Patch me if you can:
Analyzing the potential of data sources to reveal the Internet’s AS topology
Lars Prehn and Anja Feldmann
\{lprehn, anja\}@mpi-inf.mpg.de

Motivation
Various disciplines require accurate AS topology information, e.g. Resiliency Assessment, Replica Placement, (Protocol-) Simulation, or Peer/Provider Selection. AS Graph known to be incomplete.

Questions
1. Contribution of different data sources?
2. Necessary to use all of them?
3. How incomplete is a given source?
4. (with regards to the retrieved topology)
5. How much adds a Tier-1’s perspective?

Approach
We first gather topology information from AS Paths. Afterward, we apply well-known methodologies to various other data sources in order to find additional links. We compare the retrieved topology with a Tier-1’s perspective.

Cleaning Process
Squeeze Prepending
Remove AS Sets with multiple entries E.g. A-B-(C-D)
Extend single entry AS Sets A-B-(C) becomes A-B-C
Remove Paths with Cycles E.g. A-B-C-B
Remove Paths with Bogon Prefix or ASN Using CDR report’s Bogon lists
Split paths at ASN 23456 E.g. A-B-23456-C-D becomes A-B and C-D

Community Values

MLP Inferences

IP-to-AS Mapping

IRR Link inference

Publicly available topology
- = 527K (100.0%) Links
= 475K (90.1%) Links
= 381K (72.3%) Links
= 578/622 (92.9%) Nbhs
= 534/622 (85.9%) Nbhs

Table 1: Comparison of the number of undirected links in data sets for the 16th. May 2018

<table>
<thead>
<tr>
<th>[A – B]</th>
<th>0</th>
<th>MRT</th>
<th>MLP</th>
<th>IP Paths</th>
<th>IRR</th>
<th>LGs</th>
<th>Tier-1</th>
<th>Others</th>
</tr>
</thead>
<tbody>
<tr>
<td>MRT</td>
<td>380704</td>
<td>0</td>
<td>356334</td>
<td>330289</td>
<td>378795</td>
<td>378535</td>
<td>313723</td>
<td>280339</td>
</tr>
<tr>
<td>MLP</td>
<td>119104</td>
<td>949734</td>
<td>0</td>
<td>117916</td>
<td>119104</td>
<td>118818</td>
<td>118825</td>
<td>91373</td>
</tr>
<tr>
<td>IP Paths</td>
<td>816934</td>
<td>45686</td>
<td>79124</td>
<td>0</td>
<td>79233</td>
<td>79834</td>
<td>59388</td>
<td>44627</td>
</tr>
<tr>
<td>MLP</td>
<td>4653</td>
<td>2744</td>
<td>4653</td>
<td>3574</td>
<td>4653</td>
<td>3574</td>
<td>4653</td>
<td>2449</td>
</tr>
<tr>
<td>Tier-1</td>
<td>527</td>
<td>3102</td>
<td>4985</td>
<td>4739</td>
<td>5262</td>
<td>4820</td>
<td>2871</td>
<td>937</td>
</tr>
<tr>
<td>Total</td>
<td>68094</td>
<td>1113</td>
<td>67815</td>
<td>47710</td>
<td>66535</td>
<td>67643</td>
<td>0</td>
<td>937</td>
</tr>
</tbody>
</table>


