000024), and the Overseas Academic Training Funds (OATF) of USTC.

[1] By China we mean Mainland China.

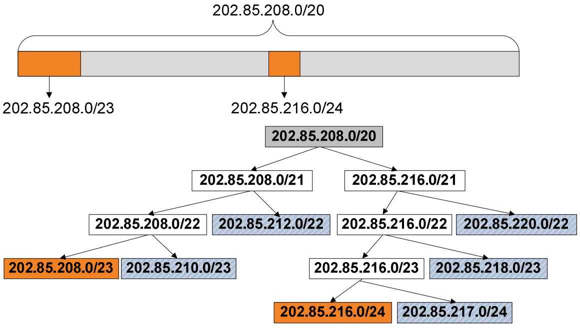

Fig. 1. Nested-block Partitioning

blocks is shown in the top graph in Fig. 1. In the graph, three blocks are obtained from BGP tables, i.e., 202.85.208.0/20, 202.85.208.0/23, and 202.85.216.0/24, where the latter two blocks are nested in the first one. Clearly, the smaller nested blocks suggest the existence of different subnets.

We design a tree-based method to partition the Chinese IP address space with a minimal number of blocks while preserving the nested blocks obtained from the BGP tables. The blocks from the BGP tables are nodes in trees. We consider a block encompassing other blocks as the root of a binary tree, and all the nested blocks as leaves. With this tree the problem becomes: given the root node and a number of leaf nodes, construct a binary tree with the fewest leaves. After the tree is obtained, we use all the blocks corresponding to the leaf nodes (including the original nested blocks) to replace the root. For example, in the case mentioned above, the corresponding binary tree is shown in the bottom graph in Fig. 1.

After the partitioning, we further evenly divide any blocks that are larger than our granularity, while reserving the smaller blocks for traceroute probing. For the example in Fig. 1, 7 blocks are probed instead of the 4 or 16 blocks that would be generated by evenly dividing 202.85.208.0/20 into /22 or /24 blocks. Thus, with nested-block partitioning, we can fully exploit the small nested blocks, suggesting different subnets, without naively dividing all the large blocks, which would geometrically increase the probing workload.

### B. Collaborative Tracerouting

To efficiently probe the large number of targets for China's Internet, we propose a mechanism for having the vantage points collaboratively and dynamically determine their traceroute targets, thereby avoiding redundant probes, which recently have been widely observed in Ark and iPlane [9].

In our measurement, the IP blocks obtained in Section II-A (which partition the Chinese IP space) are the basic probe units. For each block, we always use its second IP address (i.e., a.b.c.1) as the traceroute target. In the collaborative tracerouting scheme, a vantage point actively uses the results of its previous probes and other vantage points' probes to avoid redundant probes. Specifically, each vantage point keeps a set, $reach\_set$, of all the addresses the vantage point has observed during its previous probes; and each IP block keeps a set, $source\_set$, containing all the IP addresses that lead to this block from previous traceroutes from all the vantage points. When a vantage point $v$ encounters an IP block $B$ it has not probed before, it examines $v$'s $reach\_set$ and $B$'s $source\_set$; if the two sets overlap, then an interface path can be found from $v$ to the block $B$ from previous traceroutes, so the vantage point $v$ doesn't probe the block $B$.

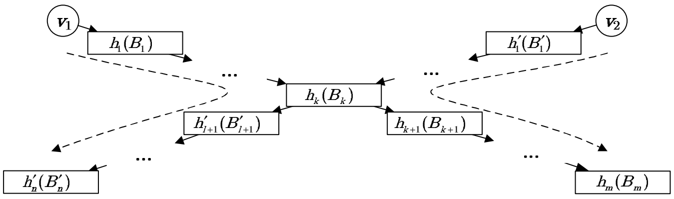

Fig. 2. An example of collaborative tracerouting

As an example, suppose a vantage point $v_1$ probes a target with the traceroute path $h_1, ..., h_k, h_{k+1}, ..., h_m$, where the interfaces are in the blocks $B_1$, ..., $B_k$, $B_{k+1}$, ..., $B_m$, as shown in Fig. 2. After this probing, $v_1$ can skip $B_1$, ..., $B_{m-1}$ in future measurements, as $v_1$'s $reach\_set$ overlaps with the $source\_set$ for each of these blocks. Moreover, suppose another vantage point $v_2$ has a traceroute path $h'_1, ..., h_k, h'_{l+1}, ..., h'_n$, that traverses the blocks of $B'_1$, ..., $B_k$, $B'_{l+1}$, ..., $B'_n$. As a result of this probe, $v_1$ can skip blocks of $B'_{l+1}$, ..., $B'_n$, as an interface path has already been found from $v_1$ to them via $h_k$, similarly, $v_2$ can also skip the blocks of $B_{k+1}$, ..., $B_m$, shown as the dotted lines in the left and the right of the figure respectively.

