Analytical Alias Resolution

Mehmet Hadi Gunes and Kamil Sarac

UTD

ICC 2006 Istanbul
Internet Mapping

- Internet measurement studies require availability of a representative Internet map
- Understanding the topological and functional characteristics of Internet is essential to verifying, correcting, and improving desirable properties, i.e.
  - Robustness
  - Reliability
  - Efficiency
  - Security
- An accurate Internet map can reveal the topology and help in designing new services or applications
There are several projects that collect large scale Internet maps. These projects utilize traceroute like tools for topology collection. From a source node (e.g. A), traceroute probes all nodes on a path towards a given destination (e.g. F).

Traceroute uses TTL-scoped probe packets to obtain ICMP error messages from the routers on the path. ICMP error message includes the IP address of one of the interfaces (incoming, outgoing, fixed) of intermediate routers as its source. Merging end-to-end path traces yields the network map, but we have some issues.
Each interface of router has a unique IP address.
A router may respond with different IP addresses to different queries.

**Alias Resolution** is the process of grouping the interface IP addresses of each router into a single node.

Inaccuracies in alias resolution may result in a network map that
- includes artificial links/nodes
- misses existing links/nodes
Alias Resolution: Existing Approaches

- **Address Based Method** (Mercator and iffinder) [J.Pansiot98]
  - Send probe packets to different IP addresses (i.e. IP_A and IP_B)
  - If replies from have the same source IP address, then classify (IP_A and IP_B) as Alias.
  - (—) Relies on a particular implementation of ICMP error generation.

- **IP Identification Based Method** (ally) [N.Spring04]
  - Send probe packets to two IP addresses (i.e. IP_A and IP_B), and a third probe to the first responding one.
  - If IP identifier values in responses are nearby and monotonic, then classify (IP_A and IP_B) as Alias.
  - (—) Few of the router respond to direct probing packages.

- **DNS Based Method** [N.Spring04]
  - Relies on similarities in the host name structures.
  - e.g. sl-bb21-lon-14-0.sprintlink.net & sl-bb21-lon-8-0.sprintlink.net.
  - (—) Works when a systematic naming is used.
Analytical Alias Resolution: Approach

- We propose an analytical approach to the Alias Resolution problem.
- If routers A and B connected by a point-to-point link then they are within a /30 or /31 subnet.
- Identifying point-to-point links between routers, we can identify IP alias pairs by proper alignment of reverse traces.
- The Analytical Alias Resolution
  - focuses on the structural connections between routers,
  - benefits from IP address assignment of point-to-point links,
  - does not require explicit probing,
  - can be used along with probing based techniques.
Some Observations:

IP address Assignment Practices

- The IP address assignment mechanism adheres to the guidelines presented in the Internet Registry IP Allocation Guidelines (RFC-2050).
- IP addresses of a domain or a network are divided into subnet ranges.
- Each subnet has a network address, and devices get an IP address from the range.
- For example, subnet 192.168.0.240 / 28
  - can identify at most 14 devices,
  - 192.168.0.240 is network address,
  - 192.168.0.255 is broadcast address.
- Smallest subnets in practice are /30 and /31 subnets for point-to-point links.
Some Observations: Point-to-point Links

- For point-to-point links between two router interfaces, a /30 subnet or /31 subnet can be defined.

\[
\begin{align*}
A & : \text{IP} = X.xxxxxx01, \\
B & : \text{IP} = X.xxxxxx10
\end{align*}
\]

- The interface IP addresses on the link are consecutive and are within /30 subnet or /31 subnet (i.e. \( \text{IP}_A \leftrightarrow \text{IP}_B \)).

- Identifying the relation \( \leftrightarrow \) between two IP addresses within reverse traces, we can realize that IP addresses belong to interfaces connected by a point-to-point link.

- This helps identifying IP aliases by proper alignment of reverse traces.
Analytical Alias Resolution: Using Point-to-point Links

For example, if trace

- from A-to-F is \((\text{IP}_{B_1}, \text{IP}_{C_1}, \text{IP}_{D_1}, \text{IP}_{E_1}, \text{IP}_F)\)
- from F-to-A is \((\text{IP}_{E_4}, \text{IP}_{D_4}, \text{IP}_{C_4}, \text{IP}_{B_4}, \text{IP}_A)\)

then observing \(\text{IP}_{B_4} \leftrightarrow \text{IP}_{C_1}, \text{IP}_{C_4} \leftrightarrow \text{IP}_{D_1}, \text{IP}_{D_4} \leftrightarrow \text{IP}_{E_1}\), we can align both traces as follows:

Finally, we can identify IP alias pairs

\((\text{IP}_{B_1}, \text{IP}_{B_4}), (\text{IP}_{C_1}, \text{IP}_{C_4}), (\text{IP}_{D_1}, \text{IP}_{D_4}), \text{and} (\text{IP}_{E_1}, \text{IP}_{E_4}).\)
Analytical Alias Resolution: Point-to-point Links in a Traceroute Pair

<table>
<thead>
<tr>
<th></th>
<th>SMU to Yale</th>
<th>Yale to SMU (reversed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>129.119.39.1</td>
<td>129.119.223.249</td>
</tr>
<tr>
<td>2</td>
<td>129.119.0.249</td>
<td>129.119.0.250</td>
</tr>
<tr>
<td>3</td>
<td>206.223.141.89</td>
<td>206.223.141.90</td>
</tr>
<tr>
<td>4</td>
<td>206.223.141.70</td>
<td>206.223.141.69</td>
</tr>
<tr>
<td>5</td>
<td>198.32.8.34</td>
<td>198.32.8.33</td>
</tr>
<tr>
<td>6</td>
<td>198.32.8.66</td>
<td>198.32.8.65</td>
</tr>
<tr>
<td>7</td>
<td>198.32.8.84</td>
<td>198.32.8.85</td>
</tr>
<tr>
<td>8</td>
<td>192.5.89.9</td>
<td>192.5.89.10</td>
</tr>
<tr>
<td>9</td>
<td>192.5.89.33</td>
<td>192.5.89.34</td>
</tr>
<tr>
<td>10</td>
<td>192.5.89.70</td>
<td>192.5.89.69</td>
</tr>
<tr>
<td>11</td>
<td>130.132.1.19</td>
<td>130.132.1.100</td>
</tr>
<tr>
<td>12</td>
<td>130.132.252.244</td>
<td>130.132.23.1</td>
</tr>
</tbody>
</table>

- We observe correlations between IP addresses in the 2\textsuperscript{nd} row till 10\textsuperscript{th} row.
- The arrangement in figure can be used to detect IP aliases
  - 129.119.0.249 and 206.223.141.90 are aliases,
  - 206.223.141.89 and 206.223.141.69 are aliases,
  - and so on
Evaluations:
Comparison to Existing Approaches

- We used AMP traceroute data between 27 vantage points with total of 351 path pairs and 503 unique IPs.
- Table shows results of AAR and other approaches’ decision for alias pairs identified by AAR.
- Figure compares aliases identified by AAR and ally, obtained by probing all possible 126,253 pairs.

<table>
<thead>
<tr>
<th>Method</th>
<th>Alias</th>
<th>Not Alias</th>
<th>Unknown</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAR</td>
<td>180</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>Source-Address (SA)</td>
<td>42</td>
<td>40</td>
<td>98</td>
</tr>
<tr>
<td>IP Identification (IPI)</td>
<td>72</td>
<td>6</td>
<td>102</td>
</tr>
<tr>
<td>Ally (uses SA&amp;IPI)</td>
<td>77</td>
<td>5</td>
<td>98</td>
</tr>
</tbody>
</table>
Evaluations: Verification of AAR

- We verified AAR for the Abilene backbone network.
- AAR had no false positives.
- Ally could not find any alias pairs because the Abilene backbone routers did not respond to probes.

<table>
<thead>
<tr>
<th>Node Location</th>
<th>From the Map</th>
<th>From AAR</th>
<th>From Ally</th>
</tr>
</thead>
<tbody>
<tr>
<td>Atlanta</td>
<td>3</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Chicago</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Denver</td>
<td>3</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>Houston</td>
<td>3</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Indianapolis</td>
<td>3</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>Los Angeles</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Kansas City</td>
<td>3</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>New York</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Seattle</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Sunnyvale</td>
<td>3</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>Washington</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>23</td>
<td>18</td>
<td>0</td>
</tr>
</tbody>
</table>
Conclusion & Future work

- Alias resolution is an important step of Internet measurement studies.
- AAR is a passive approach, i.e. does not introduce additional measurement traffic to network, to resolve IP aliases in a given traceroute data set.
- AAR does not depend on explicit probing of IP addresses where most of the routers do not respond.

- We are working to improve AAR by considering multi-access links as well.
- We are also studying integration of AAR with probing based methods.