### C. Measurement Results

TABLE I
TRACEROUTE MEASUREMENT RESULTS

| | iPlane (1 day) | iPlane (2 days) | cTrace | In both |
|---|---|---|---|---|
| Traceroutes | 1,244,667 | 2,381,482 | 106,580 | |
| Interfaces | 17,308 | 17,761 | 71,047 | 10,023 |
| Links | 76,120 | 82,791 | 146,542 | 27,735 |

Using nested-block partitioning and collaborative tracerouting, we perform a traceroute measurement on China's Internet with 15 vantage points (from 9 different cities and in 4 ISPs) in China. We applied the nested-block partitioning algorithm on the IP blocks from 8 BGP snapshots from Route Views and RIPE RIS and further divided them to prefix /22 blocks for obtaining the target addresses. The measurement was performed from Dec. 12, 2010 to Jan. 2, 2011. We also downloaded iPlane's traceroute data on Dec. 19 and Dec. 20, 2010 for comparison. For each path in iPlane, we extract the segment that is within China's Internet.

Table I compares the iPlane data with our measurement results (referred to as *cTrace*). We present the results for both one and two days of iPlane's measurement. As compared with iPlane, our approach employs only 5% of the number of traceroute probes but finds four times as many interfaces and twice as many interface links. This experiment therefore shows that using vantage points in China is much more efficient in exposing China's Internet, and collaborative tracerouting can effectively eliminate redundant probes.

In the rest part of this paper, we combine cTrace with the 2-day iPlane data, and use the combined data for further study.

## III. GEOLOCATING THE INTERFACE TOPOLOGY

With the combined traceroute data obtained in Section II, we have obtained a separate interface topology for major Chinese ISPs of China Telecom (a.k.a. ChinaNet and henceforth referred to as Telecom) and China Unicom (henceforth referred to as Unicom). In this section, we seek to geolocate the interfaces in both interface topologies.

For a given interface topology $T$, we say a set of interfaces $S$ forms a *cluster* if (a) all the interfaces in $S$ belong to the same city, and (b) the subgraph of $T$ induced by $S$ is weakly connected. We further say that a cluster $S$ is a *maximal cluster* if it is not possible to create a larger cluster by adding more interfaces to it. Our goal is to determine the maximal clusters in the interface topology.

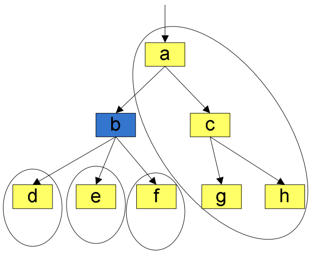

Fig. 3. Erroneous clusters example

A naive method to create the clusters is to directly use the city information provided by the geoIP databases on face value. However, by examining three best Chinese geoIP databases of IP138[10], QQWry[11] and IPcn[12], we find that they are only moderately accurate for end host geolocating, and substantially less accurate for router interfaces [13]. Due to missing and erroneous entries in the geoIP databases, the naive approach leads to a large number of small and disconnected erroneous clusters.

Fig. 3 provides an example. All the interfaces (in boxes) on the graph are at the same location, and should be included in one cluster. However, if interface $b$'s location from geoIP database is wrong or missing, four instead of one cluster is formed. On the other hand, note that all the interfaces adjacent to $b$ are at the same location, we can conclude that $b$ is likely located at the same location as all the other interfaces on the graph. Inspired by this observation, we propose a heuristic by combining the information in the geoIP databases with the topological information, for accurately determining the maximal clusters in interface topologies.

### A. Geo-Clustering Heuristic

We have developed a heuristic that could be used for any ISP with a hierarchical structure (not just Chinese ISPs). Due to space constraints, we only provide a summary of the heuristic here; for further details, please see [13].

For each of these ISPs, using the traceroute data, we first obtain an interface topology that expands from the ISP's backbone network to the traceroute targets in that ISP. For each of the resulting interface topologies, and for each of the databases, we infer the interfaces' city-level locations and cluster them through four steps. We refer to this four-step heuristic as the *geo-clustering* heuristic.

In Step 1, we select the interfaces that are at the edge of the interface topology, and form singleton clusters for each of them. For each singleton cluster formed at this step, we use the interface's location in the database (referred to as the interface's DB location) as the cluster's location.

Step 2 consists of a sequence of rounds. At the beginning of each round, we select the unclustered interfaces that are one step closer to the backbone network as candidates for clustering. For each candidate, all it's out-linked interfaces use their DB or cluster locations to *vote* to decide the candidate's cluster location. After the candidate is assigned a cluster location, it merges all the clusters it links to that have the same location to form a new large cluster. After all the candidate interfaces are processed, the round is finished. We then continue with the next round by selecting new candidates. However, for a candidate interface, if more than one province appears in the voting, we abort the voting-based inference without forming or merging any clusters, and move on to the next candidate. The Step 2 heuristic stops when we can't form or merge any clusters during a round.

Step 3 in the heuristic works similarly as Step 2 by first selecting a set of candidate interfaces, inferring their cluster locations, and merging the clusters with the same cluster location. However, unlike Step 2, in Step 3 nearly all the candidate interfaces are on backbone routers, which usually connect to many routers at different locations. Here we apply four different rules to infer an interface's cluster location by *combining the link delay* with the voting based approach [13].

After applying Steps 2 and 3, all the interfaces in the topology are clustered. Careful examination on the resulting clusters shows that for nearly all the cities, there are one or two large clusters containing most of the interfaces, as well as a number of singleton or small clusters. We categorize the clusters as mergeable *small clusters* and *large clusters* according to their sizes. For a small cluster, if it is only connected to one large cluster, then the location information given in the database for the small cluster is likely to be wrong; we therefore merge it into the large cluster, regardless of its original cluster location.

### B. Geo-Clusters

We applied the geo-clustering heuristic on the Telecom and Unicom's interface topologies using the three Chinese geoIP databases. By geo-clustering, we can group most of the interfaces on the interface topology into clusters with detailed city-level location information. We refer to a cluster with a city-level location as a *geo-cluster*. For example, for Telecom's interface topology using the geoIP database of IP138, after four steps, 532 of the final geo-clusters containing 98.2% of

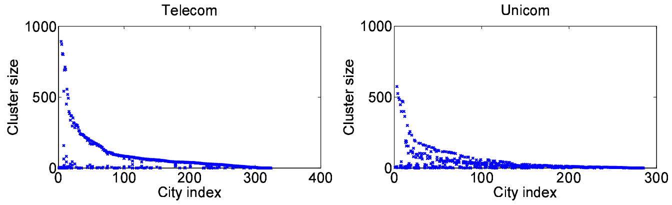

Fig. 4. Distribution of the sizes of geo-clusters across cities in major ISPs

the total interfaces have been formed. (The remaining clusters are singleton clusters for which the heuristic did not assign to a city since there was no clear majority winner in the voting.) Similar results were observed for Unicom and for using the two other geoIP databases. We omit them due to lack of space.

By examining the 532 geo-clusters obtained on Telecom's interface topology, we find they are located in 324 different cities, which are nearly all the cities in China. We show the sizes of the geo-clusters for each city for Telecom and Unicom in Fig. 4, where each point on the figure corresponds to a geo-cluster. For each ISP, the cities are indexed according to the total number of IP addresses across all geo-clusters in the city. From Telecom's figure, we can see that for many cities, there is only one geo-cluster. For a small fraction of the cities, multiple clusters are found, with one cluster containing the majority of the interfaces. There are two possible reasons for multiple clusters in a city: ($i$) the ISP has multiple networks serving different purposes in that city; and more likely ($ii$) some of the singleton and small clusters cannot be merged into large clusters in Step 4. Note that the Unicom's geo-cluster distribution is distinctly different from Telecom. In particular, for Unicom in many cities there are two large geo-clusters of comparable size. Our heuristic is consistent with the fact that in 2008 Unicom merged with China Netcom (a.k.a. CNCGroup), which used to be the second largest ISP in China. As a result, in many cities we can observe one large geo-cluster for the former Unicom network, and another large geo-cluster for the former Netcom network.

## C. *The Hierarchical Structure*

TABLE II
STATISTICS OF INTER-CLUSTER LINKS

| | Same province | | Different province | | |
|---|---|---|---|---|---|
| | Cap-Other | Other-Other | Cap-Cap | Cap-Other | Other-Other |
| Telecom | 3,236 | 2,097 | 169 | 283 | 42 |
| Unicom | 1,504 | 1,281 | 199 | 25 | 0 |

We now study the internal structure of each ISP. Table II categorizes inter-cluster links based on the locations of the two endpoints of the links. In this table we have removed the links with both endpoints on the backbone, and use "Cap" for provincial capital cities, and "Other" for non-capital cities. From the table we can see that there are many intra-province links, and more than half of them are between capital and non-capital cities. There are relatively few inter-province links, and the majority of them connect to at least one capital city. *We can therefore conclude that the major Chinese ISPs are highly hierarchical following China's provincial organization, and that the provincial capital cities are not only government centers but also serving as hubs in the ISPs' networks. This strikingly contrasts with flattening trends in the international Internet [14].*

## D. *Locating Interfaces with Null Replies*

TABLE III
NULL REPLY RATIOS

| | DB province | Cluster province | DB city | Cluster city |
|---|---|---|---|---|
| IP138 | 7.7% | 0.99% | 21.7% | 1.51% |
| QQWry | 6.2% | 1.00% | 18.7% | 1.64% |
| IPcn | 8.0% | 0.93% | 26.4% | 1.66% |

After geo-clustering, each interface in an ISP's interface topology has two locations: the geoIP database location and its cluster location (with the clusters derived from the same database). In this section, we show that the cluster locations are significantly more complete and accurate.

We first examine the completeness by comparing the null reply ratios[2]. In this comparison, all the IP addresses of the interfaces on Telecom and Unicom's interface topologies are included. Table III shows the null reply ratios at the province and the city levels for both DB and cluster locations. Observe that the ratios for cluster locations are much smaller than those for the DB locations. The geoIP services give a high-level of null replies because many router addresses do not have city-level or province-level locations in the database. However, the cluster locations for many of these router interfaces have been inferred at the city level.

TABLE IV
NUMBER OF THE INTERFACES THAT HAVE CONSISTENT LOCATIONS

| | Telecom | Unicom |
|---|---|---|
| 3DB identical | 25,625 (67.1%) | 15,794 (63.7%) |
| 3Cluster identical | 35,376 (92.7%) | 21,938 (88.5%) |

We now show that clustering approach is substantially more accurate than the geoIP databases for interfaces using cross validation. For an interface, if the locations from the three databases are the same, we say that the location is likely to be correct; if, however, all three databases do not give the same location, then we have a low level of confidence on the location information. Similarly, using the three sets of geo-clusters based on the three different geoIP databases, we can cross-validate the cluster locations. Table IV shows for each of the two ISPs, the number of the addresses that have consistent locations for the two approaches. We see that the three geoIP databases agree only for 65.8% of the interfaces (average across the two ISPs), but after applying the geo-clustering heuristic, as many as 91.1% interfaces have the same cluster locations.

---

[2] A database's null-reply ratio is defined as the fraction of the cases for which the database fails to provide location information [15].

## IV. IMPROVING GEOLOCATION SERVICES

In this section, we develop a methodology for accurately geolocating arbitrary Chinese IP addresses. Our goal here is to provide a significant improvement over the existing Chinese geoIP databases.

### A. Geolocating an Arbitrary IP Address

For a given IP address $p$ that we wish to geolocate, we first determine the ISP to which it belongs (e.g., by first determining the AS to which it belongs from BGP tables). This ISP has an interface topology, say $T$, which we obtained from our traceroute data.

To apply the geolocating algorithm in Section III to an arbitrary IP address $p$, we first augment $T$ to reach $p$ by conducting additional traceroute probes. We choose a subset of existing vantage points, and probe the target $p$, as well as any unprobed IP addresses between $T$ and $p$ that are not separated from $T$ by anonymous routers, from each of these vantage points. With the new traceroutes, we then augment the topology $T$ to create a new interface topology $T'$. Applying the geo-clustering heuristic to the new augmented topology $T'$, we obtain a new set of geo-clusters. The location of $p$ is then determined from these new geo-clusters using one of the following three cases:

- Case 1: $p$ is in the topology $T'$ and therefore is in one of the geo-clusters. In this case, we simply set $p$'s location to the location of the cluster that encompasses it.
- Case 2: $p$ can be reached by at least one traceroute path, but $p$ is not in $T'$ due to anonymous routers. In this case, we find the geo-cluster that is closest to $p$, which we refer to as the *last-hop geo-cluster*. If the distance between the last-hop geo-cluster and $p$ is no larger than a threshold (2 hops in our evaluation), we set $p$'s location to the location of the last-hop geo-cluster.
- Case 3: If we don't set $p$'s location in Case 1 and 2, the location from the geoIP database is used.

### B. Evaluation

We use a number of landmarks as the ground truth for evaluating the accuracies of the geoIP databases and of our methodology. For collecting landmarks, we leverage the numerous Internet datacenters (IDC) located in many cities in China. We skip our methodology of collecting landmarks here for space reason, interested readers can refer to our technical report [13]. We have successfully collected 305 landmarks – 199 on Telecom and 106 on Unicom – with their ground-truth locations detailed to the city level.

TABLE V
EVALUATION USING TELECOM & UNICOM LANDMARKS

| | | | Case 1 | Case 2 | Case 3 | Total |
|---|---|---|---|---|---|---|
| Telecom | IPcn | DB | 102/117 | 11/15 | 57/67 | 170/199 |
| | | Cluster | 111/117 | 14/15 | 57/67 | 182/199 |
| Unicom | IPcn | DB | 44/55 | 8/10 | 28/41 | 80/106 |
| | | Cluster | 52/55 | 9/10 | 28/41 | 89/106 |

We use ten vantage points located in seven different cities to geolocate the 305 landmarks. For each landmark, we compare the location determined by our geo-clustering methodology and the location from the corresponding geoIP database with the landmark's ground-truth location. The number of the landmarks that are accurately located by the geoIP database of IPcn and by geo-clusters (based on IPcn) are shown in Table V.

From Table V, we see that for both ISPs, our geo-clustering methodology can accurately geolocate more landmarks than can the geoIP databases. For the landmarks in Case 1 and Case 2, we are able to accurately geolocate over 9% more Telecom landmarks and over 13% more Unicom landmarks on average. In addition, more than 60% of the landmarks under evaluation fall into Case 1 and Case 2. For the databases of IP138 and QQWry, similar observation could be made, suggesting that our methodology can improve the geolocation services for many IP addresses in the Chinese Internet.

## V. CONCLUSION

In this paper, we carried out a large-scale topology mapping and geolocation study for China's Internet. With techniques of nested-block partitioning and collaborative tracerouting, we comprehensively and efficiently probe China's Internet from a small number of vantage points inside China. By further exploiting the hierarchical structure of China's Internet, we proposed a geo-clustering heuristic that clusters interfaces within the same city. Finally, we demonstrate that the geo-clustering heuristic can be used to improve the accuracy of commercial geoIP databases for geolocating arbitrary IP addresses.

## REFERENCES

[1] China Internet Network Information Center, "Statistical report on Internet development in China," 2011.

[2] N. Spring, R. Mahajan, and D. Wetherall, "Measuring ISP topologies with rocketfuel," in *Proc. of SIGCOMM'02*, Pittsburgh, PA, 2002.

[3] "iPlane Project," http://iplane.cs.washington.edu/.

[4] Y. Wang, D. Burgener, M. Flores, A. Kuzmanovic, and C. Huang, "Towards street-level client-independent IP geolocation," in *Proc. of NSDI'11*, Boston, MA, 2011.

[5] D. Li, J. Chen, C. Guo, Y. Liu, J. Zhang, Z. Zhang, and Y. Zhang, "IP-geolocation mapping for involving moderately-connected Internet regions," Microsoft, Tech. Rep., 2009.

[6] "Oregon Routeviews," http://www.routeviews.org/.

[7] "RIPE RIS," http://www.ripe.net/data-tools/stats/ris.

[8] Y. Zhu, J. Rexford, S. Sen, and A. Shaikh, "Impact of prefix-match changes on IP reachability," in *Proc. of IMC'09*, Chicago, IL, 2009.

[9] R. Beverly, A. Berger, and G. G. Xie, "Primitives for active internet topology mapping: Toward high-frequency characterization," in *Proc. of IMC'10*, Melbourne, Australia, 2010.

[10] "IP138," http://www.ip138.com/.

[11] "QQWry," http://www.cz88.net/.

[12] "IPcn," http://www.ip.cn/.

[13] Y. Tian, R. Dey, Y. Liu, and K. W. Ross, "China's Internet: Topology mapping and geolocating," University of Sci. & Tech. of China, Tech. Rep., 2011, http://staff.ustc.edu.cn/~yetian/pub/ChinaInternet.pdf.

[14] C. Labovitz, S. Iekel-Johnson, D. McPherson, J. Oberheide, and F. Jahanian, "Internet inter-domain traffic," in *Proc. of SIGCOMM'10*, New Delhi, India, 2010.

[15] Y. Shavitt and N. Zilberman, "A study of geolocation databases," 2010, preprint, arXiv:1005.5674v3 [cs.NI].